

Aix-Marseille Université

UFR Arts, Lettres, Langues, Sciences Humaines

Mémoire de Master 2 : Théories Linguistiques :

Terrain et Expérimentation (ThéLiTE<sub>x</sub>)

Parcours : Langues en Contact et Typologie (LCT)



**La variation en phonologie : le cas des emprunts abrégés  
composés en japonais**

**Mémoire présenté par : Rémi LAMARQUE**

**Sous la direction de James GERMAN**

**Année universitaire 2015-2016**

# La variation en phonologie : le cas des emprunts

## abrégés composés en japonais

### Table des matières

Remerciements :.....	4
INTRODUCTION :.....	5
1 - Travaux antérieurs sur les DAC et les emprunts japonais.....	7
1.1 - Structure du lexique et phonologie du japonais.....	7
1.1.1 - Strates lexicales du japonais.....	7
1.1.2 - Système phonologique du japonais.....	9
1.1.2.1 - Les voyelles.....	9
1.1.2.2 - Les consonnes.....	10
1.1.2.3 - Les segments moriques spéciaux /R/, /N/, et /Q/.....	10
1.1.3 - Réalisation des phonèmes.....	11
1.1.3.1 - /t/ et /d/.....	11
1.1.3.2 - /s/ et /z/.....	12
1.1.3.3 - /h/ et /p/.....	13
1.1.3.4 - /w/ et /j/.....	14
1.1.3.5 - /R/, /N/ et /Q/.....	15
1.1.4 - Particularités des gairaigo.....	16
1.2 - Processus de formation des abréviations d'emprunts composés.....	21
1.2.1 - Analyse statistique : le corpus de Laurence Labrune.....	21
1.2.1.1 - Distinction entre DAC et DAS.....	21
1.2.1.2 - Principe général de formation des DAC.....	22
1.2.1.3 - Analyse des exceptions et de leurs régularités.....	25
1.2.2 - Comparaison de corpus.....	31
1.2.2.1 - Hypothèse de l'évolution des DAC.....	31
1.2.2.2 - Modifications préalables de la hiérarchie des contraintes.....	32
1.2.2.3 - Données du corpus jeux-vidéo.....	35
2 - Aspects méthodologiques du questionnaire en ligne.....	40

2.1 - Sélection des items.....	40
2.1.1 - Les premiers constituants .....	41
2.1.2 - Items de contrôle (items 91 à 100).....	41
2.1.3 - Traitement de /R/ en deuxième more de C2 (items 1 à 20).....	42
2.1.4 - Traitement des bilabiales à l'initiale de C2 (items 21 à 40).....	43
2.1.5 - Traitement de /Q/ en deuxième more de C2 (items 41 à 60).....	44
2.1.6 - Traitement du hiatus entre le C1 et le C2 (items 61 à 80).....	45
2.1.7 - Traitement de /R/ en C1 et de /Q/ en C2 (items 81 à 90).....	47
2.2 - Aspects techniques.....	47
3 - Résultats.....	50
3.1 - Traitement de /R/ en deuxième more de C2 (items 1 à 20).....	50
3.2 - Traitement des bilabiales à l'initiale de C2 (items 21 à 40).....	52
3.3 - Traitement de /Q/ en deuxième more de C2 (items 41 à 60).....	54
3.4 - Bilan pour les items 1 à 60 et 91 à 100.....	56
3.5 - Traitement du hiatus entre le C1 et le C2 (items 61 à 80).....	57
3.6 - Traitement de /R/ en C1 et de /Q/ en C2 (items 81 à 90).....	59
4 - Discussion.....	60
4.1 - Traitement de /R/ en deuxième more de C2.....	61
4.2 - Traitement des bilabiales à l'initiale de C2.....	61
4.3 - Traitement de /Q/ en deuxième more de C2.....	62
4.4 - Bilan des items 1 à 60 et 91 à 100.....	62
4.5 - Analyse en OT stochastique.....	63
4.5.1 - Le traitement de /R/ et son évolution.....	64
4.5.2 - Le traitement de /Q/ et son évolution.....	65
4.5.3 - L'influence des bilabiales dans l'analyse OT stochastique.....	66
4.6 - Le traitement du hiatus.....	67
4.7 - La superposition de /Q/ sur /R/.....	69
4.8 - Bilan général.....	70
Conclusion :.....	73
Bibliographie :.....	75

## **Remerciements :**

J'adresse mes remerciements aux personnes qui m'ont aidé dans la réalisation de ce mémoire. En premier lieu, je remercie M. German, mon directeur de mémoire, qui m'a guidé dans mon travail et m'a permis de trouver des solutions pour avancer.

Je remercie également toute l'équipe enseignante du master ThéLiTE<sub>x</sub> ainsi que la Laboratoire Parole et Langage qui ont fourni un cadre de travail idéal pour la réalisation de ce mémoire.

J'adresse également mes remerciements à Mme Vittrant, qui m'a guidé dans mes précédents travaux et m'a poussé à toujours aller de l'avant.

Enfin, je remercie Mme Labrune pour avoir mis ses corpus à disposition de la communauté scientifique et les participants anonymes à mon questionnaire sans qui ce travail n'aurait pas pu exister.

## INTRODUCTION:

En japonais, les mots d'emprunt constituent une couche du lexique dont l'étude peut s'avérer particulièrement intéressante pour la compréhension de la phonologie de cette langue. En effet, les particularités phonologiques que l'on peut observer dans les mots d'emprunt sont autant d'infractions aux contraintes phonologiques auxquelles sont soumises les autres strates lexicales du japonais. De plus, l'étude des contraintes que doivent respecter les mots d'emprunt nous permet de comprendre quelles sont les limites que les mots japonais ne peuvent franchir pour être considérés comme tels. En d'autres termes, les mots d'emprunt nous donnent un aperçu non seulement des frontières du lexique japonais mais aussi des contraintes phonologiques qui s'appliquent à celui-ci. Il n'est donc pas étonnant de trouver plusieurs travaux sur le sujet, notamment dans l'optique de la Théorie de l'Optimalité (Itô & Mester, 1995, 2008 ; Itô, Kubozono & Mester, 2009 ; Lovins, 1973 ; Ohso, 1971 ; Lamarque, 2013, 2015b ; Shinohara, 1996, 1997).

Dans ce travail, nous nous intéresserons à la phonologie des emprunts, et plus particulièrement au processus de formation d'abréviations à partir de plusieurs constituants étrangers. Bien que ces abréviations soient formées à partir de mots empruntés, le processus qui régit leur formation est quant à lui spécifique au japonais. Cela apparaît clairement si l'on compare les manières d'abrégier les mêmes mots dans leur langue d'origine et en japonais. Par exemple, les mots 'personal computer' sont abrégés en anglais par 'PC', selon un processus de siglaison. Toutefois, en japonais, bien que les mots 'personal computer' aient été empruntés, leur abréviation n'est pas 'PC' mais [paso-koN]. Cela vient tout d'abord de l'adaptation de ces mots à la phonologie du japonais : 'personal computer' est prononcé [pa:sonaru-koNɸjuta:]. Cette forme adaptée est ensuite abrégée selon un processus qui s'apparente davantage à l'acronymie<sup>1</sup>. Ce processus d'abréviation est particulièrement productif en japonais et nous désignerons par DAC (Dérivé Abrégé Composé)<sup>2</sup> les abréviations de ce type. Précisons que nous employons le terme DAC uniquement pour désigner les abréviations formées à partir d'exactly deux constituants d'origine étrangère, comme dans les deux exemples ci-dessus.

L'exemple [paso-koN] est composé de quatre mores, ce qui semble être un modèle prosodique assez apprécié dans le lexique japonais de manière générale. Cependant, on trouve également une quantité assez importante de DAC ne contenant que trois mores. Par exemple, l'abréviation de 'potato chips', prononcé [poteto-t̃eippusu], est [pote-t̃eï]. Il semblerait donc

1 Notons que le français procède également par siglaison, comme l'anglais. Voir Kreidler (2000) pour plus d'explications concernant les différents procédés morphologiques employés dans les abréviations.

2 Terme emprunté à Labruno (2007).

qu'il existe une certaine variabilité dans la formation des DAC. Dans notre travail, nous essayerons donc de déterminer si cette variation est conditionnée par la grammaire ou si, au contraire, elle apparaît de manière libre. Pour cela, nous avons choisi d'appuyer notre étude sur un cadre théorique qui, comme nous l'avons vu, est largement employé dans le domaine : la Théorie de l'Optimalité (Prince & Smolensky, (1993/)/2004 ; Mc.Carthy, 1993)<sup>3</sup>.

Selon l'opinion de certains locuteurs que nous avons interrogés au sujet des DAC, les abréviations trimoriques seraient plus fréquentes dans les DAC récents. L'un des principaux objectifs de notre travail sera donc de vérifier si la variation observée au niveau de la formation des DAC n'aurait pas pour origine l'évolution de leur processus de formation. Cela se traduirait, en termes de contraintes, par des modifications au niveau du rang d'une ou plusieurs contraintes dans la hiérarchie.

Notre travail est composé de deux parties : dans un premier temps, nous reviendrons sur les travaux antérieurs portant sur la phonologie des emprunts et des DAC japonais. Dans cette partie, nous nous appuyerons sur nos précédents mémoires de recherche (Lamarque, 2013 et 2015a). Nous nous familiariserons avec les particularités phonologiques des mots d'emprunt, qui servent de constituants aux DAC, en commençant par une brève introduction à la phonologie du japonais. Puis nous nous intéresserons plus précisément aux DAC en proposant une analyse de leur processus de formation dans l'optique de la Théorie de l'Optimalité. Cette analyse est une révision de celle formulée dans Lamarque (2015a) et repose sur la comparaison statistique de deux corpus.

Dans un second temps, nous aborderons l'enquête que nous avons menée afin de mettre à l'épreuve les conclusions des travaux antérieurs. Nous décrirons tout d'abord cette enquête qui s'est effectuée sous la forme d'un questionnaire en ligne. Puis nous présenterons nos résultats et nous les discuterons.

---

3 Voir également Kager (1999) pour une introduction à la Théorie de l'Optimalité.

# **1 - Travaux antérieurs sur les DAC et les emprunts japonais**

## **1.1 - Structure du lexique et phonologie du japonais**

### **1.1.1 - Strates lexicales du japonais**

Avant de nous concentrer sur les DAC, il nous semble essentiel de rappeler brièvement le fonctionnement de la phonologie du japonais. On reconnaît traditionnellement trois couches lexicales en japonais qui se caractérisent, entre autres, par des comportements morphologiques et phonologiques particuliers (Labrune, 2006a). Nous pouvons ainsi distinguer :

- Le lexique *yamato* ou *wago* : il s'agit du lexique indigène japonais. C'est le noyau du lexique, l'ensemble des contraintes phonologiques de la langue y sont respectées. Les mots *yamato* s'écrivent à l'aide des *hiragana*, l'un des deux syllabaires du japonais, complétés par les caractères chinois.
- Le lexique sino-japonais ou *kango* : ce sont des morphèmes empruntés au chinois ancien depuis le IV<sup>ème</sup> siècle. Cette catégorie de mots tolère certaines distributions de segments qui ne sont pas admises en *yamato*. Les mots sino-japonais s'écrivent à l'aide des caractères chinois.
- Le lexique des mots d'emprunt ou *gairaigo* : il correspond aux emprunts lexicaux aux langues étrangères autres que le chinois ancien. Les premiers emprunts de ce type remontent aux premiers échanges avec l'Occident, autour du XV<sup>e</sup> siècle. Il s'agissait essentiellement d'emprunts à l'espagnol, au portugais ou au hollandais. De nombreux emprunts à d'autres langues, telles que l'anglais, font leur apparition depuis le XIX<sup>e</sup> siècle. De nos jours, environ 80% du lexique *gairaigo* est composé de mots empruntés à l'anglais. Cette couche du lexique autorise un nombre important de combinaisons de phonèmes interdites dans toutes les autres couches. Les mots du lexique *gairaigo* s'écrivent en *katakana*, un autre syllabaire japonais.<sup>4</sup>

De plus, il existe au sein de la catégorie des mots *yamato* divers mots mimétiques, tels que les onomatopées (*giseigo*) ou les idéophones (*gitaigo*), qui présentent quelques particularités phonologiques. En nous basant sur la théorie de Itô & Mester (1995), la

---

4 Cette description des différentes strates lexicales est reprise de Lamarque (2013)

structure du lexique japonais peut être représentée par une série de cercles concentriques correspondant chacun à une strate lexicale. Les cercles situés au centre de la structure et qui constituent son noyau correspondent aux strates lexicales dans lesquelles les contraintes phonologiques de la langue sont davantage respectées. Au contraire, plus on s'éloigne vers la périphérie de la structure et plus les strates lexicales sont permissives concernant l'infraction de certaines contraintes phonologiques. Or, on trouve dans les mots mimétiques un certain nombre d'infractions aux contraintes phonologiques du japonais qui ne sont pas admises dans les strates *yamato* ou sino-japonaise. Dans ce cadre théorique, nous considérerons donc les mots mimétiques comme une catégorie à part, située entre les strates sino-japonaise et *gairaigo*. Le schéma 1 ci-dessous illustre de quelle manière la structure du lexique japonais peut être représentée en suivant le raisonnement de Itô & Mester.



Schéma 1 : Les strates lexicales du japonais.<sup>5</sup>

Les DAC que nous traiterons par la suite sont des abréviations formées à partir de deux constituants appartenant à la catégorie des *gairaigo*. Notons qu'il existe des catégories de mots mixtes, pouvant intégrer des morphèmes issus de strates lexicales différentes. C'est le cas par exemple du mot *karaoke* qui correspond à l'abréviation du mot *kara* appartenant à la strate *yamato*, signifiant 'vide'; et de *ookesutora*, mot *gairaigo* emprunté de l'anglais 'orchestra'. Nous considérons cependant que les DAC doivent être uniquement formés à partir de mots *gairaigo*, ce qui exclu donc toutes formes de mélange entre différentes strates lexicales.

<sup>5</sup> Ce schéma est repris de Lamarque (2015b)



Dans les sections qui suivent, nous passerons brièvement en revue le système phonologique du japonais ainsi que les principales contraintes phonologiques qui ont cours dans cette langue. Nous soulignerons les spécificités de la catégorie *gairaigo*, que l'on retrouve donc par la suite dans les DAC.

## 1.1.2 - Système phonologique du japonais

### 1.1.2.1 - Les voyelles

Le japonais possède un système vocalique assez simple constitué de cinq voyelles : /a/, /i/, /u/, /e/ et /o/. Elles se distinguent selon trois traits : [haut], [bas] et [avant]. Le schéma 2 et le tableau 1 ci-dessous illustrent respectivement le système vocalique du japonais et la distinction des voyelles en fonction de leurs traits :

Avant		Arrière		
i			u	Haut
	e	o		
		a		Bas

Schéma 2 : le système vocalique japonais

	i	u	e	o	a
Haut	+	+	-	-	-
Bas	-	-	-	-	+
Avant	+	-	+	-	-

Tableau 1 : Traits distinctifs des voyelles

Les durées moyennes des réalisations sonores de ces voyelles sont classées de la plus longue à la plus brève de la manière suivante : a > e > o > i > u (Shimizu Han, 1962). Concernant leur degré d'aperture, les voyelles se classent ainsi de la plus ouverte à la plus fermée : a > e > o > u > i (Mabuchi, 1971). Notons pour la suite que les voyelles hautes sont les plus brèves et les plus fermées du système. Par ailleurs, le caractère fermé et bref de /u/ lui confère, entre autres propriétés remarquables, le statut de voyelle d'épenthèse par défaut (Labrune, 2006 ; Lamarque, 2015b). Cela apparaît de manière assez claire dans la strate *gairaigo* dans laquelle /u/ est très fréquemment employé comme voyelle épenthétique. Cela permet, par exemple, de remédier à la présence de groupes consonantiques ou de consonnes

finale dans le mot d'origine, ceux-ci étant illicites en japonais.

Notons enfin que l'allongement vocalique est pertinent pour toutes les voyelles. L'allongement vocalique est traditionnellement représenté dans la phonologie du japonais par le segment /R/ que nous aborderons dans la section §1.1.2.3.

### 1.1.2.2 - Les consonnes

Les consonnes et semi-consonnes du japonais sont présentées dans le tableau ci-dessous :

	Labiales	Alvéolaires	Palatales	Vélaires	Glottales
Occlusives	p b	t d		k g	
Fricatives		s z			h
Nasales	m	n			
Glides			j	w	
Liquide		r			

Tableau 2 : Les consonnes et semi-consonnes du japonais.

Les consonnes obstruantes peuvent être gémées à l'aide du segment spécial /Q/ que nous aborderons dans la section §1.1.2.3 ci-dessous. Plusieurs de ces consonnes présentent des variantes allophoniques que nous traiterons dans la section §1.1.3.

### 1.1.2.3 - Les segments moriques spéciaux /R/, /N/, et /Q/

On reconnaît habituellement trois segments moriques spéciaux en japonais :

- /R/ : l'allongement vocalique
- /N/ : la nasale homorganique
- /Q/ : la première partie d'une consonne gémée

Il est important de noter que ces trois segments sont « moriques », c'est-à-dire qu'ils constituent des mores à part entière. En japonais, la more est l'unité prosodique la plus pertinente<sup>6</sup>. Plusieurs travaux, notamment en typologie rythmique des langues, classent le japonais parmi les langues moriques et soulignent l'importance de cette unité prosodique,

---

<sup>6</sup> Mc.Cawley (1968) est le premier à souligner l'importance de la more en japonais, il sera suivi par de nombreuses autres études, telles que Kubozono (1992).

certaines allant jusqu'à remettre en cause la pertinence du découpage syllabique dans l'étude de cette langue (Labrune, 2001, 2005, 2012 ; Pamies Bertránd, 1999 ; Pike, 1945 ; Ramus, 1999 ; Shinohara, 1996). Par ailleurs, la more est la seule unité reconnue dans la réflexion linguistique autochtone, occultant complètement la pertinence du découpage en syllabes.

Rappelons que la more est une unité prosodique inférieure à la syllabe et permet de mesurer le poids de celle-ci. Ainsi, une syllabe légère, dépourvue de coda, correspond à une more tandis qu'une syllabe lourde, pourvue d'une coda, a un poids de deux mores. Or, en japonais, seuls les trois segments moriques spéciaux peuvent apparaître en position de coda.

La structure de la syllabe japonaise, si toutefois cette unité était pertinente, suit donc l'un des schémas ci-dessous :

- Syllabe légère : (C)(j)V
- Syllabe lourde : (C)(j)VR / (C)(j)VN / (C)(j)VQ

Notons pour la suite que la présence d'une syllabe de type (C)(j)VR ou (C)(j)VQ à l'initiale d'un des constituants d'un DAC semble avoir une influence considérable sur la formation de l'abréviation. Par souci de clarté, nous transcrivons systématiquement les segments moriques spéciaux à l'aide des lettres R, N et Q, que ce soit dans les transcriptions phonologiques ou phonétiques.

### **1.1.3 - Réalisation des phonèmes**

Dans les sections qui suivent, nous expliquerons quelles sont les différentes règles d'allophonie que l'on peut trouver en japonais. Nous nous focaliserons ici sur les strates lexicales centrales (*yamato* et sino-japonaise) afin de montrer quelles sont les contraintes phonologiques qui ont cours en japonais. La section §1.1.4 sera consacrée aux infractions de ces contraintes qui sont admises dans les strates lexicales les plus périphériques (*gairaigo*).

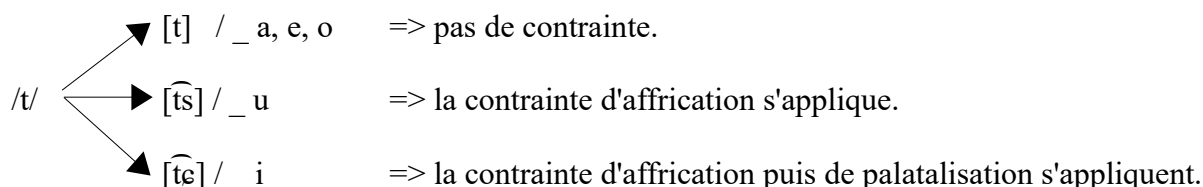
Nous ne reviendrons donc pas sur le cas des voyelles, ni des consonnes /b/, /k/, /g/, /m/, /n/ et /t/, dont les réalisations sont plus ou moins identiques dans toutes les positions.

#### 1.1.3.1 - /t/ et /d/

Les consonnes /t/ et /d/ possèdent toutes deux plusieurs allophones. Cela est lié à l'existence de deux contraintes phonologiques en japonais. La première, que nous qualifierons

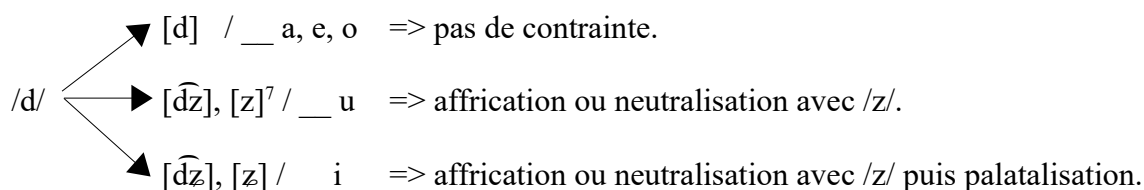
de « contrainte d'affrication », entraîne l'affrication des occlusives alvéolaires lorsqu'elles précèdent une voyelle haute (/i/ et /u/). La seconde, que nous appellerons « contrainte de palatalisation », implique la palatalisation des fricatives alvéolaires devant /i/. Ces deux contraintes entraînent donc les règles d'allophonie suivantes pour /t/ :

°/t/ { t ~ ts̄ ~ tɛ̄ }



Le cas de /d/ est plus complexe car, en japonais moderne, on constate une neutralisation des phonèmes /d/ et /z/ devant les voyelles hautes. On trouve donc, dans ces positions, des variantes libres de ces phonèmes (voir également le cas de /z/ dans la section §1.1.3.2 ci-dessous) :

°/d/ { d ~ d̄z̄ ~ z ~ d̄z̄ ~ z }

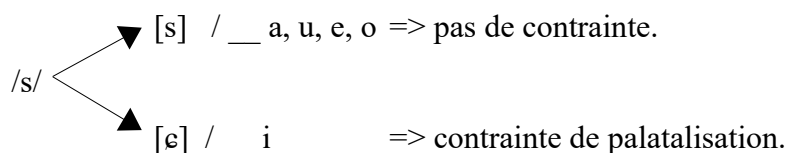


Ce phénomène de neutralisation s'explique par une mutation en japonais moderne des réalisations affriquées de /d/ vers une simple fricative.

### 1.1.3.2 - /s/ et /z/

La fricative alvéolaire sourde /s/ ne possède que deux allophones. Cette allophonie est encore une fois provoquée par la contrainte de palatalisation devant /i/. Ainsi :

°/s/ { s ~ ɕ }



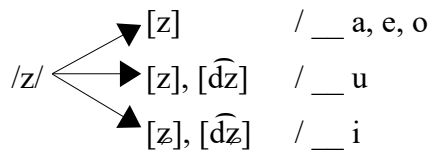
Cependant, les règles d'allophonie de /z/ sont plus complexes puisque, comme nous l'avons vu précédemment, il existe un phénomène de neutralisation entre /d/ et /z/ devant les

<sup>7</sup> Lorsque deux réalisations sont indiquées, cela signifie qu'elles sont en variantes libres dans ce contexte. Ce type de schéma est emprunté à *La phonologie du japonais* de Laurence Labrune (2006).

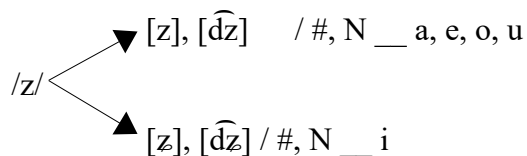
voyelles hautes. Le cas de /z/ est d'autant plus particulier que la neutralisation est encore plus prononcée à l'initiale de mot ou après le segment spécial /N/. Ainsi, dans ces positions, /z/ peut se réaliser [z] ou [dz̄] devant toutes les voyelles sauf /i/. Cela est résumé dans les schémas suivants :

°/z/ { z ~ dz̄ ~ z ~ dz̄ }

- En position interne de mot (sauf après /N/) :



- A l'initiale de mot ou après /N/ :

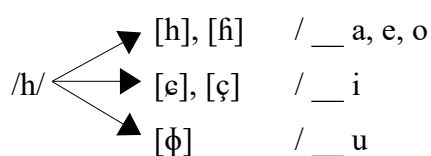


### 1.1.3.3 - /h/ et /p/

La consonne /h/ est définie comme une fricative laryngale ou glottale. Son point d'articulation n'est pas clairement déterminé et ces appellations indiquent qu'elle ne possède pas de point d'articulation supra-glottique. Par conséquent, sa réalisation effective est fortement influencée par la voyelle qu'elle précède. Ainsi, sa réalisation pourra varier entre [h] et [h̄] devant les voyelles /a/, /e/ et /o/. Devant /i/, /h/ peut-être réalisé [ɛ] ou [ç]. Dans le premier cas, on peut donc constater un phénomène de neutralisation entre /s/ et /h/ qui se réalisent tous deux [ɛ] devant /i/. Enfin, /h/ est réalisé [ϕ] lorsqu'il précède /u/. Notons par ailleurs que la voyelle /u/ n'est que légèrement arrondie en japonais. Aussi est-elle parfois transcrite à l'aide de ce symbole : /u/. Toutefois, nous ne saurions expliquer l'assimilation d'un point d'articulation bilabial pour /h/ lorsqu'il précède /u/ si cette voyelle n'était pas arrondie. Nous préférons donc le symbole /u/ à /u/.

En résumé, les règles d'allophonie pour /h/ sont les suivantes :

°/h/ { h ~ h̄ ~ ɛ ~ ç ~ ϕ }



Contrairement aux autres obstruantes sourdes, la consonne /h/ n'est jamais gémignée, du moins, dans les lexiques *yamato* et sino-japonais. On pourrait supposer que les suites /Qh/ ne sont pas autorisées en japonais. Toutefois, des études diachroniques montrent que /h/ a évolué depuis /p/ (Labrunne & Takayama, 2004) . Il est intéressant de noter que dans les deux strates lexicales centrales, /p/ apparaît presque toujours de manière gémignée ou précédée de /N/. La consonne /p/ n'apparaît donc jamais en initiale de mot. De plus, on trouve plusieurs exemples où [Qp] est une alternative plus expressive de [h]. Par exemple : [jahari] 'comme on pouvait s'y attendre' VS [jaQpari] 'comme on pouvait s'y attendre !'. Cela semble indiquer que /h/, bien qu'ayant évolué depuis /p/ conserve la réalisation [p] en cas de gémination. Ainsi, /Qh/ se réalise [Qp]. Toutefois, nous considérerons /p/ comme un phonème à part entière car il peut apparaître non gémigné en japonais moderne : devant toutes les voyelles lorsqu'il est précédé de /N/ ; dans les mots mimétiques (qui sont étymologiquement d'origine *yamato* pour la plupart) ; et dans les mots d'emprunts.

#### 1.1.3.4 - /w/ et /j/

Les semi-consonnes /w/ et /j/ apparaissent uniquement devant certaines voyelles. /w/ n'apparaît que devant la voyelle /a/. Sa réalisation se situe entre les symboles API [ɰ] et [w]. Par exemple : /wa/ [wa] 'moi'. Quant à la semi-consonne /j/, elle n'apparaît que devant les voyelles /a/, /u/ et /o/ et sa réalisation dépend du fait qu'elle soit précédée ou non d'une consonne.

Nous avons vu dans la section §1.1.2.3 que la structure des syllabes japonaises, si l'on exclut les segments spéciaux qui peuvent apparaître en coda, suivait le modèle suivant : (C)(j)V. Cela signifie que /w/ n'est jamais précédé d'une consonne et sert donc d'attaque consonantique lorsqu'il apparaît devant /a/. La semi-consonne /j/ peut également remplir ce rôle lorsqu'elle n'est pas précédée d'une consonne et se réalise alors [j], comme dans /joi/ [joi] 'bien'. En revanche, il est possible que /j/ suive une consonne, auquel cas elle palatalise celle-ci. En d'autres termes, les suites /Cj/ se réalisent [C<sup>j</sup>], comme dans /kjoka/ [k<sup>j</sup>oka] 'autorisation'.

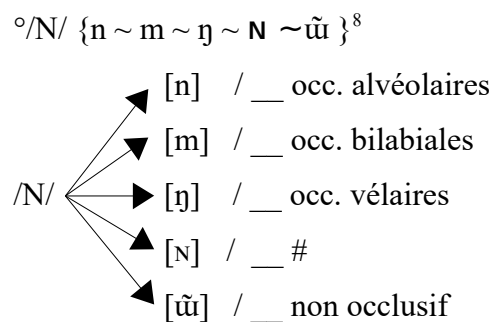
Cela permet d'expliquer que [ɛ̃], [z̃]-[d̃z̃] et [t̃ɛ̃], respectivement des allophones de /s/, /z/-/d/ et /t/, n'ont pas le statut de phonème bien qu'on puisse les trouver devant toutes les voyelles sauf /e/. Par exemple, si l'on trouve bien des séquences [ɛ̃i], [ɛ̃a], [ɛ̃u] et [ɛ̃o] en

japonais, il faut distinguer d'une part [ei] qui correspond phonologiquement à /si/, de [ea], [eu] et [eo] qui correspondent respectivement à /sja/, /sju/ et /sjo/. Cela est confirmé non seulement par le système d'écriture mais aussi par le fait qu'il n'existe pas de suite \*[ee]. En effet, puisque nous savons qu'il n'existe pas de suite \*/je/, il semble logique qu'il n'existe pas de suite \*/sje/ qui se réaliserait \*[ee].

### 1.1.3.5 - /R/, /N/ et /Q/

La réalisation des segments moriques spéciaux /R/, /N/ et /Q/ dépend du contexte phonologique dans lequel ils apparaissent. La more d'allongement vocalique /R/ peut se placer après n'importe quelle voyelle et se réalise par l'allongement de celle-ci. Une même voyelle peut ainsi être allongée à une durée de deux mores.

Concernant les mores /N/ et /Q/, il est intéressant de noter que dans le lexique *yamato*, celles-ci apparaissent en distribution complémentaire. En effet, /N/ n'apparaît que devant des consonnes voisées tandis que /Q/ n'est suivi que par des consonnes sourdes. La réalisation de /Q/ est une gémation totale de la consonne qu'il précède (par exemple : /Qt/ se réalise [tt]). En revanche, /N/ est une consonne nasale homorganique, cela signifie qu'elle assimile le point d'articulation de la consonne qui suit, auquel vient s'ajouter le trait de nasalité. Voici un schéma récapitulatif de ses réalisations possibles :



On peut supposer que /N/ et /Q/ correspondent étymologiquement à un seul et même phonème visant à renforcer une consonne, bien que ce ne soit plus le cas en japonais moderne. En effet, en dehors de la strate lexicale *yamato* la nasale organique peut apparaître devant une consonne sourde, comme dans /geNki/ [genki] 'en forme' (issu de la strate sino-japonaise). Toutefois, la gémation reste limitée aux consonnes sourdes dans l'ensemble du lexique à l'exception de la strate *gairaigo*. Il semble donc que le licenciement correct de /Q/, par une

<sup>8</sup> Les réalisations de /N/ sont très variées. Il peut encore se réaliser, entre autres, par la nasalisation de la voyelle qu'il suit dans les suites /VNV/. Ce schéma est donc simplifié.

consonne sourde à sa droite, soit une contrainte assez importante en japonais. Cela a notamment des répercussions sur le processus de formation des DAC que nous aborderons dans la section §1.2.

Par la suite, nous nous intéresserons également aux rapports qu'entretiennent les mores spéciales /R/ et /Q/ dans une analyse auto-segmentale du japonais (voir section §2.1.7)<sup>9</sup>. Ces deux mores spéciales ont pour particularité commune d'opérer un branchement dans la structure de la syllabe, ce qui les lie intrinsèquement.

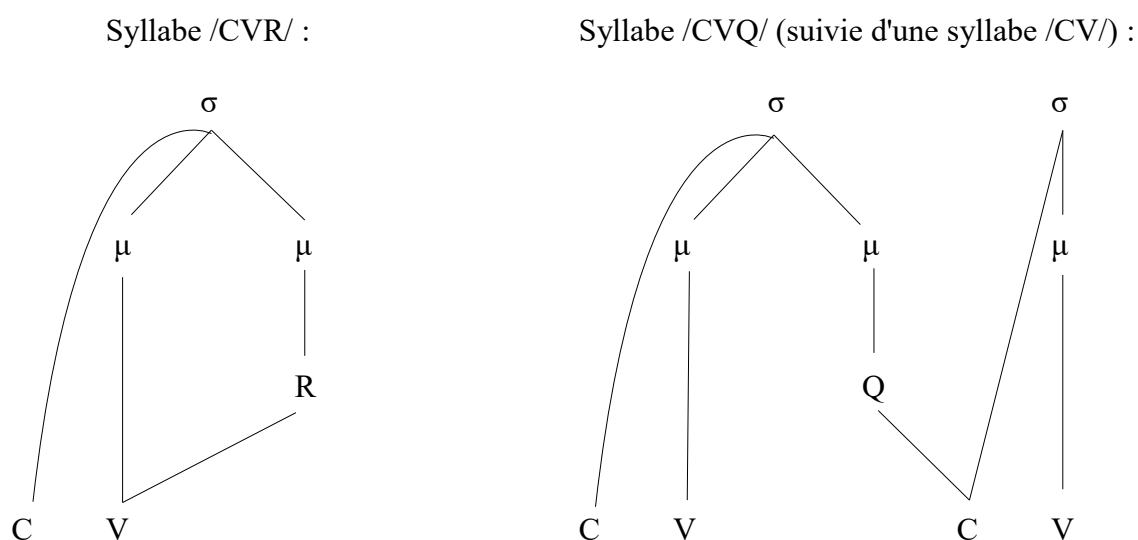


Schéma 3 : la structure des syllabes de type /CVR/ et /CVQ/

De ce point de vue, la principale différence entre /R/ et /Q/ réside dans le type de branchement que la more en coda va effectuer. Dans le cas de /R/, le branchement se fait vers le segment vocalique qui précède, tandis que dans le cas de /Q/, le branchement s'effectue avec le segment consonantique qui suit. Il n'est donc pas étonnant de retrouver une certaine proximité entre ces deux segments spéciaux. Labrune (2006a) présentent quelques exemples de mots du lexique japonais, issus de Jôo (1977) dans lesquels /R/ et /Q/ peuvent alterner, en raison, entre autres, d'une variation de style ou de dialecte. Par exemple, le nom de la lettre H peut être réalisé [eQtẽi] ou [eRtẽi]<sup>10</sup>.

#### 1.1.4 - Particularités des *gairaigo*

En passant en revue les différentes règles d'allophonie du japonais, nous avons évoqué

<sup>9</sup> Pour une introduction à la phonologie auto-segmentale, voir Gussenhoven & Jacobs (2011)

<sup>10</sup> Ces exemples montrent en réalité qu'il existe des alternances non seulement entre /R/ et /Q/ mais également avec le troisième segment spécial /N/ et la voyelle /i/. Par ailleurs, la lettre H peut aussi être réalisée [eitẽi].



plusieurs fois l'existence de contraintes phonologiques. Cette notion de contrainte renvoie inévitablement au cadre théorique de la Théorie de l'Optimalité (désormais OT)(Prince & Smolensky, (1993/2002)2004 ; Mc.Carthy, 1993)<sup>11</sup>. Dans ce cadre théorique, les tendances phonologiques s'expriment en termes de contraintes violables plutôt que de règles absolues. Cela permet de modéliser et d'analyser de manière convaincante bon nombre de phénomènes en japonais et notamment ceux concernant les mots d'emprunt de manière générale. Si l'on trouve sur ce sujet de recherche des travaux antérieurs au développement de OT (Lovins, 1973), on trouve également, de nos jours, plusieurs travaux portant sur les emprunts japonais dont l'analyse se situe dans une optique OT, que ce soit sur l'adaptation phonologique des mots étrangers (Shinohara 1997 ; Lamarque 2013, 2015b) ; l'abréviation de mots d'emprunt simples<sup>12</sup> (Labrune 2002) ; ou encore l'abréviation de mots d'emprunts composés (DAC) (Labrune, 2006b, 2007, 2008 ; Lamarque, 2015a).

De plus la théorie sur la structure lexicale des langues proposée par Itô & Mester (1995) est également profondément ancrée dans le courant OT. Nous avons vu brièvement dans la section §1.1.1 que les strates centrales étaient celles qui respectaient davantage les contraintes phonologiques de la langue. Selon OT, les contraintes phonologiques sont en opposition aux contraintes de fidélité qui imposent que l'output soit le plus proche possible de l'input. L'ensemble de ces contraintes en opposition, supposément universelles, sont hiérarchisées de manière différente selon les langues, ce qui explique que toutes les langues ne sont pas soumises aux mêmes contraintes phonologiques.

On reconnaît habituellement au moins deux contraintes de fidélité : MAX et DEP. Celles-ci imposent respectivement que les segments de l'input se retrouvent tous dans l'output (pas de suppression) et que les segments de l'output correspondent tous à un segment de l'input (pas d'épenthèse). Cependant, dans le cas présent, nous pouvons réunir ces deux contraintes en une seule, plus générale :

FID : l'output doit être conforme à l'input (ni suppression, ni épenthèse)

Selon la théorie de Itô & Mester (1995), la contrainte de fidélité gagnerait en importance dans la hiérarchie des contraintes à mesure que l'on s'éloigne du centre de la structure du lexique. En d'autres termes, FID serait à son niveau hiérarchique le plus bas dans la catégorie des *yamato*, ce qui explique que cette couche lexicale soit soumise à un grand

---

11 Voir également Kager (1999) pour une introduction à la Théorie de l'Optimalité.

12 Les Dérivés Abrégés Simples (DAS) se distinguent des DAC par le fait qu'ils ne sont des formes abrégées que d'un seul constituant étranger. Les processus de formation des DAS et des DAC suivent des logiques totalement différentes, ce qui rend leur distinction indispensable (voir §1.2.1.1).

nombre de contraintes phonologiques. Tandis que dans la strate des *gairaigo*, FID occuperait une place beaucoup plus haute dans la hiérarchie, dominant ainsi la plupart des contraintes phonologiques qui s'appliquent aux autres strates lexicales. Une autre manière d'arriver à des conclusions similaires serait de considérer que chaque strate lexicale possède une contrainte FID spécifique et que les différentes contraintes de fidélité occuperaient des places différentes dans la hiérarchie. Pour le cas du japonais, on pourrait imaginer une hiérarchie de contraintes telle que celle présentée ci-dessous, qui justifierait que les mots *gairaigo* puisse enfreindre une contrainte X bien que celle-ci s'applique sur l'ensemble des autres strates lexicales :

FID[GAIRAIGO] >> CONTRAINTE-X >> FID[MIMÉTIQUE], FID[SINO-JAPONAIS], FID[YAMATO]

Ce point de vue nous semble adéquat pour décrire les particularités phonologiques des mots d'emprunt japonais. Ainsi, nous allons proposer dans cette section une brève analyse en termes de contraintes de ces particularités. Par ailleurs, nous proposerons dans la section §1.2 de ce travail une analyse selon OT du processus de formation des DAC.

La plupart des règles d'allophonie que nous avons vues jusqu'à présent peuvent également être comprises en termes de contraintes. C'est le cas notamment pour les consonnes /s/, /z/, /t/ et /d/ qui sont soumises à des contraintes d'affrication et de palatalisation que nous pouvons décrire comme suit :

AFFRIC° : Les occlusives alvéolaires doivent être affriquées devant les voyelles hautes.

PALAT° : Les fricatives alvéolaires doivent être palatalisées devant /i/.

De même, les interdictions de créer certaines suites phonologiques, telles que /N/ devant une consonne sourde ou /Q/ devant une consonne voisée, peuvent également être considérées comme des contraintes :

\*/NT/ : la nasale homorganique ne doit pas précéder une consonne sourde.

\*/QD/ : les consonnes voisées ne peuvent être gémées

Les contraintes peuvent également porter sur la structure des syllabes puisque nous avons vu que le japonais n'acceptait ni groupes consonantiques, ni consonnes finales (voir §1.1.2.1). On peut ainsi en dégager deux contraintes :

\*CC : pas de groupes consonantiques

\*C# : pas de consonnes finales

Bien que nous ne l'ayons pas abordé précédemment, certaines contraintes portent sur la structure des morphèmes, comme la loi de Lyman, qui stipule qu'un morphème ne peut contenir qu'une seule obstruante voisée :

LYMAN : pas plus d'une obstruante voisée par morphème

Enfin, nous pourrions considérer tout un ensemble de contraintes plus spécifiques, qui limitent l'apparition des allophones à certains contextes phonologiques. Par exemple, le fait que [ϕ] soit l'allophone de /h/ devant /u/ implique qu'il ne peut apparaître devant une autre voyelle.

Sur l'ensemble de ces contraintes, on peut constater que les seules qui soient respectées dans la strate *gairaigo* sont les trois suivantes : \*CC, \*C# et PALAT°. Par exemple, les adaptations des mots 'bed' ou 'bag', respectivement [beQdo] et [baQgu], enfreignent à la fois la loi de Lyman et la contrainte \*/QD/. De même, l'adaptation de 'identity' [aideNtiti] enfreint à la fois les contraintes \*/NT/ et AFFRIC°.

Par ailleurs, en suivant le principe de redistribution phonémique (Haugen, 1950), la plupart des allophones peuvent apparaître dans des positions beaucoup plus variées dans les emprunts que dans le lexique autochtone. Si l'on prend l'exemple de /t/ et de ses allophones, on peut considérer que chacune de ses variantes a le statut de phonème dans la strate *gairaigo*. On constate tout d'abord que /t/ peut être réalisé [t] devant /u/ et /i/, ce qui enfreint la contrainte AFFRIC°. On peut citer par exemple les mots 'tattoo' [tatuR] et 'party' [paRtiR]. Concernant l'affriquée [tʃ], elle n'est plus limitée à précéder la voyelle /u/, comme le montrent les exemples 'tsar' [tʃaR], 'Elsine' [eritʃiN]<sup>13</sup>, 'konzern' (allemand) [koNtʃeruN] et 'canzone' (italien) [kaNtʃone]. Elle peut donc être considérée comme un phonème à part entière dans la strate *gairaigo*. Enfin concernant l'affriquée [tɕ], nous savons qu'elle est normalement l'allophone de /t/ devant /i/ mais que les suites [tɕa], [tɕu] et [tɕo] existent également et correspondent phonologiquement à /tja/, /tju/ et /tjo/ respectivement. Or, on trouve également des suite [tɕe] dans certains mots d'emprunt, faisant de l'affriquée /tɕ/ un phonème à part entière dans cette strate lexicale (par exemple 'check' [tɕeQku]).

Ainsi, bien que [t], [tʃ], [tɕ] soient des variantes conditionnées de /t/ dans la plupart des strates lexicales, elles correspondent à trois phonèmes, /t/, /tʃ/ et /tɕ/ respectivement, dans la strate *gairaigo*. En suivant le même raisonnement, on peut remarquer les redistributions

<sup>13</sup> On peut noter là une infraction à la contrainte PALAT° normalement inviolable dans l'ensemble du lexique. Les exemples de ce type sont néanmoins extrêmement rares.

phonémiques suivantes dans cette strate :

- [ɸ], [ɛ] et [z], allophones respectifs de /h/ devant /u/ et de /s/ et /z/ devant /i/, sont entièrement phonologisés.
- /d/ peut apparaître devant les voyelles /u/ et /i/.
- /w/ peut apparaître devant la voyelle /o/ et/ou précéder les consonnes /k/ ou /g/, formant des syllabes de types CwV, non admises dans les autres strates lexicales.
- /Qh/ peut être réalisé [hh] et non [pp].
- /v/, réalisé [v] est un phonème entièrement emprunté.<sup>14</sup>

Si l'on compare simplement le lexique *gairaigo*, le plus périphérique, au lexique *yamato* le plus central, on peut s'apercevoir que leurs caractéristiques phonologiques sont sensiblement différentes. Cela se reflète dans la hiérarchie de contraintes ci-dessous<sup>15</sup> :

\*CC, \*C#, PALAT<sup>o</sup> >> FID[GAIKAIKO] >> AFFRIC<sup>o</sup>, \*/NT/, \*/QD/, LYMAN >> FID[YAMATO]

Les mots *gairaigo*, qui vont servir de constituants aux DAC qui nous intéressent, possèdent déjà quelques particularités phonologiques remarquables. On pourrait s'attendre à ce que les DAC, eux-mêmes constitués de mots *gairaigo*, se situent également à l'extrême périphérie de la structure du lexique. Pourtant, nous verrons par la suite que par certains aspects, les DAC semblent soumis à des contraintes qui n'ont cours que dans les strates lexicales les plus centrales. Il est intéressant de noter que l'abréviation est une forme d'appropriation du lexique et, dans le cas des DAC, d'un lexique étranger. Or il arrive que certains mots fréquents glissent progressivement de la périphérie vers le centre, passant d'une strate lexicale à l'autre indépendamment de l'origine étymologique de ces mots. Nous pouvons donc envisager qu'il existe différents processus d'appropriation du lexique qui rendent perméables les frontières entre les différentes couches du lexiques. La formation de DAC et l'abréviation de manière générale pourraient bien faire partie de ces processus. Il est donc difficile de placer les DAC de manière catégorique dans la structure du lexique telle que proposée par Itô & Mester (1995).

---

14 Ce phonème a été emprunté assez récemment. Les consonnes /v/ des langues d'emprunt étaient auparavant rendues à l'aide du /b/ japonais (ex : 'video' [bideo]). Toutefois, la réalisation [b] au lieu de [v] persiste chez certains locuteurs, notamment les plus âgés.

15 Notons que ce schéma ne prend pas en compte tous les phénomènes de redistributions phonémiques.

Quoi qu'il en soit, nous verrons dans la section suivante que le processus de formation des DAC répond à des contraintes spécifiques. Nous proposerons donc une analyse de celui-ci dans une perspective OT.

## **1.2 - Processus de formation des abréviations d'emprunts composés**

### **1.2.1 - Analyse statistique : le corpus de Laurence Labrune**

#### **1.2.1.1 - Distinction entre DAC et DAS**

Les DAC suivent un processus d'abréviation qui peut paraître simple *a priori*. En effet, la formation d'un DAC canonique suit le principe suivant : sont conservées dans l'abréviation les deux mores initiales de chaque constituant. Toutefois, il existe un nombre assez important d'exceptions à ce principe, ce qui pourrait indiquer que le processus de formation des DAC est en réalité plus complexe qu'il n'y paraît aux premiers abords. Une analyse statistique de la phonologie des DAC permet en effet de mettre en lumière certaines régularités dans le formation de DAC *a priori* exceptionnels.

C'est ce que nous proposent les travaux de Laurence Labrune portant sur les abréviations d'emprunts en japonais. Labrune a collecté aux alentours des années 2000 un large corpus de 1314 abréviations contenant aussi bien des DAC que des DAS (Dérivés Abrégés Simples). A partir de ces données, elle a pu effectuer une analyse statistique de ces deux types d'abréviations, proposant ainsi une description de ces derniers. Cette analyse permet tout d'abord de constater que ces deux types d'abréviations fonctionnent selon des principes totalement différents. Dans la plupart des cas, les DAS sont tronqués après la more accentuée de la base, ce qui résulte généralement en une abréviation de 3 mores. En revanche, les DAC suivent généralement le principe évoqué ci-dessus, sans tenir compte d'aucune caractéristique prosodique de la base. Aucune description des abréviations d'emprunts en japonais ne peut donc s'avérer pleinement satisfaisante sans effectuer cette distinction entre DAS et DAC.

Par la suite, Labrune proposera donc une analyse selon la Théorie des correspondances (Benua, 1995, 1997) des DAS d'une part (Labrune, 2002) et des DAC d'autre part (Labrune 2007, 2008). Dans Labrune (2002:3), la Théorie des correspondances est décrite de la manière suivante : il s'agit d'un cadre théorique développé à la suite de OT par McCarthy & Prince (1995) afin de rendre compte de phénomènes tels que la reduplication ou la troncation. La correspondance s'effectue entre deux « correspondants » (« strings » en anglais), S1 et S2, qui

sont deux formes existant de manière autonome dans la langue. Ils sont donc liés par un rapport d'output à output, plutôt qu'input à output (forme sous-jacente / forme de surface). Contrairement au lien input-output, qui est régi par des contraintes de fidélité, cette relation output-output est soumise à des contraintes d'identité. Benua (1995) propose le modèle de correspondances suivant pour le processus morphologique de l'abréviation.

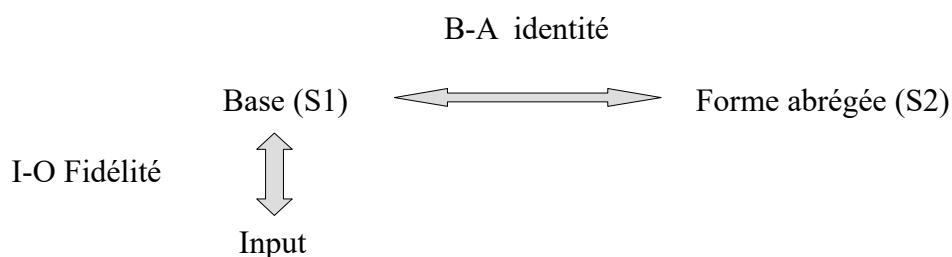


Schéma 4 : l'abréviation<sup>16</sup>

En principe, la forme abrégée est uniquement liée à la forme de surface qu'est l'output de sa base (S1) et non à son input d'origine. Les contraintes d'identité sont ensuite dominées par diverses contraintes prosodiques ou phonologiques qui expliquent la perte de matériel segmental au cours de l'abréviation.

Nous ne nous attarderons pas davantage sur l'analyse que Labrune (2002) propose du processus de formation des DAS. En revanche, nous prendrons pour point de départ dans notre étude des DAC l'analyse qu'elle en propose dans ses articles de 2007 et 2008.

#### 1.2.1.2 - Principe général de formation des DAC

En partant du principe évoqué plus haut, selon lequel les DAC sont formés en maintenant les deux mores initiales de deux constituants étrangers, Labrune (2007) propose une analyse en termes de contraintes du processus de formation des DAC. Les contraintes qui y entrent en jeu sont décrites ci-dessous. La hiérarchisation de ces contraintes ainsi que les ajouts à cette liste que nous verrons dans la section §1.2.1.3 reposent sur l'analyse statistique des caractéristiques phonologiques des 799 DAC que comporte le corpus de Laurence Labrune. Dans nos précédents travaux, nous avons cependant exclu 88 des DAC de ce corpus provenant d'un dictionnaire intitulé *Wakamono kotoba jiten* (dictionnaire du langage des jeunes), afin de réduire autant que possible la proportion de DAC récents dans le corpus de

<sup>16</sup> Ce schéma apparaît dans Labrune (2002) mais est adapté de celui de Benua (1995), lui-même inspiré de Mc.Carthy & Prince (1995)

Laurence Labrune. Comme nous le verrons plus en détails dans la section §1.2.2, le but était de comparer deux corpus d'époques aussi différentes que possible. Pour éviter toute confusion, nous présenterons uniquement nos propres données chiffrées (basées sur 711 DAC). Bien entendu, cela n'invalide en rien les descriptions proposées par Labrune (2007, 2008). Voici donc la liste des contraintes qu'elle propose :

- PIEDBINAIRE : les pieds sont binaires en mores.
- MOTPROSODIQUEBINAIRE : le mot prosodique (= le dérivé) est binaire en pieds.
- MAXLEX : chaque lexème de la base est présent dans le dérivé.
- ANCRAGEGAUCHE : chaque pied du dérivé est ancré à gauche du lexème de la base dont il provient.
- CONTIG : les mores prélevées dans chaque pied forment une chaîne contiguë dans la base.
- DEP : tout segment présent dans le dérivé est présent dans la base (pas d'épenthèse).
- IDENTTRAITS : il y a identité de traits distinctifs entre les segments en correspondances dans la base et dans le dérivé.

Ces contraintes sont organisées selon la hiérarchie suivante :

MOTPROSODIQUEBINAIRE >> CONTIG, ANCRAGEGAUCHE >> PIEDBINAIRE,  
MAXLEX, DEP, IDENTTRAITS

Par ailleurs, Labrune (2007:5) suppose qu'en japonais, les mores appartiennent toutes à un pied (parsage exhaustif). Les mores « orphelines » (non intégrées dans un pied) sont rendues illicites par la présence d'une contrainte inviolable PARSE $\mu$ , placée au-dessus de MOTPROSODIQUEBINAIRE.

La contrainte MOTPROSODIQUEBINAIRE impose que les DAC soient binaires en pieds, soit un pied pour le premier constituant (désormais C1) et un autre pour le deuxième (désormais C2). S'il est possible de trouver dans certains DAC des constituants irréguliers d'une ou de trois mores, en plus de constituants réguliers de deux mores, on ne trouve de constituant de quatre mores, soit deux pieds, dans aucun DAC. Cette infraction à MOTPROSODIQUEBINAIRE n'étant visiblement pas permise, il semble logique de placer cette contrainte au sommet de la hiérarchie.

Le fait que ce soient les deux mores initiales de chaque constituant qui sont conservées dans les DAC est assuré par les trois contraintes : CONTIG, ANCRAGEGAUCHE et PIEDBINAIRE. CONTIG impose que les mores retenues soient contiguës dans la base, ANCRAGEGAUCHE qu'elles se situent à l'initiale, et PIEDBINAIRE qu'elles soient exactement deux.

Les trois contraintes restantes (MAXLEX, DEP et IDENTTRAITS) assurent une conformité entre l'abréviation et sa base. MAXLEX impose que l'abréviation contienne au moins un élément de chacun des deux constituants de la base, DEP limite l'ajout d'élément non présent dans la base et IDENTTRAITS sanctionne les éventuelles modifications que les segments pourraient subir lors de l'abréviation.

Le tableau 3 ci-dessous présente la dérivation d'un DAC canonique [meta-kiN], issu de l'anglais 'metal king' réalisé [metaru-kiNgu]. On peut y observer quel type de candidats sont exclus par quelle contrainte.

	[metaru-kiNgu]	PARSE $\mu$	MPBIN	CONTIG	ANCRG	PIEDBIN	MAXLEX	DEP	IDENT
1.	(meta)		*!				*		
2.	(me-ki)		*!	*					
3.	(me)-(ki)					**!			
4.	(meta)-(ki)					*!			
5.	(me)-(kiN)					*!			
6. =>	(meta)-(kiN)								
7.	(meta)-(kiN)-(gu)		*!			*			
8.	(meta)-(kiN)-gu	*!	*						
9.	(meru)-(kiN)			*!					
10.	(taru)-(kiN)				*!				
11.	(meR)-(kiN)							*!	
12.	(meta)-(giN)								*!

Tableau 3 : [metaru-kiNgu] => [meta-kiN] 'metal king'

Les candidats 1 et 2 ne sont pas binaires en pieds, ce qui constitue une infraction à la contrainte MOTPROSODIQUEBINAIRE. Cette contrainte se situant vers le sommet de la hiérarchie, ces deux candidats ne peuvent être sélectionnés. De plus, le candidat 1 ne comportant qu'un seul constituant, il enfreint la contrainte MAXLEX. Concernant le candidat 3, les deux mores du DAC sont parsées dans deux pieds différents. Ces deux pieds unaires sont donc irréguliers, ce qui entraîne une double infraction à la contrainte PIEDBINAIRE. Contrairement au candidat 3, Les candidats 4 et 5 ne contiennent qu'un pied unaire qui



correspond respectivement à C1 et C2. Même s'ils n'enfreignent qu'une fois PIEDBINAIRE, cela suffit à les exclure. Le candidat 7 possédant trois pieds, il est éliminé pour les mêmes raisons que le candidat 1. Notons qu'il enfreint également PIEDBINAIRE une fois. Le candidat 8 en plus d'enfreindre MOTPROSODIQUEBINAIRE, contient une more non parsée, ce qui n'est pas admis en japonais. Enfin, les candidats 9, 10, 11 et 12 enfreignent respectivement les contraintes CONTIG, ANCRAGEGAUCHE, DEP et IDENTTRAITS. C'est donc le candidat 6, n'enfreignant aucune contrainte, qui sera retenu.

### 1.2.1.3 - Analyse des exceptions et de leurs régularités

L'intérêt du travail de Labrune (2007, 2008) ne réside pas seulement dans la description du principe général de formation des DAC, mais aussi dans les possibilités d'analyse qu'offre son corpus. En réunissant un grand nombre de DAC, et parmi eux, un grand nombre d'exceptions, Labrune a pu identifier des régularités dans celles-ci. En effet, il semblerait que certains contextes phonologiques favorisent la formation d'exceptions, et ce dans des proportions trop importantes pour qu'elles soient ignorées.

Plusieurs de ces régularités justifient l'ajout de nouvelles contraintes à la hiérarchie, afin que celle-ci rende mieux compte du fonctionnement du processus de formation des DAC. Ainsi, certaines des abréviations qui semblaient faire figure d'exception sont en fait régulières si l'on considère le processus de formation des DAC dans toute sa complexité. Par souci de clarté, nous distinguerons à présent les trois termes suivants :

- DAC canoniques : ces abréviations de quatre mores suivent strictement le principe général de formation. Ils représentent plus de 60% du corpus.
- DAC réguliers : ces abréviations ne respectent pas tout à fait le principe général de formation mais leur traitement particulier est causé par un certain contexte phonologique. Ils fonctionnent donc de manière régulière. Ils représentent environ un dixième du corpus.
- DAC irréguliers : ces abréviations ne respectent pas le principe général de formation et il ne semble pas que cela soit causé par un contexte phonologique particulier. On trouve dans cette catégorie des exceptions aux diverses contraintes de la hiérarchie (celles déjà présentées et celles que nous aborderons dans cette section). Nous verrons par la suite qu'il arrive que de petits groupes d'exceptions

semblent partager certaines caractéristiques, sans que cela apparaisse de manière assez régulière pour justifier l'intégration de nouvelles contraintes. Dans ce cas, ces abréviations sont également considérées comme irrégulières. Les irrégularités représentent un peu moins d'un tiers du corpus.

Le corpus de Labrune contient donc un peu moins de 40% de DAC qui ne respectent pas le principe général de formation. Parmi ceux-ci, une très large majorité possède une caractéristique commune : la deuxième more d'au moins un des constituants de leur base est un segment spécial /R/ ou /Q/ (vu en §1.1.3.5). En d'autres termes, un grand nombre d'exceptions est créé lorsque /R/ ou /Q/ doivent apparaître en position finale d'un des constituants du DAC. Lorsque ce cas de figure se présente, le segment spécial peut être traité de quatre manières différentes :

- Le maintien : la more problématique est maintenue, comme s'il s'agissait de n'importe quelle autre more. Un DAC canonique est ainsi formé. Par exemple<sup>17</sup> : [biRto-mania] => [biR-mani] 'beat mania'.
- Le maintien avec more supplémentaire : la more problématique est maintenue mais la more qui la suit l'est également, de manière à ce que /R/ ou /Q/ ne se retrouve pas en position finale de constituant. Cela entraîne la création d'un pied ternaire, ce qui constitue une infraction à PIEDBINAIRE. Par exemple : [purasutiQku-keRsu] => [pura-keRsu] 'plastic case'
- La suppression : la more problématique est simplement supprimée. Cela entraîne la création d'un pied unaire, ce qui constitue également une infraction à PIEDBINAIRE. Par exemple : [akua-d̄zeQto] => [aku-d̄ze] 'aqua jet'
- Le remplacement : la more problématique est remplacée par celle qui la suit dans le constituant. Cela constitue une infraction à la contrainte CONTIG. Par exemple : [baQku-suteQpu] => [baku-sute] 'back step'

Dans les tableaux suivants, on peut observer quels sont les traitements qui sont choisis dans les DAC du corpus de Laurence Labrune<sup>18</sup> :

17 Ces exemples servent d'illustration pour ces quatre traitements possibles et ne sont pas nécessairement issus du corpus de Laurence Labrune.

18 La ligne « Maintenu + » correspond au deuxième type de traitement : le maintien avec more supplémentaire. Les rares cas d'ajout ne sont pas pertinents ici, il s'agit de DAC dans lesquels /R/ ou /Q/ ont été ajoutés dans l'abréviation alors qu'ils ne sont pas présents dans la base, ce qui constitue une infraction à DEP.

/R/ en deuxième more : 193 cas			
En C1 : 81 cas (41,54%)		En C2 : 112 cas (58,03%)	
Formation canonique :		Formation canonique :	
Maintenu	53 cas ( <b>65,43%</b> )	Maintenu	42 cas (37,5%)
Formations non-canoniques :		Formations non-canoniques :	
Maintenu +	4 cas (4,94%)	Maintenu +	11 cas (9,82%)
Supprimé	5 cas (6,17%)	Supprimé	49 cas ( <b>43,75%</b> )
Remplacé	18 cas ( <b>22,22%</b> )	Remplacé	8 cas (7,14%)
Ajouté	1 (1,23%)	Ajouté	2 cas (1,79%)

Tableau 4 : /R/ dans le corpus de Labrune

/Q/ en deuxième more : 51 cas			
En C1 : 18 cas (35,3%)		En C2 : 33 cas (64,7%)	
Formation canonique :		Formation canonique :	
Maintenu	7 cas ( <b>38,89%</b> )	Maintenu	0 cas ( <b>0%</b> )
Formations non-canoniques :		Formations non-canoniques :	
Maintenu +	2 cas (11,11%)	Maintenu +	13 cas ( <b>39,4%</b> )
Supprimé	3 cas (16,67%)	Supprimé	15 cas ( <b>45,45%</b> )
Remplacé	5 cas (27,78%)	Remplacé	5 cas (15,15%)
Ajouté	1 cas (5,56%)	Ajouté	0 cas (0%)

Tableau 5 : /Q/ dans le corpus de Labrune

On peut tout d'abord constater que, malgré quelques préférences pour certains types de traitements, ces données sont assez partagées. Aucun des traitements possibles ne domine de manière écrasante l'ensemble des données. Seuls les cas de maintiens de /R/ en C1 obtiennent un score supérieur à 50%.

On peut également remarquer que la formation canonique n'est jamais majoritaire en C2, que ce soit pour /R/ ou /Q/. Dans le cas de /Q/, il n'y a même aucun cas de maintien en C2. Nous avons vu dans la section §1.1.3.5 que le licenciement correct de /Q/ par une consonne sourde à sa droite était une contrainte phonologique assez importante en japonais. En réalité, il n'existe pas de réalisation possible de /Q/ s'il n'est pas suivi d'une consonne, ce qui explique qu'il ne puisse pas apparaître en position finale de DAC.

Un premier pas dans l'amélioration de la description du processus de formation des DAC serait donc de prendre en compte cela. Ainsi, Labrune (2007) propose l'ajout d'une contrainte LICENCE/Q/ à la hiérarchie des contraintes :

LICENCE/Q/ : /Q/ doit être correctement licencié par une consonne sourde à sa droite.

Cette contrainte explique non seulement l'impossibilité de maintenir /Q/ en C2 mais aussi la possibilité de le maintenir en C1 dans les cas où la consonne initiale de C2, qui va se retrouver à son contact, est sourde. Dans la hiérarchie, LICENCE/Q/ étant assez importante en japonais, elle se situe au-dessus de CONTIG. Cependant, elle se situe en dessous de MOTPROSODIQUEBINAIRE qui est inviolable.

En C2, le traitement de /Q/ est essentiellement partagé entre le maintien avec more supplémentaire et la suppression, qui constituent tous deux une infraction à la contrainte PIEDBINAIRE. La grammaire ne semble pas avoir de franche préférence pour l'une ou l'autre des possibilités, limitant ainsi le développement d'une analyse en termes de contraintes.

On constate cependant que parmi l'ensemble des formations non-canoniques, les cas de remplacements, en infraction à CONTIG, sont numériquement bien moins importants que les deux types d'infraction possibles à PIEDBINAIRE. Cela confirme la supériorité de CONTIG sur PIEDBINAIRE dans la hiérarchie.

Le tableau 6 ci-dessous illustre de quelle manière les contraintes LICENCE/Q/, CONTIG et PIEDBINAIRE doivent être organisées pour justifier les données du corpus. Il s'agit de l'exemple de la dérivation de [aku- $\widehat{d}ze$ ] à partir de 'aqua jet' [akua- $\widehat{d}zeQto$ ] :

	[akua- $\widehat{d}zeQto$ ]	LICENCE/Q/	CONTIG	PIEDBIN
1.	(akua)-( $\widehat{d}zeQ$ )	*!		
2.	(akua)-( $\widehat{d}zeQto$ )			*!
3.=>	(akua)-( $\widehat{d}ze$ )			*
4.	(akua)-( $\widehat{d}zeto$ )		*!	

Tableau 6 : [akua- $\widehat{d}zeQto$ ] => [aku- $\widehat{d}ze$ ] 'aqua jet'

Avec les contraintes que nous avons définies jusqu'à présent, nous savons que le maintien et le remplacement ne sont pas favorisés par la grammaire, ce qui exclut les candidats 1 et 4, qui enfreignent respectivement LICENCE/Q/ et CONTIG, placées plus haut que PIEDBINAIRE, dans la hiérarchie. En revanche, en l'état, la grammaire ne permet pas de départager les candidats 2 et 3. Ici, c'est la suppression, donc le candidat 3, qui est retenu.

Le traitement de /R/ est un peu plus complexe que celui de /Q/. On peut tout d'abord constater qu'il est maintenu dans plus de la majorité des cas lorsqu'il apparaît en deuxième

more de C1. C'est seulement lorsqu'il apparaît en deuxième more de C2 que son traitement est problématique. Labrune (2007) attribue cela à une contrainte qui s'applique de manière générale en japonais et qui limite l'apparition de voyelles longues en position finale de mot<sup>19</sup>.

\*V:# : pas de voyelle longue en finale de mot.

Parmi les trois formations non canoniques qui permettraient de respecter cette contrainte, c'est la suppression qui est choisie dans la grande majorité des cas. Nous savons que CONTIG est placé plus haut dans la hiérarchie que PIEDBINAIRE, ce qui explique que le remplacement ne soit pas l'option privilégiée. En revanche, l'ensemble de contraintes établi jusqu'à présent ne permet pas de trancher entre les deux infractions possibles à PIEDBINAIRE : le maintien avec more supplémentaire (pied ternaire), ou la suppression (pied unaire). Cela nécessite l'introduction d'une nouvelle contrainte d'alignement prosodique.

MORESAGAUCHE : dans un pied, toutes les mores sont alignées à gauche.

Cette contrainte implique que toutes les mores contenues dans un pied soient alignées avec le bord gauche de ce dernier. Or, une seule more peut effectivement remplir cette position, faisant des pieds unaires les seuls qui ne soient pas en infraction avec cette contrainte. Cela exclut donc l'option du maintien avec more supplémentaire, qui entraîne la création de pieds ternaires enfreignant deux fois MORESAGAUCHE. Par ailleurs, même les pieds binaires, que l'on trouve entre autres dans les DAC canoniques, sont en infraction à cette contrainte, ce qui impose qu'elle soit placée tout en bas de la hiérarchie.

Labrune remarque également que le traitement de /R/ dépend du timbre de la voyelle qu'il suit : les voyelles hautes, plus brèves par nature (voir §1.1.2.1), permettent généralement le maintien de /R/, tandis que les voyelles non-hautes, plus longues, entraînent plus souvent sa suppression. Le diagramme 1 ci-dessous présente les données relatives au traitement de /R/ en fonction de la voyelle qu'il suit.

---

19 Contrainte également employé dans l'analyse de Nishihara (2001)

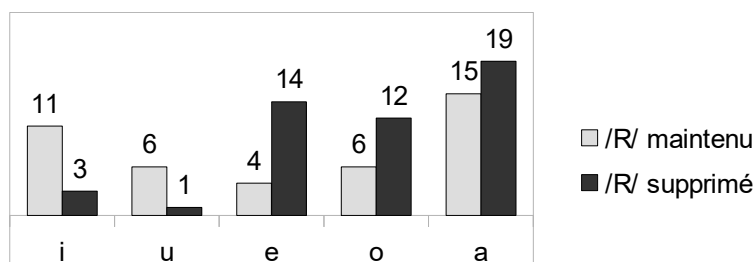


Diagramme 1 : traitement de /R/ en C2 en fonction du timbre de la voyelle.<sup>20</sup>

Afin de tenir compte de cette différence de traitement en fonction de la hauteur de la voyelle, Labrune propose de scinder la contrainte \*V:# en deux :

\*V-HAUT:# : pas de voyelle non haute allongée en position finale de mot.

\*V+HAUT:# : pas de voyelle haute allongée en position finale de mot.

Tandis que la contrainte \*V-HAUT:# serait située au-dessus de PIEDBINAIRE, ce qui impose de recourir à la suppression, \*V+HAUT:# serait située en dessous, permettant le maintien de /R/. Ceci est illustré dans les tableaux 7 et 8 ci-dessous, présentant respectivement la dérivation de [poke-bo], issu de 'pocket board' [pokeQto-boRdo] et de [njuR-ziR], issu de 'New Zealand' [njuR-ziRrando] :

	[pokeQto-boRdo]	*V-HAUT:#	CONTIG	PIEDBIN	*V+HAUT:#	MORESÀG
1.	(poke)-(boR)	*!				*
2.	(poke)-(boRdo)			*		**!
3.=>	(poke)-(bo)			*		
4.	(poke)-(bodo)		*!			*

Tableau 7 : [pokeQto-boRdo] => [poke-bo] 'pocket board'

	[njuR-ziRrando]	*V-HAUT:#	CONTIG	PIEDBIN	*V+HAUT:#	MORESÀG
1.=>	(njuR)-(ziR)				*	*
2.	(njuR)-(ziRra)			*!		**
3.	(njuR)-(zi)			*!		
4.	(njuR)-(zira)		*!			*

Tableau 8 : [njuR-ziRrando] => [njuR-ziR] 'New Zealand'

Dans le tableau 7, la voyelle qui précède /R/ en C2 est non-haute. La contrainte \*V-HAUT:# étant haut placée dans la hiérarchie, son infraction est éliminatoire pour le candidat 1

<sup>20</sup> Diagramme adapté de Labrune (2007:11). Nos données pour /a/ sont moins disparates que celles de Laurence Labrune car parmi les 88 DAC retirés, on trouve 9 cas de suppression avec cette voyelle et seulement 2 cas de maintien. Le corpus total contenait donc 17 cas de suppression contre 28 de maintien.

et la suppression (candidat 3) est retenue. En revanche, dans le tableau 8, la voyelle qui précède /R/ en C2 est haute. L'infraction à \*V+HAUT:# étant considérée comme moins coûteuse, le candidat 1 peut être retenu.

La hiérarchie de contraintes proposées au final par Labrune (2007) est donc la suivante :

PARSE<sub>μ</sub> >> MOTPROSODIQUEBINAIRE >> \*V-HAUT:#, LICENCE/Q/ >> CONTIG,  
ANCRAGEGAUCHE >> PIEDBINAIRE, MAXLEX, DEP, IDENTTRAITS >>  
\*V+HAUT:#, MORESÀGAUCHE

Cette hiérarchie permet donc de régulariser plusieurs types de DAC : ceux dans lesquels /R/ est supprimé après les voyelles non hautes lorsqu'il apparaît en deuxième more de C2 ; ceux dans lesquels la présence de /Q/ dans cette même position entraîne la formation d'un DAC unaire ; ainsi que ceux dans lesquels /Q/ est remplacé en C1, lorsque l'initiale de C2 n'est pas une consonne sourde. Cette analyse plus fine du processus de formation des DAC permet de réduire sensiblement la proportion d'exceptions.

## 1.2.2 - Comparaison de corpus

### 1.2.2.1 - Hypothèse de l'évolution des DAC

Bien que l'étude sur corpus de Labrune ait permis de mettre en lumière l'influence de certains contextes phonologiques dans la formation de DAC exceptionnels, son travail n'avait pas pour but de vérifier si le processus d'abréviation des DAC avait évolué au cours du temps. En d'autres termes, grâce à l'analyse statistique d'une grande quantité d'abréviations, Labrune a pu identifier certains facteurs phonologiques ou prosodiques qui pourraient expliquer la variation dans la formation des DAC, mais son attention ne s'est pas portée sur d'éventuels facteurs diachroniques.

Or, lorsque nous nous sommes intéressés à l'étude de la variation dans les DAC, l'une des premières étapes de notre réflexion nous a amené à interroger les locuteurs sur le sujet, notamment en ce qui concerne la différence entre les DAC quadrimoriques et trimoriques. Bien que la plupart d'entre eux ne voyaient aucune explication particulière à ce phénomène, certains nous ont répondu que le fait d'abréger en trois mores plutôt que quatre était une tendance récente. En outre, aucune autre forme d'explication ne nous a été proposée. Il se peut

que ce type d'affirmations n'ait pas de fondement concret sur le plan linguistique et qu'il s'agisse simplement d'une idée reçue transmise par certains locuteurs. Toutefois, si elle s'avérait exacte, cela pourrait signifier que certains facteurs diachroniques entrent en compte dans la variation observée dans les DAC. Dans notre étude précédente (Lamarque, 2015a), nous avons donc souhaité explorer cette piste suggérée par l'opinion des locuteurs.

En suivant l'exemple des travaux de Labrune, nous avons à notre tour constitué un corpus de 300 DAC. Afin de collecter des DAC aussi récents que possibles, nous nous sommes focalisés sur le média jeux-vidéo, qui reste à ce jour majoritairement utilisé par les plus jeunes générations. De plus, les jeux-vidéo contiennent une quantité considérable de mots d'emprunt divers et originaux, qui n'avaient donc que peu de chance de figurer dans la liste de ceux déjà collectés par Labrune.

Comme nous l'avons évoqué précédemment, le but était de comparer ce nouveau corpus (désormais « corpus jeux-vidéo ») au corpus plus ancien de Labrune (désormais « corpus non-jeune ») afin de vérifier, entre autres, si le traitement des mores spéciales /R/ et /Q/ en deuxième more de constituant avait évolué. Rappelons que le corpus non-jeune a été constitué aux alentours des années 2000 et contient des DAC d'époques et de sources très diverses. Il est impossible de déterminer précisément l'époque d'apparition de chacune des abréviations collectées. Tous les DAC du corpus jeux-vidéo ne sont pas nécessairement récents et vice-versa. Toutefois, afin de nous assurer que les DAC des deux corpus provenaient d'époques aussi différentes que possible, nous avons exclu du corpus non-jeune les abréviations clairement identifiées comme appartenant au langage des jeunes. C'est pourquoi nous avons retiré les 88 DAC issus du dictionnaire *Wakamono kotoba jiten* (dictionnaire du langage des jeunes)<sup>21</sup>.

#### 1.2.2.2 - Modifications préalables de la hiérarchie des contraintes

Nous verrons dans la section suivante (§1.2.2.3) que la comparaison des deux corpus nous a mené à modifier la hiérarchie des contraintes proposée par Labrune (2007, 2008) afin de la mettre à jour concernant les évolutions observées dans le corpus jeux-vidéo. Toutefois, il nous a semblé que certains ajustements pouvaient être apportés à la hiérarchie établie par Labrune en dehors de toute considération des aspects diachroniques. Nous souhaitons donc apporter quelques modifications à celle-ci. Précisons que certains de ces ajustements ne

---

<sup>21</sup> Par ailleurs, l'analyse de ces 88 DAC isolés nous a permis d'obtenir de premiers éléments de confirmation de notre hypothèse avant de nous lancer dans la création du corpus jeux-vidéo (voir Lamarque, 2015a).



figurent pas dans Lamarque (2015a). Ainsi nous proposons une version légèrement révisée de la hiérarchie de contraintes établie dans nos travaux précédents.

Le premier ajustement concerne la contrainte MAXLEX. Pour rappel :

MAXLEX : Chaque lexème de la base est présent dans le dérivé.

Cette contrainte permet d'assurer la distinction essentielle qui doit être effectuée entre les DAC et les DAS (voir §1.2.1.1). En effet, même si une base contient plusieurs lexèmes, si l'abréviation n'en retient qu'un seul, elle relève du processus de formation des DAS et non des DAC. Par exemple 'convenience (store)' [koNbinieNsu-(sutoa)] est abrégé en [koNbini] selon le processus de formation des DAS car, bien que la base contienne deux lexèmes, l'abréviation ne retient que le premier : 'convenience'.

Cependant, un problème se pose lorsqu'une base contient trois lexèmes ou plus. En postulant l'existence de cette contrainte, Labrune intègre les abréviations contenant trois constituants ou plus dans la catégorie des DAC. Il est vrai que la formation d'abréviations de DAC à trois constituants ou plus repose sur des principes similaires à ceux de deux constituants. Néanmoins, la présence d'un troisième constituant entraîne inévitablement l'infraction de certaines contraintes, notamment celles liées à la binarité des unités prosodiques. Si tous les DAC à trois constituants doivent figurer parmi les exceptions, nous préférons les considérer comme un groupe d'abréviations à part, suivant son propre processus d'abréviation. Nous sommes convaincu qu'une étude approfondie de ce type d'abréviations permettrait de mettre à jour des contraintes spécifiques à leur processus de formation.

De plus, nous trouvons dans le corpus jeux-vidéo plusieurs DAC dont la base contient plus de deux lexèmes tandis que l'abréviation n'en retient que deux. Ceci constituerait une infraction à MAXLEX, mais ces DAC sont par ailleurs formés de manière tout à fait canonique. Considérons l'exemple suivant :

'Funky Kong / super Bowser' > [ϕaNkiR-koNngu-suRpaR-**bauzaR**] > [ϕaN-bau]

Bien que la base contienne quatre lexèmes, seuls deux d'entre eux sont conservés dans l'abréviation. Si l'on considère uniquement ces deux lexèmes, la formation de ce DAC est canonique. Il nous semble donc contre-productif d'intégrer une contrainte dans la hiérarchie qui rendrait ce type d'abréviations irrégulières. Dans l'exemple ci-dessus, l'abréviation renvoie à l'association de deux concepts, eux-mêmes constitués de deux lexèmes. L'abréviation se

limite à conserver des éléments de chacun de ces deux concepts plutôt que de chacun des lexèmes qui les constituent. Pour prendre en compte ce type de phénomènes, nous considérerons comme DAC uniquement les abréviations qui retiennent exactement deux lexèmes de la base. Les abréviations contenant un nombre différent de lexèmes relevant d'un processus de formation différent. Plutôt que MAXLEX, nous postulons donc l'existence d'une contrainte LEX=2, inviolable, située au-dessus de MOTPROSODIQUEBINAIRE qui garantit que les DAC ne contiennent ni plus ni moins de deux lexèmes.

LEX=2 : l'abréviation est constituée d'exactly deux lexèmes.

La seconde modification que nous souhaitons apporter porte sur les contraintes PIEDBINAIRE et MORESAGAUCHE. Comme nous l'avons vu, la faiblesse de PIEDBINAIRE par rapport à CONTIG permet d'exclure l'option du remplacement du traitement de /R/ et /Q/ en deuxième more de C2. En revanche, la contrainte PIEDBINAIRE à elle seule ne permet pas de différencier l'option de la suppression de celle du maintien avec more supplémentaire, ce qui rend indispensable l'ajout de la contrainte MORESAGAUCHE. Or, cette contrainte d'alignement prosodique MORESAGAUCHE nous semble être un moyen quelque peu artificiel d'expliquer un phénomène assez simple : la suppression, créant des pieds unaires, est préférée au maintien avec more supplémentaire, qui crée des pieds ternaires. De plus, MORESAGAUCHE est systématiquement enfreinte par la très grande majorité des DAC qui contiennent des pieds binaires, c'est-à-dire, réguliers. Bien que cela ne soit pas un problème dans l'absolu avec une approche OT, étant donné qu'il s'agit d'une infraction mineure, nous pensons qu'il est préférable d'envisager une solution qui n'intègre pas de contrainte qui soit enfreinte de manière quasi-systématique.

Pour cela, nous retirons la contrainte MORESAGAUCHE de notre analyse et scindons en deux contraintes distinctes PIEDBINAIRE :

\*PIED>2 : un pied ne doit pas contenir plus de deux mores.

\*PIED<2 : un pied ne doit pas contenir moins de deux mores.

Tandis que la première est enfreinte dans le cas d'un maintien avec more supplémentaire, la seconde est enfreinte en cas de suppression. La préférence pour la suppression s'explique alors par le placement de \*PIED>2 au-dessus de \*PIED<2 dans la hiérarchie. Ainsi, si l'on reprend l'exemple du 'pocket board' vu dans le tableau 7, sa dérivation correspond au tableau suivant :

	[pokeQto-boRdo]	*V-HAUT:#	CONTIG	*PIED>2	*PIED<2	*V+HAUT:#
1.	(poke)-(boR)	*!				
2.	(poke)-(boRdo)			*!		
3.=>	(poke)-(bo)				*	
4.	(poke)-(bodo)		*!			

Tableau 9 : [pokeQto-boRdo] => [poke-bo] 'pocket board'

Cette fois, le candidat 3 est sélectionné parce qu'il commet l'infraction la moins coûteuse, c'est-à-dire, celle à \*PIED<2 par le choix de la suppression.

La hiérarchie que nous proposons, avant même de comparer les deux corpus, est donc la suivante :

PARSEμ, **LEX=2** >> MOTPROSODIQUEBINAIRE >> \*V-HAUT:#, LICENCE/Q/ >> CONTIG,  
 ANCRAGEGAUCHE >> \*PIED>2, DEP, IDENTTRAITS >> \*PIED<2 >> \*V+HAUT:#

Ainsi, l'ordre « \*V-HAUT:# >> CONTIG >> \*PIED>2 >> \*PIED<2 >> \*V+HAUT:# » favorise la suppression de /R/ lorsqu'il suit une voyelle non haute, puisque, \*V+HAUT:# n'étant pas concernée, \*PIED<2 est la contrainte la plus faible. De même, le maintien est préféré lorsque /R/ est suivi d'une voyelle haute car \*V+HAUT est la contrainte la plus faible. Cela correspond aux données du corpus non-jeune observées dans la section §1.2.1.3.

### 1.2.2.3 - Données du corpus jeux-vidéo

Si l'on s'intéresse à présent aux données du corpus jeux-vidéo, on peut constater que les traitements de /R/ et /Q/ lorsqu'ils apparaissent en deuxième more de C2 n'est pas le même que dans le corpus non-jeune.

/R/ : 91 cas			
En C1 : 37 cas (40,66%)		En C2 : 54 cas (59,34%)	
Formation canonique :		Formation canonique :	
Maintenu	10 cas (27,02%)	Maintenu	6 cas (11,11%)
Formations non-canoniques :		Formations non-canoniques :	
Maintenu +	0 cas (0%)	Maintenu +	0 cas (0%)
Supprimé	1 cas (2,70%)	Supprimé	42 cas (77,78%)
Remplacé	26 cas (70,27%)	Remplacé	4 cas (7,40%)
Ajouté	0 cas (0%)	Ajouté	2 cas (3,70%)

Tableau 10 : /R/ dans le corpus jeux-vidéo

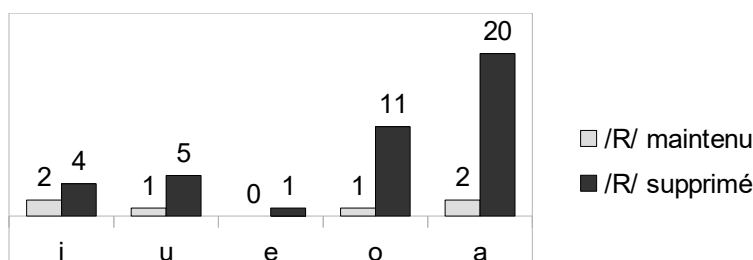


Diagramme 2 : traitement de /R/ en C2 en fonction du timbre de la voyelle. Corpus jeux-vidéo.

/Q/ : 33 cas			
En C1 : 12 cas (33,33%)		En C2 : 21 cas (66,66%)	
Formation canonique :		Formation canonique :	
Maintenu	6 cas (50%)	Maintenu	0 cas (0%)
Formations non-canoniques :		Formations non-canoniques :	
Maintenu +	0 cas (0%)	Maintenu +	0 cas (0%)
Supprimé	1 cas (8,33%)	Supprimé	18 cas (85,71%)
Remplacé	5 cas (41,67%)	Remplacé	3 cas (14,29%)
Ajouté	0 cas (0%)	Ajouté	0 cas (0%)

Tableau 11 : /Q/ dans le corpus jeux-vidéo

Si l'on observe tout d'abord le tableau 11, on s'aperçoit que le traitement de /Q/ est plus régulier encore que dans le corpus non-jeune. Lorsqu'il apparaît en deuxième more de C1, il est tantôt maintenu, tantôt supprimé. Conformément aux exigences de la contrainte LICENCE/Q/, les 6 cas dans lesquels il est maintenu correspondent à des abréviations dans lesquels le C2 débute par une consonne sourde, ce qui procure un licenciement correct pour /Q/. On remarque cependant que lorsqu'il n'est pas maintenu, il est remplacé, plutôt que supprimé comme l'aurait prédit notre hiérarchie de contraintes. En revanche, lorsque /Q/

apparaît en deuxième more de C2, il est supprimé dans la grande majorité des cas, ce qui est en accord avec la hiérarchie établie jusqu'à présent.

Concernant le traitement de /R/, on trouve de manière similaire à /Q/ une préférence pour le remplacement lorsqu'il apparaît en deuxième more de C1 et pour la suppression quand il apparaît en deuxième more de C2. Par ailleurs, le diagramme 2 montre que, contrairement aux faits observés dans le corpus non-jeune, le timbre de la voyelle qui précède /R/ n'influe que légèrement sur le traitement qu'il subit, la suppression étant préférée dans tous les cas. Cela remet en question la pertinence de la distinction entre les contraintes \*V+HAUT:# et \*V-HAUT:#. De plus, notre jeu de contraintes ne permet pas d'expliquer pourquoi /R/ est remplacé plutôt que maintenu en C1 dans le corpus jeux-vidéo, puisque ces contraintes ne concernent que la position finale de mot et non de constituant.

Afin de prendre en compte ces différences, nous proposons non seulement de réunir les contraintes \*V+HAUT:# et \*V-HAUT:#, mais d'élargir leur portée à la position finale de constituant plutôt que de mot. Ainsi, nous postulons l'existence de la contrainte :

\*V:#CONST : pas de voyelle allongée en finale de constituant.

Nous pouvons constater qu'au cours du temps, l'influence des contextes phonologiques identifiés par Labrune s'est régularisée. En effet, bien que les données du corpus non-jeune soit assez disparates, celles du corpus jeux-vidéo sont davantage concentrées sur une option. En C2, la suppression est l'option privilégiée dans la très grande majorité des cas, pour /R/ comme pour /Q/. Cela était déjà prévisible avec la hiérarchie de contraintes proposées par Labrune (2007, 2008). Seule la régularité avec laquelle s'appliquent ces contraintes a évolué au cours du temps.

Toutefois, en C1, les choses sont légèrement différentes, bien que /R/ était accepté en finale de ce constituant dans le corpus non-jeune, ce n'est plus le cas dans le corpus jeux-vidéo. Cela peut sembler être une évolution logique si l'on prend en compte le fait que le timbre de la voyelle a un moindre effet sur le traitement de /R/ dans le corpus jeux-vidéo. Tout cela nous incite à penser que le traitement de /R/ se généralise et se simplifie : dans le corpus non-jeune, son traitement dépendait non seulement du constituant dans lequel il apparaissait mais aussi du timbre de la voyelle qu'il suivait. En revanche, dans le corpus jeux-vidéo, il entraîne une formation non canonique dans tous les cas.

Il reste à expliquer pourquoi le traitement non canonique de /R/, comme celui de /Q/, est différent en C1 et en C2. En effet, le premier constituant semble favoriser le remplacement tandis que le second constituant préfère la suppression. Il semblerait donc que la binarité du pied de C1 soit plus importante que celle du C2. Cela avait déjà été observé par Labruno (2007) qui remarquait que la grande majorité des pieds non binaires apparaissaient en C2.

Pour rendre compte de cela, nous pouvons scinder les contraintes \*PIED>2 et \*PIED<2 en deux contraintes propres à chaque constituant. Ainsi :

\*PIEDC1>2 : les pieds de C1 ne doivent pas contenir plus de deux mores.

\*PIEDC1<2 : les pieds de C1 ne doivent pas contenir moins de deux mores.

\*PIEDC2>2 : les pieds de C2 ne doivent pas contenir plus de deux mores.

\*PIEDC2<2 : les pieds de C2 ne doivent pas contenir moins de deux mores.

Les deux premières de ces contraintes sont nécessairement placées au-dessus de CONTIG, ce qui permet d'expliquer la préférence pour le remplacement lorsqu'une more problématique apparaît en deuxième more de C1. Quand aux deux autres contraintes, elles prennent la place de \*PIED>2 et \*PIED<2, en suivant le même raisonnement qu'en §1.2.2.2.

Les tableaux 12 et 13 ci-dessous justifient l'ordre de ces nouvelles contraintes. Ils correspondent respectivement à la dérivation de [supa-koN] à partir de 'super combo' [suRpaR-koNbo] et de [tori-ru] à partir de 'trick room' [toriQku-ruRmu].

	[suRpaR-koNbo]	*V:#CONST	*PC1>2	*PC1<2	CONTIG	*PC2>2	*PC2<2
1.	(suR)-(koN)	*!					
2.	(suRpa)-(koN)		*!				
3.	(su)-(koN)			*!			
4.=>	(supa)-(koN)				*		

Tableau 12 : [suRpaR-koNbo] => [supa-koN] 'super combo'

	[toriQku-ruRmu]	*V:#CONST	*PC1>2	*PC1<2	CONTIG	*PC2>2	*PC2<2
1.	(tori)-(ruR)	*!					
2.	(tori)-(ruRmu)					*!	
3.=>	(tori)-(ru)						*
4.	(tori)-(rumu)				*!		

Tableau 13 : [toriQku-ruRmu] => [tori-ru] 'trick room'

Dans le tableau 12, la more problématique apparaît en C1. Les candidats 2 et 3 sont donc éliminés du fait de leur infraction respective à \*PIEDC1>2 et \*PIEDC1<2. Le candidat 4 est retenu malgré son infraction à CONTIG qui est alors considérée comme minimale. En revanche, dans le tableau 13, la more problématique apparaît en C2. Puisque la contrainte \*PIEDC2<2 est celle qui est placée le plus bas, c'est la suppression (candidat 3) qui est alors considérée comme optimale.

La hiérarchie que nous proposons est donc la suivante :

PARSEM, LEX=2 >> MOTPROSODIQUEBINAIRE >> \*V:#CONS, LICENCE/Q/, \*PIEDC1>2 >> \*PIEDC1<2 >> CONTIG, ANCRAGEGAUCHE >> \*PIEDC2>2, DEP, IDENTTRAITS >> \*PIEDC2<2

De manière générale, cette étude par comparaison de corpus a permis de montrer que le traitement de /Q/ et de /R/ en deuxième more de constituant s'était régularisé au cours du temps. En C2, cette régularisation a entraîné une forte augmentation de la proportion de suppressions. Or, c'est précisément les cas de suppressions qui entraînent la formation de DAC trimoriques. En toute logique, le corpus jeux-vidéo devrait donc contenir davantage de DAC trimoriques que le corpus non-jeune. Cela est confirmé par la figure ci-dessous :

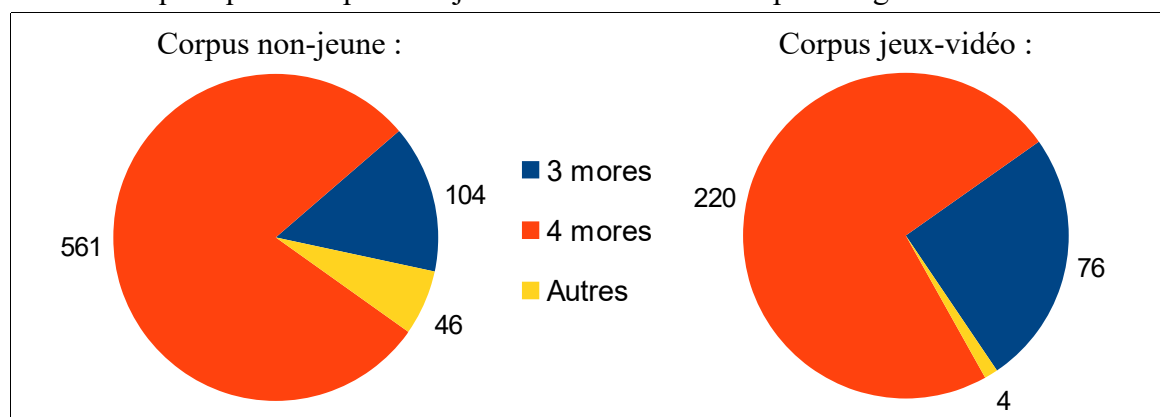


Diagramme 3 : Proportion de DAC en fonction du nombre de mores dans chaque corpus

Nous avons pu ainsi confirmer notre hypothèse selon laquelle les DAC trimoriques seraient plus fréquents parmi les abréviations les plus récentes. Cette hypothèse nous ayant été proposée par les locuteurs, il est possible que certains d'entre eux soient sensibles à cette évolution, sans pour autant être conscients des facteurs phonologiques qui entrent en jeu dans celle-ci. Quoi qu'il en soit, il semblerait que la variation dans la formation des DAC ne puisse

s'expliquer entièrement selon des critères phonologiques ou prosodiques, mais que certains aspects diachroniques doivent également être pris en compte.

Pour résumer, les précédents travaux par comparaison de corpus nous ont permis de comprendre qu'elles étaient les raisons phonologiques qui entraînent la formation de DAC trimoriques ainsi que de tester notre hypothèse de l'évolution de l'organisation hiérarchique de certaines contraintes qui entrent en jeu dans la formation des DAC. Toutefois, il n'est pas impossible que les faits observés sur le corpus jeux-vidéo soient en réalité spécifiques à une certaine communauté de locuteurs. Nous souhaitons donc vérifier si ces observations pouvaient véritablement être attribuées à une évolution de la phonologie des DAC en choisissant une approche plus expérimentale. Nous avons donc opté pour un questionnaire en ligne, qui nous permettrait de collecter un maximum de données auprès de locuteurs d'horizons aussi différents que possible.

Le but de ce questionnaire est de vérifier l'impact de divers contextes phonologiques sur la formation des abréviations et d'évaluer si l'âge des locuteurs a une influence sur leur traitement de ces contextes phonologiques particuliers. Dans les sections qui suivent, nous verrons tout d'abord les aspects méthodologiques de la création de ce questionnaire (§2), puis nous présenterons les résultats obtenus (§3) avant de les discuter (§4).

## ***2 - Aspects méthodologiques du questionnaire en ligne***

### ***2.1 - Sélection des items***

Les mots d'emprunt qui ont été retenus pour notre questionnaire n'ont pas été choisis au hasard mais ont évidemment été sélectionnés selon les critères phonologiques qui nous intéressent. Dans la mesure du possible, nous avons sélectionné des constituants qui apparaissent dans les deux corpus à notre disposition en les réarrangeant avec d'autres constituants que ceux avec lesquels ils sont associés dans les corpus. Toutefois, pour certains contextes phonologiques, les corpus ne contenaient pas assez d'items. Nous avons donc recherché sur des dictionnaires spécialisés en ligne pour compléter notre liste de 100 items.



### **2.1.1 - Les premiers constituants**

Ce questionnaire se focalise essentiellement sur l'influence du contenu phonologique du deuxième constituant sur la création de l'abréviation. Bien que nous ayons vu dans la section §1.2 que le premier constituant pouvait également avoir une influence sur la formation du DAC, c'est essentiellement au niveau du traitement du deuxième constituant que la formation de DAC non quadrimoriques se réalise. Ainsi, dans la majorité des cas, nous avons sélectionné pour premiers constituants des mots d'emprunt ne présentant *a priori* aucune particularité susceptible d'entraîner une quelconque irrégularité dans la formation du DAC. Les C1 sont donc théoriquement interchangeables les uns avec les autres sans que cela n'ait d'influence sur nos résultats. Afin de limiter au maximum l'effet que pourrait avoir le C1, nous avons interverti, dans deux des quatre versions du questionnaire, les C1 de quatre groupes de 10 items. Les quatre groupes en question traitent de contextes phonologiques similaires ce qui rendait d'autant plus indispensable le fait d'écarter toute influence éventuelle des premiers constituants. Par exemple, les items 1 à 20 portent tous sur le traitement de /R/ en deuxième more de C2 (voir §2.1.3), toutefois, dans les items 1 à 10, /R/ est précédé d'une voyelle haute, tandis que dans les items 11 à 20, il est précédé d'une voyelle non haute. Ainsi, en inversant les C1 des items 1 à 10 avec ceux des items 11 à 20 dans deux versions du questionnaire, nous pouvons neutraliser toute influence éventuelle de ces constituants sur la formation du DAC. De même, les premiers constituants des items 21 à 30 ont été intervertis avec ceux des items 31 à 40 (voir §2.1.4) dans deux versions.

En résumé, bien que les versions 1, 2, 3 et 4 possèdent chacune un ordre d'apparition des items différents, les C1 des items 1 à 40 sont différents dans les versions 1 et 2 d'une part et les versions 3 et 4 d'autre part. Puisque les premiers constituants peuvent varier pour certains items, nous définissons ces derniers uniquement par leur deuxième constituant. Par exemple, l'item 6 correspond au même C2 'cheat', qu'il soit précédé par le C1 'millionnaire', comme dans les versions 1 et 2, ou par 'voice', comme dans les versions 3 et 4.

### **2.1.2 - Items de contrôle (items 91 à 100)**

Les items 91 à 100 ne possèdent pas de caractéristiques phonologiques particulières susceptibles d'entraîner la formation d'exceptions. Ils servent ainsi de contrôle pour déterminer si les participants ont correctement effectué la tâche qui leur été demandée. Ils permettent également de vérifier si les participants présentent des habitudes d'abréviation

inattendues dans des contextes qui ne posent *a priori* aucun obstacle à la formation d'un DAC canonique. De manière générale, les résultats obtenus sur ces items peuvent servir de référence pour diverses comparaisons avec les autres groupes d'items.

### **2.1.3 - Traitement de /R/ en deuxième more de C2 (items 1 à 20)**

Les items 1 à 20 sont consacrés au traitement de /R/ lorsqu'il apparaît en deuxième more de C2. Tout comme dans les études sur corpus décrites dans la section §1.2, nous nous intéressons ici principalement à quatre possibilités pour le traitement de /R/. Ces quatre traitements possibles commettent tous des infractions aux contraintes que nous avons définies dans la section §1.2. À titre d'exemples, nous donnerons pour chacun des traitements le résultat qu'il produit sur l'item 2 : 'cancel' [kjaNseru] + 'super' [suRpaR]. Les quatre traitements qui retiendront notre attention sont :

- Le maintien : en infraction à \*V#CONST. (ex : [kjaN-suR])
- Le maintien avec more supplémentaire : en infraction à \*PIED>2<sup>22</sup>. (ex : [kjaN-suRpa])
- La suppression : en infraction à \*PIED<2. (ex : [kjan-su])
- Le remplacement : en infraction à CONTIG. (ex : [kjan-sup])

Les résultats du questionnaire pour ces items devraient donc nous permettre de vérifier la validité de la hiérarchie entre les contraintes concernées. De plus, nous pourrions vérifier si l'âge des participants a une influence sur le traitement qu'ils favorisent.

Par ailleurs, nous avons souhaité inclure dans cette partie du questionnaire une distinction entre les voyelles hautes et les voyelles non-hautes. Cela nous permettra de vérifier si le timbre de la voyelle a une influence sur le traitement de /R/, comme suggéré par les travaux de Labrune (2007, 2008), ou si, au contraire, son traitement reste sensiblement le même peu importe la voyelle qu'il suit, comme c'était le cas dans le corpus jeux-vidéo. Ainsi, dans les items 1 à 10, /R/ est toujours précédé d'une voyelle haute (/i/ ou /u/), tandis que dans les items 11 à 20, il est précédé d'une voyelle non-hautes (/a/, /o/ ou /e/)

Rappelons que les premiers constituants des items 1 à 10 sont intervertis avec ceux des items 11 à 20 dans deux des quatre versions du test, afin de limiter l'influence de ces derniers

---

<sup>22</sup> Le questionnaire ne portant que sur le C2, nos résultats ne pourront confirmer la pertinence d'une distinction en deux contraintes relatives à chaque constituant des contraintes \*PIED>2 et \*PIED<2. Aussi préférons ne pas faire cette distinction ici.

sur nos résultats.

#### **2.1.4 - Traitement des bilabiales à l'initiale de C2 (items 21 à 40)**

La présence d'une bilabiale à l'initiale de C2 est un contexte phonologique particulièrement intéressant pour l'étude de la variation dans la formation des DAC. En effet, dans notre étude par comparaison de corpus, nous avons pu constater que la présence d'une bilabiale à l'initiale de C2 pouvait avoir une légère influence sur la formation d'exceptions, notamment trimoriques. Cette influence n'étant que très légère dans les corpus, il nous était impossible d'envisager d'inclure de nouvelles contraintes à la hiérarchie qui justifieraient l'apparition de ces exceptions car, dans la majorité des cas, une bilabiale à l'initiale de C2 n'empêche pas la création d'un DAC canonique. Toutefois, les différents corpus n'offrent que peu de données sur lesquelles se baser pour arriver à des conclusions satisfaisantes. C'est pourquoi nous avons souhaité tester ce contexte phonologique également dans notre questionnaire.

Les items 21 à 40 se décomposent en deux groupes de 10 :

- Les items 21 à 30, qui présentent non seulement une bilabiale à l'initiale de C2 mais également /R/ en deuxième more de C2 (ex : 'ball' [**bo.R.ru**]).
- Les items 31 à 40, qui présentent seulement une bilabiale à l'initiale de C2 (ex : 'volt' [**bo.ru.to**]).

De plus, nous avons sélectionné pour chaque série de 10, deux items débutant par chacune des cinq consonnes bilabiales du japonais (/b/, /p/, /m/, /w/ et /h/<sup>23</sup>), que nous avons fait suivre par les mêmes voyelles dans les deux listes. Les items 21 et 31 débutent par /bo/ ; les 22 et 32 par /bi/, les 23 et 33 par /pa/ ; les 24 et 34 par /pi/ ; etc. Notons que l'on trouve, pour chaque paire, une voyelle non-haute et une voyelle haute, afin de limiter l'influence éventuelle de ce facteur sur nos résultats.

Ainsi, en comparant d'une part les items 21 à 30 avec les items 1 à 20 et d'autre part les items 31 à 40 avec les items 91 à 100, nous pourrions vérifier si la présence d'une bilabiale à l'initiale de C2 entraîne la formation de DAC exceptionnels de manière significative, et ce

---

<sup>23</sup> Nous comptons /h/ parmi les consonnes bilabiales pour différentes raisons étymologiques, notamment évoquées en §1.1.3.3.

dans deux contextes différents. Bien entendu, l'âge des locuteurs sera également pris en compte dans l'analyse des résultats.

### **2.1.5 - Traitement de /Q/ en deuxième more de C2 (items 41 à 60)**

Les items 41 à 60 sont consacrés à l'étude du traitement de /Q/ lorsqu'il apparaît en deuxième more de C2. Dans ce contexte, il est impossible qu'un DAC canonique soit formé, car /Q/ se retrouverait en position finale de l'abréviation. Or, /Q/ ne peut pas être réalisé s'il n'est pas suivi d'une consonne. Nous trouvons cependant quelques exemples pour lesquels les participants ont, d'une certaine manière, choisi de maintenir /Q/ en position finale avec la réalisation [t̥su]. Cela ne peut s'expliquer qu'à travers le prisme du système d'écriture du japonais : dans les syllabaires, le caractère pour noter /Q/ est ツ en *katakana* et っ en *hiragana*. Or, ces caractères correspondent au caractère < tu > [t̥su] réduit à la moitié de sa taille, soit ツ en *katakana* et っ en *hiragana*<sup>24</sup>. Nous avons donc, comme pour /R/, quatre traitements possible pour la more /Q/. Chacun de ces traitements est présenté avec l'exemple du résultat qu'il produit pour l'item 51 : 'multi' [maru<sup>t̥</sup>ei] + 'god' [goQdo].

- Le maintien : en infraction à DEP. (ex : [maru-gotsu])
- Le maintien avec more supplémentaire : en infraction à \*PIED>2. (ex : [maru-goQdo])
- La suppression : en infraction à \*PIED<2. (ex : [maru-go])
- Le remplacement : en infraction à CONTIG. (ex : [maru-godo])

Tout comme les items 1 à 20, les items 41 à 60 nous permettrons de vérifier l'ordre hiérarchique entre les contraintes concernées. De plus, nous avons séparé les items portant sur le traitement de /Q/ en deuxième more de C2 en deux groupes de 10. Dans les items 41 à 50, le deuxième constituant, dont la deuxième more est /Q/, est un mot contenant plus de trois mores. En revanche, dans les items 51 à 60, tous les C2 sont des mots trimoriques correspondant à une syllabe lourde en anglais. Labrune (2007) remarquait que dans ce contexte, la proportion de maintien avec more supplémentaire était plus importante. Toutefois, cette observation n'était pas valable pour le corpus jeux-vidéo. Cette distinction devrait donc nous permettre de vérifier si cela a une influence ou non sur la formation des DAC. Il s'agit là d'un phénomène particulièrement intéressant car c'est le seul cas de figure dans lequel la

---

<sup>24</sup> On trouve dans le corpus de Labrune de très rares exceptions de ce type qui attestent de la réalisation [t̥su]. Par exemple, 'world cup' [waRudo-kaQpu] est abrégé en [waR-ka<sup>t̥</sup>su].

prosodie de la langue d'origine pourrait avoir une influence sur la formation du DAC. Nous vérifierons également si l'âge entre en jeu dans le choix du traitement de /Q/ et notamment dans le cas de la présence d'un C2 trimorique.

### **2.1.6 - Traitement du hiatus entre le C1 et le C2 (items 61 à 80)**

La présence d'une voyelle à l'initiale du deuxième constituant est un contexte phonologique qui n'a été que peu étudié dans les études précédentes sur corpus, bien que son influence ait été remarquée. Dans ce cas de figure, la création du DAC entraîne l'apparition d'un hiatus entre la voyelle finale du C1 et celle à l'initiale de C2. Or, le hiatus est évité, de manière générale, dans les langues du monde. Il n'est donc pas étonnant que ce contexte phonologique ait une influence sur la formation des DAC. Par conséquent, les items 61 à 80 sont consacrés à l'étude du traitement du hiatus.

Pour éviter la formation du hiatus, les locuteurs doivent trouver une solution pour se débarrasser de la voyelle à l'initiale de C2. Il existe pour ce faire deux possibilités : soit enfreindre la contrainte ANCRAGEGAUCHE et ne pas conserver la more initiale de C2, ou bien procéder à un télescopage. Le télescopage est un type d'abréviation qui fonctionne de manière totalement différente des DAC. Tout d'abord, contrairement aux DAC, qui conservent les mores initiales de chaque constituant selon un processus proche de l'acronymie, les télescopages conservent d'une part le début du premier constituant et d'autre part la fin du deuxième constituant. Par ailleurs, le processus de formation des télescopages ne tient pas nécessairement compte du nombre de mores retenues dans l'abréviation. Ainsi en trouve-t-on de taille très variable. Dans notre étude, nous désignons par « télescopage » plusieurs types d'abréviations qui ont pour point commun de conserver le début du C1 et la fin du C2. Nous illustrerons ces différents types avec des exemples de réponses données pour l'item 68 : 'customize' + 'italian' [kasutamaizu] + [itariaN] :

- Télescopage avec superposition totale d'une more : le C1 et le C2 contiennent une more identique qui va servir de jonction entre les deux constituants. Ainsi, on ne peut déterminer si la more en question appartient au C1 ou au C2. (ex : [kasutariaN])
- Télescopage avec superposition partielle : la consonne d'attaque d'une more du C1 se retrouve suivie d'une voyelle d'une more de C2. Il n'y a pas de chevauchement au niveau des phonèmes, mais la more qui sert de jonction entre le C1 et le C2 est

partagée entre les deux constituants. (ex : [kasut**am**itariaN]).

- Télescopage sans superposition : la distinction entre le C1 et le C2 est clairement identifiable. (ex: [kasutama-riaN] ; [ka-tariaN] ; [kasuta-N] ; [kasutama-N] ; [kasu-riaN] )
- Télescopes irréguliers : l'abréviation contient une forme de superposition, typique des télescopes, mais certaines mores de C1 qui précèdent la jonction ou certaines mores de C2 qui suivent la jonction ne sont pas retenues dans l'abréviation. Le statut de télescopage de certaines de ces formes est discutable puisqu'elles ne respectent pas toujours le critère définitoire que nous avons retenu : conserver le début de C1 et la fin de C2. Toutefois, nous préférons compter ces formes parmi les télescopes pour les dissocier des DAC, dans lesquels la superposition n'est pas supposée intervenir. (ex : [kasut**am**ita])

Notons qu'il est parfois impossible de définir si l'abréviation observée relève du télescopage ou du DAC en infraction à ANCRAGEGAUCHE. C'est le cas par exemple lorsque le C2 est trimorique dans la base et que seules deux mores du C1 sont conservées. Par exemple, pour l'item 63 'flower + office' [ɸurawaR] + [oɸisu], si le participant répond [ɸura-ɸisu], nous ne pouvons déterminer avec certitude s'il s'agit d'un télescopage ou d'une infraction à ANCRAGEGAUCHE. Quoi qu'il en soit, l'une des deux stratégies a été employée pour éviter la formation d'un hiatus.

Pour les items 61 à 80 nous nous intéresserons donc principalement aux deux questions suivantes : 1) Le hiatus est-il maintenu ou évité ? 2) S'il est évité, par quel procédé (infraction à ANCRAGEGAUCHE ; télescopage ; ou indéterminé) ?

Pour répondre à ces questions, nous étudierons trois configurations différentes :

- Dans les items 61 à 70 : le C1 est un mot d'emprunt sans particularité et le C2 débute par une voyelle.
- Dans les items 71 à 75 : le C1 et le C2 débutent tous les deux par une voyelle identique. Nous supposons que cette ressemblance entre les initiales des deux constituants pourrait encourager les participants à créer davantage de télescopes.
- Dans les items 76 à 80 : le C1 et le C2 débutent tous les deux par une voyelle identique suivie de /R/ (une voyelle allongée). Dans ce cas de figure, l'infraction à ANCRAGEGAUCHE est d'autant plus défavorisée qu'une double infraction serait

nécessaire pour éviter le hiatus ou un enchaînement de type /VRR/.

Comme pour tous les items de notre questionnaire une attention particulière sera portée à l'âge des participants et à l'influence qu'il pourrait avoir sur le choix du maintien ou non du hiatus ainsi que du processus employé pour éviter l'apparition de celui-ci.

### **2.1.7 - Traitement de /R/ en C1 et de /Q/ en C2 (items 81 à 90)**

Pour finir, nous nous intéresserons, avec les items 81 à 90, à un phénomène que nous avons pu observer de manière assez rare dans les corpus. Il s'agit de télescopages particuliers dans lesquels il semblerait que les mores /R/ et /Q/ se superposent. Cela se produit de manière exceptionnelle lorsque C1 contient une more /R/ et C2 une more /Q/. Par exemple, l'item 81 'universal' + 'pops' [junibaR<sup>s</sup>aru] + [poQ<sup>p</sup>usu] peut être abrégé en [junibaQ<sup>p</sup>usu]. Dans ce cas, on peut supposer que la more /Q/ s'est superposée sur la more /R/ du C1 bien qu'il ne s'agisse pas de mores identiques.

Ce phénomène ne nous semble pas particulièrement pertinent pour la compréhension du fonctionnement des DAC, car ce type d'abréviation reste extrêmement rare. Toutefois, il nous semble intéressant à étudier pour comprendre les liens qui peuvent exister entre /R/ et /Q/ dans une analyse auto-segmentale du japonais (voir section §1.1.3.5).

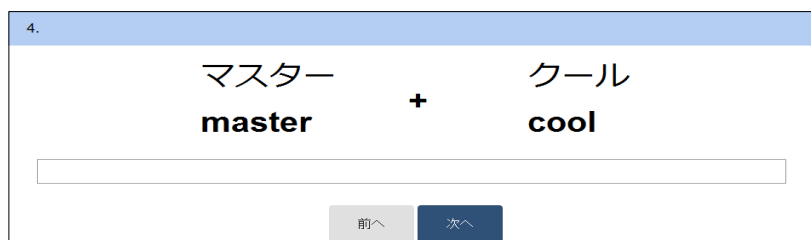
Le cas des télescopages avec superposition de /Q/ sur /R/ pourrait être un autre exemple intéressant nous permettant d'étudier les liens entre ces deux segments spéciaux. Dans les items 81 à 90, la place de /R/ dans le premier constituant est variable. Cependant, dans les items 81 à 85, /Q/ est toujours suivi de deux mores dans le C2, tandis que dans les items 86 à 90, /Q/ est suivi d'une seule more. Nous pourrions donc vérifier si ces caractéristiques ont une influence sur la formation de télescopages avec superposition de /Q/ sur /R/. Bien entendu, l'âge des locuteurs sera également un facteur pris en compte dans l'analyse de ce phénomène.

## **2.2 - Aspects techniques**

Afin d'obtenir un maximum de réponses auprès de locuteurs d'origines aussi différentes que possibles, il était indispensable que notre questionnaire soit diffusé via internet. Nous avons donc choisi dans un premier temps un outil spécialisé dans la conception d'enquêtes et de sondages sur internet. Nous avons opté pour la plate-forme SurveyMonkey.

La tâche principale de ce questionnaire consiste à former des abréviations à partir de

deux mots empruntés à l'anglais et qui, d'ordinaire, n'apparaissent pas de manière conjointe dans la langue japonaise. Pour cela, les locuteurs se sont vus présentés une centaine de paire de mots d'emprunt, écrits sous leur forme empruntée, en *katakana*<sup>25</sup>, et sous leur forme d'origine en anglais, en alphabet latin<sup>26</sup>. L'illustration 1 ci-dessous montre de quelle manière les items étaient présentés aux participants.



The screenshot shows a questionnaire item. At the top left, there is a small blue box with the number '4.'. Below this, the word 'マスター' (katakana) is displayed above the word 'master' (Latin script). To the right of 'master' is a plus sign '+', followed by the word 'クール' (katakana) above the word 'cool' (Latin script). Below these words is a long, empty text input field. At the bottom of the input area, there are two buttons: a grey button labeled '前へ' (Previous) and a blue button labeled '次へ' (Next).

Illustration 1 : exemple d'un item du questionnaire en ligne

Avant de débiter cette tâche, les participants devaient tout d'abord lire et accepter un formulaire de consentement. Puis ils devaient fournir quelques informations personnelles : âge ; sexe ; occupation ; plus haut niveau d'étude ; lieu de résidence ; langue(s) maternelle(s) ; langue(s) parlée(s) ; séjours de longue durée à l'étranger.

Les instructions pour répondre à la tâche leur étaient ensuite exposées, en voici une traduction :

« Des paires de mots empruntés à l'anglais vont vous être présentées en alphabet et en *katakana*. Veuillez créer une abréviation en vous basant sur ces deux mots. Certaines combinaisons de mots peuvent n'avoir aucun sens. Ne vous préoccupez pas du sens et créer l'abréviation qui vous semble la plus naturelle.

Vous êtes libre de créer n'importe quel type d'abréviation, néanmoins, vous devez respecter ces deux règles :

- 1 – L'ordre des mots ne doit pas être modifié.
- 2 – L'abréviation doit contenir des éléments des deux mots présentés.

Si possible, veuillez écrire vos réponses en *katakana*.

Le questionnaire va maintenant commencer. »

Les participants voyaient ensuite s'afficher les trois premiers items, accompagnés du message suivant :

<sup>25</sup> Les *katakana* constituent le syllabaire japonais utilisé principalement dans l'écriture des mots d'emprunts.

<sup>26</sup> A l'exception des formes d'origine en anglais, l'ensemble du questionnaire était rédigé en japonais.

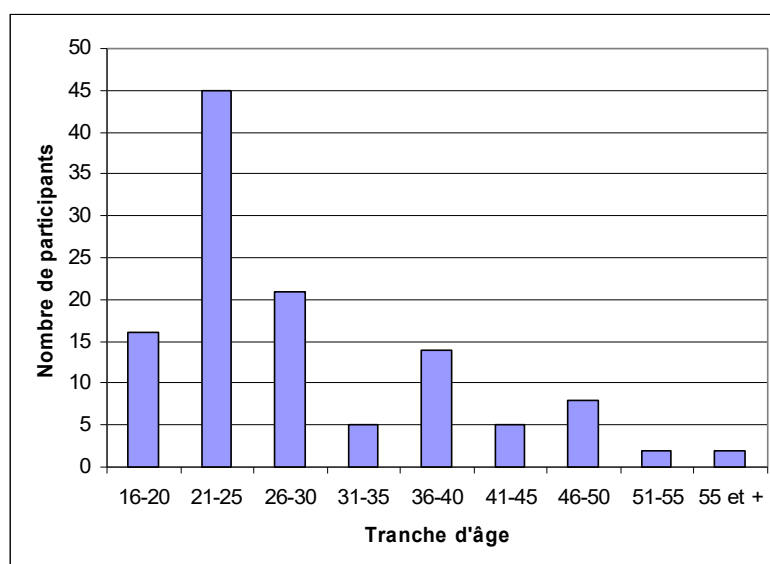


« Commençons le questionnaire avec ces trois questions. Les questions apparaîtront ensuite une par une et vous ne pourrez pas revenir en arrière après avoir cliqué sur le bouton "suivant" »

En dehors de ces trois premières questions, il n'était donc pas possible pour les participants de modifier leurs réponses.

Comme nous l'avons vu plus dans la section §2.1., les 100 items sélectionnés sont répartis en 10 groupes de 10 items contenant un contexte phonologique particulier. Ainsi, pour éviter que les locuteurs ne s'habituent à un contexte particulier, l'ordre d'apparition des 100 items a été rendu aléatoire. L'outil SurveyMonkey ne permettant pas d'établir un ordre aléatoire pour chaque participant, nous avons créé quatre versions du questionnaire dans lesquelles les items apparaissent dans des ordres différents (définis aléatoirement). Chaque participant a ensuite été dirigé vers l'une de ces quatre versions de manière aléatoire.

Le questionnaire a été diffusé en ligne via e-mails et les réseaux sociaux, ce qui permettait à n'importe qui de participer. Les réponses ont été collectées sur une durée d'environ un mois (du 08/03/2016 au 10/04/2016). Durant cette période, 118 participants ont répondu à notre questionnaire. L'histogramme ci-dessous présente la répartition des participants en fonction de leur âge :



Histogramme 1 : répartition des participants en fonction de l'âge

Comme nous pouvions nous y attendre, nous avons eu de nombreuses réponses de participants situés dans les tranches d'âge les plus jeunes et moins de réponses de la part de participants plus âgés. Notons enfin que tous les participants n'ont pas fourni de réponses à

l'ensemble des 100 items. Les participants étaient libres de passer des questions sans y répondre et d'interrompre le questionnaire à tout moment.

### ***3 - Résultats***

Dans les sections qui suivent, nous présenterons les résultats pour chaque groupe d'items décrits dans les sections précédentes. Nous souhaitons tout d'abord représenter visuellement les réponses de manière individuelle, en considérant le facteur de l'âge comme un continuum. Cependant, comme plusieurs participants n'ont fourni que peu ou pas du tout de réponses pour certains groupes d'items, cela entraînait certains biais dans les résultats. Nous avons donc décidé de regrouper les locuteurs par tranches d'âge. Nous distinguons quatre groupes : les 16-25 ans ; les 26-35 ans ; les 36-45 ans et les 46 ans et plus. Notons que cette distinction n'est utile que pour fournir une représentation visuelle des données dans les graphiques et n'a pas été prise en compte dans nos analyses statistiques.

Nous ne reviendrons pas spécifiquement sur les résultats des items 91-100 qui servent d'items de contrôle.

#### ***3.1 - Traitement de /R/ en deuxième more de C2 (items 1 à 20)***

Pour rappel, les items 1 à 20 portent sur le traitement de /R/ lorsqu'il apparaît en deuxième more de C2. Dans les items 1 à 10, /R/ est suivi d'une voyelle haute, tandis que dans les items 11 à 20, il est suivi d'une voyelle non-haute. Nous nous intéressons principalement aux quatre options de traitement que sont : le maintien ; le maintien avec more supplémentaire (maintien +) ; la suppression ; et le remplacement. Les figures 1 et 2 ci-dessous présentent les résultats obtenus pour les items 1-20.

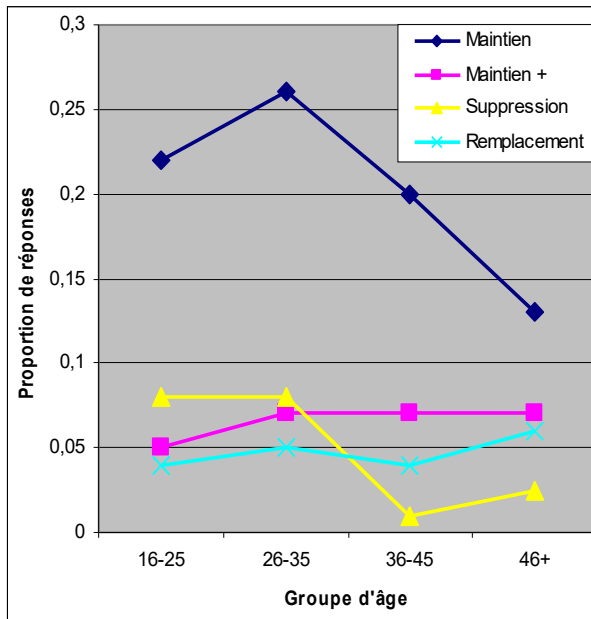


Fig. 1 : Traitement de /R/ dans les items 1-10 (/R/ précédé d'une voyelle haute)

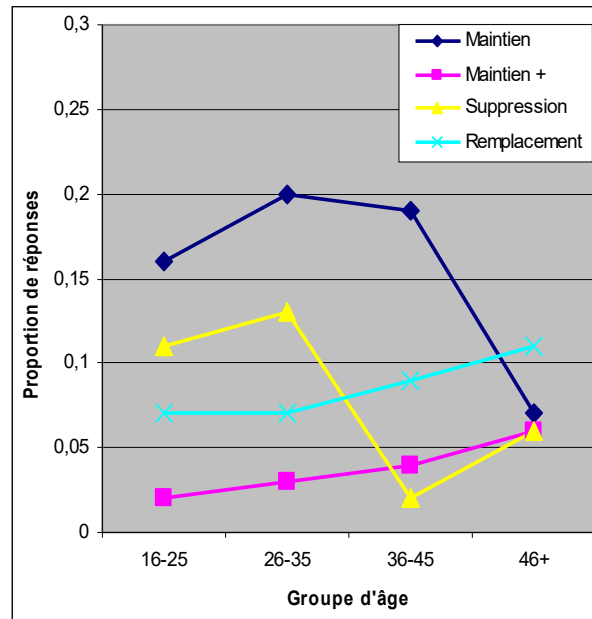


Fig. 2 : Traitement de /R/ dans les items 11-20 (/R/ précédé d'une voyelle non-haute)

Ces deux graphiques nous permettent d'observer que la proportion de maintiens et de suppressions parmi l'ensemble des DAC créés pour les items 1 à 20 diminue avec l'âge des locuteurs. Au contraire, on peut observer une légère augmentation du maintien avec more supplémentaire et du remplacement. En dehors de la diminution du nombre de suppressions, ces observations ne vont pas dans le sens des faits observés précédemment sur corpus.

Toutefois, l'analyse statistique de ces données montre que la plupart de ces changements ne sont pas significatifs<sup>27</sup>. Nous obtenons néanmoins une valeur p inférieure à 0,001 pour la diminution de la proportion de suppression en fonction de l'âge, que ce soit pour les items 1 à 10 ou 11 à 20. Cela semble indiquer une forte corrélation entre l'âge des participants et leur préférence pour cette option.

De plus, pour les items 11-20, l'augmentation de la proportion de maintiens avec more supplémentaire est également significative ( $p < 0,05$ ). Cette légère augmentation pourrait bien servir à contrebalancer la diminution du nombre de suppression, bien que nous nous attendions à ce que le maintien simple remplisse ce rôle.

Si l'on compare les deux graphiques l'un avec l'autre, on peut évaluer l'influence du

<sup>27</sup> Ces tendances ont été évaluées statistiquement en utilisant un modèle linéaire mixte généralisé à l'aide du pack lme4 du logiciel R (Bates & al., 2015). Ici, l'âge est le facteur fixe tandis que le choix des différents traitements est la variable dépendante. Le numéro d'item est un facteur aléatoire. L'ensemble des analyses statistiques de ce travail ont été réalisées selon le même modèle et avec le même facteur aléatoire. Désormais, nous précisons uniquement en notes de bas de page quels étaient le facteur fixe et la variable dépendante pour chaque analyse statistique à venir.

timbre de la voyelle sur le traitement de /R/. Comme nous le supposions, on peut observer que /R/ est plus rarement maintenu et plus fréquemment supprimé lorsqu'il suit une voyelle non-haute (items 11 à 20). Néanmoins, cette différence n'est pas statistiquement significative<sup>28</sup>.

### 3.2 - Traitement des bilabiales à l'initiale de C2 (items 21 à 40)

Les items 21 à 40 avaient pour objectifs de mesurer l'influence de la présence d'une bilabiale à l'initiale de C2 sur la formation du DAC. Rappelons que, dans les items 21 à 30, la deuxième more de C2 est /R/, tandis que dans les items 31 à 40, il s'agit d'une more de type (C)V. Dans notre analyse, nous avons donc comparé d'une part les items 21 à 30 aux items 1 à 20, et d'autre part les items 31 à 40 aux items 91 à 100 (items de contrôle).

Voyons ce qu'il en est de cette première comparaison à l'aide des figures 3 et 4 ci-dessous :

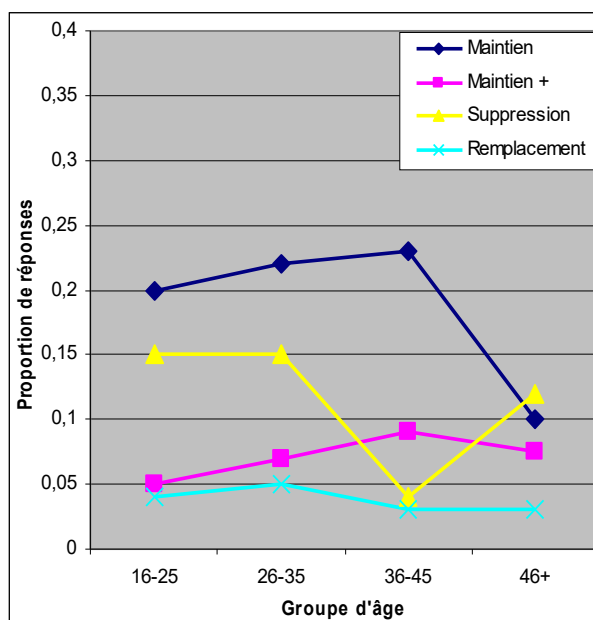


Fig. 3 : Traitement de la deuxième more de C2 dans les items 21-30 (bilabiale à l'initiale de C2 suivie de /R/)

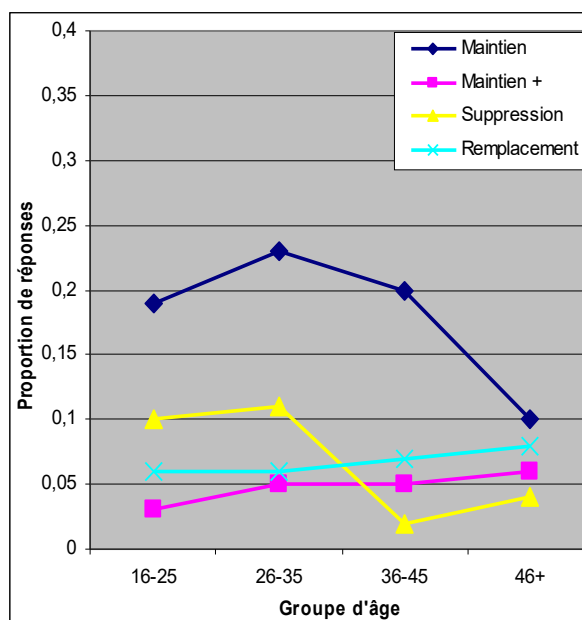


Fig. 4 : Traitement de la deuxième more de C2 dans les items 1-20 (/R/ suivi de voyelles hautes et non-hautes)

Penchons-nous tout d'abord sur la figure 4. Celle-ci regroupe les données des items 1 à 10 ainsi que 11 à 20 étudiés dans la section précédente. Il est donc logique qu'elle présente des données similaires. On trouve ainsi une diminution de la proportion de maintiens et de suppressions avec l'âge et, au contraire, une augmentation de la proportion de maintiens avec more supplémentaire et de remplacement.

28 Facteur fixe : groupe d'items

Variable dépendant : nombre de maintiens/suppressions

L'analyse statistique de ces données montre une fois de plus que la diminution du nombre de suppressions avec l'âge est significative ( $p < 0,001$ ). Il en va de même pour la légère augmentation du nombre de maintiens avec more supplémentaire ( $p < 0,05$ )<sup>29</sup>.

La figure 3 présente des résultats assez similaires si ce n'est pour le remplacement qui semble diminuer légèrement plutôt que d'augmenter. Cette observation n'est, quoi qu'il en soit, pas significative statistiquement. En revanche, comme pour les items 1 à 20, la diminution de la proportion de suppressions et l'augmentation de celle du maintien avec more supplémentaire le sont (respectivement  $p < 0,001$  et  $p < 0,05$ )<sup>30</sup>.

Nous souhaitons par ailleurs vérifier si la présence d'une consonne bilabiale à l'initiale de C2 pouvaient augmenter la proportion de suppressions. L'observation des courbes en jaune sur ces deux graphiques montrent effectivement une légère augmentation du nombre de suppressions dans les items 21 à 30, mais cela n'est pas statistiquement significatif.<sup>31</sup>

Passons à présent à la comparaison des items 31 à 40 avec les items de contrôle :

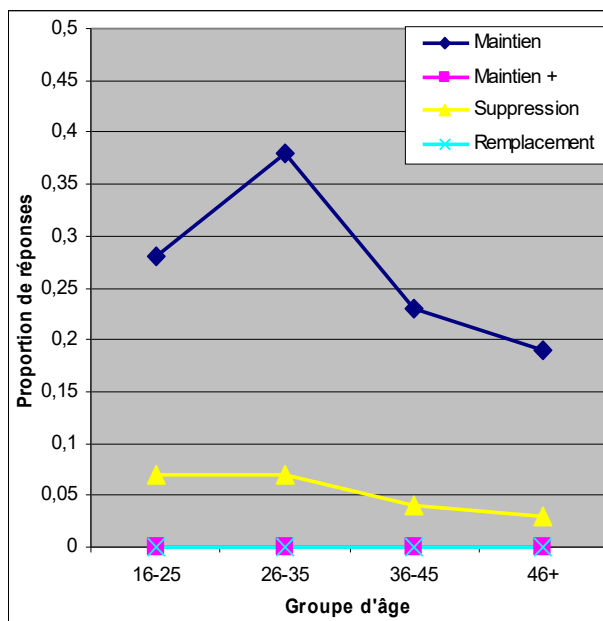


Fig. 5 : Traitement de la deuxième more de C2 dans les items 31-40 (bilabiale à l'initiale de C2 non suivie de /R/)

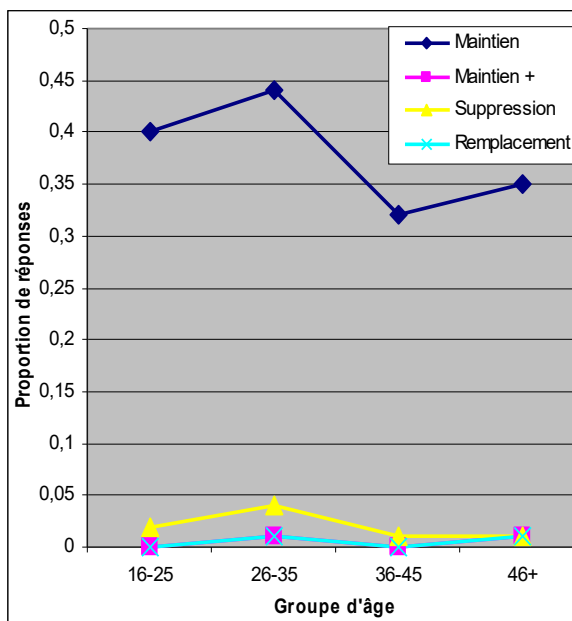


Fig. 6 : Traitement de la deuxième more de C2 dans les items 91-100 (items de contrôle)

Lorsque la deuxième more de C2 est de type (C)V, il n'y a, *a priori*, aucun obstacle à la formation d'un DAC canonique. C'est pourquoi on trouve dans ces deux figures des résultats

29 Facteur fixe : âge

Variable dépendante : traitement choisi

30 Facteur fixe : âge

Variable dépendante : traitement choisi

31 Facteur fixe : groupe d'items

Variable dépendante : nombre de suppressions

similaires : une grande quantité de maintiens et très peu de formations non canoniques.

La proportion de maintiens semble cependant diminuer avec l'âge des participants, et ce de manière significative ( $p < 0,05$  pour les deux groupes d'items)<sup>32</sup>. Pourtant, cette diminution du nombre de maintiens n'est pas compensée par les autres options, qui ne sont que rarement employées. Cela signifie que dans les tranches d'âge les plus élevées, les participants ont créé davantage d'exceptions diverses qui n'entrent pas dans les quatre options qui nous intéressent ici.

Nous souhaitons par ailleurs mesurer l'impact de la présence d'une consonne bilabiale à l'initiale de C2, indépendamment de l'âge des participants. Si l'on compare ces deux graphiques l'un avec l'autre, on peut constater une légère baisse de la proportion de maintiens ainsi qu'une légère hausse de la proportion de suppressions dans les items 31 à 40. Ces deux observations sont aussi significatives statistiquement ( $p < 0,05$ )<sup>33</sup>.

### **3.3 - Traitement de /Q/ en deuxième more de C2 (items 41 à 60)**

Les items 41 à 60 nous permettent d'observer le traitement réservé à /Q/ lorsqu'il apparaît en deuxième more de C2. Pour rappel, dans les items 41 à 50, le C2 n'est pas un constituant trimorique. Tandis que dans les items 51 à 60, C2 est un constituant trimorique correspondant à une syllabe lourde en anglais. Une fois de plus, nous nous intéressons principalement aux quatre mêmes options pour le traitement de la deuxième more de C2. Les figures 7 et 8 présentent les résultats obtenus pour les items 41 à 60.

---

32 Facteur fixe : âge                      Variable dépendante : nombre de maintiens

33 Facteur fixe : groupe d'items              Variable dépendante : nombre de maintiens/suppressions

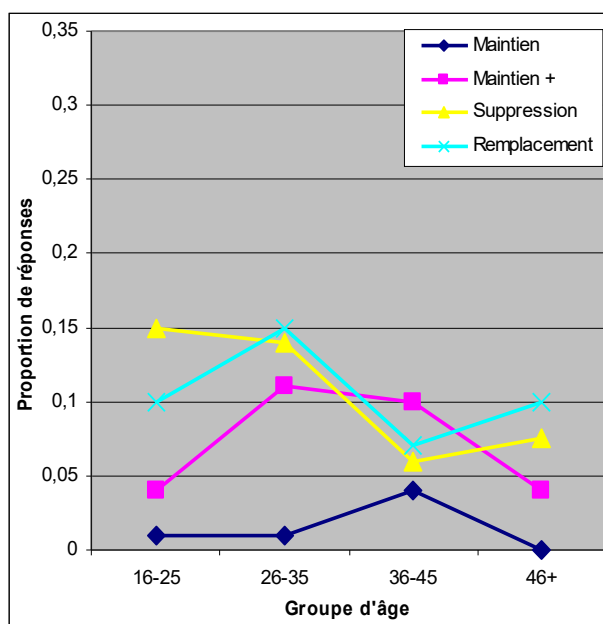


Fig. 7 : Traitement de la deuxième more de de C2 dans les items 51-60 (/Q/ dans un C2 non-trimorique)

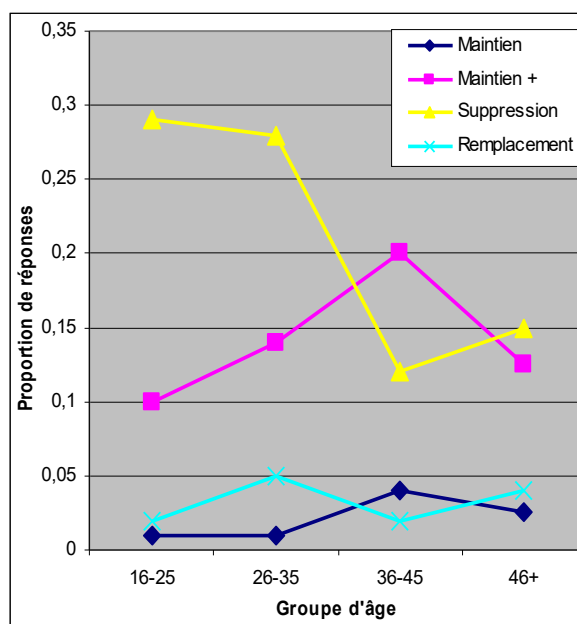


Fig. 8 : Traitement de la deuxième more de de C2 dans les items 61-70 (/Q/ dans un C2 trimorique)

Notons tout d'abord que l'on trouve très peu de maintiens dans ces deux groupes d'items. Pour rappel, nous considérons ici comme maintien les cas très exceptionnels dans lesquels /Q/ a été maintenu avec la réalisation [tsu] pour des raisons liées à la graphie. En outre, le pic que l'on peut observer pour le groupe 36-45 ans est fortement lié à la présence d'un seul participant de 39 ans qui a procédé à ce type d'abréviation de manière quasi-systématique. Il nous paraît donc difficile d'établir une corrélation entre l'âge et le choix de cette option, bien que l'on obtienne un valeur  $p < 0,05$  pour les items 51-60.<sup>34</sup>

Le choix du remplacement ne semble pas non plus être lié à l'âge des participants, puisque les résultats oscillent d'une tranche d'âge à l'autre. Il en va de même pour le maintien avec more supplémentaire dans les items 41 à 50. En revanche, pour les items 51 à 60, on trouve une hausse significative de ce dernier avec l'âge ( $p < 0,01$ ).<sup>35</sup>

Enfin, concernant la suppression, on peut constater qu'elle diminue avec l'âge dans les deux graphiques. Dans les items 41 à 50, cette diminution est moins importante mais néanmoins significative ( $p < 0,01$ ). Dans les items 51 à 60, la forte baisse de la proportion de suppressions semble étroitement liée à la hausse de celle du maintien avec more supplémentaire. Cette baisse est également statistiquement significative ( $p < 0,001$ ).<sup>36</sup>

34 Facteur fixe : âge

Variable dépendante : nombre de maintiens

35 Facteur fixe : âge

Variable dépendante : nombre de maintiens avec more supp.

36 Facteur fixe : âge

Variable dépendante : nombre de suppressions

### 3.4 - Bilan pour les items 1 à 60 et 91 à 100

Puisque nous nous intéressons principalement à l'influence que peut avoir l'âge sur la formation de DAC trimoriques, il nous a semblé intéressant de faire un bilan de l'ensemble des items évoqués jusqu'à présent. Les groupes d'items que nous traiterons par la suite (61 à 90) se focalisent essentiellement sur les facteurs phonologiques qui peuvent entraîner la formation d'un télescopage plutôt que d'un DAC. Ils n'ont donc pas de rapport avec le nombre de mores dont sont constituées les abréviations.

En étudiant le traitement qui est réservé à la deuxième more de C2 dans les items 1 à 60, qui regroupent plusieurs contextes phonologiques favorables à la formation d'exceptions, mais aussi dans les items 91 à 100, qui ne présentent aucune particularité phonologique, nous pourrions avoir une idée plus générale de l'influence de l'âge sur la formation de DAC trimoriques. En effet, rappelons que la proportion de DAC trimoriques est étroitement liée à la proportion de suppressions. La figure 9 ci-dessous présente les données pour l'ensemble des items étudiés jusqu'à présent.

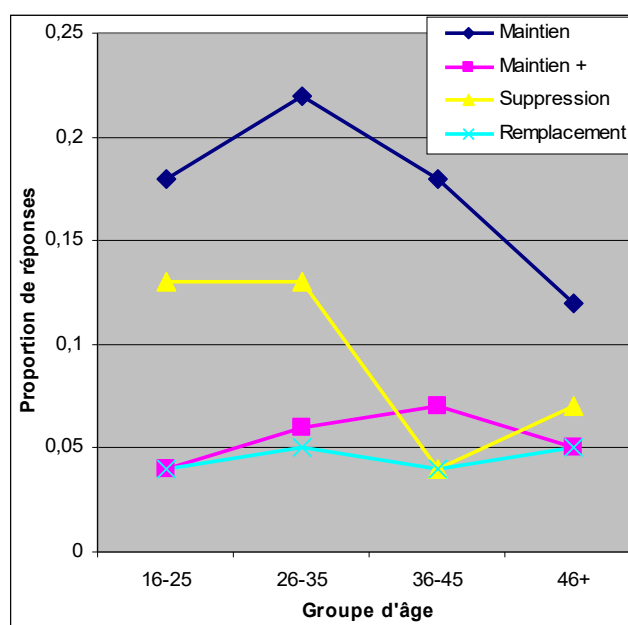


Fig. 9 : Traitement de la deuxième more de C2 dans les items 1-60 + 91-100

On retrouve dans ce graphique des résultats semblables à ceux que nous avons déjà vu dans les différents groupes d'items. La proportion de maintiens et de suppressions décroît avec l'âge tandis que celle du maintien avec more supplémentaire augmente légèrement. Cependant, le fait de réunir l'ensemble de ces items nous permet d'effectuer une analyse statistique plus précise. En effet, toutes ces observations sont statistiquement significatives :



$p < 0,01$  pour le maintien ; et  $p < 0,001$  pour la suppression et le maintien avec more supplémentaire.<sup>37</sup> Concernant la proportion du remplacement, elle reste plus ou moins constante dans toutes les tranches d'âge, ce qui exclut toute corrélation entre ces deux facteurs.

Contrairement à nos attentes, on peut également constater que le maintien reste l'option largement privilégiée même lorsque le contexte phonologique est favorable à l'apparition d'exceptions. Toutefois, on voit que les participants des deux tranches d'âge les plus jeunes ont, après le maintien, une préférence notable pour la suppression, tandis que les participants des deux autres groupes présentent des résultats plutôt partagés entre les trois formations non canoniques.

### 3.5 - Traitement du hiatus entre le C1 et le C2 (items 61 à 80)

Les items 61 à 80 sont consacrés à l'étude des effets de la présence d'un hiatus entre le C1 et le C2, ce qui se produit lorsqu'une voyelle se trouve à l'initiale de C2. Nous souhaitons savoir d'une part si les hiatus sont évités et d'autre part, comment ils sont évités. Pour cela, voyons tout d'abord si la présence d'un hiatus a une influence sur la formation de télescopes plutôt que de DAC. La figure 10 ci-dessous présente la proportion de télescopes parmi l'ensemble des réponses fournies pour différents groupes d'items.

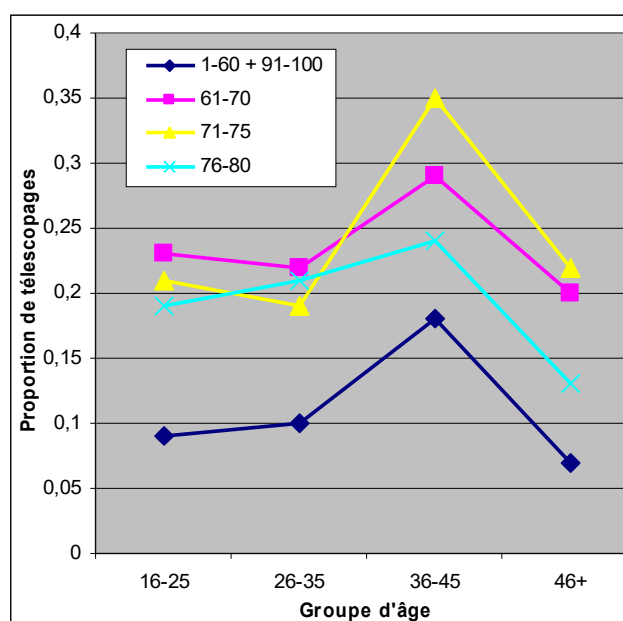


Fig. 10 : Proportion de télescopes dans différents groupes d'items

37 Facteur fixe : âge

Variable dépendante : traitement choisi

Pour rappel, les items 1 à 60 et 91 à 100 correspondent à tous les items vus jusqu'à présent et ne présentent aucun facteurs favorables à la formation de télescopages. Dans les items 61 à 70, le C2 débute par une voyelle. Dans les items 71 à 75 les C1 et C2 débutent par la même voyelle brève. Enfin dans les items 76 à 80, les deux constituants débutent par la même voyelle longue.

Sur ce graphique, on peut constater que les courbes pour chaque groupe d'items sont assez similaires, oscillant en fonction de la tranche d'âge. Cela indique qu'il n'y a pas de corrélation entre l'âge et la propension à créer des télescopages. En revanche, la proportion de télescopages est effectivement généralement plus élevée dans les items comportant un hiatus. Ces différences sont statistiquement pertinentes, avec une valeur  $p < 0,001$  pour les items 61 à 70 ainsi que 71 à 75 et  $p < 0,01$  pour les items 76 à 80.<sup>38</sup>

Le télescopage n'est cependant pas la seule solution pour éviter le hiatus. Il est également possible pour les participants d'enfreindre la contrainte ANCRAGEGAUCHE en ne conservant pas la more initiale du C2. Les trois figures ci-dessous nous présentent la proportion d'infractions à ANCRAGEGAUCHE et de télescopages dans les trois groupes d'items portant sur le hiatus. Les cas où l'on ne peut déterminer laquelle des deux solutions a été choisie sont également indiqués.

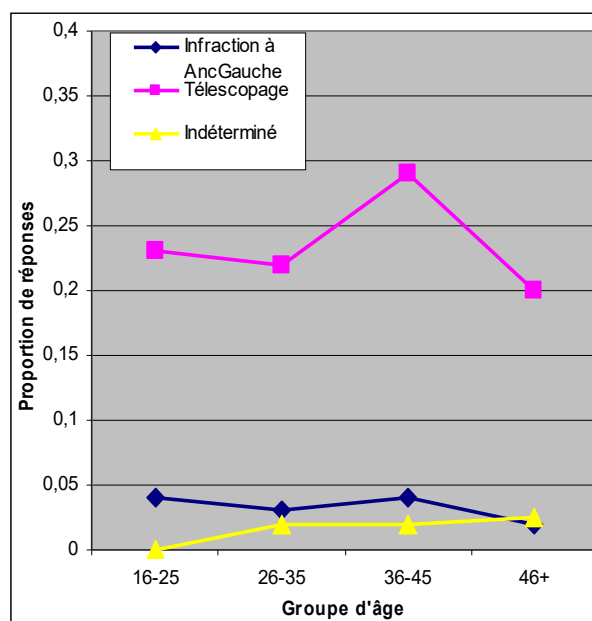


Fig.11 : Traitement du hiatus dans les items 61-70 (voyelle à l'initiale de C2)

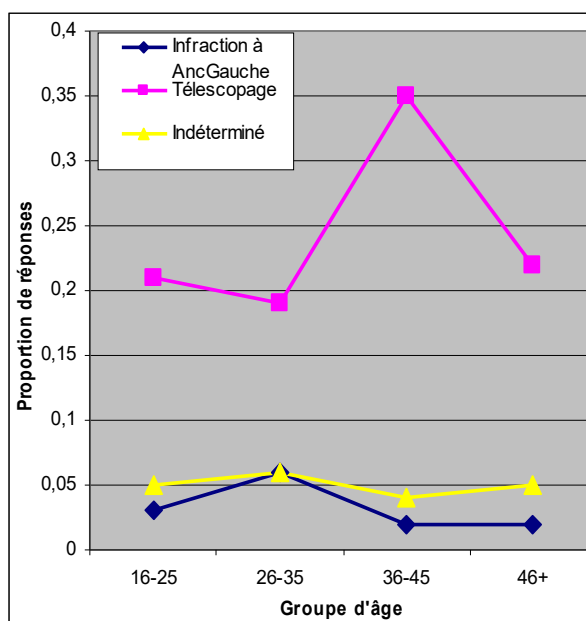


Fig. 12 : Traitement du hiatus dans les items 71-75 (même voyelle brève à l'initiale de C1 et C2)

38 Facteur fixe : groupe d'items

Variable dépendante : nombre de télescopages

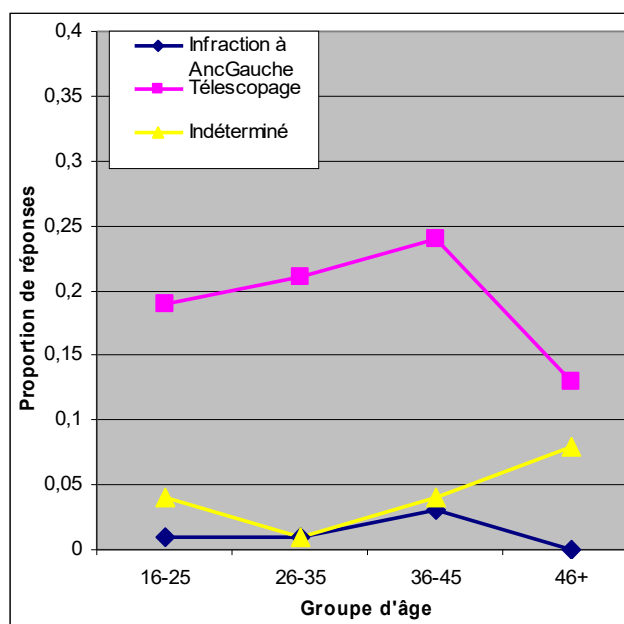


Fig. 13 : Traitement du hiatus dans les items 76-80 (même voyelle longue à l'initiale de C1 et C2)

Comme nous l'avons vu précédemment, il ne semble pas y avoir de corrélation entre l'âge des participants et leur propension à créer des télescopages. Au vu de ces trois nouvelles figures on peut constater que la proportion d'infractions à ANCRAGEGAUCHE reste également sensiblement la même pour toutes les tranches d'âge.

En revanche, on peut constater que le télescopage est l'option largement privilégiée pour le traitement du hiatus dans l'ensemble des groupes d'items. Par ailleurs, la proportion de télescopages ne semble pas varier de manière significative entre les différents groupes d'items. Notons toutefois que pour les items 76 à 80, les personnes âgées de 46 ans et plus ont formé davantage d'abréviations de type indéterminé. Comme le télescopage semble être préféré de manière très claire à l'infraction à ANCRAGEGAUCHE, on peut supposer que la majeure partie de ces cas indéterminés sont en réalité des télescopages. Quoi qu'il en soit, retenons que l'option privilégiée pour éviter la formation de hiatus est le télescopage.

### **3.6 - Traitement de /R/ en C1 et de /Q/ en C2 (items 81 à 90)**

Pour finir, dans les items 81 à 90, nous nous intéressons à un type particulier de télescopage : lorsqu'une more /Q/ présente dans le C2 semble se superposer sur une more /R/ présente dans le C1. Nous souhaitons vérifier si le fait de trouver /R/ dans le C1 et /Q/ dans le C2 pouvait être un facteur favorisant la formation de télescopages et si le nombre de mores qui suivaient /Q/ dans le C2 avait une influence sur la création de ce type de télescopages. Les

figures 14 et 15 ci-dessous présentent les résultats obtenus pour les items 81 à 90.

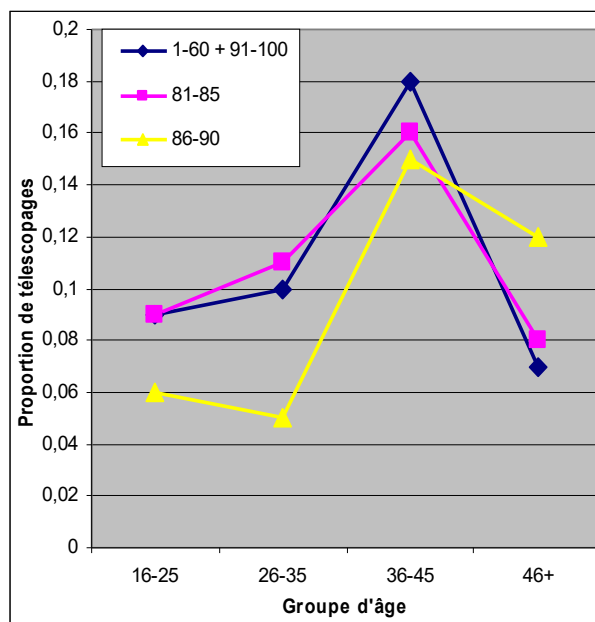


Fig. 14 : Proportion de télescopages dans différents groupes d'items.

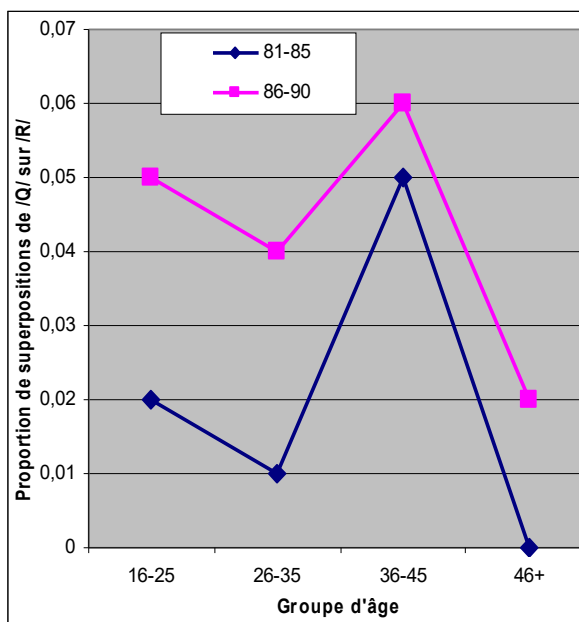


Fig. 15 : Proportions de superpositions de /Q/ sur /R/ dans les items 81-90

Sur la figure 14, on peut constater que la présence de /R/ en C1 et /Q/ en C2 ne fait pas augmenter la proportion de télescopages, que ce soit pour les items 81 à 85 ou les items 86 à 90. La figure 15 nous permet de constater que le nombre de superpositions de /Q/ sur /R/ est légèrement plus important dans les items 86 à 90, dans lesquels /Q/ est suivi d'une seule more, que dans les items 81 à 85, dans lesquels /Q/ est suivi de deux mores. Toutefois, cette différence n'est pas statistiquement significative.<sup>39</sup>

De manière générale, ce type de télescopages reste assez rare et ne semble pas être privilégié à toutes les autres formes possibles de télescopages avec ou sans superposition de mores. Ni l'âge ni le nombre de mores qui suit /Q/ ne semblent avoir d'influence sur la formation de ce type d'abréviations. Néanmoins, il reste intéressant de noter que l'on en repère quelques occurrences.

#### ***4 - Discussion***

Les résultats que nous avons observés dans les sections précédentes nous procurent une nouvelle analyse du fonctionnement des abréviations d'emprunts composés en japonais. Ils devraient nous permettre de vérifier l'exactitude des observations formulées grâce aux

<sup>39</sup> Facteur fixe : groupe d'items

Variable dépendante : nombre de superpositions

études sur corpus réalisées précédemment. Dans la section qui suit, nous discuterons ces résultats, notamment en les comparant aux données des différents corpus étudiés dans la section §1.2. Nous réfléchissons aux implications que ces nouvelles informations pourraient avoir sur notre analyse en termes de contraintes du processus de formation des DAC.

#### **4.1 - Traitement de /R/ en deuxième more de C2**

Intéressons-nous tout d'abord aux résultats obtenus pour les items 1 à 20, portant sur le traitement de /R/. Nous nous attendions à trouver davantage de suppressions chez les participants les plus jeunes et c'est ce que nous avons obtenu. En revanche, contrairement à ce qui avait été observé sur corpus, les cas de suppressions sont loin d'être majoritaires de manière générale. Les figures 1 et 2 montrent que le maintien reste l'option privilégiée pour le traitement de /R/, peu importe la tranche d'âge et le timbre de la voyelle qui précède.

Bien que les données du corpus jeux-vidéo remettent déjà en question la pertinence de la distinction entre les contraintes \*V-HAUT:# et \*V+HAUT:# proposée par Labrune (2007), les résultats du questionnaire remettent en cause l'existence même de toutes les contraintes de type \*V:#, limitant l'apparition de /R/ en position finale de mot ou de constituant. Le fait que les participants les plus âgés aient une propension légèrement plus importante à opter pour le maintien avec more supplémentaire est également inattendu du point de vue de la hiérarchie de contraintes établie jusqu'à présent. Du moins, cela suggère que la contrainte \*PIED>2 aurait perdu en importance avec le temps pour arriver à sa place actuelle dans la hiérarchie.

Quoi qu'il en soit, les différents travaux s'accordent sur un point : la présence de /R/ en deuxième more de C2 reste un facteur phonologique propice à la variation dans la formation des DAC. De plus l'influence de ce facteur est plus importante parmi les tranches d'âge les plus jeunes, qui ont tendance à opter plus fréquemment pour la suppression et ainsi à créer davantage de DAC trimoriques, ce que nous souhaitons démontrer.

#### **4.2 - Traitement des bilabiales à l'initiale de C2**

Le cas des items 21 à 40 est particulièrement intéressant car nous avons remarqué dans les études sur corpus que la présence d'une bilabiale à l'initiale de C2 n'était pas un critère suffisant pour la formation de DAC exceptionnels, bien que ce facteur ait une certaine influence. Les résultats du questionnaire confirment ces observations puisque la proportion de suppression augmente légèrement et celle du maintien diminue légèrement lorsque C2 débute

par une bilabiale. Rappelons que ces différences sont statistiquement significatives pour les items 31 à 40 dans lesquels la deuxième more de C2 n'est pas /R/.

Contrairement au traitement de /R/, pour lequel le facteur phonologique est amplifié par un facteur diachronique, l'âge ne semble pas avoir d'influence sur le traitement de la deuxième more de C2 lorsque celui-ci débute par une bilabiale. Il s'agit néanmoins également d'un facteur phonologique qu'il faut prendre en compte pour comprendre la variation dans les DAC, même si son effet est plus limité.

Bien que le questionnaire nous ait permis de confirmer cet effet de la présence d'une bilabiale à l'initiale de C2, il reste trop irrégulier pour pouvoir être inclus sous forme de contrainte dans notre analyse en OT classique.

### **4.3 - Traitement de /Q/ en deuxième more de C2**

En ce qui concerne le traitement de /Q/ lorsqu'il apparaît en deuxième more de C2, le questionnaire confirme plusieurs de nos hypothèses. Tout d'abord, le maintien avec la réalisation [t̃su], bien que plus présente que dans les corpus, reste très exceptionnelle. Ensuite, la proportion de suppressions diminue avec l'âge des locuteurs. Enfin, pour les items 51 à 60, dans lesquels le C2 est trimorique, la proportion de maintiens avec more supplémentaire augmente chez les participants les plus âgés.

Cela reflète précisément ce que nous avons pu observer dans les corpus. Labrune (2007) remarquait que certains locuteurs avaient tendance à utiliser le maintien avec more supplémentaire lorsque le C2 était trimorique et correspondait à une syllabe lourde en anglais. En revanche, nous ne trouvons aucun maintien avec more supplémentaire dans le corpus jeux-vidéo.

Les résultats obtenus dans le questionnaire correspondent assez bien à ceux obtenus dans notre étude par comparaison de corpus. Aussi le traitement de /Q/ en deuxième more de C2 semble suivre l'analyse en termes de contraintes que nous avons alors proposée.

### **4.4 - Bilan des items 1 à 60 et 91 à 100**

Si l'on observe l'ensemble des items qui ne concernent pas la formation de télescopages (1 à 60 et 91 à 100, figure 9), on s'aperçoit que, de manière générale, la proportion de suppressions diminue avec l'âge, ce qui correspond à nos attentes. Le fait que

les jeunes locuteurs aient plus souvent recours à la suppression permet d'expliquer que la proportion de DAC trimoriques soit plus importante parmi les abréviations récentes.

Il est en revanche assez inattendu de trouver également une diminution de la proportion de maintiens avec l'âge et, au contraire, une augmentation de la proportion de maintiens avec more supplémentaire. Ces observations vont à l'encontre des conclusions formulées à partir de l'étude sur corpus. Si la légère augmentation des maintiens avec more supplémentaire peut s'expliquer par une forme de compensation de la diminution des autres options, il est plus difficile d'expliquer la diminution du nombre de maintiens. Toutefois, cela nous permet de confirmer l'une de nos observations sur les données des différents corpus, qui montraient que le traitement des mores spéciales /R/ et /Q/ en deuxième more de C2 étaient plus éparées, moins régulières, dans le corpus de Labruna que dans le corpus jeux-vidéo.

D'une manière ou d'une autre, les abréviations récentes, ou formées par les locuteurs les plus jeunes, affichent des préférences plus marquées pour certaines options et donc un traitement plus régulier. Cependant, les différences de résultats entre l'étude sur corpus et par questionnaire montrent que notre compréhension de la variation dans la formation des DAC est encore incomplète et mérite d'être approfondie.

#### ***4.5 - Analyse en OT stochastique***

Nous avons jusqu'à présent proposé une analyse en termes de contraintes selon un cadre OT classique. Cependant, le questionnaire semble indiquer que la variation ne serait pas aussi systématique qu'on pouvait le supposer grâce à l'étude sur corpus. Peut-être qu'un modèle OT prenant davantage en compte la variation serait plus approprié à la description du processus de formation des DAC, tels que les modèles en OT stochastique (Boersma & Hayes, 2001)<sup>40</sup>. Ainsi plutôt que de considérer que la hiérarchie de contraintes que nous avons proposée est entièrement erronée, on pourrait envisager que le placement de certaines contraintes n'est simplement pas aussi catégorique que nous l'avions supposé. Les rapports hiérarchiques entre plusieurs contraintes pourraient ne pas être clairement déterminés.

L'objectif principal de ce travail étant de vérifier l'influence de l'âge sur la variation dans les DAC, nous ne pouvons présenter ici une description complète en OT stochastique du processus de formation de ces abréviations. Toutefois, bien que nous ne puissions pas déterminer par un analyse statistique détaillée à quelle fréquence les rapports hiérarchiques

---

<sup>40</sup> Pour les différents modèles OT proposant d'intégrer la variation, voir aussi Antilla (1997), Coetzee & Pater (2011) et German, Pierrehumbert & Kaufmann (2006)

entre contraintes semblent s'inverser, nous pouvons proposer une première ébauche schématique de ce vers quoi devrait s'orienter une analyse en OT stochastique.

Selon ce modèle, les contraintes n'occupent pas un point fixe dans la hiérarchie mais s'appliquent avec une sévérité qui varie plus ou moins, de manière aléatoire, à chaque évaluation d'un nouveau candidat. Les contraintes peuvent alors être représentées par des courbes gaussiennes sur une échelle de sévérité. Il est possible, lorsque deux courbes se chevauchent, que deux contraintes proches dans la hiérarchie voient leur rapport hiérarchique inversé dans certaines évaluations. Cela peut être le cas si la contrainte faible s'applique avec une forte sévérité tandis que la contrainte forte s'applique avec une faible sévérité.

Ce modèle permet d'intégrer la variation à des analyses OT, puisque bien que la contrainte forte domine la faible dans la plupart des cas, il n'est pas exclu qu'elle soit aléatoirement dominée par la contrainte faible dans quelques cas.

#### 4.5.1 - Le traitement de /R/ et son évolution

Prenons l'exemple des quatre contraintes qui entrent en jeu dans le traitement de /R/ en deuxième more de C2 dans les DAC : \*V:#CONST, qui interdit le maintien ; \*PIED>2, qui interdit le maintien avec more supplémentaire ; \*PIED<2, qui interdit la suppression ; et CONTIG, qui interdit le remplacement. Nous avons vu que pour les jeunes participants l'ordre de préférence pour les différents traitements possibles était le suivant : maintien > suppression > maintien avec more supplémentaire > remplacement. Ce qui nous laisse supposer que les contraintes relatives à ces traitements sont hiérarchisées dans l'ordre inverse :

CONTIG >> \*PIED<2 >> \*PIED>2 >> \*V:#CONST

Toutefois, nous savons que le traitement de /R/ est soumis à beaucoup de variation et qu'aucun de ces traitements n'est totalement exclu. Les courbes de ces quatre contraintes doivent donc se superposer sur l'échelle de sévérité. Nous obtenons donc le schéma suivant :

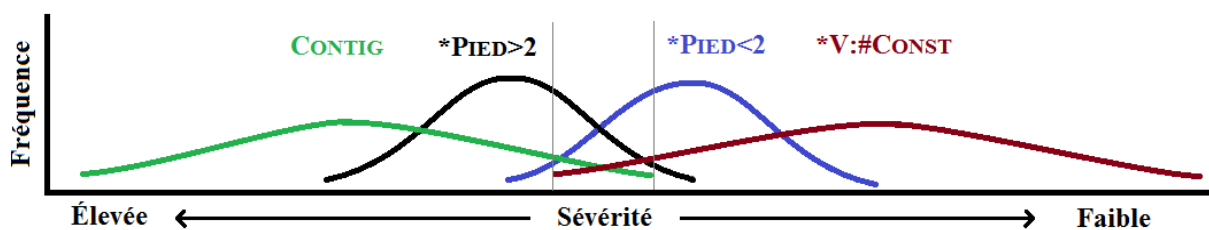


Schéma 5 : Le traitement de /R/ en deuxième more de C2 chez les jeunes locuteurs



Sur ce schéma, on peut voir que les pics de chacune des courbes qui représentent les contraintes sont organisés selon la hiérarchie que nous avons établie. Toutefois, comme les courbes se chevauchent, il n'est pas impossible qu'une contrainte globalement plus faible en vienne à dominer une autre globalement plus forte. Par exemple, si \*V:#CONST s'applique avec une forte sévérité tandis que \*PIED<2 s'applique avec une faible sévérité, \*V:#CONST peut dominer \*PIED<2. Cela signifierait que l'option de la suppression serait retenue dans ce cas, plutôt que le maintien. Au centre du schéma, entre les deux barres grises, les quatre courbes se chevauchent, ce qui rend possible, bien que statistiquement peu probable, l'inversion de tous les rapports hiérarchiques et donc le choix de n'importe quel traitement possible pour /R/.

Par ailleurs, nous savons que les participants les plus âgés ont une préférence pour le maintien avec more supplémentaire plutôt que la suppression. Cela signifie simplement que la position des contraintes \*PIED>2 et \*PIED<2 sur l'échelle de sévérité s'est inversée au cours du temps. Ainsi, nous serions passé du schéma 6 ci-dessous vers le schéma 5 ci-dessus :

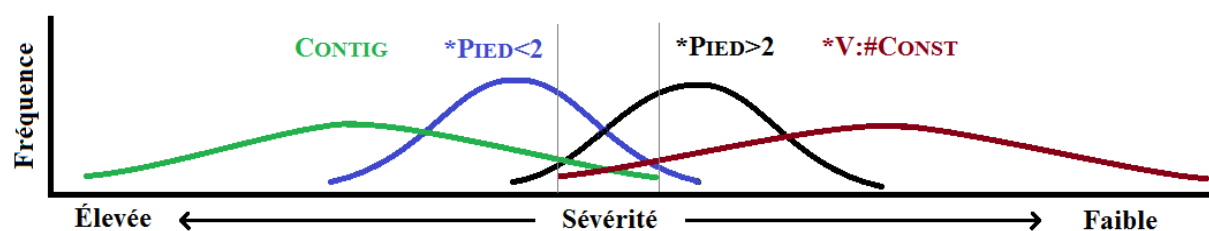


Schéma 6 : Le traitement de /R/ en deuxième more de C2 chez les locuteurs les plus âgés

#### 4.5.2 - Le traitement de /Q/ et son évolution

Le traitement de /Q/ est très similaire à celui de /R/. La principale différence réside dans le fait que /Q/ ne peut pas être maintenu en fin de mot, si ce n'est dans les rares cas où il est réalisé [tsu]. Toutefois, ce type de traitement est en infraction à DEP. Les résultats montrent que ce type de maintien n'est que rarement employé, tout comme le remplacement. Nous pouvons donc imaginer que DEP se situe à un niveau proche de CONTIG sur l'échelle de sévérité. Par souci de clarté, et puisque nous ne fournissons ici qu'une version schématique d'une analyse en OT stochastique, nous placerons DEP exactement au même niveau que CONTIG. Pour les jeunes locuteurs, nous obtenons donc le schéma suivant :

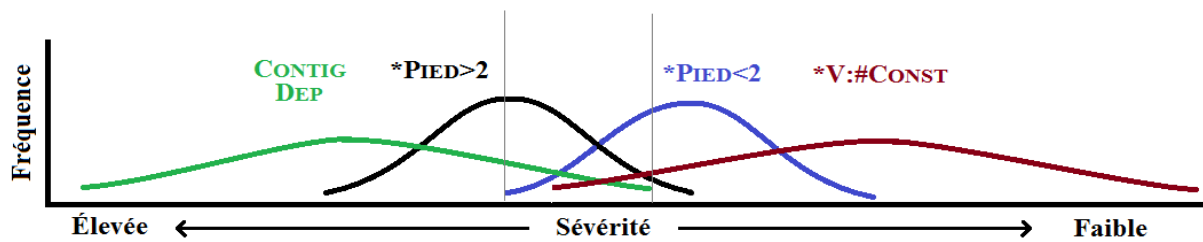


Schéma 7 : Le traitement de /Q/ en deuxième more de C2 chez les jeunes locuteurs

La more /Q/ n'est pas concernée par la contrainte \*V:#CONST. Toutefois, comme pour le traitement de /R/ précédemment, on peut constater que les courbes des quatre contraintes qui concernent son traitement se chevauchent, au centre du schéma, entre les deux barres grises.

Puis, on observe comme pour le traitement de /R/ que les participants les plus âgés ont une préférence pour le maintien avec more supplémentaire plutôt que la suppression. En d'autres termes, on observe bien la même inversion des positions de \*PIED>2 et \*PIED<2. La hiérarchie de contraintes aurait donc évolué du schéma 8 vers le schéma 7.

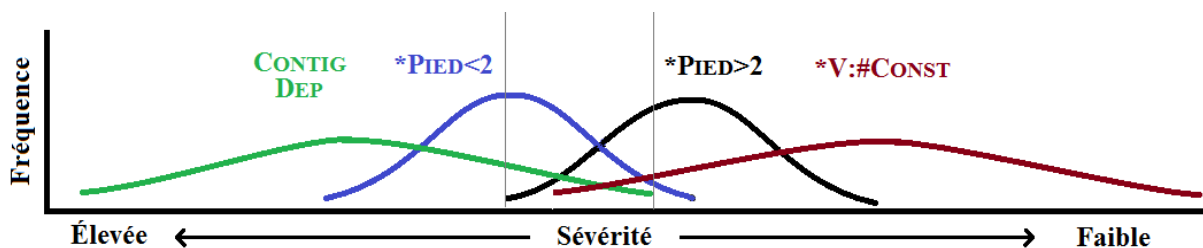


Schéma 8 : Le traitement de /Q/ en deuxième more de C2 chez les locuteurs les plus âgés

#### 4.5.3 - L'influence des bilabiales dans l'analyse OT stochastique

Dans notre analyse en OT classique, bien que nous ayons remarqué l'influence que pouvait avoir la présence d'une bilabiale à l'initiale de C2 sur la formation des DAC, nous ne pouvions inclure de nouvelle contrainte dans notre hiérarchie car cette influence n'agissait pas de manière assez régulière. Or dans une analyse en OT stochastique, il est possible d'intégrer cette légère variation à notre modèle par l'ajout d'une nouvelle contrainte n'ayant que peu de chance de dominer les contraintes qui lui sont supérieures. En d'autres termes, dans notre schéma, les courbes représentant ces contraintes ne se chevaucheraient que légèrement.

Nous postulons donc l'existence d'une nouvelle contrainte, placée bas sur l'échelle de sévérité, qui interdit la formation de pieds binaires en C2 lorsque celui-ci débute par une

consonne bilabiale. Nous nommerons cette contrainte \*BILAB+PIEDBIN.

\*BILAB+PIEDBIN : le deuxième constituant ne doit pas être binaire s'il débute par une bilabiale.

Puisque la présence d'une bilabiale à l'initiale de C2 entraîne une légère diminution du nombre de maintiens et une légère augmentation du nombre de suppressions, les courbes de \*BILAB+PIEDBIN et \*PIED<2 doivent se chevaucher. En revanche, la présence d'une bilabiale ne semble pas avoir d'effet sur les autres traitements possibles, ce qui signifie que les courbes des autres contraintes n'ont pas à interagir avec celle de \*BILAB+PIEDBIN. Par ailleurs, l'âge n'ayant pas non plus d'effet sur l'influence des bilabiales à l'initiale de C2, nous nous concentrerons ici sur la hiérarchie des contraintes pour les jeunes participants :

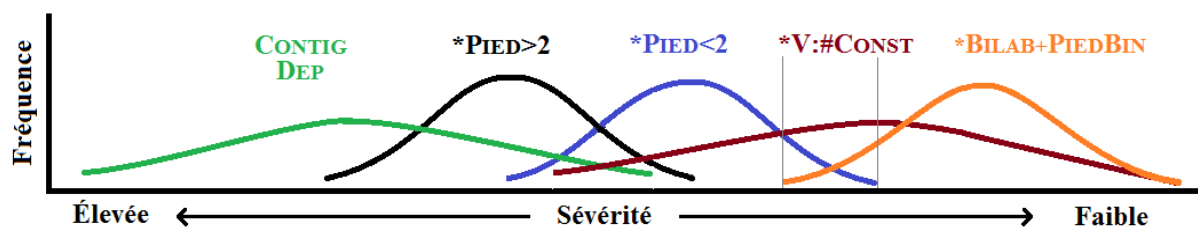


Schéma 9 : Le traitement de la deuxième more de C2 lorsqu'il débute par une bilabiale, chez les jeunes locuteurs

L'espace entre les deux lignes grises est celui dans lequel les courbes de \*PIED<2 et \*BILAB+PIEDBIN se chevauchent. En d'autres termes, sur cette portion de l'échelle de sévérité, le rapport hiérarchique entre ces deux contraintes peut être inversé, ce qui favoriserait l'emploi de la suppression plutôt que du maintien. Lorsque le C2 débute par une bilabiale et que sa deuxième more est /R/, comme dans les items 21 à 30, les facteurs de variations sont multiples puisque les deux contraintes \*V:#CONST et \*BILAB+PIEDBIN entrent en considération. En revanche, lorsque la deuxième more de C2 n'est pas /R/, la contrainte \*V:#CONST n'est pas prise en compte. Cependant, cela n'empêche pas une légère variation due au fait de la superposition des courbes de \*PIED<2 et \*BILAB+PIEDBIN.

#### 4.6 - Le traitement du hiatus

Le fait de s'intéresser au traitement du hiatus avec les items 61 à 80 nous permet de nous interroger sur le rapport entre les DAC et les télescopages. Tout au long de nos travaux, nous avons insisté sur la distinction entre ces deux types d'abréviations qui nous semble cruciale. Labrune (2002) avait remarqué l'importance de la distinction entre DAC et DAS en soulignant le fait que les études antérieures sur le sujet des abréviations d'emprunt en japonais

avaient échoué à produire une description satisfaisante de ces dernières à cause, notamment, de l'absence de cette distinction. Il est évidemment crucial, lorsque l'on tente de décrire un processus morphologique, de s'assurer que l'ensemble des phénomènes observés, sur lesquels se fonde notre analyse, soient issus d'un seul et même processus de dérivation. Il s'avère que la formation des DAC et celle des DAS suivent des logiques totalement différentes, ce qui nous impose de les traiter séparément pour en faire une description pertinente.

Or, lors de la constitution des différents corpus, la distinction entre DAC et télescopages n'avaient pas été envisagée. Par conséquent, le principal critère de sélection pour les abréviations qui ont été collectées était d'être constituées de deux constituants d'origine étrangère. Toutefois, nous savons que les DAC relèvent d'un processus morphologique proche de l'acronymie qui fonctionne d'une manière sensiblement différente du processus de télescopage.

La distinction entre télescopages et DAC est donc théoriquement tout aussi cruciale que celle entre DAS et DAC. Toutefois, dans les faits, il s'avère que le nombre de télescopages dans l'ensemble des deux corpus est très faible, ce qui rend leur influence sur les analyses proposées négligeable. Labrune n'effectue aucune distinction entre DAC et télescopages, considérant ces derniers parmi les infractions à ANCRAGEGAUCHE. Or cette solution ne nous paraît pas pleinement satisfaisante car les télescopages, en plus de maintenir les mores finales, plutôt qu'initiales, du C2, ont la particularité de ne pas être soucieux du nombre de mores que contient l'abréviation. Or, nous savons que les DAC sont particulièrement sensibles à ce critère. Il nous semble contre intuitif de considérer les télescopages comme des DAC en infraction à ANCRAGEGAUCHE ainsi qu'aux contraintes relatives à la binarité des pieds, parfois de manière multiple, plutôt que comme un groupe d'abréviations à part suivant sa propre logique.

Pour toutes ces raisons, il nous semblait particulièrement intéressant d'essayer d'identifier des facteurs phonologiques qui auraient une influence sur le choix du processus d'abréviation employé. Labrune (2007) remarquait que la présence d'une voyelle à l'initiale de C2 entraînait fréquemment des infractions à ANCRAGEGAUCHE. Toutefois, ces DAC exceptionnels pourraient être en réalité des télescopages, puisque Labrune ne faisait pas de distinction entre ces deux types d'abréviations. C'est pourquoi nous avons souhaité tester ce contexte phonologique particulier dans les items 61 à 80, en effectuant cette fois cette distinction, selon la définition des télescopages vue en §2.1.6.

Si l'on se réfère aux résultats de la figure 10, on peut constater que la présence d'un

hiatus entre le C1 et le C2 d'un DAC augmente effectivement la proportion de télescopages de manière significative. Rappelons que l'âge n'a aucune influence statistiquement significative sur le choix du processus d'abréviation. Le hiatus est donc évité de manière générale, peu importe l'âge des participants.

Par ailleurs, les figures 11 à 13 nous montrent que les cas clairement identifiés comme des infractions à ANCRAGEGAUCHE sont beaucoup plus rares que les télescopages. La présence d'un hiatus entre les deux constituants n'est donc pas seulement un facteur de variation dans la formation des DAC mais agit à un autre niveau, sur la sélection du processus d'abréviation qui sera privilégié.

Afin de tenir compte de cela dans notre analyse OT, nous pourrions envisager que le processus de formation des abréviations à partir de mots d'emprunt suit deux étapes différentes : dans un premier temps, le grammaire sélectionne le type d'abréviation qui est le plus adapté en fonction de la base ; puis, dans un seconde temps, l'abréviation est formée conformément aux contraintes relatives au processus d'abréviation retenu.

On peut alors imaginer qu'il existe plusieurs contraintes qui régulent la première étape. Par exemple, une contrainte qui impose que l'abréviation soit un DAS si la base ne contient qu'un seul constituant ; une autre contrainte qui impose la formation d'un DAC lorsque la base contient deux constituants ; etc.

Dans une perspective en OT stochastique, on peut supposer qu'il existe également une contrainte qui impose qu'un télescopage soit formé lorsqu'un hiatus est présent entre les deux constituants. La courbe représentative de cette contrainte chevaucherait celle qui impose la formation d'un DAC lorsque la base contient deux constituants. Cela permettrait d'expliquer pourquoi le processus d'abréviation retenu dans ce contexte phonologique est tantôt le DAC, tantôt le télescopage.

#### **4.7 - La superposition de /Q/ sur /R/**

Dans l'ensemble des items du questionnaire, nous avons pu constater que le nombre de télescopages était bien plus important que ce que l'on pouvait attendre en nous basant sur les données des corpus, qui n'en contiennent que quelques uns. L'un des facteurs propices à la formation de télescopages que nous n'avons pas pu étudier en profondeur est la présence d'une more identique dans les deux constituants. En effet, cette situation permet la création d'un télescopage avec superposition totale d'une more, ce qui semble être une condition plutôt

favorable à l'abréviation sous forme de télescopage plutôt que de DAC.

Nous avons supposé que les liens étroits qui existent entre les segments /R/ et /Q/ pourrait permettre que ces deux mores, bien que différentes, soient superposées dans un télescopage. C'est pourquoi nous nous sommes intéressés au contexte phonologique particulier de la présence de /R/ en C1 et de /Q/ en C2. Toutefois, les résultats de la figure 14 portant sur les items 81 à 90 montrent que ce contexte phonologique, contrairement au hiatus, n'a aucune incidence sur la proportion de télescopages.

Cela semble indiquer que la superposition de /Q/ sur /R/ n'est pas une solution particulièrement appréciée. Cependant, nous avons pu, grâce au questionnaire, obtenir plusieurs occurrences de ce type d'abréviations qui était par ailleurs extrêmement rare dans l'ensemble des deux corpus. L'objectif principal de notre travail étant assez éloigné d'une analyse auto-segmentale de la phonologie du japonais, nous ne pouvons pas tirer de conclusions à partir de nos données. Toutefois, nous pensons avoir montré que les abréviations d'emprunts pouvaient éclairer la relation entre /R/ et /Q/ sous un angle nouveau, qui mériterait des études plus approfondies.

#### ***4.8 - Bilan général***

De manière générale, les résultats obtenus dans le questionnaire confirment certaines tendances observées dans l'étude par comparaison de corpus sans pour autant permettre de conclusions aussi catégoriques.

Tout d'abord, contrairement à nos attentes, nous observons que l'option du maintien reste largement employée, pour toutes les tranches d'âge dans la plupart des groupes d'items. Cela remet en question l'ensemble des observations formulées jusqu'à présent sur l'influence de certains contextes phonologiques sur la formation des DAC. Le corpus de Labruno présentait une certaine régularité dans ses exceptions, puis le corpus jeux-vidéo affichait un traitement presque systématique de ces contextes phonologiques, ce qui laissait penser que la hiérarchie de contraintes avait évolué avant de se stabiliser. Or les résultats du questionnaire présente beaucoup de variation qui ne semble pas être conditionnée par la grammaire. Plutôt qu'une hiérarchie de contrainte plus précise, ces résultats nous incitent à envisager une hiérarchie plus flexible et variable telle que celle proposée dans une optique OT stochastique.

Il n'en demeure pas moins que, globalement, la proportion de suppressions diminue avec l'âge, ce qui est cohérent à la fois avec les données des deux corpus mais aussi avec

l'intuition des locuteurs sur le sujet. Dans les contextes phonologiques problématiques que sont la présence de /R/ ou /Q/ en deuxième more de C2, ou encore la présence d'une bilabiale à l'initiale de C2, la suppression est plus fréquemment employée par les jeunes locuteurs et donc dans les abréviations les plus récentes. Cela a des répercussions directes sur la proportion de DAC trimoriques, qui augmente en conséquence. Ceci est la principale hypothèse que nous souhaitons vérifier.

Par ailleurs, notre questionnaire nous a permis de mettre en lumière quelques pistes pour améliorer la description des DAC. Il est crucial de les dissocier des télescopes tout comme il était indispensable de les dissocier des DAS. Une étude plus approfondie des télescopes en japonais permettrait probablement d'améliorer également notre compréhension des DAC, notamment en identifiant les facteurs qui influent sur le choix du type d'abréviations. De manière intéressante, le hiatus entre le C1 et le C2 semble clairement faire partie de ces facteurs mais il doit être possible d'en trouver d'autres, tels que, par exemple, la présence d'une more commune dans les deux constituants, pouvant servir de point de superposition pour le télescope.

Enfin, un autre fait étonnant est la faible proportion de DAC créés dans l'ensemble des items. Comme nous l'avons vu, les différents corpus réunissent des abréviations diverses formées à partir de deux constituants étrangers. Or, parmi celles-ci, on trouve une très grande majorité de DAC et une quantité infime d'autres types d'abréviations

Nous nous attendions donc à ce que la formation de DAC ne soit pas un problème et que ce type d'abréviations domine largement l'ensemble de nos données. Or nous avons vu que la proportion de télescopes était plus importante qu'attendu. De plus, viennent s'ajouter à cela toutes sortes d'autres abréviations : sigles, traductions de certains constituants en japonais avant d'abrégé, inversion de la place des constituants (bien que cela soit interdit par les règles du questionnaire), mélange de plusieurs types d'abréviations (un constituant en sigle et le second en DAC par exemple), etc.

Ces différences entre les résultats du questionnaire et les faits observés sur corpus nous interrogent sur les conditions dans lesquelles nous avons demandé à nos participants de créer des abréviations et les différences qu'elles présentent avec les conditions d'apparition

naturelles d'abréviations au sein de communautés de locuteurs. En effet, puisque les deux corpus réunis comportent une grande quantité de DAC attestés, nous pouvons supposer que les résultats du questionnaires soit proches d'au moins l'une ou l'autre des tendances observées. Le fait que les résultats du questionnaire soient si différents des abréviations effectivement attestées semble indiquer que certains facteurs qui entrent en jeu dans la création des DAC ont été négligés dans cette approche par questionnaire.

Dans le cas de notre questionnaire, chaque participant est probablement seul face à son ordinateur et exposé à des mots inattendus. Tout d'abord, cela implique qu'il ne crée pas une abréviation à partir de mots qu'il emploie fréquemment, par nécessité, dans un souci d'économie. De plus, les participants n'interagissent pas avec d'autres locuteurs pour décider de la manière d'abrégier et l'acceptabilité de leur abréviation n'est évaluée que par eux-mêmes. Autrement dit, aucun feedback d'une quelconque communauté de locuteurs n'entre en compte dans ce processus.

Ainsi, nous pensons qu'un certain nombre de facteurs sociolinguistiques ont été négligés dans cette nouvelle approche. Une abréviation, comme tout autre signe linguistique, nécessite d'être comprise et acceptée non pas par un seul locuteur mais par une certaine communauté afin de se maintenir. Nous évoquions précédemment la possibilité que les différences observées entre les deux corpus ne soient pas dues à des facteurs diachroniques mais plutôt sociolinguistiques. Il se pourrait, en effet, que les particularités du corpus jeux-vidéo ne proviennent pas du fait qu'il contient des abréviations récentes mais plutôt qu'il contient des abréviations formées principalement par la communauté des joueurs. Nous avons vu que le média jeu-vidéo, en plus d'être utilisé principalement par des jeunes personnes, contenait un nombre particulièrement élevé d'emprunts à l'anglais. Il se pourrait que le rapport à l'anglais des gens de cette communauté, indépendamment de leur âge, soit différent de celui d'autres locuteurs du japonais.

Il est particulièrement intéressant de noter que c'est précisément cette communauté, qui est fréquemment confrontée à des emprunts, qui présente les traitements les plus systématiques des contextes phonologiques particuliers évoqués tout au long de ce travail. Cela n'est pas sans rappeler certains principes élémentaires des Théories des exemplaires, selon laquelle l'acquisition de la phonologie se fait par l'analyse de la récurrence statistique de certaines propriétés acoustiques (Pierrehumbert, 2003). La Théorie des exemplaires, bien que développée pour décrire l'acquisition de la phonologie, peut s'appliquer de manière plus générale à son évolution, ainsi qu'à d'autres domaines de la langue.



Ici, nous tentons de décrire l'évolution du processus morphologique qu'est l'abréviation sous forme de DAC. Nous sommes, dans le cadre du corpus jeux-vidéo, en présence d'une communauté fortement exposée à ce processus relativement nouveau dans la langue et qui, par conséquent, comporte bon nombre de variations *a priori* libres. La manière dont le traitement des DAC exceptionnels se systématisé au sein de cette communauté nous indique que de nouvelles contraintes émergent avec l'augmentation de la fréquence à laquelle les membres de cette communauté sont exposés à ces exceptions.

Par ailleurs, dans les Théories des exemplaires, le rôle des réseaux de locuteurs et du feedback de la communauté est un élément essentiel au développement de la phonologie, ou de tout autre aspect de la langue qui nous intéresse. Avec notre questionnaire, nous avons effacé totalement ces aspects là, ce qui explique la différence observée entre les résultats du questionnaire et ceux de l'étude sur corpus. Il n'est donc pas impossible que les mêmes participants, dans des conditions différentes, favorisant l'échange entre locuteurs et ainsi un certain feedback mutuel, puissent produire des abréviations bien différentes. C'est précisément vers cette piste que nous souhaitons mener nos futurs travaux.

## **Conclusion :**

Le variation qui existe dans le processus de formation des DAC est un phénomène intéressant pour plusieurs raisons. Non seulement il est possible d'identifier des facteurs phonologiques qui provoquent cette variation, mais des facteurs diachroniques sont également à prendre en compte. Il semble par ailleurs que les locuteurs soient sensibles à ces derniers, puisque certains d'entre eux suggèrent que les tendances d'abréviation ont évolué au cours du temps.

Notre première étude sur le sujet, par comparaison de corpus, nous a permis de confirmer l'influence des facteurs phonologiques sur la formation de DAC exceptionnels mais aussi de vérifier l'hypothèse d'une évolution du processus d'abréviation, formulée à partir de l'intuition des locuteurs. À ce stade, nous pouvions supposer que le traitement des exceptions s'était régularisé avec le temps. Il n'est pas étonnant qu'un processus morphologique dont la productivité a fortement augmenté relativement récemment se soit régularisé. Si l'on trouve des DAC assez anciens, datant des premiers contacts avec les langues occidentales, ceux-ci se sont multipliés après la Seconde Guerre mondiale et les emprunts lexicaux massifs à l'anglais qui s'en sont suivis. Les locuteurs japonais ont alors été exposés à davantage de DAC

contenant des contextes phonologiques particuliers et ont ainsi pu régulariser leur traitement.

Toutefois, cette nouvelle approche par questionnaire montre que la variation est loin d'avoir été effacée à un niveau individuel. Si dans les faits les DAC qui sont attestés dans la langue tendent vers plus de régularité, les réponses de nos participants montrent au contraire une grande variabilité. On peut alors supposer que le maintien d'une abréviation dans la langue dépend également de critères d'ordre sociolinguistiques, relatifs à l'interaction des locuteurs, à leur jugement d'acceptabilité et au feedback mutuel qui peut exister au sein des communautés. En excluant ces facteurs là dans notre étude par questionnaire, nous avons pu montrer que les individus ne partagent pas tous les mêmes préférences dans le traitements des contextes phonologiques particuliers. Autrement dit, concernant la formation des DAC, la grammaire reste assez variable à un niveau individuel.

Dans notre analyse du point de vue OT, il nous a donc semblé pertinent de nous orienter vers un modèle qui tient davantage compte de la variation, tel que l'OT stochastique. La représentation des contraintes par des courbes sur une échelle de sévérité nous a permis de rendre compte de la variation dans le traitement des contextes phonologiques particuliers ainsi que de leur évolution.

Le fait que la grammaire soit assez variable à un niveau individuel ouvre un champ de recherche intéressant pour la suite de nos travaux. En effet, nous pourrions à présent nous interroger sur les processus sociolinguistiques qui entrent en jeu dans l'apparition de régularités au niveau des DAC attestés concernant des contextes phonologiques pourtant traités de manière variable à un niveau individuel. Cela pourrait nous éclairer sur l'influence que peut avoir l'interaction des locuteurs au sein de réseaux ou de communautés sur l'émergence de nouvelles contraintes phonologiques dans la grammaire.

## **Bibliographie :**

- ANTILLA, Arto, « Deriving variation from grammar : a study of Finnish genitive », in HINSKENS, Frans, VAN HOUT, Roeland, WETZELS, Leo (éds.), *Variation, Change and Phonological Theory*, Amsterdam/Philadelphia : John Benjamins, 1997, pp.35-68.
- BATES, Douglas, MAECHLER, Martin, BOLKER, Ben, WALKER, Steve, « Fitting Linear Mixed-Effects Models Using lme4 », in *Journal of Statistical Software*, 67(1), 2015, pp.1-48.
- BENUA, Laura, « Identity effects in morphological truncation » in *Papers in Optimality Theory*. BECKMAN, J., WALSH DICKEY, L., URBANCZYK, S. (éds), Amherst : University of Massachussets, 1995, pp.77-136.
- BENUA, Laura, « Transderivational Identity, Phonological Relations between Words », Thèse de doctorat, Amherst : University of Massachussets, 1997.
- BOERSMA, Paul, HAYES, Bruce, « Empirical tests of the gradual leaning algorithm », in *Linguistic Inquiry*, 32, 2001, pp.45-86.
- COETZEE, Andries, PATER, Joe, « The place of variation in phonological theory », in GOLDSMITH, John, RIGGLE, Jason, YU, Alan (éds.), *The Handbook of Phonological theory. 2nd edition*. Oxford : Blackwell, 2011, pp. 401-434.
- GERMAN, James, PIERREHUMBERT, Janet KAUFMANN, Stefan, « Evidence for phonological constraints on nuclear accent placement. » in *Language*, 82, 2006, pp.151-168.
- GUSSENHOVEN, Carlos, JACOBS, Haike, *Understanding Phonologie*, 3ème édition, [S.I] : Hodder Education, 2011, 319p.
- HAUGEN, Einar, « The analysis of borrowings », *Language*, 26, 1950, pp.210-231.
- ITÔ Junko, KUBOZONO Haruo, MESTER, Armin, « Consonant gemination in Japanese loanword phonology », in The Linguistic Society of Korea (éd.), *Current issues in unity and diversity of languages. collection of papers selected from the 18th international congress of linguists*. Republic of Korea: Dongam Publishing Company, 2009, pp. 953–973.
- ITÔ Junko, MESTER, Armin, « Japanese Phonology, Constraint Domains and Structure Preservation », in GOLDSMITH, John (ed.), *The handbook of phonological theory*, Oxford : Blackwell, 1995, pp. 817-838.
- ITÔ Junko, MESTER, « Lexical classes in phonology », in MIYAGAWA Shigeru, SAITO Mamoru (éds.), *The Oxford handbook of Japanese linguistics*. New York : Oxford University Press, 2008, pp. 84-106.
- JÔO Hakutarô, « gendai nihongo no on'in » 現代日本語の音韻 (Phonologie du japonais moderne), in *Iwanami Kôza Nihongo 5 on'in 岩波講座日本語5 音韻* (Académie d'Iwanami : le japonais 5 : phonologie), ÔNO S., SHIBATA T. (éds.), Tokyo : Iwanami shoten, 1977, pp.107-145.
- KAGER, René, *Optimality Theory*, New York : Cambridge University Press, 1999, 452 p.

- KREIDLER, Charles W., « Clipping and acronymy », in BOOIJ, Geert, LEHMANN, Christian, MUGDAN, Joachim (éds.), *Morphology. An International Handbook on inflection and word formation*. Berlin ; New York : Walter de Gruyter, 2000.
- KUBOZONO Haruo, « Nihongo no moora : sono yakuwari to tokusee » 日本語のモーラ:その役割と特性 (La more en japonais : son rôle et ses caractéristiques), *Nihongo no moora to onsestukoozoo ni kansuru soogooteki kenkyuu*, 日本語のモーラと音節構造に関する総合的研究 (études sur la structure de la syllabe et de la more en japonais),1. Tokyo : Ministère de l'éducation, 1992, pp. 48-61.
- LABRUNE, Laurence, TAKAYAMA Tomoaki, « Aspects de la phonologie de /p/ », in GRIOLET, Pascal, LUCKEN, Michael (éds.), *Japon pluriel 5 : actes du cinquième colloque de la Société française des études japonaises*. Arles : Philippe Picquier, 2004 , pp. 365-377. (colloque du 19 au 21 Décembre 2002)
- LABRUNE, Laurence, « Structure de la syllabe japonaise », in BLIN, Raoul, TAMBA, Irène (éds.), *Faits de langues coréen – japonais*. Paris : Ophrys, 2001, pp. 111-122.
- LABRUNE, Laurence, « The prosodic structure of simple abbreviated loan-words: a constraint based account », *Onsei kenkyû* 音声研究 (recherche en phonétique) / 外来語における促音の挿入について Journal of the Phonetic Society of Japan. 2002, pp. 98-120.
- LABRUNE, Laurence, « Autour de la syllabe:les constituants prosodiques mineurs en phonologie », in *Phonétique et phonologie, approches contemporaines*, NGUYEN, N., WAUQUIERS, S., DURAND, J. (éds), Hermès, 2005, pp. 95-116.
- LABRUNE, Laurence, *La phonologie du japonais*. Leuven : Peeters, 2006a, 305 p.
- LABRUNE, Laurence, « Patterns of phonemic preferences in Japanese non-headed binary compounds : what waa-puro, mecha-kucha and are-kore have in common », in *Gengo Kenkyû, Journal of the Linguistic Society of Japan*, 129, 2006b, pp. 3-41.
- LABRUNE, Laurence, « Morphophonologie des emprunts abrégés en japonais : aspects prosodiques et mélodiques », in *Des sons et des sens : données et modèles en phonologie et en morphologie*, DELAIS-ROUSSARIE, E., & LABRUNE, L., (éds), 2007, Paris : Hermès, 28p.
- LABRUNE, Laurence, « Principes d'organisation phonémique des emprunts occidentaux composés abrégés », in *Revue d'Etudes Japonaises*, Paris : Université Paris 7, 2008, 14p.
- LABRUNE, Laurence, « Questioning the universality of the syllabe : evidence from Japanese », in *Phonology 29:1*, Cambridge University press, 2012, pp.113-152.
- LAMARQUE, Rémi, « Théorie des contraintes et phonologie des mots d'emprunt en japonais ». Mémoire de master sous la direction de VITTRANT Alice, Aix-Marseille Université, 2013, 64p.
- LAMARQUE, Rémi, « Morphophonologie des emprunts abrégés japonais dans les jeux vidéo : évolution du processus d'abréviation ». Mémoire de master sous la direction de VITTRANT Alice, Aix-Marseille Université, 2015a, 132p.

- LAMARQUE, Rémi, « Contraintes phonologiques et mots d'emprunt en japonais » in *Actes de la 5<sup>e</sup> édition des JéTou*. Toulouse : Université Toulouse – Jean Jaurès, 2015b, pp.21-30.
- LOVINS, Julie Beth, « Loanwords and the phonological structure of Japanese », Thèse de doctorat, Chicago : University of Chicago, 1973, 164 p.
- MABUCHI Kazuo, *Kokugo On'inron 国語音韻論*(Phonologie du japonais), Kasama Shoin : Tokyo, 1971, 158 p.
- MC.CARTHY, John, « A case of surface constraint violation » in PARADIS, Carole, LACHARITÉ, Darlene (éds.), *Constraint-based theories in multilinear phonology*, special issue of *Canadian Journal of Linguistics* 38, 1993, pp.169-195.
- MC.CARTHY, John, ALAN, Prince, « Faithfulness and Reduplicative Identity », in BECKMAN, J.N., WALSH DICKEY, L., URBANCZYK, S. (éds.), *Papers in Optimality Theory*, Amherst : University of Massachussets, 1995, pp.249-384.
- MC.CAWLEY, James D., *The Phonological Component of a Grammar of Japanese*. La Haye : Mouton, 1968, 208p.
- NISHIHARA Tetsuo, VAN DE WEIJER, Jeroen, NANJO Kensuke, « against headedness in compound truncation: English compounds in Japanese », in *Issues in Japanese Phonology and Morphology*, VAN DE WEIJER J., NISHIHARA T. (eds.), Berlin : Mouton de Gruyters, 2001, pp.299-326.
- OHSO, Mieko, « A Phonological Study of Some English Loan Words in Japanese », Thèse de doctorat, Columbus : Ohio State University, 1971.
- PAMIES BERTRAND, Antonio, « Prosodic Typology: On the Dichotomy between Stress-Timed and Syllable-Timed Languages », in *Language Design* 2, 1999, pp. 103-130.
- PIKE, Kenneth Lee, *The Intonation of American English*. Ann Arbor: University of Michigan Press. 1945, 200p.
- PIERREHUMBERT, Janet, « Phonetic Diversity, Statistical Learning, and Acquisition of Phonology », in *Language and speech*, 2003, 46, Northwestern University, pp. 115-154.
- PRINCE, Alan, SMOLENSKY, Paul, *Optimality theory constraint interaction in generative grammar*. Malden : Blackwell, (1993/2002)2004, 289p.
- RAMUS, Frank, « La discrimination des langues par la prosodie: Modélisation linguistique et études comportementales. », in *De la caractérisation à l'identification des langues, Actes de la 1<sup>ère</sup> journée d'étude sur l'identification automatique des langues*, PELLEGRINO, F. (éd.), Lyon : Éditions de l'Institut des Sciences de l'Homme, pp. 186-201, 1999.
- SHINOHARA Shigeiko, « The roles of the syllable and the mora in Japanese adaptations of French words », *Cahiers de Linguistique - Asie Orientale*. 25-1,1996, pp. 87-112.

SHINOHARA Shigeko, « Analyse phonologique de l'adaptation japonaise de mots étrangers ». Thèse de doctorat sous la direction de RIALLAND Annie, Paris : Université de la Sorbonne Nouvelle - Paris III, 1997, 219 p.

SHIMIZU HAN Mieko, « The feature of duration in Japanese », *Onsei Kenkyuu* 音声研究 (Recherche en phonétique). 10, 1962, pp. 65-80.