



## **Le rôle de la fréquence lexicale dans l'acquisition et la production de liaisons en L1 et L2**

**Nadine de Moras**, Brescia University College (Western University) - Canada

### **Résumé :**

Afin de tester le scénario d'acquisition des langues (L1 et L2) proposé par le modèle basé sur l'usage (Bybee, 2001 et Ellis, 2002), cet article évalue les effets des différents types de fréquence sur la production de liaisons catégoriques (obligatoires) de francophones et apprenants FL2. 12 francophones majoritaires, 8 francophones minoritaires et 37 apprenants FL2 de niveau intermédiaire ont lu un texte contenant 51 liaisons catégoriques. Les productions ont été corrélées aux informations syntaxiques, lexicales et fréquentielles des séquences liaisonnantes, en utilisant les banques de données de fréquence *Vocabprofil* et *Lexique 3*, et *GoldVarb* pour les statistiques. Les francophones majoritaires ont produit 95,6% de liaisons catégoriques, les francophones minoritaires en ont produit 85,5%, et les apprenants FL2 60.7%. Plus la fréquence du mot1 et du Mot2 et la fréquence du taux de cooccurrence sont élevées, plus les productions de liaisons sont élevées pour tous les groupes.

**Mots-clés :** Liaisons catégoriques – enchaînements – répétitions – fréquence

### **Abstract:**

In order to test the (L1 and L2) language acquisition scenario proposed by the Usage-Based Model (Bybee, 2001 and Ellis, 2002), this article evaluates the

effects of different types of frequency on the production of categorical (obligatory) liaisons of Francophones and L2 learners. 12 Majority Francophones, 8 Minority Francophones and 37 L2 learners of intermediate level read aloud a text with 51 categorical liaisons. Their oral productions were correlated with syntactic, lexical and frequency information of linking sequences, using frequency databases *Vocabprofil* and *Lexique*, and *Goldvarb* for statistical analysis. Majority Francophones produced 95.6% of categorical liaisons, Minority Francophones 85.5%, and L2 learners 60.7 %. The higher the Word1, the Word2 and the co-occurrence rate, the more the liaisons were produced by all groups.

**Keywords:** Obligatory liaisons – linking – repetitions – frequency

## 1. Introduction

L'égalité syllabique, la syllabation, et la resyllabation du français, les liaisons et les enchaînements sont difficiles à maîtriser pour tous les apprenants de français L2 (Charliac & Motron, 1998 : 7-9). La liaison se produit quand une consonne latente est prononcée et enchaînée au mot suivant, si celui-ci commence par une voyelle ou un H muet.

*Exemple 1* : les (Mot1)\_(z)enfants (Mot2) se prononce [le.zã.fã].

Les liaisons peuvent être catégoriques (dorénavant abrégées en LC), variables ou erronées, selon les contextes syntaxiques (LC dans certains contextes tels que [déterminant + nom ou adjectif]), phonétiques (la liaison est plutôt variable avec /r/), lexicales (certaines expressions figées), et stylistiques (on produit les liaisons variables plus rares dans un style plus soutenu).

Pour maîtriser la prononciation des liaisons, il faut connaître les contextes morphosyntaxiques, et aussi connaître les règles phonétiques telles que les règles de prononciation des consonnes finales et leurs exceptions, les règles de changement de timbre des consonnes de liaison (*neuf heures*) et de dénasalisation, (*bon amis*). Quant aux expressions figées, elles contredisent les autres règles (*Comment allez-vous ?* = LC, vs *Comment allez-vous lui parler ?* = liaison erronée). De plus, les liaisons sont régies par des contraintes lexicales et fréquentielles.

D'après Wauquier (2009 : 125), les stratégies d'acquisition de la liaison sont différentes en L1 et L2 et les apprenants L2 ont une « une conception lexicaliste de la liaison [...] un apprentissage basé essentiellement sur l'usage et conditionné par la fréquence des informations lexicales présentes dans l'input [...] » tandis que d'après Harnois-Delpiano, *et al.* (2012) les apprenants se basent sur des généralisations phonologiques. D'après Bybee (2001), la cohésion syntaxique est la fréquence de cooccurrence des séquences, et elle détermine la force de l'association entre le premier élément et le second. Ces connections sont stockées en mémoire et sont renforcées par leur usage fréquent. Bybee (2001,

2005) soutient que la fréquence de cooccurrence et la fréquence des mots jouent un rôle majeur dans la production de liaisons. Dans l'étude de Adda-Decker, Boula de Mareuil et Lamel (1999), avec les 256 mots lieurs les plus fréquents, 75% des liaisons sont réalisées, alors qu'avec les 2560 mots les plus fréquents, le taux de réalisation diminue à 64%. Les études menées dans le cadre du Modèle Basé sur l'Usage suggèrent que les locuteurs traitent la langue à partir d'une série de probabilités, et d'une banque de données établie d'après l'input reçu (Ellis 2002).

Cet article analyse l'hypothèse selon laquelle les locuteurs natifs comme les apprenants du français L2 (dorénavant AFL2) utilisent des stratégies lexicales basées sur la fréquence dans le traitement des liaisons.

Après avoir présenté la classification des liaisons utilisée ici, nous ferons un bref survol des recherches en acquisition et productions des liaisons en L1 et L2. Nous présenterons la méthodologie de l'étude et analyserons ensuite les résultats empiriques quant au rôle des différents types de fréquence sur les productions de LC dans le groupe nominal.

## **2. Catégorisation des liaisons**

Si les linguistes s'entendent sur les liaisons erronées (H aspiré, par exemple), la classification des contextes catégoriques et variables n'est pas unanime. Dans leur analyse du corpus PFC, Durand et Lyche (2008) observent que les liaisons sont systématiquement réalisées dans quatre contextes : [déterminant + nom], [proclitique + verbe], [verbe + enclitique] et expressions figées.

Le contexte [préposition + nom] longtemps considéré comme catégorique est variable d'après les analyses du corpus PFC. La raison de cette différence de catégorisations tient au fait que les catégories ne dépendent pas seulement des contextes syntaxiques, mais aussi de facteurs phonologiques, lexicaux et fréquentiels.

Durand et Lyche (2008) relèvent que la liaison est systématique après *en* (100%), et quasi systématique après *dans* (95%), mais moins fréquente avec d'autres prépositions (chez, sans).

Certains adjectifs prénominaux se placent avant le nom parce qu'ils sont courts et fréquents (grand, gros). Il est possible de placer des adjectifs moins fréquents (fin, hautain...) avant les noms pour des effets stylistiques ; dans ce cas les locuteurs natifs ne font pas la liaison (Sampson, 2001). De même, Durand *et al.* (2011 : 116) observent que la séquence *grand honneur* (93.77%) a un taux de réalisation de la liaison (sans erreurs) supérieur à *grand émoi* (81.25%) parce que *grand honneur* est cinq fois plus fréquent que *grand émoi*. Autrement dit, la liaison après un adjectif dépend du Mot1 et du Mot2 et de leur taux de co-occurrence.

La classification utilisée dans cette étude (tableau 1), est notamment utilisée par Mastromonaco (1999), Boula de Mareüil, *et al.* (2003), et Dugua (2005).

<b>Liaisons catégoriques</b>	<b>Liaisons erronées</b>	<b>Liaisons variables</b>
Déterminant + nom : <i>un_arbre</i>	Déterminant + nom avec un h aspiré:	Liaisons qui ne sont ni catégoriques ni erronées <i>les hiboux</i>
Adjectif + nom : <i>gros_effort</i>	Déterminant + nom avec y: <i>un yoyo</i>	Nom + adjectif pluriel : <i>Des enfants_intelligents</i>
Après un pronom : <i>nous_avons</i>	Après « ET » : <i>beau et intelligent</i>	Aux. + participe passé : <i>Ils sont_arrivés</i>
Après préposition monosyllabique : <i>en_avance</i>	Nom sujet + verbe : <i>Jean arrive</i>	Adverbe polysyllabique +...part. passé, adjectif...) <i>beaucoup_intéressés</i>
Expressions figées : <i>États-Unis</i>	Nom + adjectif singulier : <i>enfant intelligent</i>	Négation + participe passé : <i>pas_arrivé</i>

**Tableau 1. Classification des liaisons**

### **3. L'acquisition de la liaison en L1 et en L2**

#### ***3.1. L'acquisition de la liaison en L1***

Chabanal et Liégeois (2014) observent « une combinatoire limitée » et la fréquence de certaines structures de l'input parental qui facilitent l'acquisition des items et des structures des jeunes enfants. Les LC des parents sont généralement produites sans erreurs et les liaisons variables ne sont réalisées qu'entre 2.7% et 13.8% selon les parents. Les enfants reproduisent en majorité l'input parental.

Les enfants francophones apprennent d'abord les liaisons les plus fréquentes, les LC (Chabanal, 2003), et à l'intérieur de cette catégorie, ils apprennent d'abord les contextes les plus fréquents. Dugua (2006) a montré que les erreurs d'enfants âgés de 2 à 6 ans étaient corrélées avec la fréquence des formes entendues. S'ils avaient le plus souvent entendu « un n-avion », ils faisaient des erreurs du type « les n-avions ».

Les enfants issus de milieux socioéconomiques plus élevés ont un input plus riche leur permettant de développer un plus grand vocabulaire (Hoff, 2002). Dans l'étude de Chevrot, Dugua et Fayol (2005) l'input langagier des enfants joue un rôle déterminant dans les productions justes de LC ainsi que sur les erreurs. De même, Chabanal (2007) a montré que les enfants de milieu favorisé reconnaissent comme justes et produisent plus de liaisons variables que ceux de milieu défavorisé.

Nicoladis et Paradis (2011) relèvent l'effet de la fréquence et de la maîtrise des mots. Le vocabulaire français d'enfants monolingues et bilingues (3;1 et 5;7) est une meilleure mesure que l'âge, car il correspond à la fréquence de l'input. Ainsi, les enfants ayant eu un plus grand input connaissent plus de mots et s'ils connaissent plus de mots ils produisent plus de liaisons.

La fréquence en réception et en production et la qualité des productions présentes dans l'input jouent un rôle majeur dans l'acquisition des liaisons chez les enfants.

### ***3.2. L'acquisition de la liaison en L2***

Les études de productions de liaisons des AFL2 suggèrent que leurs productions se rapprochent de celles des natifs et qu'ils maîtrisent les LC à un niveau natif ou proche des natifs. Les étudiants coréens de l'étude de Harnois-Delpiano, Cavalla, et Chevrot (2012) produisent 70% de LC après 400 heures d'étude du français, ce qui est beaucoup après si peu d'heures.

Les AFL2 canadiens de Mastromonaco (1999) produisent 93.32% de LC (dont 7% non enchaînées); ceux de Thomas (2004) 86.7%-93.9% (dont 8.5% non enchaînées). Les AFL2 irlandais de Howard (2005) 82.5%- 95%. Les AFL2 de Howard (2013) produisent même plus de liaisons facultatives (18-33%) que les francophones (10% Ahmad, 1993 ; 11% Ashby, 2003). Les AFL2 anglophones de Tennant (2015) en produisent 90% (natifs : 99%). Selon la classification de Durand et Lyche (2008), (12 sites de liaison potentielle), les apprenants FL2 japonophones sans séjour de Detey, Kawaguchi, et Kondo (2015) produisent 75.28% de LC et les apprenants avec séjour 86.03% (natifs : 99.17%). Cette étude met en relief le rôle du séjour linguistique dans la production des apprenants, mais également l'importance de la tâche et des outils utilisés, tels que le système de classification des liaisons.

Dans l'étude de Racine (2015), le comportement des apprenants hispanophones est identique à celui des natifs pour la lecture d'un texte. Elle remarque toutefois que 7.83% des liaisons sont réalisées sans enchaînement (p.156) bien que l'espagnol soit, comme le français, une langue à syllabation ouverte.

Dans l'étude de Racine et Detey (2015), les apprenants japonophones produisent moins de liaisons (72.48% sans séjour, 85.10% avec séjour) que les hispanophones (90% sans séjour, 92.29% avec séjour), mais font plus de progrès pour les liaisons non enchaînées après un séjour (1.68 %) que les hispanophones (7.51 %). On voit ici le rôle de la langue maternelle dans la production de liaisons.

Ces études indiquent une maîtrise native ou quasi native des liaisons mais soulignent également quelques difficultés communes aux différents apprenants

FL2 : les liaisons non enchaînées, la prononciation des voyelles nasales, les productions de consonnes finales muettes, et la prononciation « orthographique » (prononciation du d de *grand* en liaison au lieu de t, par exemple). Les chercheurs observent l'effet positif d'un séjour en milieu francophone. Ils signalent des différences de productions dépendant des tâches (lecture/ production semi-guidée), et l'effet de la langue maternelle sur la maîtrise de certains traits phonologiques.

#### **4. Méthodologie**

Cette partie explique la méthodologie avec une description des participants, de la tâche (de lecture), du texte, de la collecte et du traitement des données, et une description des deux banques de données de fréquence utilisées.

##### **4.1. Les participants**

20 francophones (âge moyen : 38 ans) ayant tous fait au moins deux ans d'études après le lycée se sont enregistrés lisant le texte spécialement conçu pour cette étude à haute voix. Parmi ces 20 francophones il y avait :

- a. 12 francophones majoritaires, dorénavant appelés FL1maj (francophones vivant dans une région où le français est la langue officielle parlée par la majorité des habitants), dont 4 Français, 1 Belge, 5 Québécois et 2 sujets bilingues qui ont appris le français avec au moins un parent francophone en Belgique et en Ontario. Les deux sujets bilingues ont appris et pratiqué le français à la maison, et à l'école.
- b. 8 francophones minoritaires, dorénavant appelés FL1min (francophones vivant dans une région où le français n'est pas la langue maternelle de la majorité des habitants) dont 4 Franco-Ontariens et 4 Africains (venant du Burundi, du Congo, du Mali, et du Sénégal).

Ces deux groupes sont différents puisque le français est une langue maternelle pour les Franco-Ontariens, et une L2 pour les Africains. Toutefois, les



deux groupes ont été rassemblés car ils ont chacun eu un input avoisinant la moitié de celui des FL1maj. Les Franco-Ontariens de cette étude parlaient le français seulement à la maison, et l'anglais le reste du temps ; les Africains parlaient le français à l'école et une langue africaine le reste du temps. Les deux groupes parlent couramment le français, et ont tous deux eu un contact linguistique restreint, ce qui permet d'évaluer le rôle de l'input dans cette étude.

37 apprenants français L2, dorénavant appelés apprenants AFL2 (âge moyen: 19 ans) inscrits dans un cours de première année de français dans une université ontarienne anglophone se sont enregistrés lisant le même texte. Pour s'inscrire dans ce cours universitaire, ils devaient avoir étudié le français au moins 9 ans : de la 4<sup>ème</sup> année (9 ans) à la 12<sup>ème</sup> année (17 ans). Sinon, ils ont passé un test de placement qui a déterminé que leur niveau de français correspondait à ce cours.

Il existe au Canada différents systèmes scolaires permettant d'apprendre le français L2 :

- *Le programme « cadre »* : approche traditionnelle où le français est enseigné en tant que matière, 20 à 50 minutes par jour, selon les écoles et les niveaux.
- *Le programme intensif de français* : le français est enseigné en tant que matière, et au moins 25% de l'enseignement est en français à l'école élémentaire.
- *Le programme d'immersion en français* : le français est enseigné en tant que matière, et au moins 50% de l'enseignement est en français à l'école élémentaire.

Selon Tennant (2015 : 69), les AFL2 ayant étudié le français en programme cadre jusqu'en 12<sup>ème</sup> année ont un niveau A2 d'après le CEFRL (Conseil de l'Europe, 2001), et ceux ayant étudié le français en immersion ont au moins un niveau B1.

Les AFL2 de cette étude avaient appris le français dans différents systèmes (cadre, immersion, programme intensif) et avaient étudié un nombre d'heures différents. De plus, beaucoup d'entre eux avaient étudié dans plusieurs

systèmes (cadre et immersion). Dans cet article, tous les AFL2 ont été réunis en un seul groupe car les comparaisons de ces groupes font l'objet d'une étude séparée.

## **4.2. La tâche et les analyses**

### *4.2.1. Choix de tâche (lecture)*

Afin de comparer les mêmes contextes de productions des groupes natifs et non natifs, la lecture à haute voix de phrases paraissait être le choix le plus judicieux. La parole semi-guidée permet d'évaluer ce que les apprenants maîtrisent, mais ne permet pas d'évaluer ce qu'ils ne maîtrisent pas. De plus, Ils parlent plus lentement que les francophones et produisent ainsi moins de sites liaisonnants, avec un vocabulaire et des structures limitées (moins ou pas de pronoms et peu ou pas d'expressions figées). Comme l'indique Howard (2013 : 227) :

« L'occurrence peu fréquente de certains contextes dans nos données indique ainsi l'importance d'études futures impliquant d'autres tâches linguistiques pour analyser l'acquisition de la liaison en L2 [...] Par exemple, des recherches fondées sur la lecture de séquences de mots en contexte liaisonnant [...] Enfin, les résultats que nous avons présentés soulignent l'importance d'une approche lexicale dans les recherches futures. »

### *4.2.2. Le texte*

Pour évaluer les connaissances des francophones et des AFL2 de certains contextes de liaisons, un texte comprenant 51 LC et 251 a été conçu à cet effet (voir Annexe). La lecture d'un texte permet d'examiner certains contextes morphosyntaxiques particuliers (déterminant, adjectif, préposition, etc.), phonologiques (nasales, liaison avec /r/), lexicaux (expressions figées), et

fréquentiels. C'est surtout la seule façon de comparer exactement les mêmes séquences et contextes d'un groupe à l'autre.

Le texte a été conçu pour tester différents contextes phonétiques (h muets : des histoires ; h aspirés : les handicaps ; nasales : aucun effort ; semi-voyelles : aux yeux ; /r/ : premier amour) ; syntaxiques (déterminant + nom : les affaires ; adjectif + nom : bon ami ; expressions figées : accent aigu) ; lexicaux (belles acclamations) ; et l'effet d'une consonne « orthographique » (grand, second, neuf, six, dix). Les mots du texte étaient aussi fréquents que possible, afin que les AFL2 en reconnaissent et comprennent la plupart. S'ils étaient plus rares, ils avaient un équivalent anglais ou un mot apparenté : *accent (accent)*, *yogourt (yogurt)*.

Les participants disposant de peu de temps au laboratoire, le texte est aussi court que possible constitué de phrases séparées.

#### 4.2.3. Collecte et traitement des données

Les participants se sont enregistrés eux-mêmes lisant seuls le texte (les phrases) à haute voix, texte qu'ils pouvaient préparer. Ils avaient une copie papier, sur laquelle ils pouvaient faire des annotations.

12 codes possibles ont été attribués pour la prononciation des consonnes de liaison : a = bien réalisée, bien enchaînée dès la première tentative ; b = non réalisée ; c = consonne liaison non enchaînée ; c = consonne liaison mal prononcée ; e = reprise avec bonne liaison ; f = reprise sans liaison ; g = reprise avec mauvaise liaison ; h = fausse liaison (pataquès): consonne rajoutée pour faire une liaison ; i = fausse liaison + reprise + correction (et bonne liaison) ; j = phonème du mot d'avant mal prononcée (par exemple, vieille) ; k = phonème ou syllabe du mot d'après mal prononcée ; l = fausse liaison + reprise pas de liaison.

Dans ce système de codage, les liaisons sont considérées comme « correctement » réalisées (a) seulement si elles ont été enchaînées, si la consonne de liaison était bien prononcée, si le phonème précédant et suivant la consonne de liaison et la séquence ont été réalisées conformément aux normes orthoépiques du français standard. Dans les autres études sur la liaison, le critère était

généralement que la consonne de liaison soit prononcée. Par exemple, si la consonne de liaison est non enchaînée, les auteurs considèrent que c'est la bonne consonne de liaison - non enchaînée ; tandis dans cette étude une consonne non enchaînée est considérée comme la prononciation fautive de la consonne finale du Mot1 (et non une liaison correctement réalisée). Ce système de codage est plus strict que les autres études.

Les productions ont ensuite été regroupées dans deux catégories générales: (i) Liaisons correctement réalisées ; (ii) Liaisons incorrectement réalisées. Le but de ces 12 codes possibles était d'évaluer les causes des productions non conformes à la norme pour une analyse plus détaillée, notamment les liaisons non enchaînées dont nous parlerons brièvement ici. Une évaluation de locuteur natif semblait suffisante pour le codage des données. Dans les rares cas de doutes, les enregistrements ont été écartés de l'étude.

Le logiciel *GoldVarb* a été utilisé pour calculer les pourcentages et statistiques. Les productions de liaisons ont été corrélées aux renseignements des questionnaires (âge, années d'étude du français et autres questions linguistiques) et à des informations lexicales, syntaxiques et fréquentielles relatives aux Mot1 et au Mot2. *GoldVarb* est un logiciel conçu pour l'analyse de la variation linguistique (voir Sankoff, Tagliamonte et Smith, 2005, pour des explications plus détaillées). Il permet de sélectionner et d'analyser certaines variables ayant un effet significatif sur la variable linguistique (ici la liaison), et d'effectuer une analyse de régression multiple afin de déterminer les effets de certaines variables indépendantes (la fréquence du Mot1, par exemple) ici sur la réalisation de la liaison. Un poids de facteur supérieur à 0.500 indique un effet positif sur la variable dépendante (la réalisation de la liaison est favorisée), et un poids inférieur à 0.500 un effet négatif (la réalisation de la liaison n'est pas favorisée).

Puisque cette étude se concentre sur le rôle de la fréquence, des analyses *GoldVarb* ont été effectuées pour vérifier la signification statistique des différents types de fréquence (Mot1, Mot2, et taux de cooccurrence entre le Mot1 et le Mot2 avec *Lexique 3*), et pour comparer les productions des différents groupes.

### 4.3. Les banques de données *Lexique 3* et *Vocabprofil*

*Lexique 3* est une banque de données de la langue orale basée sur les recherches de New, Pallier, Ferrand et Matos (2001) sur les sous-titres de 9474 films et séries de télévisions francophones et de films sous-titrés en français, ce qui constitue un total de 1.9 million de mots. Le logiciel indique le nombre de fois qu'un mot apparaît dans la banque de données.

Nous avons d'abord cherché la fréquence des Mots1, des Mots2 et de la cooccurrence de ces mots avec *Lexique 3* (dorénavant abrégé en *Lex3*). Puis, nous avons divisé ces groupes en tranches de fréquences (mots moins fréquents : 0-40 ; moyennement fréquents : 41-500 ; fréquents : 501+). Ces indices de fréquence ont ensuite été codés pour chaque mot et corrélés aux productions de liaisons.

Ensuite, nous avons établi un indice de fréquence avec *Vocabprofil*, version francophone de la version originale anglophone *Vocabprofile* de Laufer et Nation (1995). *Vocabprofil* calcule la fréquence d'utilisation des mots d'après les critères suivants :

- K1 : les 1000 mots les plus fréquemment utilisés (1 à 1000) ;
- K2 : les 1000 mots suivants les plus utilisés (1001 à 2000) ;
- K3 : les 1000 mots suivants les plus utilisés (2001 à 3000) ;
- Mots Hors-liste : les mots absents des trois catégories précédentes (à partir de 3001).

Ces deux banques de données fournissent des informations différentes en se basant sur des corpus différents. Dans *Lex3*, la fréquence des items est calculée à partir d'un million de mots de sous-titrages de films ; dans *Vocabprofil*, la fréquence est calculée selon la fréquence relative à d'autres mots de l'écrit. Par exemple, le mot *elle* dans *Lex3* a un taux de fréquence de 4520, et *il* a un taux de fréquence de 13 222 ; *il* est donc trois plus fréquemment utilisé que *elle*. Par contre, dans *Vocabprofil*, *il* et *elle* appartiennent tous deux à la même catégorie K1. C'est la raison pour laquelle l'utilisation des deux banques de données permet d'évaluer et de comparer différentes informations.

Dans une séquence de deux mots, le Mot1 fournit des informations lexicales mais aussi morphosyntaxiques en déterminant le genre et le nombre. Le Mot2 représente principalement la fréquence lexicale. Enfin, la fréquence de cooccurrence détermine la « cohésion » de la fréquence de la séquence et joue un rôle important au niveau du traitement des données, surtout avec les expressions figées. Ces trois types de fréquence lexicale sont interreliés et jouent ensemble un rôle déterminant pour l'acquisition et la réalisation des liaisons. Si une séquence est entendue et utilisée très fréquemment, elle est bien ancrée dans le lexique du locuteur, et il y a plus de chances pour que la liaison soit produite.

## **5. Résultats des analyses de fréquence des Mots1 et des Mots2**

Le rôle de la fréquence lexicale sera analysé dans l'ordre suivant : la fréquence du mot1 basée sur *Vocabprofil* et sur *Lex3*, la fréquence du mot2 basée sur *Vocabprofil* et sur *Lex3*, et la fréquence de cooccurrence des mot1 et mot2 basée sur *Lex3*.

### **5.1. Analyse du Mot1 basée sur *Vocabprofil* et *Lex3***

#### *5.1.1. Analyse du Mot1 basée sur *Vocabprofil**

Dans l'analyse des Mots1 avec *Vocabprofil* (tableau 2), il n'y a pas de mots hors-liste car ce sont des articles et autres mots très fréquents.

Les FL1maj réalisent les LC avec les mots les plus fréquents à 97.4%, avec les moyennement fréquents à 70.8%, et avec les moins fréquents à 58.3%. Les FL1min réalisent les LC avec les mots les plus fréquents à 87.8%, avec les moyennement fréquents à 50%, et avec les moins fréquents à 50%. Les AFL2 réalisent les LC avec les mots les plus fréquents à 62.7%, avec les moyennement fréquents à 24.3%, et avec les moins fréquents à 37.8%.

Les tendances des trois groupes sont globalement les mêmes. Les Mots1 les plus fréquents ont le taux de réalisation le plus élevé pour tous les locuteurs. Toutefois, les productions des FL1min sont les mêmes pour les mots

moyennement fréquents (50%) et les mots moins fréquents (50%). Les AFL2 produisent plus de liaisons avec les mots moins fréquents (37.8%) qu'avec les mots moyennement fréquents (24.3%).

Fréq.Mot1	FL1maj			FL1min			AFL2		
	LR	LP	%	LR	LP	%	LR	LP	%
<i>Vocabprofil</i>									
3 (K1)	561	576	<b>97.4</b>	337	384	<b>87.8</b>	1113	1776	<b>62.7</b>
2 (K2)	17	24	<b>70.8</b>	8	16	<b>50</b>	18	74	<b>24.3</b>
1 (K3)	7	12	<b>58.3</b>	4	8	<b>50</b>	14	37	<b>37.8</b>
0 (hors-liste)	0	0	<b>0</b>	0	0	<b>0</b>	0	0	<b>0</b>
Total	585	612	<b>95.6</b>	349	408	<b>85.5</b>	1145	1887	<b>60.7</b>

**Tableau 2. LC et fréquence du Mot1 basée sur *Vocabprofil***  
**LR : liaisons réalisées - LP : liaisons potentielles**

Il y a une corrélation entre la fréquence des Mots1 (tableau 3) et la production de LC des trois groupes. Les FL1maj réalisent les LC avec les mots les plus fréquents à 99.5%, avec les moyennement fréquents à 98%, et avec les moins fréquents à 85.4%. Les FL1min réalisent les LC avec les Mots1 les plus fréquents à 98.6%, à 81% avec les moyennement fréquents, et à 74% avec les moins fréquents. Les AFL2 réalisent les LC avec les mots les plus fréquents à 74.2%, avec les moyennement fréquents à 59.8%, et avec les moins fréquents à 41.9%.

Plus les Mots1 sont fréquents, plus les trois groupes réalisent la liaison. Avec les mots les plus fréquents, il y a peu de différences entre les groupes majoritaires (99.5%) et minoritaires : (98.6%). Les FL1maj produisent 98% de liaisons avec les mots moyennement fréquents, alors que les FL1min en produisent 81%. Enfin, avec les mots les plus rares, les francophones majoritaires produisent 85.4% de liaisons, et les FL1min en produisent 74%. Ce sont avec les mots les moins fréquents qu'un observe les plus grandes différences entre les groupes. Les liaisons des AFL2 sont produites en plus faibles proportions dans toutes les catégories. On observe un continuum entre les groupes. Les liaisons avec les mots les moins fréquents sont produites par les FL1min à 74%, et les liaisons avec les mots les plus fréquents sont produites dans les mêmes proportions (74.2%) par les AFL2.

Le contact linguistique restreint des FL1min a pour conséquences des productions de liaisons plus basses, en particulier avec les mots les moins fréquents. La fréquence du Mot1 joue donc un rôle important dans la production de liaisons pour les trois groupes. La fréquence du mot1 et du type de structure (*type frequency*) qui lui est associé a un impact sur la production de liaison.

Fréq.Mot1	FL1maj			FL1min			AFL2		
	LR	LP	%	LR	LP	%	LR	LP	%
<i>Lex3</i>									
501+	215	216	<b>99.5</b>	142	144	<b>98.6</b>	494	666	<b>74.2</b>
41-500	247	252	<b>98</b>	136	168	<b>81</b>	465	777	<b>59.8</b>
0-40	123	144	<b>85.4</b>	71	96	<b>74</b>	186	444	<b>41.9</b>
Total	585	612	<b>95.6</b>	349	408	<b>85.5</b>	1145	1887	<b>60.7</b>

**Tableau 3. LC et fréquence du Mot1 basée sur *Lex3***  
**LR : liaisons réalisées - LP : liaisons potentielles**

## 5.2. Analyse du Mot2 basée sur *Vocabprofil* et *Lex3*

### 5.2.1. Analyse du Mot2 basée sur *Vocabprofil*

L'analyse de la fréquence des Mots2 d'après *Vocabprofil* (tableau 4) suggère des tendances similaires à celles obtenues avec l'analyse de fréquence des Mots1 d'après *Vocabprofil*. Avec les Mots2 les plus fréquents, les liaisons sont réalisées en plus grande proportion par tous les groupes. Les liaisons avec les Mots2 de la catégorie K1 sont réalisées à 97.7% par les FL1maj, à 91.6% par les FL1min, et à 67.2% par les AFL2. Les liaisons avec les mots de la catégorie K2 sont réalisées à 97.2% par les FL1maj, à 83.3% par les FL1min et à 58.1% par les AFL2.

Par contre, les FL1min produisent deux fois plus de liaisons avec les Mot2 hors-liste (68.8%) qu'avec les mots de la catégorie K3 (31.3%) ; de même, les AFL2 produisent plus de Mot2 hors-liste (34.7%) qu'avec les Mot2 de la catégorie K3 (25.7%).



Fréq. Mot2	FL1maj			FL1min			AFL2		
	LR	LP	%	LR	LP	%	LR	LP	%
<i>Vocabprofil</i>									
3 (K1)	434	444	<b>97.7</b>	271	296	<b>91.6</b>	920	1369	<b>67.2</b>
2 (K2)	70	72	<b>97.2</b>	40	48	<b>83.3</b>	129	222	<b>58.1</b>
1 (K3)	22	24	<b>91.7</b>	5	16	<b>31.3</b>	19	74	<b>25.7</b>
0 (Hors-Liste)	59	72	<b>81.9</b>	33	48	<b>68.8</b>	77	222	<b>34.7</b>
Total	585	612	<b>95.6</b>	349	408	<b>85.5</b>	1145	1887	<b>60.7</b>

**Tableau 4. LC et fréquence du Mot2 basée sur *Vocabprofil***

### 5.2.2. Analyse du Mot2 basée sur *Lex3*

Les mots avaient au départ été divisés en 13 catégories de fréquence. Toutefois, lorsqu'il n'y a aucune production pour certaines catégories, cela empêche l'analyse *Goldvarb*. Les mots ont donc été reclassés en trois catégories : les mots les plus fréquents, moyennement fréquents, et les mots plus rares.

Plus les Mots2 sont fréquents, plus les liaisons sont réalisées par les trois groupes (tableau 5). Les FL1maj réalisent les liaisons avec les Mots2 les plus fréquents à 99.3%, avec les Mots2 moyennement fréquents à 95.1%, et avec les mots les plus rares à 90.6%. Les FL1min réalisent les liaisons avec les Mots2 les plus fréquents à 95.7%, avec les Mots2 moyennement fréquents à 79.2%, et avec les mots les plus rares à 75.8%. Les AFL2 suivent les mêmes tendances que les groupes de francophones. Ils réalisent les liaisons avec les Mots2 les plus fréquents à 75%, avec les Mots2 moyennement fréquents à 51.1%, et avec les mots les plus rares à 47.3%. C'est avec les mots moyennement fréquents et les mots les plus rares que les différences entre les groupes sont les plus grandes.

Fréq. Mot2	FL1maj			FL1min			AFL2		
	LR	LP	%	LR	LP	%	LR	LP	%
<i>Lex3</i>									
101+	274	276	<b>99.3</b>	176	184	<b>95.7</b>	638	851	<b>75</b>
21-80	137	144	<b>95.1</b>	76	96	<b>79.2</b>	227	444	<b>51.1</b>
0-20	174	192	<b>90.6</b>	97	128	<b>75.8</b>	280	592	<b>47.3</b>
Total	585	612	<b>95.6</b>	349	408	<b>85.5</b>	1145	1887	<b>60.7</b>

**Tableau 5. LC et fréquence du Mot2 basée sur *Lex3***

### 5.3. Analyse de cooccurrence Mots1-Mots2 basée sur *Lex3*

Les FL1maj réalisent un plus grand nombre de liaisons avec les mots au taux de cooccurrence le plus élevé (99.2%) (tableau 6) qu'avec les mots au taux de cooccurrence moyennement élevé (98.8%), et plus de liaisons avec les mots au taux de cooccurrence moyennement élevé qu'avec les mots au taux de cooccurrence plus faible (92%). Les différences entre les groupes majoritaires et minoritaires sont plus grandes avec les séquences au taux moyennement fréquent (98.9% vs 90%), et au taux le plus bas (92% vs 77.5%).

Les AFL2 suivent la même tendance que les groupes francophones. Ils font plus de liaisons avec les séquences au taux de cooccurrence très élevé (75.9%) qu'avec celles au taux de cooccurrence moyennement élevé (71.2%), et font plus de liaisons avec les séquences au taux de cooccurrence moyennement élevé (71.2%) qu'avec celles au taux de cooccurrence plus bas (47.7%). Il y a relativement peu de différences dans la production des AFL2 entre le taux de fréquence élevé (75.9%) et moyennement élevé (71.2%). Par contre, il y a une nette différence entre les taux de fréquence les plus bas (47.7% pour les AFL2, comparativement aux 92% des FL1maj). C'est dans cette catégorie que les différences sont les plus grandes entre les francophones et les AFL2.

Fréq. cooc.	FL1maj			FL1min			AFL2		
	LR	LP	%	LR	LP	%	LR	LP	%
<i>Lex3</i>									
500+	131	132	<b>99.2</b>	86	88	<b>97.7</b>	309	407	<b>75.9</b>
60-449	178	180	<b>98.8</b>	108	120	<b>90</b>	395	555	<b>71.2</b>
0-59	276	300	<b>92</b>	155	200	<b>77.5</b>	441	925	<b>47.7</b>
Total	585	612	<b>95.6</b>	349	408	<b>85.5</b>	1145	1887	<b>60.7</b>

Tableau 6. LC et fréquence de cooccurrence des Mot1-Mot2 basée sur *Lex3*

## 6. Analyses statistiques *GoldVarb*

L'étude statistique permet de vérifier les pourcentages de productions de liaisons les plus pertinents, selon les différents types de fréquence. Un poids de facteur supérieur à 0.500 signifie que statistiquement la production est favorisée, et inférieur à 0.500, qu'elle est défavorisée. Cette analyse (tableau 7) confirme que la

liaison est favorisée chez les FL1maj (0.891), et les FL1min (0.692) mais défavorisée chez les AFL2 (0.437). Les résultats plus bas des AFL2 de cette étude sont différents de ceux des études précédentes (dans la partie 3.2.) en raison notamment des difficultés de la tâche (lecture à haute voix de contextes plus variés) et d'un codage plus strict.

	<b>Liaisons réalisées</b>	<b>Liaisons possibles</b>	<b>%</b>	<b>Poids du facteur</b>
Input: .725				
<b>Groupe de participants :</b>				
FL1maj	585	612	<b>95.6%</b>	0.891
FL1min	349	408	<b>85.5%</b>	0.692
AFL2	1145	1887	<b>60.7%</b>	0.437
Ecart				45
TOTAL	2079	2907	<b>71,5%</b>	

**Tableau 7. Taux de réalisation des LC, francophones et AFL2**

### **6.1. La fréquence des Mots1 avec Lex3**

Pour l'analyse *Goldvarb*, les statistiques du tableau 7 indiquent des différences significatives entre les FL1maj et les FL1min, mais des tendances toutefois proches ; réunir les deux groupes permet d'établir les tendances générales des francophones.

Les groupes de fréquence ont été recodés en trois catégories : mots les plus fréquents (3), moyennement fréquents (2) et moins fréquents (1) à partir de la banque de données de *Lex3*. Statistiquement (tableau 8), la liaison est favorisée pour les mots les plus fréquents (0.857), et défavorisée pour les mots moyennement fréquents (0.343) et les moins fréquents (0.175) pour les francophones.

	<b>Liaisons réalisées</b>	<b>Liaisons possibles</b>	<b>%</b>	<b>Poids du facteur</b>
Input : .952				
<b>Fréquence Mot1 Lex3 :</b>				
3 : mots plus fréquents	357	360	<b>99.2%</b>	0.857
2 : mots moyennement fréquents	383	420	<b>91.2%</b>	0.343
1 : mots moins fréquents	194	240	<b>80.8%</b>	0.175
Écart				68
TOTAL	934	1020	<b>91.6%</b>	

**Tableau 8. LC de tous les francophones. Fréquence Mot1 d'après Lex3**

En ce qui concerne le rôle de la fréquence des Mots1 (tableau 9) sur la réalisation de LC chez les AFL2, la liaison est favorisée pour les mots les plus fréquents (0.644) et défavorisée pour les mots moyennement fréquents (0.475) et les moins fréquents (0.329). La tendance est similaire pour les francophones et les AFL2. La fréquence des Mots1 a un impact sur leur traitement de la liaison.

	<b>Liaisons réalisées</b>	<b>Liaisons possibles</b>	<b>%</b>	<b>Poids du facteur</b>
Input: .673				
<b>Fréquence Mot1 Lex3:</b>				
3: mots plus fréquents	1576	1998	<b>78.9%</b>	0.644
2 : mots moyennement fréquents	1516	2331	<b>65.0%</b>	0.475
1 : mots moins fréquents	668	1332	<b>50.2%</b>	0.329
Écart				31
TOTAL	3760	5661	<b>66.4%</b>	

**Tableau 9. LC des AFL2. Fréquence Mot1 d'après Lex3**

### **6.2. La fréquence des Mots2 avec Lex3**

En ce qui concerne le rôle de la fréquence des Mots2 (tableau 10) sur la réalisation de liaisons chez les *francophones*, la liaison est favorisée pour les mots les plus fréquents (0.830) et moyennement fréquents (0.707) et elle est défavorisée pour les mots les moins fréquents (0.263). Ainsi, plus les Mots2 sont fréquents, plus les

liaisons ont tendance à être réalisées. La fréquence des Mots2 est donc un facteur qui détermine la production de liaisons des francophones.

	<b>Liaisons réalisées</b>	<b>Liaisons possibles</b>	<b>%</b>	<b>Poids du facteur</b>
Input: .942				
<b>Fréquence Mot2 Lex3:</b>				
3: mots plus fréquents	79	80	98.8%	0.830
2 : mots moyennement fréquents	429	440	97.5%	0.707
1 : mots moins fréquents	426	500	85.2%	0.263
Écart				57
TOTAL	934	1020	91.6%	

**Tableau 10. LC de tous les francophones. Fréquence Mot2 d'après Lex3**

En ce qui concerne le rôle de la fréquence des Mots2 (tableau 11) sur la réalisation de LC des AFL2, la liaison est favorisée pour les mots les plus fréquents (0.583) et moyennement fréquents (0.629) et défavorisée pour les mots les moins fréquents (0.373). La fréquence des Mots2 a donc un impact dans le traitement des liaisons des AFL2.

	<b>Liaisons réalisées</b>	<b>Liaisons possibles</b>	<b>%</b>	<b>Poids du facteur</b>
Input: .674				
<b>Fréquence Mot2 Lex3:</b>				
3: mots plus fréquents	330	444	<b>74.3%</b>	0.583
2 : mots moyennement fréquents	1901	2442	<b>77.8%</b>	0.629
1 : mots moins fréquents	1529	2775	<b>55.1%</b>	0.373
Écart				26
TOTAL	3760	5661	<b>66.4%</b>	

**Tableau 11. LC des AFL2. Fréquence Mot2 d'après Lex3**

### **6.3. La fréquence de cooccurrence des Mots1 et Mots2 avec Lex3**

Pour le rôle de la fréquence de cooccurrence des mots1 et Mots2 avec Lex3 dans la réalisation des LC chez les *francophones* (tableau 12), la liaison est favorisée dans les séquences au taux de cooccurrence le plus élevé (0.848), elle est neutre

pour les séquences au de cooccurrence moyennement élevé (0.498), et elle est défavorisée pour le taux le plus bas (0.269). Ainsi, plus le taux de cooccurrence est élevé, plus les liaisons ont tendance à être réalisées par les francophones.

	<b>Liaisons réalisées</b>	<b>Liaisons possibles</b>	<b>%</b>	<b>Poids du facteur</b>
Input : .943				
<b>Fréquence cooccurrence Lex3</b>				
3 : mots plus fréquents	277	280	<b>98.9%</b>	0.848
2 : mots moyennement fréquents	245	260	<b>94.2%</b>	0.498
1 : mots moins fréquents	412	480	<b>85.8%</b>	0.269
Écart				55
TOTAL	934	1020	<b>91.6%</b>	

**Tableau 12. LC des francophones. Fréquence cooccurrence Lex3**

Pour le rôle de la fréquence de cooccurrence (tableau 13) sur la réalisation de LC chez les **AFL2**, la liaison est favorisée pour les taux de fréquence de cooccurrence les plus élevés (0.655) et moyennement élevés (0.599) et elle est défavorisée pour les taux de cooccurrence les plus bas (0.356). La fréquence de cooccurrence a donc un impact dans le traitement des liaisons des AFL2.

	<b>Liaisons réalisées</b>	<b>Liaisons possibles</b>	<b>%</b>	<b>Poids du facteur</b>
Input : .676				
<b>Fréquence cooccurrence Lex3 :</b>				
3 : mots plus fréquents	1241	1554	<b>79.9%</b>	0.655
2 : mots moyennement fréquents	1093	1443	<b>75.7%</b>	0.599
1 : mots moins fréquents	1426	2664	<b>53.5%</b>	0.356
Écart				30
TOTAL	3760	5661	<b>66.4%</b>	

**Tableau 13. LC des AFL2. Fréquence cooccurrence d'après Lex3**

## 7. Discussion

Cette étude suggère que les francophones majoritaires et minoritaires et les AFL2 se basent sur des stratégies lexicales dans leur traitement de la liaison avec la

fréquence du Mot1, la fréquence du Mot2 et la fréquence de cooccurrence du Mot1 et du Mot2. L'utilisation de stratégies basées sur la fréquence lexicale n'exclue pas nécessairement un traitement phonologique. Les stratégies lexicales pourraient justement permettre l'acquisition de principes abstraits ; ou bien des stratégies lexicales pourraient être utilisées quand les stratégies phonologiques sont moins pertinentes.

Les résultats de cet article s'accordent avec Wauquier et Shoemaker (2013) pour proposer un modèle qui serait un continuum entre des stratégies lexicales et phonologiques dans le traitement des données en L1 ou bien un modèle multidimensionnel. Plus les structures sont fréquentes et comprennent un grand nombre d'items, plus les locuteurs L1 se baseraient sur des stratégies phonologiques abstraites. Moins les structures sont fréquentes, moins elles contiennent d'items, plus les locuteurs se baseraient sur des stratégies de fréquence lexicale.

Si des structures syntaxiques sont plus rares telles que [adjectif + nom], les locuteurs L1 et L2 se basent sur la fréquence de cooccurrence et sur les items eux-mêmes. Ce phénomène se retrouve dans cette étude, celles présentées en parties 3.1. et 3.2. et dans l'étude du corpus PFC, par exemple les liaisons dans les séquences *grand honneur* et *grand émoi*.

Si l'input des AFL2 est plus restreint, leur banque de données incomplète ne leur permet pas de mettre en place des règles abstraites. C'est la raison pour laquelle ils traitent les informations au coup par coup, surtout avec des stratégies lexicales, tout comme le font les enfants L1 en début d'acquisition. c'est ainsi qu'ils produisent une majorité de liaisons correctes, mais ils ne sont pas en mesure d'appliquer les règles aux cas jamais ou moins entendus, même pour les catégories morphosyntaxiques les plus fréquentes [déterminant + nom].

## **8. Conclusion**

Les trois types de fréquence lexicale : la fréquence du Mot1 (la fréquence structure morphosyntaxique), la fréquence du Mot2, et la fréquence de

cooccurrence du Mot1 et du Mot2 ont un impact sur le traitement de la liaison des francophones majoritaires et minoritaires. Les trois types de fréquence jouent tous un rôle statistiquement significatif (0.857, 0.848, 0.830) sur leur production de liaisons, et l'importance de chaque type de fréquence est semblable ( $\geq 0.830$ ). Les trois types de fréquence pourraient se renforcer les unes les autres pour un traitement efficace de la liaison.

Les AFL2 traitent eux aussi les liaisons selon le même modèle, en utilisant les trois types de fréquence. En L2, l'input se réduit souvent à la salle de classe. Il est donc nécessaire que les enseignants fournissent aux apprenants un input le plus riche possible en classe, comprenant des répétitions des mêmes items et structures (avec des exercices variés qui ne paraissent pas répétitifs) et si possible en dehors de la classe, avec des exercices d'écoute, de discrimination, et de production d'un grand nombre d'items, pour faciliter l'apprentissage des items et l'acquisition des structures correspondantes.

### ***8.1 Limites de l'étude***

Il y avait ici 20 locuteurs francophones avec peu de sujets dans chaque groupe. Il serait utile de reproduire une telle étude avec un plus grand nombre de participants dans chaque groupe et un plus grand nombre d'apprenants AFL2. Une étude qui contrôlerait davantage l'input des apprenants et en évaluerait les effets, ajouterait aux recherches basées sur l'usage. Enfin, dans un contexte canadien, il serait pertinent d'utiliser des banques de données de l'oral plus susceptibles de correspondre à la fréquence de l'input réel des apprenants canadiens.

### **Références**

- Adda-Decker, M., Boula de Mareüil, P., & Lamel, L. (1999). Pronunciation variants in French: schwa & liaison, *ICPhS*, 2239-2242, San Francisco.
- Ashby, W. J. (2003). La liaison variable en français parlé tourangeau : une analyse en temps réel *Colloque AFLS Tours : 25-27 septembre 2003*.



- Bybee, J. (2001). Frequency effects on French liaison. In J. Bybee & P. Hopper (Eds.), *Frequency and the Emergence of Linguistic Structure* (pp. 337-359). Amsterdam: Benjamins.
- Bybee, J. (2005). Liaison: effets de fréquence et constructions. *Langages*, 158, 24-37.
- Boula de Mareuil, P., Adda-Decker, M., & Gendner, V. (2003). Liaisons in French: a corpus-based study using morpho-syntactic information. Communication au 15<sup>ème</sup> congrès annuel de ICPHS, Barcelone, 1329-1332.
- Chabanal, D. (2003). *Un aspect de l'acquisition du français oral : la variation socio-phonétique chez l'enfant francophone*. Doctorat de Sciences du langage, Université Paul Valéry, Montpellier III.
- Chabanal, D. et Liégeois, L. (2014). Production de liaisons dans l'input parental. In C. Soum- Favaro, A. Coquillon & J.-P. Chevrot (eds.), *La liaison : approches contemporaines*. 263-282. Berne: Peter Lang.
- Charliac, L., & Motron, A.-C. (1998). *Phonétique progressive du français avec 600 exercices*. Paris : CLE International.
- Chevrot, J.-P., Dugua, C., et Fayol, M. (2005). Liaison et formation des mots en français; un scénario développemental. *Langages*, 158, 38-52.
- Conseil de l'Europe (2001). *Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer*. Paris : Didier.
- Delattre, P. (1955). Les facteurs de la liaison facultative en français. *The French Review*, 29 (1), 42-49.
- Detey, S., Kawaguchi, Y. & Kondo, N. (2015). La liaison chez les apprenants japonophones avancés de FLE. *Bulletin VALS- ASLA*, 102, 123-145.
- Dugua, C. (2005). De la liaison à la formation du lexique chez les jeunes enfants francophones. *Le langage et l'homme*, 40 (2), 163-182.
- Dugua, C. (2006). “un nous”/“des nous” ou le rôle de la fréquence sur l'acquisition de la liaison en français. In Loiseau et al. (Eds.). *Autour des langues et du langage : perspective pluridisciplinaire*, Papiers sélectionnés du Colloque international des étudiants chercheurs en didactique des langues et en linguistique, Grenoble : PUG.

- Durand, J., et Lyche, C. (2008). French liaison in the light of corpus data. *Journal of French Language Studies*, 18 (1), 33-66.
- Durand, J., Laks, B., Calderone, B. & Tchobanov, A. (2011). Que savons-nous de la liaison aujourd'hui? *Langue française*, 169, 103-135.
- Ellis, N. C. (2002). Frequency effects in language acquisition: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24, 143-188.
- Harnois-Delpiano, M., Cavalla, C., & Chevrot, J.-P. (2012). L'acquisition de la liaison en L2: étude longitudinale chez des apprenants coréens de FLE et comparaison avec enfants francophones natifs. *SHS Web of Conferences*, 1, 1575-1589.
- Hoff, E. (2002). Causes and consequences of SES-related differences in parent-to-child speech. In: M.H. Bornstein and R.H. Bradley (dir.), *Socioeconomic Status, Parenting, and Child Development* (pp. 147–160). Mahwah, NJ: Lawrence Erlbaum Associates.
- Howard, M. (2005). L'acquisition de la liaison en français langue seconde: Une analyse quantitative d'apprenants avancés en milieu guidé et en milieu naturel. *Corela*, Numéros spéciaux: Colloque AFLS.
- Howard, M. (2013). La liaison en français langue seconde, une étude longitudinale préliminaire. *Language, Interaction and Acquisition*, 4(2), 190–231. doi: 10.1075/lia.4.2.04how
- Laufer, B. & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16, 316-322.
- Mastromonaco, S. M. (1999). *Liaison in French as a second language*. Dissertation Abstracts International. (UMI No. DANQ41234).
- Nicoladis, E. & J. Paradis (2011). Learning to liaise and elide *comme il faut*: evidence from bilingual children. *Journal of Child Language*, 38(4), 701-730
- New, B., C. Pallier, L. Ferrand, & R. Matos (2001). Une base de données lexicales du français contemporain sur Internet : LEXIQUE. *L'Année Psychologique*, 101, 447-462.

- Racine, I. (2015). La liaison chez les apprenants hispanophones avancés de FLE. *Bulletin VALS- ASLA*, 102, 147-167.
- Racine, I. & Detey, S. (2015). La liaison dans un corpus d'apprenants : Le projet « Interphonologie du Français Contemporain » (IPFC), Corpus [En ligne], 15 | 2016, mis en ligne le 15 janvier 2017, consulté le 19 décembre 2018. URL : <http://journals.openedition.org/corpus/3028>
- Sankoff, D., Tagliamonte, S.A, & Smith, E. (2005). *GoldVarb X: A variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto.
- Sampson, R. (2001). Liaison, nasal vowels and productivity. *Journal of French Language Studies*, 11 (2), 241-258.
- Tennant, J. (2015). Canadian anglophone learners' realization of French liaison. *Bulletin suisse de linguistique appliquée* 102: 65–85.
- Thomas, A. (2002). La variation phonétique en français langue seconde au niveau universitaire avancé. *Aile: Acquisition et Interaction en Langue Étrangère*, 17, 101-121.
- Thomas, A. (2004). Phonetic norm versus usage in advanced French as a Second Language. *International Review of Applied Linguistics in Language Teaching (IRAL)*, 42(4), 365-382. doi: 10.1515/iral.2004.42.4.365
- Tomasello, M. (1992) *First Verbs: A Case Study in Early Grammatical Development*. Cambridge University Press.
- Wauquier, S. (2009). Acquisition de la liaison en L1 et L2 : stratégies phonologiques ou lexicales ? ». *Acquisition et interaction en langue étrangère*, AILE, 1, 93-130.
- Wauquier, S. & Shoemaker, E. (2013). Convergence and divergence in the acquisition of French liaison by native and non-native speakers. *Language, Interaction and Acquisition*, 4, 2, 160-188.

## Annexe 1. Le texte de l'étude

1. Cette Anglaise a demandé aux invités les affaires des étudiants.
2. Tes assistants, mes ouvriers et leurs enfants sont tous nos amis.
3. Un bon ami a dit qu'au moyen-âge, à un certain âge, on chantait en plein air.
4. À mon avis, ton enfant ne fait aucun effort pour s'adapter à son école.
5. Son premier amour l'a mise devant le fait accompli, ce qui est un léger ennui.
6. Ils sont allés au dernier étage de l'ancien édifice, mais ne sont pas allés au premier.
7. De nouveaux étudiants ont attendu au second étage du grand immeuble.
8. Ses vieux écrits et ses nouvelles idées lui ont valu de belles acclamations.
9. Un oiseau aux yeux bleus regarde les oies manger des yogourts dans les yachts.
10. Ces beaux Allemands ont eu de folles aventures grâce à leurs faux-airs de gigolos.
11. Il n'y a pas de sous-entendu : on met un accent aigu sur le « e » d'États-Unis.
12. Cet homme a parlé d'un hôpital où il y avait de vieux habits et de vieilles éponges.
13. Des personnes ont écrit des histoires sur ces hôpitaux.
14. Les handicaps des Hollandais et des Hongrois font d'eux des héros.
15. Deux amis de dix ans se sont vus à six heures devant les trois arbres.
16. Il a vingt-trois ans et elle a vingt-cinq ans. A neuf heures, il aura cent ans.
17. Un héros, un garçon intelligent, a eu cet accident affreux dans un bois immense.