



La Résolution de la liaison par des locuteurs natifs et non-natifs

Ellenor Shoemaker, David Birdsong

► To cite this version:

Ellenor Shoemaker, David Birdsong. La Résolution de la liaison par des locuteurs natifs et non-natifs. Acquisition et Interaction en Langue Etrangère, 2008, 27, pp.43-62. halshs-00668888

HAL Id: halshs-00668888

<https://shs.hal.science/halshs-00668888>

Submitted on 5 Mar 2012

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

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

LA RÉOLUTION DE LA LIAISON EN FRANÇAIS PAR DES LOCUTEURS NATIFS ET NON-NATIFS

Ellenor SHOEMAKER & David BIRDSONG

(Université du Texas, Austin, E-U) ¹

RÉSUMÉ

Les processus de perception et de compréhension de la parole nécessitent la segmentation d'un signal acoustique continu où il y a souvent une absence de marques acoustiques explicites de frontière de mot ou de frontière de syllabe, à plus forte raison à l'intérieur des groupes rythmiques en français où se produisent les phénomènes de liaison et d'enchaînement. Considérons les cas où la resyllabation due à ces phénomènes crée une ambiguïté (par exemple, la séquence [œ.nɛʁ] correspond également à *un air* ou à *un nerf*). Cependant, des recherches récentes (Spinelli *et al.*, 2003 ; Wauquier-Gravelines, 1996) ont mis en évidence des différences de durée entre les consonnes de liaison et les consonnes initiales. La tâche d'amorçage de Spinelli *et al.* suggère que ces variations peuvent fournir un indice acoustique dont les auditeurs se serviraient pour identifier des mots dans le flux sonore. Par exemple, le /n/ de *un nerf* serait plus long que celui de *un air*, ce qui produirait un effet d'amorçage qui favoriserait la reconnaissance de 'nerf' par rapport à 'air'. En employant une tâche d'identification à choix obligatoire, notre étude vise à tester directement les capacités perceptives de locuteurs natifs du français et d'apprenants tardifs de français L2 dans les cas où la localisation des frontières est ambiguë. Les résultats suggèrent que, bien que ces variations de durée soient systématique-

-
1. ellishoe@gmail.com ; birdsong@ccwf.cc.utexas.edu
Department of French and Italian, University of Texas at Austin, 1 Univ Station
B7600, Austin, Texas 78712 USA

ment présentes dans le signal acoustique, elles ne sont pas assez nettes pour permettre aux auditeurs de distinguer entre deux options. Même si les deux groupes de participants n'ont pas témoigné d'une sensibilité aux variations acoustiques, leurs réponses se sont avérées fortement similaires, ce qui suggère que les apprenants ont employé des stratégies perceptives comparables à celles des locuteurs natifs. On a donc pris en compte des facteurs relevant des processus du traitement descendant (*top-down*) tels que la fréquence lexicale et la plausibilité des phrases, mais aucun de ces facteurs ne s'est avéré jouer un rôle systématique dans les réponses.

(Mots-clefs : français L2, liaison, resyllabation, perception/ reconnaissance/ segmentation de la parole.)

1. Introduction

La perception et la compréhension de la parole débutent par l'activation automatique des mots possibles correspondant à l'entrée auditive, ou en d'autres termes au signal physique. Dans la mesure où les mots s'enchaînent sans interruption dans un signal auditif continu ², comment les auditeurs parviennent-ils à segmenter le flux en mots discrets ? En l'absence de marque explicite de frontière de mot, les auditeurs doivent souvent utiliser des indices ³ dans le signal acoustique pour apparier l'entrée auditive avec les entrées stockées dans le lexique mental.

Il existe de nombreux modèles qui visent à expliquer le mécanisme mental qui sous-tend la reconnaissance d'un mot selon l'entrée auditive. Les modèles TRACE (McClelland & Elman, 1986) et Shortlist, (Norris, 1994) par exemple, proposent que tous les mots correspondant à l'entrée auditive sont activés conjointement et sont en concurrence avant que l'un d'entre eux ne soit sélectionné. Selon ces modèles, lorsqu'on entend la séquence [œ.nẽ...] en français, tous les mots et toutes les combinaisons de mots qui commencent par

2. Hormis des cas de dictée lente, de scansion, ou lorsque le débit est particulièrement soigné.
3. Les évaluateurs de cette contribution ont demandé une précision sur l'ensemble des indices dont l'auditeur se sert pour comprendre la parole. Cette étude visait à examiner l'exploitation des indices ascendants (*bottom-up*) dans l'activation des candidats du lexique mental selon l'entrée sonore, hors tout contexte sémantique et avant traitement descendant (*top-down*). C'est pour cette raison que les facteurs descendants sont considérés seulement dans les analyses *post-hoc* de cet article, bien qu'ils jouent un rôle incontestable dans la compréhension orale.

cette séquence (*un nain, un inconnu, un imbécile*, etc.) sont simultanément activés. En conséquence, il s'agit au cours du traitement de la parole d'éliminer successivement les nombreux candidats potentiels qui partagent un seul ensemble phonique. Se situant dans le cadre de ces modèles, Content *et al.* (2000) suggèrent que les auditeurs adoptent une stratégie *Syllable Onset Segmentation Hypothesis* (SOSH) selon laquelle ils font coïncider les débuts de syllabes avec ceux des mots pour segmenter sur cette base le flux sonore. Autrement dit, les attaques de syllabes constitueraient des points d'alignement à partir desquels l'accès au lexique mental s'effectue avec le plus d'efficacité. Plusieurs études ont corroboré cette hypothèse en montrant que les coûts de traitement sont élevés lorsque les frontières des syllabes ne sont pas alignées sur celles des mots, ce qui est le cas dans le processus de resyllabation. La resyllabation se produit lorsque la consonne finale d'un mot est (re)-syllabée avec la voyelle initiale du mot suivant au travers d'une frontière de mots.

Vroomen & De Gelder (1997) ont montré que la resyllabation peut être problématique en néerlandais. Les résultats d'une tâche de détection de phonèmes suggèrent qu'il est plus difficile d'identifier un phonème lorsqu'il est resyllabé que lorsqu'il coïncide avec la frontière de mot. Par exemple, dans cette étude le phonème /t/ dans *boot* 'bateau' a été reconnu plus vite dans *de boot die gezonken is* 'le bateau est coulé' (où le /t/ coïncide avec la frontière de mot) que dans *de boot is gezonken* 'le bateau a coulé' (où le /t/ a été resyllabé comme l'attaque du mot suivant *is*).

De manière analogue, Dumay *et al.* (1998) ont montré à l'aide d'une tâche de détection de mots en français que les auditeurs détectent par exemple le mot *lac* plus rapidement dans le non-mot *ZUN.LAC* (où *lac* correspond à l'attaque d'une syllabe) que dans le non-mot *ZU.GLAC* (où la frontière du mot et celle de la syllabe ne sont pas alignées).

Cependant cette stratégie de segmentation pose de nombreux problèmes lorsque les attaques de syllabes ne correspondent pas aux frontières de mots- ce qui est souvent le cas en français parlé où la resyllabation est assez fréquente. Ce phénomène résulte d'une tendance marquée en français pour une structure syllabique consonne-voyelle et de l'existence de processus phonologiques pour retenir cette structure, à savoir l'enchaînement et la liaison. L'enchaînement se produit lorsqu'un mot qui commence par une voyelle est précédé d'un mot qui se termine par une consonne. Cette consonne finale est (re)-syllabée avec la voyelle initiale du mot suivant au travers d'une frontière de mot. Par exemple *une amie* [yn + a.mi] se prononce [y.na.mi]. Par contre, dans le cas de la liaison, on a affaire à une consonne muette (c'est-à-dire, qui ne se prononce pas) en

isolation ou devant un mot à consonne initiale. Devant un mot à voyelle initiale cette consonne sous-jacente se prononce alors et s'attache à la voyelle suivante par enchaînement. Par exemple *un ami* [œ + a.mi] devient [œ.na.mi] où le /n/ se réalise et devient l'attaque de la syllabe suivante. C'est ainsi que ces phénomènes créent de nouvelles frontières syllabiques en français. En conséquence, les frontières syllabiques peuvent facilement se confondre avec celles des mots.

Les phénomènes de la liaison et de l'enchaînement sont très fréquents en français. Comment les locuteurs natifs du français arrivent-ils à reconnaître des mots dans le flux sonore si les frontières syllabiques ne coïncident pas toujours avec celles des mots?

1.1. La perception de la liaison

Il existe plusieurs études sur les effets de la resyllabation due à la liaison sur la perception du français parlé. Gaskell *et al.* (2002) montrent que, bien que la syllabe soit l'unité de base pour la perception chez les locuteurs du français (Mehler *et al.*, 1981), la resyllabation n'empêche pas la perception de la parole continue. Celle-ci peut même faciliter la reconnaissance d'un mot qui commence par une voyelle quand ce mot a subi la resyllabation. En se servant d'un paradigme d'amorçage sémantique intermodal (auditif-visuel), Spinelli *et al.* (2003) montrent aussi que la liaison ne réduit pas l'efficacité de la reconnaissance de la parole. Dans cette étude ils ont sélectionné 27 mots commençant par une voyelle d'une base de données françaises (Trésor de la Langue Française, TLF; Imbs, 1971) en fonction de leur homophonie avec un autre mot français dans des environnements de liaisons (par exemple *air* /ɛʁ/ précédé par *aucun* /o.kœ/ crée une séquence qui est superficiellement identique à celle de *aucun nerf* [o.kœ.nɛʁ]). Ces auteurs ont enregistré 27 paires de phrases ambiguës qui incluaient ces paires minimales (par exemple *Il n'a aucun air* et *Il n'a aucun nerf* [il.na.o.kœ.nɛʁ]). Ces 54 phrases ont été ensuite mélangées avec des phrases contenant des néologismes. Après avoir entendu une de ces phrases, les participants devaient décider si une cible visuelle était un vrai mot français. Les auteurs ont mesuré le temps de réponse et ont trouvé un effet d'amorçage pour les mots à consonne initiale ainsi que pour les mots à voyelle initiale, ce qui suggère que malgré la resyllabation, l'accès aux représentations lexicales n'est pas perturbé et que les auditeurs peuvent différencier entre une consonne de liaison et une consonne initiale selon l'intention du locuteur. Autrement dit, les auditeurs semblent sensibles aux indices acoustiques dans la parole. Dans ces deux études, les auteurs font

l'hypothèse que les auditeurs exploitent des indices acoustiques pour marquer les frontières, mais ne la testent pas explicitement.

Quoique plusieurs études aient établi que d'autres langues exploitent des indices acoustiques afin de segmenter le flux sonore (par exemple Nakatani & Dukes, 1977 ; Warner *et al.* 2004), le rôle des facteurs acoustiques dans la compréhension du français parlé n'a pas été encore à cette date suffisamment exploré. Bien qu'il existe un très large accord chez les chercheurs travaillant sur le français pour considérer que les consonnes de liaison et les consonnes initiales sont identiques au niveau acoustique (Grammont, 1960 ; Passy, 1906), des études plus récentes ont mis en évidence l'existence de différences de durée entre ces deux types de consonnes. Spinelli *et al.* (2003) ont soumis aux analyses acoustiques des échantillons de français parlé qui contenaient des phrases rendues ambiguës par la liaison (i.e. *dernier oignon/ dernier rognon*). Même si les transcriptions phonétiques de ces deux phrases sont censément identiques, cette étude a montré que les consonnes de liaison sont en moyenne 17 % plus courtes que les consonnes initiales. Il en est de même chez Yersin-Besson & Grosjean (1996) qui notent que les consonnes de liaison sont en moyenne 10 % plus brèves que les consonnes initiales.

Ces variations de durée concordent avec celles relevées dans plusieurs recherches sur la production de la parole. En effet, dans beaucoup de langues, y compris le français et l'anglais, les consonnes initiales se révèlent plus longues que la même consonne en position finale (Fougeron, 2001 ; Fougeron & Keating, 1997 ; Gow & Gordon, 1995). Ces résultats confirment la tendance relevée chez les locuteurs à renforcer l'articulation des segments aux frontières des groupes prosodiques (Cho & Keating, 2001 ; Fougeron, 2001).

La présence de ces variations de durée appariée avec les résultats de Spinelli *et al.* (2003) soulève donc la question suivante : dans quelle mesure les auditeurs exploitent-ils ces variations dans la segmentation de la parole ?

1.2. Les apprenants de français comme langue seconde

Le débat concernant la possibilité pour un apprenant tardif d'une deuxième langue d'atteindre un niveau de compétence comparable à celui d'un locuteur natif est loin d'être tranché. Dans le cadre de l'Hypothèse de la Période Critique (voir par exemple Lenneberg, 1967) il est stipulé que la maîtrise complète d'une deuxième langue (L2) par les apprenants qui commencent son étude après la puberté est totalement hors de portée. Mais des recherches plus récentes montrent que cette hypothèse s'avère trop réductrice (voir par exemple Birdsong, 2006 pour un compte-rendu des études sur cette question).

En ce qui concerne la reconnaissance de la parole en particulier, il est prédit que les apprenants tardifs auront d'importantes difficultés à atteindre un niveau d'efficacité équivalent à celui des locuteurs natifs (dorénavant LN). Les résultats de plusieurs études font état des difficultés sérieuses que connaissent les apprenants tardifs dans la compréhension d'une L2 (Cutler, 2002 ; Dupoux *et al.* 1997). Cutler (2002), par exemple, suggère que ces difficultés sont dues au fait que les langues emploient des modalités de segmentations différentes, et qu'il est peu probable qu'un apprenant puisse apprendre et incorporer parfaitement de nouvelles stratégies perceptives dans une langue seconde. Mais, les apprenants tardifs sont-ils toujours voués à l'échec dans leur perception d'une langue étrangère ou leur est-il possible d'atteindre un niveau de traitement comparable à celui d'un natif ?

Flege et ses collègues ont montré que la perception en L2 n'est pas toujours si déficiente et que l'âge d'acquisition n'est pas le seul facteur influant sur les capacités perceptives en L2. MacKay *et al.* (2001) et Flege & MacKay (2004) ont étudié la perception par des italophones des voyelles et des consonnes de l'anglais. L'ensemble de leurs résultats a montré que l'âge⁴ n'est qu'un des facteurs influents dans les capacités perceptives des catégories phonétiques de la L2. Ces auteurs ont trouvé de fortes corrélations entre les capacités des apprenants à percevoir et discriminer des sons en L2, et la fréquence d'emploi de la L2 par rapport à la L1. Autrement dit, si un apprenant utilise fréquemment sa L1 au quotidien, cela peut peser sur sa perception en L2. Les participants à ces études qui utilisaient plus fréquemment la L2 que la L1 ont démontré de meilleures capacités perceptives.

Ces deux études ont examiné des discriminations catégoriques et contrastives en L2. Darcy *et al.* (2007) ont étudié la perception et l'exploitation de la variation allophonique en L2. Dans une série d'expériences, ces auteurs ont testé la perception de l'assimilation chez des apprenants du français (de langue maternelle anglaise). En français parlé, l'assimilation régressive se produit généralement lorsqu'il y a un contraste de voisement. Par exemple, *botte* [bɔt] devient [bɔd] quand cette séquence est suivie d'une consonne voisée comme dans *botte grise* [bɔd.griz]. En anglais par contre l'assimilation résulte beaucoup plus souvent de contrastes de point d'articulation. Par exemple le mot *sweet* 'sucré' [swit] devient [swip] devant une consonne labiale comme dans *sweet melon* 'melon sucré' ou [swik] devant une consonne vélaire comme dans *sweet grapes* 'raisins sucrés'. En se fondant sur une tâche de détection de mot,

4. Dans ces enquêtes, l'âge représente l'âge d'arrivée de l'apprenant dans le pays de la L2 (Canada ici) et non l'âge d'acquisition.

ces auteurs ont trouvé que les apprenants très avancés du français étaient non seulement sensibles aux processus de l'assimilation du français, mais qu'ils pouvaient aussi inhiber les modalités d'assimilation propres à l'anglais lorsqu'ils étaient à l'écoute du français. Les apprenants débutants par contre ont appliqué les modalités d'assimilation de l'anglais en français. Les résultats de cette étude vont à l'encontre de la notion selon laquelle il n'existe qu'un seul système phonologique chez les auditeurs et suggèrent que la plasticité du système perceptif n'est pas remise en cause après la puberté et ce, même dans le cas de variations non-contrastives en L2.

Il n'existe que peu de recherches sur les effets de la resyllabation et de la segmentation du français parlé par des apprenants tardifs du français. Dans une étude effectuée par Dejean de la Bâtie & Bradley (1995), les auteurs ont comparé les stratégies de segmentation de la parole de natifs avec celles de non-natifs. Dans des phrases commençant par *C'est un...* suivi de deux mots, les participants devaient identifier le phonème /t/ au début du deuxième mot. La moitié des stimuli commençait par /t/ et l'autre moitié par une voyelle (*C'est un excellent tableau/excellent acteur ; C'est un vrai tableau/vrai acteur*). Les résultats ont montré que les francophones ne faisaient pas d'erreurs dans l'identification du /t/ initial dans le cas de resyllabation, alors que les réponses des apprenants étaient aléatoires.

De même Stridfeldt (2003) a testé des locuteurs natifs du suédois apprenant le français comme L2. L'expérience était basée sur des paires minimales de néologismes utilisant des consonnes de liaison (*un avas/un navas, les avas/ les zavas*, etc.) enregistrées par un locuteur natif du français qui ignorait l'objectif du test. La tâche des apprenants du français L2 dans ce test était d'écrire ce qu'ils avaient entendu après avoir écouté des phrases commençant par *Je vois* suivies d'un de ces néologismes. Les apprenants n'ont pas réussi à distinguer les phrases et en plus ils ont montré une forte tendance à choisir le mot à voyelle initiale.

On constate donc un manque de recherches sur la perception de la liaison par des apprenants L2. Les deux études citées n'ont testé que des apprenants débutants, ce qui soulève la question de l'identification chez les apprenants très avancés du français L2. Cela nous conduit à aborder la question suivante : les apprenants de français L2 qui sont à un niveau très avancé dans leur apprentissage de la langue peuvent-ils parvenir à reconnaître des mots qui ont subi la resyllabation avec la même efficacité que les LN ?

2. La présente enquête

2.1. Participants

Le groupe expérimental était composé de 15 locuteurs natifs de l'anglais apprenant le français comme seconde langue (âge moyen = 30 ans ; dispersion = 24-55 ans). Au moment de l'expérience, les participants anglophones avaient étudié le français pendant une période allant de 3,5 à 25 ans (moyenne 14 ans) et avaient vécu dans un pays francophone entre 2 mois à 10 ans (moyenne 2 ans). Ils étaient tous professeurs de français L2 ou étudiants de troisième cycle en français à l'Université du Texas. Les apprenants se répartissaient entre 12 femmes et 3 hommes. Le groupe de contrôle était composé de 15 LN de français (âge moyen = 27 ans ; dispersion = 21-39) dont 7 femmes et 8 hommes. Tous les participants habitaient Austin, Texas aux Etats-Unis ou ses alentours au moment du test.

2.2. Stimuli

Les résultats d'amorçage de Spinelli *et al.* (2003) ont conduit à suggérer que les auditeurs (natifs) sont sensibles aux variations acoustiques dues à la resyllabation dans les environnements de liaison et qu'ils exploitent ces variations pour découper le signal continu en unités discrètes. Afin de tester plus directement les capacités perceptives des LN de français ainsi que celles des apprenants tardifs de français comme seconde langue sur la résolution des ambiguïtés dues à la liaison, cette étude a utilisé une tâche d'identification à choix obligatoire. Les phrases à la base de cette étude contenaient des éléments qui rendaient ambigu l'établissement des frontières de mots en raison de la resyllabation (par exemple *Il n'a aucun air* et *Il n'a aucun nerf* [il.na.o.kœ.nɛʁ]). Les mots expérimentaux et les enregistrements ont été empruntés à Spinelli *et al.* (2003). Dans leur étude le premier auteur a enregistré 54 phrases contenant des séquences ambiguës en environnement de liaison et nous en avons sélectionné 24 dans notre étude. Ces 12 paires de stimuli comprennent quatre consonnes de liaison (/g/, /n/, /p/, /t/). La distribution des consonnes est la suivante : /g/- une paire, /n/- trois paires, /t/- trois paires, /p/- cinq paires. Ces phrases ont été présentées oralement, à titre isolé sans contexte sémantiquement contraignant. (Voir Annexe pour la liste complète des stimuli.)

2.3. Passation de l'expérience

La passation s'est effectuée individuellement dans un endroit calme et silencieux. Les stimuli ont été présentés au moyen d'un casque audio Koss UR20 relié à un ordinateur portable Dell Inspiron 600m. Tous les participants ont suivi une séquence d'entraînement avant de commencer l'expérience. Les participants ont entendu une des phrases cibles et immédiatement après deux phrases se sont affichées sur l'écran de l'ordinateur. Les participants devaient indiquer laquelle de ces deux phrases ils venaient d'entendre en appuyant sur une touche de l'ordinateur. L'intervalle interstimulus était de 2,5 secondes. Chacune des 24 phrases a été présentée deux fois pendant l'expérience d'une façon aléatoire. Le test a duré environ 25 minutes.

3. Résultats

Les résultats suggèrent que, bien qu'il existe des variations systématiques de durée entre les consonnes de liaison et les consonnes initiales, ces différences ne sont pas assez marquées pour permettre aux auditeurs de désambigüiser des phrases rendues ambiguës par la liaison. Aucun des deux groupes de participants n'a réussi à distinguer les paires minimales d'une façon systématique, comme le montrent les taux d'identification proches du niveau aléatoire. Le taux d'identification correct des LN de français était de 53.2 % et de 52.7% pour les apprenants, ce qui n'entraîne aucune différence entre les deux groupes ($t = 0.236$, $df = 28$, NS).

Les résultats de Spinelli *et al.* suggéraient que ces variations acoustiques de durée pouvaient provoquer l'activation dans le lexique mental de candidats commençant par une voyelle ainsi que par une consonne. Cependant nos résultats suggèrent que dans les cas d'ambiguïté due à la liaison, cette information acoustique ne suffit pas en soi pour permettre aux auditeurs d'éliminer le candidat impropre une fois activé.

On a voulu tester aussi la possibilité qu'il y ait une corrélation entre les réponses et les durées de certains stimuli pris individuellement, même si globalement le taux d'identification n'était pas significatif. Plus précisément, on a voulu voir si les stimuli contenant les consonnes plus courtes provoqueraient plus de réponses correctes *liaison* alors que les stimuli contenant les consonnes plus longues provoqueraient plus de réponses correctes *non liaison*. Le Tableau 1 recense les durées des consonnes dans les stimuli que nous avons utilisés ainsi que dans l'étude de Spinelli *et al.* (2003).

Tableau 1. Durées des consonnes (en millisecondes) pour les mots commençant par une consonne, pour les mots commençant par une voyelle et différence entre les deux (consonne initiale moins consonne de liaison) des stimuli utilisés dans Spinelli *et al.* (2003) ainsi que dans notre enquête.

Longueur des consonnes initiales en msec		Longueur des consonnes de liaison en msec		Différence entre les deux consonnes en msec
aucun <u>N</u> erf	93	aucu <u>N</u> air	77	16
aucun <u>N</u> ectar	49	aucu <u>N</u> hectare	39	10
aucun <u>N</u> euf	81	aucu <u>N</u> oeuf	83	-2
brillant <u>T</u> act	115	brillan <u>T</u> acte	94	21
grand <u>T</u> amis	109	gran <u>T</u> ami	89	20
grand <u>T</u> ermite	105	gran <u>T</u> ermite	81	24
long <u>G</u> ommage	30	lon <u>G</u> hommage	37	-7
trop <u>P</u> arqués	84	tro <u>P</u> arqués	76	8
trop <u>P</u> artisan	74	tro <u>P</u> artisan	47	27
trop <u>P</u> oureux	124	tro <u>P</u> heureux	94	30
trop <u>P</u> unis	143	tro <u>P</u> unis	109	34
beaucoup <u>P</u> osé	94	beauco <u>P</u> osé	71	23
MOYENNE	91,8		74,8	17

Nous avons d'abord analysé les pourcentages de réponses correctes des LN pour les stimuli à consonne initiale. Si les différences de durée avaient influencé les réponses, on aurait observé une corrélation positive entre la durée des stimuli et le taux de réponse correcte pour les items *non liaison*. Or, nous n'avons pas trouvé de corrélation significative ; en fait, la corrélation s'est avérée légèrement négative (Spearman rho, $r = -.187$, NS). La corrélation était aussi négative et non significative dans une régression simple ($r = -.101$, NS).

En revanche, en ce qui concerne les consonnes les plus courtes on s'attendrait à ce que les pourcentages corrects soient plus élevés pour les items cibles à voyelle initiale. C'est-à-dire que l'on observerait une corrélation négative entre la durée des stimuli particuliers et le taux d'exactitude pour les items *liaison*. Cependant, nous n'avons pas trouvé de corrélation significative entre la durée des consonnes et le pourcentage de réponses correctes pour les items *liaison*. La corrélation était légèrement positive dans une analyse Spearman rho ($r = .261$, NS) et aussi pour une régression simple ($r = .186$, NS).

Les résultats se sont avérés semblables chez les apprenants. Il n'y avait pas non plus de corrélation significative entre les durées et les pourcentages corrects pour les items à consonne initiale dans une corrélation Spearman rho et, contrairement à notre attente, la corrélation aussi était négative ($r = -.147$, NS). Il n'y avait pas non plus de corrélation significative dans une régression simple ($r = -.138$, NS).

Pour les apprenants aussi il n'y avait pas de corrélation significative ni négative entre la durée des items à voyelle initiale et les pourcentages corrects dans une corrélation Spearman rho ($r = .117$, NS). Nous avons constaté les mêmes résultats dans une régression simple ($r = .066$, NS).

Par conséquent, il apparaît que ni les LN ni les apprenants n'ont témoigné d'une sensibilité aux indices acoustiques entre les consonnes de liaison et les consonnes initiales. Ces données suggèrent que la durée des consonnes de liaison ne constitue pas à elle seule une information suffisante pour permettre aux auditeurs de trancher entre les deux options.

3.1. Résultats des LN par rapport à ceux des apprenants

Même si aucun des deux groupes n'a montré une sensibilité perceptive significative à des différences de durée, les deux groupes ont affiché les mêmes tendances pour les paires minimales. Les pourcentages moyens d'exactitude des deux groupes pour chaque stimulus étaient fortement corrélés. Une analyse Spearman rho a montré une corrélation significative entre les LN et les locuteurs non-natifs ($r = .605$, $p < .004$). La corrélation était aussi significative pour les deux groupes dans une régression simple ($r = .572$, $p < .004$).

Des analyses plus approfondies des similarités entre les deux groupes ont révélé aussi des correspondances dans la distribution des réponses. Même si les taux d'identification pour les deux groupes ne sont pas au-dessus du niveau aléatoire, nous avons observé des tendances fortes et convergentes pour opter dans un sens comme dans l'autre pour 10 paires sur 12. Autrement dit, les participants des deux groupes ont favorisé un membre de la paire dans la majorité des paires minimales. Le Tableau 2 montre les proportions de réponses correctes pour les deux groupes de participants pour chaque paire de stimuli. Les cellules ombrées indiquent les paires où les deux groupes de participants ont sélectionné le même membre de la paire dans la majorité des réponses.

Tableau 2. Proportion de réponses correctes pour les LN et les apprenants pour chaque paire de stimuli.

Paire de stimuli	Pourcentage correct pour les LN	Pourcentage correct pour les apprenants
aucun air	53%	70%
aucun nerf	40	27
aucun hectare	43	63
aucun nectar	67	27
aucun oeuf	57	57
aucun neuf	40	27
brillant acte	70	70
brillant tact	17	50
grand ami	73	50
grand tamis	40	33
grand ermite	63	63
grand termite	60	43
long hommage	20	63
long gommage	93	53
trop arqués	33	47
trop parqués	60	67
trop artisan	37	17
trop partisan	83	73
trop heureux	63	73
trop peureux	53	53
beaucoup osé	20	20
beaucoup posé	77	77
trop unis	20	20
trop punis	80	87

Dans la mesure où les différences de durée dans les cas de liaison ne semblent pas suffisamment déterminantes pour permettre l'élimination définitive des candidats dans le lexique mental, on peut présumer que d'autres facteurs ont influencé les réponses des deux groupes, facteurs qui expliqueraient les similarités dans leurs performances respectives. Les données obtenues ici

incitent à éliminer ce facteur acoustique pour envisager des facteurs relevant d'un traitement post-lexical. Les corrélations suggèrent également que les apprenants de français L2 sont sensibles à ces facteurs à l'instar des LN et qu'ils les exploitent dans la perception en L2.

3.2. Analyses Post-Hoc

Dans une série d'analyses *post hoc*, nous avons tenté d'identifier le facteur ou les facteurs post-lexicaux qui auraient pu influencer les réponses des deux groupes. Nous avons considéré d'abord les effets de la fréquence lexicale. Dans un environnement sans contexte sémantiquement contraignant, des facteurs comme la fréquence peuvent jouer un rôle important dans la reconnaissance de la parole et le choix d'un candidat lexical. Le rôle de la fréquence est postulé dans des modèles de reconnaissance de la parole comme le *Neighborhood Activation Model* (NAM) proposé par Luce & Pisoni (1998). Selon ce modèle, pour chaque candidat activé dans le lexique mental, le système de traitement effectue une série de calculs pour choisir le bon candidat. Les probabilités sont évaluées pour chaque patron phonétique qui correspond à l'entrée auditive, en prenant en compte la fréquence lexicale, le niveau d'activation du patron (c'est-à-dire le degré de correspondance phonétique), et également les niveaux d'activation et les fréquences de tous les concurrents. Tous ces facteurs étant calculés, le candidat ayant la probabilité la plus élevée en ressort gagnant. Des modèles comme celui-ci suggèrent que dans des environnements sans contexte où les indices acoustiques ne sont pas suffisamment nets pour pouvoir trancher, la fréquence devrait avoir un poids plus important.

Afin de tester l'hypothèse selon laquelle les fréquences relatives des mots dans chacune des paires minimales influent sur les réponses des participants des deux groupes, la fréquence de chaque mot-cible a été relevée à partir d'une base de données (FRANTEXT: <http://www.lib.uchicago.edu/efts/ARTFL/database/TLF/index.html>). Ces fréquences ont été calculées à partir d'un ensemble de 250 documents. Les formes au singulier ainsi qu'au pluriel de chaque mot-cible ont été incluses dans le calcul. La fréquence moyenne des mots-cibles commençant par une voyelle (les environnements de liaison) était de 1733 occurrences, tandis que la fréquence moyenne des mots-cibles commençant par une consonne était de 356 occurrences. La fréquence de chaque item cible est présentée dans le Tableau 3.

Tableau 3. Fréquence d'occurrence dans la base de données FRANTEXT pour chaque stimulus. La fréquence moyenne des deux conditions (liaison et non-liaison) est aussi présentée en bas de chaque colonne.

Stimulus commençant par une voyelle (liaison)		Stimulus commençant par une consonne (non-liaison)	
air	9084	nerf	537
hectare	52	nectar	10
oeuf	588	neuf	1858
osé	413	posé	869
acte	4592	tact	121
ami	5360	tamis	3
ermite	73	termite	32
hommage	393	gommage	0
arqués	13	parqués	20
artisan	211	partisan	232
heureux	2468	peureux	49
unis	1178	punis	166
MOYENNE	1733	MOYENNE	356

Pour chacun des deux groupes nous avons ensuite effectué des corrélations entre la fréquence de chaque item cible et le pourcentage moyen d'exactitude pour cet item en utilisant une corrélation Spearman rho. Aucune corrélation significative n'a été relevée pour les LN ($r = -.009$, NS) ni pour les apprenants ($r = .256$, NS). Une régression simple pour chacun des deux groupes n'a donné de corrélation significative ni pour les LN ($r = .141$, NS) ni pour les apprenants ($r = .205$, NS).

Ces résultats montrent que la fréquence des mots-cibles n'a pas contribué à déterminer les réponses dans aucun des deux groupes. Les items cibles les plus fréquents de chaque paire n'ont pas été systématiquement choisis par les participants. Par ailleurs, pour les items les plus fréquents, les participants n'ont pas montré non plus un taux d'exactitude supérieur.

On a ensuite envisagé le rôle de facteurs sémantiques dans l'identification des stimuli. Au début de chaque séance nous avons demandé aux participants de ne pas tenir compte de la plausibilité des phrases pendant l'expérience, mais après avoir passé le test, certains participants ont exprimé des

réticences sur l'acceptabilité de quelques unes des phrases et des questions sur l'influence de ce facteur dans leurs réponses. Par exemple, dans la paire *oeuf* /*neuf*, bien que le mot *neuf* ait un taux de fréquence trois fois plus élevé que *oeuf* (1858 et 588 occurrences, respectivement) la phrase *Il n'a aucun neuf* est beaucoup moins acceptable que *Il n'a aucun oeuf* en langage spontanée.

Afin d'examiner l'effet éventuel de la plausibilité, nous avons demandé à 20 locuteurs natifs de français qui n'avaient pas participé à l'expérience précédente de juger les 12 phrases. Ils devaient classer chaque phrase dans une échelle de un à sept, un étant une phrase '*pas du tout plausible*' et sept étant une phrase '*complètement plausible*'. Les mesures moyennes de plausibilité sont présentées dans le Tableau 4.

Tableau 4. Plausibilité moyenne de chaque phrase cible. L'échelle de plausibilité était de 1 à 7, un représentant '*pas du tout plausible*' et sept représentant '*complètement plausible*'.

Stimulus commençant par une voyelle (liaison)	Plausibilité moyenne	Stimulus commençant par une consonne (non-liaison)	Plausibilité moyenne
Il n'a aucun air	2.7	Il n'a aucun nerf	4.2
Il n'a aucun hectare	4.9	Il n'a aucun nectar	4.5
Il n'a aucun oeuf	4.6	Il n'a aucun neuf	3.8
Il a beaucoup osé	5.2	Il a beaucoup posé	5
C'est un brillant acte	2.9	C'est un brillant tact	2.3
C'est un grand ami	6.6	C'est un grand tamis	5.3
C'est un grand ermite	5.3	C'est un grand termite	4.3
C'est un long hommage	6.1	C'est un long gommage	4.5
Ils sont trop arqués	5.7	Ils sont trop parqués	4.6
Il est trop artisan	4	Il est trop partisan	5.9
Il est trop heureux	6.1	Il est trop peureux	6.4
Ils sont trop unis	5.8	Ils sont trop punis	5.6

Nous avons ensuite effectué des corrélations entre les jugements moyens de plausibilité et la proportion de réponses correctes pour chaque stimulus. L'analyse Spearman rho n'a pas montré de corrélation significative ni pour les LN ($r = .129$, NS) ni pour les apprenants ($r = .200$, NS). Aucune corrélation

significative n'a été trouvée non plus dans une régression simple pour les LN ($r = .120$, NS) ni pour les apprenants ($r = .098$, NS). Les données ci-dessus suggèrent que, tandis que plusieurs participants ont constaté qu'ils avaient été influencés par la plausibilité des phrases, celle-ci n'a pas joué un rôle systématique dans leurs décisions.

4. Discussion

Pris dans leur ensemble, ces résultats suggèrent que les différences de durée entre les deux types de consonnes ne sont pas assez nettes pour permettre aux auditeurs de désambiguïser des paires de phrases que la liaison rend ambiguës, et qu'il existe d'autres facteurs acoustiques, voire relevant du traitement post-lexical qui peuvent influencer l'identification des mots ambigus. Selon l'hypothèse de Spinelli *et al.* (2003), les auditeurs (natifs) sont sensibles aux variations acoustiques dues à la resyllabation dans les environnements de liaison et ils exploitent ces variations pour découper le signal continu en unités discrètes. Néanmoins leur étude n'a pas testé directement cette hypothèse. À cette fin, nous avons employé dans cette étude, une tâche à choix obligatoire. Nous avons constaté que ni les LN ni les apprenants avancés de français L2 n'ont montré de sensibilité systématique aux différences acoustiques. Globalement parlant, nos résultats ne vont pas nécessairement à l'encontre de ceux de Spinelli *et al.* Ces auteurs ont employé un paradigme d'amorçage sémantique qui peut entraîner des processus cognitifs différents de ceux qui sous-tendent notre tâche expérimentale.

À quoi peut-on donc attribuer ces similarités dans la distribution des réponses des deux groupes ? Dans les cas où les indices de traitement ascendants n'offrent pas assez d'informations pour résoudre des cas d'ambiguïté, les auditeurs recourent aux indices descendants. Harley (1995: 53) décrit le contexte ainsi « *all of the information not in the immediate sensory signal, including previous sensory input (prior context), as well as higher-level knowledge (e.g. world knowledge, pragmatics.)* ». L'ensemble de la recherche sur la reconnaissance de la parole a montré qu'il existe de multiples niveaux d'activation dans le processus de la compréhension du signal acoustique. Le niveau de traitement pré-lexical se base uniquement sur l'input physique, c'est-à-dire le signal acoustique lui-même. Les modèles comme TRACE et Shortlist supposent qu'une fois les phonèmes activés, ils déclenchent l'activation des mots candidats qui les contiennent. Le traitement post-lexical s'effectue dès que les mots entiers sont activés par l'entrée auditive. Tous ces mots entrent

alors en concurrence jusqu'à ce que tous les candidats sauf le gagnant soient *désactivés*. C'est au niveau post-lexical donc que le contexte entre en jeu. Ce sont les indices descendants qui jouent un rôle important dans la *désactivation* des mots candidats. Cependant, les deux facteurs post-lexicaux testés ici n'ont pas montré un effet significatif sur le taux d'exactitude de l'identification. Aucun des deux groupes n'a manifesté de sensibilité ni à la fréquence lexicale des mots-cibles ni à la plausibilité des phrases. Le facteur ou les facteurs qui ont provoqué la distribution des réponses n'ont pas été mis en évidence jusqu'ici.

En ce qui concerne l'acquisition d'une L2, dans 10 des 12 paires minimales testées dans cette expérience, le groupe expérimental a préféré la même réponse que les LN, ce qui implique que les apprenants tardifs sont sensibles à des facteurs jusqu'ici non-identifiés dans cette tâche perceptive et que les apprenants se comportent dans une certaine mesure comme les LN. Il semble également que les décisions des apprenants ne sont pas dues à des capacités liées à l'âge d'acquisition d'une deuxième langue étant donné que tous les apprenants qui ont participé à cette étude ont appris le français après la puberté. À cet égard des travaux supplémentaires seront nécessaires pour identifier les indices exploités dans la compréhension et l'identification des séquences dites *homophones* comme celles testées ici. Pour l'instant, ces résultats ne nous permettent pas de comprendre les raisons pour lesquelles nous avons constaté ces similarités de perception entre les LN du français et les apprenants du français L2. Il nous incombe donc d'examiner non seulement d'autres indices acoustiques qui pourraient entrer en jeu (par exemple la durée des autres segments, l'intonation, la transition des formants acoustiques, etc.) mais aussi les hauts niveaux de traitement.

REMERCIEMENTS

Nous tenons à exprimer notre profonde gratitude à ceux qui ont rendu possible cette étude : Elsa Spinelli à l'Université Pierre Mendès, pour les stimuli et les enregistrements ; Randy Diehl à l'Université du Texas, pour toute son aide et surtout pour ses commentaires pendant les premières étapes de l'enquête ; deux évaluateurs anonymes pour une lecture critique très attentive et des commentaires judicieux sur ce texte ; et Laura Goudet pour ses conseils stylistiques. Nous remercions aussi Edith Taïeb et Rebekah Rast d'avoir bien voulu nous permettre de participer à la rencontre *Apprendre et enseigner une langue dans un contexte multiculturel et multilingue* à l'American University of Paris en octobre 2007. Nous tenons également à remercier tous les participants francophones et anglophones qui ont généreusement donné de leur temps.

RÉFÉRENCES BIBLIOGRAPHIQUES

- BIRDSONG, D. 2006. Age and second language acquisition: A selective overview. *Language Learning* n° 56 (s1), 9-49.
- CHO, T., & P. A. KEATING. 2001. Articulatory and acoustic studies on domain- initial strengthening in Korean. *Journal of Phonetics* n° 29, 155-190.
- CONTENT, A., N. DUMAY, & U. H. FRAUENFELDER. 2000. The role of syllable structure in lexical segmentation in French: Helping listeners avoid mondegreens. In A. Cutler, J.M. McQueen & R. Zondervan (eds) *Spoken Word Access Processes*, 39-42. Nijmegen.
- CUTLER, A. 2002. Native listeners. *European Review* n° 10, 27-41.
- CUTLER, A. & D. NORRIS. 1988. The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance* n° 14, 113-121.
- DARCY, I., S. PEPERKAMP, & E. DUPOUX. 2007. Plasticity in compensation for phonological variation: the case of late second language learners. In J. Cole & J. Hualde (eds.) *Laboratory Phonology* n° 9, Mouton de Gruyter, Berlin.
- DEJEAN DE LA BATIE, B., & D. BRADLEY. 1995. Resolving word boundaries in spoken French: Native and non-native strategies. *Journal of Applied Psycholinguistics* n° 16, 59-81.
- DUMAY, N., M. H. BANEL, U. H. FRAUENFELDER & A. CONTENT. 1998. Le rôle de la syllabe : segmentation lexicale ou classification ? *Actes des XXII^e Journées d'Etude sur la Parole*, 33-36, Martigny, Suisse.
- DUPOUX, E., C. PALLIER, N. SEBASTIAN-GALLES, & J. MEHLER. 1997. A destressing 'deafness' in French? *Journal of Memory and Language* n° 36, 406-421.
- FLEGE, J. & I. MACKAY. 2004. Perceiving vowels in a second language. *Studies in Second Language Acquisition* n° 26, 1-34.
- FOUGERON, C. 2001. Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics* n° 29, 109-135.
- FOUGERON, C., & P. A. KEATING. 1997. Articulatory strengthening at edges of prosodic domain. *Journal of the Acoustical Society of America* n° 106, 3728-3740.
- FRANTEX Database
<http://www.lib.uchicago.edu/efts/ARTFL/databases/TLF/index.html>
- GASKELL, G., E. SPINELLI, & F. MEUNIER. 2002. Perception of resyllabification in French. *Memory and Cognition* n° 30, 798-810.
- GOW, D.W. & P.C. GORDON. 1995. Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human Perception and Performance* n° 21, 344-359.
- GRAMMONT, M. 1960. *Traité pratique de phonétique française*. Delagrave, Paris.
- HARLEY, T. 1995. *The Psychology of Language: From Data to Theory*. Psychology Press, East Sussex.

- LENNEBERG, E. 1967. *Biological Foundations of Language*. Wiley, New York.
- LUCE, P. A. & D. B. PISONI. 1998. Recognizing spoken words: The neighborhood activation model. *Ear & Hearing* n° 19, 1-36.
- MACKAY, I., D. MEADOR, & J. FLEGE. 2001. The identification of English consonants by native speakers of Italian. *Phonetica* n° 58, 103-125.
- MCCLELLAND, J. L., & J. L. ELMAN. 1986. Interactive processes in speech perception: The TRACE model. In D.E. Rumelhart & J.L. McClelland (eds) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 2*. MIT Press, Cambridge, MA.
- MEHLER, J., J. Y. DOMMERGUES, U. H. FRAUENFELDER, & J. SEGUI. 1981. The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, n° 20, 298-305.
- NAKATANI, L. H. & K. D. DUKES. 1977. Locus of segmental cues to word juncture. *Journal of the Acoustical Society of America* n° 62, 714-719.
- NORRIS, D. 1994. Shortlist: A connectionist model of continuous speech recognition. *Cognition*, n° 52, 189-234.
- PASSY, P. 1906. *Les sons du français* (8th ed.). Didier, Paris.
- SPINELLI, E., J. MCQUEEN & A. CUTLER. 2003. Processing resyllabified words in French. *Journal of Memory and Language* n° 48, 233-254.
- STRIDFELDT, M. 2003. French liaison and segmentation of non-words by Swedish learners of French. *PHONUM* n° 9, 169-172.
- VROOMEN, J., & B. DE GELDER. 1997. Activation of embedded words in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance* n° 23, 710-720.
- WARNER, N., A. JONGMAN, J. SERENO, & R. KEMPS. 2004. Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics* n° 32, 251-276.
- WAUQUIER-GRAVELINES, S. 1996. *Organisation phonologique et traitement de la parole continue. Contraintes prosodiques et phonologiques de l'accès lexical*. Thèse de doctorat. Université Paris 7.
- YERSIN-BESSON, C. & F. GROSJEAN. 1996. L'effet de l'enchaînement sur la reconnaissance des mots dans la parole continue. *L'Année Psychologique* n° 96, 9-30.

ANNEXE : STIMULI

1. Il n'a aucun air. / Il n'a aucun nerf.
2. Il n'a aucun hectare. / Il n'a aucun nectar.
3. Il n'a aucun œuf. / Il n'a aucun neuf.
4. C'est un brillant acte. / C'est un brillant tact.
5. C'est un grand ami. / C'est un grand tamis.
6. C'est un grand termite. / C'est un grand ermite.

7. C'est un long hommage. / C'est un long gommage.
8. Ils sont trop arqués. / Ils sont trop parqués.
9. Il est trop artisan. / Il est trop partisan.
10. Il est trop heureux. / Il est trop peureux.
11. Ils sont trop unis. / Ils sont trop punis.
12. Il a beaucoup osé. / Il a beaucoup posé

ABSTRACT

The online segmentation of the continuous speech stream requires that the listener employ strategies to identify word and syllable boundaries. In French, the phonological processes of liaison and enchaînement often render these boundaries ambiguous (consider *un air* 'a melody' and *un nerf* 'a nerve'; both are syllabified as [œ.nɛʁ]). Some research on liaison (e.g. Spinelli *et al.*, 2003; Wauquier-Gravelines, 1996) has suggested that speakers give listeners acoustic cues to word boundaries, and thus segmentation, through durational variation of pivotal consonants. Spinelli *et al.* suggest that priming of the word 'nerf' via *un nerf* versus via *un air* may be mediated by consonant length; a word-initial consonant (*un Nerf*) is longer than the corresponding consonant that surfaces in liaison (*uN air*). The current study more directly tested the perceptual capacities of speakers by employing a forced-choice identification task in which both native speakers of French and late learners of French as a second language were asked to differentiate ambiguous phonemic content. The results suggest that, though durational differences may be systematically present in the acoustic signal and may allow for the activation of vowel-initial candidates in the word recognition process, these durational differences are not robust enough to systematically guide listeners in disambiguation. Neither group performed above chance in the identification task; at the same time, the responses of the two groups were highly correlated. Factors involved in top-down processing such as frequency and plausibility were also taken into consideration, but these factors were found not to play a systematic role in the differentiation of ambiguous phonemic content.

(Keywords: L2 French, liaison, resyllabification, spoken word recognition, speech perception/ segmentation.)