

# *L'acquisition des liaisons après les clitiques préverbaux est-elle spécifique ? Apport d'une expérimentation à grande échelle*

CÉLINE DUGUA<sup>a</sup>, AURÉLIE NARDY<sup>b</sup>, LOÏC LIÉGEAIS<sup>a</sup>,  
JEAN-PIERRE CHEVROT<sup>b</sup> et DAMIEN CHABANAL<sup>c</sup>

<sup>a</sup>Université d'Orléans

<sup>b</sup>Université Grenoble Alpes

<sup>c</sup>Université Blaise Pascal, Clermont Ferrand

(Received February 2016; revised October 2016)

## RÉSUMÉ

L'acquisition des liaisons est bien documentée et a donné lieu à plusieurs théorisations, dont un modèle basé sur l'usage, essentiellement étayé sur des données expérimentales recueillies en contexte déterminant-nom. Notre but est d'examiner si ce modèle est généralisable à la liaison après les clitiques verbaux, deuxième grand contexte de liaison catégorique. Nous observons les productions de 205 enfants âgés de 3 à 6 ans recueillies dans le contexte de tâches expérimentales. L'étude permet de poser des repères sur l'acquisition de la liaison préverbale et de la comparer à la liaison prénominale.

## I. INTRODUCTION

### *1.1. Le modèle développemental de l'acquisition de la liaison basé sur l'usage*

Le modèle développemental de l'acquisition de la liaison de Chevrot, Dugua et Fayol (2009) sur lequel s'appuie cette recherche<sup>1</sup> repose sur les principes des théories basées sur l'usage (Tomasello, 2003), de l'approche exemplariste (Bybee, 2010) et des grammaires de constructions (Goldberg, 2006). Parmi ces principes, la fréquence des formes rencontrées dans l'environnement est considérée comme le principal facteur de structuration et de fonctionnement du système linguistique. D'après ce modèle, fondé sur l'acquisition de la liaison catégorique entre déterminant et nom, les enfants construiraient leur grammaire à partir de formes lexicales plus ou moins longues rencontrées dans l'input et mémorisées globalement. À partir de la mise en relation de ces premières constructions concrètes et de leur traitement analogique, émergeraient progressivement des constructions plus abstraites. À l'issue de ce processus, la connaissance du fonctionnement de la liaison prendrait la forme d'un système stratifié de constructions formant un continuum incluant des

<sup>1</sup> Nous ne revenons pas ici sur le débat qui l'oppose au modèle phonologique soutenu par Wauquier (2009). Pour une revue précise de ces deux approches, voir Chevrot et al. (2013).

séquences concrètes mémorisées, des formes mixtes impliquant à la fois des éléments phonologiques concrets et des emplacements (*slots*) représentant les premières traces d'abstraction et des formes abstraites composées de symboles grammaticaux. Ce modèle rend compte des erreurs enfantines fréquentes et de leur disparition progressive en mettant en évidence deux stades (Chevrot et al., 2009). Chacun d'eux se caractérise par des types de constructions différents qui cohabitent, interagissent et s'influencent en même temps. La spécificité de ce modèle est également de relier le processus d'acquisition de la liaison pré nominale chez l'enfant à la segmentation des déterminants et des noms ainsi qu'à la structuration du groupe nominal.

Les enfants, confrontés précocement à des séquences déterminant-nom dans leur environnement, auraient tendance à privilégier la segmentation entre les deux mots sur la base de la structure syllabique consonne-voyelle (CV) en plaçant une frontière de mot devant un élément consonantique (Content et al., 2001; Mattys et Jusczyk, 2001). En français, cette tendance permet d'isoler correctement la plupart des formes nominales. En effet, dans plus de 86% des groupes nominaux entendus par des enfants entre 2 et 3 ans, le nom commence par une consonne (Siccardi, 2015). Ainsi, la segmentation CV appliquée aux suites *les camions*, *les chats*, *les serpents* aboutit à l'extraction des formes lexicales canoniques /kamjɔ̃/, /ʃa/, /sɛʁpɑ̃/. En revanche, pour les groupes nominaux intégrant un nom à initiale vocalique précédé d'une liaison, les formes récupérées ne correspondraient pas aux formes canoniques. Par exemple, dans *les avions*, *un ours*, *petit âne*, la mise en œuvre d'une segmentation CV isole des formes à initiale consonantique telles que /zavjɔ̃/, /nuʁs/, /tan/. Du fait de la segmentation d'un même nom à initiale vocalique dans des contextes différents, l'enfant pourrait disposer de plusieurs variantes lexicales : la variante /zavjɔ̃/ récupérée dans *deux avions*, /navjɔ̃/ récupérée dans *un avion*, /tavjɔ̃/ récupérée dans *petit avion* et aussi /avjɔ̃/ récupérée dans des constructions sans liaison comme *joli avion*.

Dans les lignes qui suivent, nous présentons les deux stades du modèle d'acquisition de la liaison (Chevrot et al., 2009).

**Stade I** – Entre les âges de 2 et 4 ans environ, l'enfant disposerait dans son lexique de constructions concrètes de deux types : (a) des constructions « longues » impliquant un déterminant et un nom et mémorisées telles qu'elles ont été perçues dans l'input. Par exemple, *un avion*, *un castor*, *des ânes*, *des chiens*, etc.; (b) des formes plus courtes segmentées à partir de groupes nominaux sur la base d'un découpage syllabique de type CV comme /tan/, /zavjɔ̃/, etc. À partir de la mise en relation de ces deux types de formes, des régularités émergeraient autour du déterminant qui constitue l'élément fréquent et stable des séquences déterminant-nom. Ainsi, à partir de séquences mémorisées telles que *un avion*, *un castor*, *un doudou*, émergerait un schéma général (*un* + X) composé de la forme concrète du déterminant et d'un slot (symbolisé par « X »). Ce slot, dans lequel peut s'insérer tout type d'unité en mémoire, constitue une première trace d'abstraction. Ce schéma est général car il ne fournit aucune restriction sur la forme qui peut intégrer le slot. Il permet d'expliquer les erreurs par remplacement telles /œzɑʁbɛʁ/ (*un /z/ arbre*), /œtuʁs/ (*un /t/ ours*) qui résultent de l'insertion d'une forme segmentée /zɑʁbɛʁ/

ou /tuʁs/ dans un schéma *un* + X et les erreurs par omission (/œ̃aʁbʁ/ *un arbre*, /œ̃uʁs/ *un ours*) correspondant à l'insertion d'une forme à initiale vocalique après le déterminant.

**Stade 2** – Ce stade se situe dans la continuité du précédent (entre 4 et 6 ans environ). L'accumulation de l'input reçu et mémorisé permettrait à l'enfant d'élaborer des liens entre un déterminant donné et des formes lexicales particulières. À partir de l'accumulation de séquences telles que *un n-ours*, *un n-arbre*, *un n-ordinateur* etc., il va peu à peu apprendre que le déterminant *un* est suivi par des variantes commençant par /n/. À partir de ces associations, émergerait un autre type de schéma de type *un* + nX, où nX correspond soit à une variante en /n/ initial, soit à une forme lexicale généralisée par le schéma par analogie aux variantes mémorisées. Ce schéma permet de réaliser correctement l'ensemble des liaisons en contexte "déterminant + nom". En revanche, sa généralisation à des groupes nominaux sans liaison entraîne des erreurs, observées à la fois en situation naturelle et dans des tâches d'élicitation, dans lesquelles les enfants surgénéralisent le schéma spécifié à des noms à initiale consonantique, comme par exemple /œ̃nɛbʁ/ (*un nèbre*) pour *un zèbre*, /døzɔ̃bʁil/ (*deux zombri*) pour *deux nombrils* (Chevrot et al., 2009).

Ce modèle développemental est étayé par un nombre important de données expérimentales – tâches de dénomination d'images (Dugua, 2006; Nardy, 2008), tâches de jugements d'acceptabilité (Nardy, 2008), tâches de chronométrie mentale (Dugua et al., 2009; Siccardi, 2015) – souvent obtenues à partir d'études transversales menées auprès de larges échantillons d'enfants ainsi que par un suivi longitudinal de 20 enfants (Dugua et Chevrot, 2015) et des études de corpus (Liégeois, 2014). Toutefois, seule la liaison prénominale a été étudiée. Or, d'autres contextes donnent lieu à des réalisations catégoriques dans la langue adulte : les séquences impliquant un verbe et un pronom clitique (*nous avons*, *vient-on ?*), les séquences clitique-clitique (*nous en voulons*) et certaines expressions figées (*tout à coup*) (Durand et Lyche, 2008).

Dans cette recherche, nous verrons si ce modèle est généralisable au contexte préverbal (séquences clitique-verbe), plus fréquent dans les interactions adultes-enfants que le contexte prénominal (Chevrot et al., 2007; Liégeois, 2014).

### *1.2. Quelques résultats préalables sur la liaison préverbale*

A partir d'un recueil de 898 erreurs de liaisons tous contextes confondus chez une fillette âgée de 2;1 à 6;4 lors d'interactions familiales (Chevrot et Fayol, 2000; Dugua, 2006), nous avons pu mettre en évidence la rareté des erreurs de liaison en contexte préverbal (49/898 erreurs, soit 5,5% du corpus). Parmi celles-ci, aucune n'a été relevée après un clitique sujet et seulement 6 ont été observées entre un clitique objet et un verbe. On relève également deux principaux types d'erreurs par adjonction d'une consonne de liaison : celles qui consistent en l'ajout de la consonne /z/ dans le contexte GN pluriel + *ont* (*les garçons /z/ ont*); et celles qui

concernent l'ajout de la consonne /z/ dans le contexte *qui* + V (*qui* /z/ *attendaient*). Ces erreurs par adjonction d'une consonne de liaison laissent penser que l'enfant a segmenté une variante verbale à initiale consonantique (majoritairement en /z/), notamment dans le premier cas où l'erreur aboutit systématiquement à la séquence /z/ + *ont*. La sous-représentation des erreurs en contexte préverbal dans le corpus de Sophie semble indiquer que ce contexte facilite l'acquisition des liaisons en comparaison au contexte prénominal.

À partir de l'analyse de corpus d'enfants âgés de 3 à 12 ans en situation d'interaction avec des adultes, Chevrot et al. (2007) relèvent que les taux de liaisons préverbaux justes atteignent 91% dès l'âge de 3–4 ans et 100% à 5–6 ans alors qu'il faut attendre respectivement les âges de 5–6 ans et 10–11 ans pour que les taux de liaisons prénominales justes atteignent ces mêmes valeurs. Les raisons de ce décalage développemental peuvent être trouvées dans l'analyse distributionnelle des cadres lexicaux (Mot1–Mot2) entourant les liaisons proposée par Laks et al. (2014) à partir du corpus *Phonologie du Français Contemporain* (Durand et al., 2005). Parmi les dix cadres les plus fréquents entourant les liaisons réalisées, les six premiers sont composés d'un clitique et d'un verbe (*on a*, *on est*, *ils ont*, *en a*, *on avait*, *on était*) et couvrent à eux seuls 21% des 16 805 liaisons réalisées dans le corpus. Aucune séquence déterminant-nom n'apparaît parmi les cadres les plus fréquents relevés dans les travaux cités. Il est donc probable que les enfants rencontrent les liaisons préverbaux dans des séquences clitique-verbe fréquentes qu'ils mémorisent globalement avec la consonne de liaison. Selon le modèle basé sur l'usage, cette particularité distributionnelle est un puissant facilitateur de l'acquisition.

Ces quelques études de corpus suggèrent que la liaison préverbale est acquise plus tôt que la liaison prénominale. Le but de notre recherche est de documenter ce décalage à partir d'une étude comparative des productions enfantines de liaisons en contexte prénominal et préverbal. Pour la première fois, nous pourrions comparer l'usage des liaisons dans chacun des deux contextes sur un échantillon important d'enfants.

## 2. MÉTHODOLOGIE

### 2.1. Sujets

Deux cent cinq enfants (109 filles et 96 garçons), âgés entre 3;2 et 6;0 ont participé à deux tâches expérimentales<sup>2</sup> de dénomination d'images visant la production de liaisons. Ils ont été répartis en trois tranches d'âge comparables à celles utilisées dans les précédentes études sur la liaison prénominale (Chevrot et al., 2009; Chevrot et al., 2011) (cf. Tableau 1).

<sup>2</sup> Nous remercions l'ensemble des enfants ainsi que les étudiants de Licence 3 de Sciences du Langage d'Orléans et de Grenoble pour leur participation au recueil des données.

Tableau 1: Description de l'échantillon

	Effectif	Intervalle d'âges	Âge moyen
TA-1	52	3;2 – 4;1	3;8
TA-2	75	4;2 – 5;1	4;8
TA-3	78	5;2 – 6;0	5;7

## 2.2. Tâche de dénomination d'images en contexte verbal

Le protocole utilisé pour faire produire des liaisons préverbaux se divise en trois parties: (1) production de séquences “clitique sujet + verbe”, (2) production de séquences “clitique objet + verbe” et (3) production de séquences “groupe nominal pluriel + verbe”. Pour chaque contexte, la procédure était la même : l'expérimentateur présentait une première image à l'enfant en la lui décrivant (amorce), puis on présentait une seconde image (cible), liée à la première, que l'enfant devait décrire à son tour. Une aide pouvait être apportée à l'enfant en lui posant des questions. Par exemple, lorsque l'image représentait des garçons en train de lire, l'expérimentateur la décrivait en produisant *Ici les garçons lisent*, puis montrait l'image cible où les mêmes garçons étaient en train d'écrire et produisait l'énoncé amorce *Et là ? que font les garçons ?*. La réponse attendue était alors *Ils écrivent*.

En tout, 30 productions par sujet étaient attendues (cf. Tableau 2). Les verbes à produire ont été choisis à partir des erreurs relevées dans le corpus des erreurs de Sophie. Afin d'éviter un effet d'amorçage lexical, les images étaient présentées aléatoirement au sein de chacun des trois contextes.

Plusieurs travaux attestent que les contextes “clitique sujet + verbe” et “clitique objet + verbe” appartiennent aux liaisons catégoriques (les classements classiques (Delattre, 1947; Encrevé, 1983) ainsi que les travaux fondés sur corpus de De Jong (1994) et Durand et Lyche (2008)). En revanche, le contexte “GN pluriel + verbe”, considéré par Delattre comme facultatif, a fait l'objet d'observations plus fines dans le corpus PFC.<sup>3</sup> Relevant qu'aucun des 100 locuteurs de leur échantillon n'a réalisé la liaison dans la tâche de lecture considérée comme étant la tâche la plus formelle du protocole (l'énoncé comportant ce contexte est : *quelques fanatiques[z]auraient même entamé*), Durand et Lyche (2008) excluent ce contexte des liaisons variables. Suivant ces résultats, nous n'attendons pas la réalisation d'une consonne de liaison dans ce contexte, réalisation qui semble très rare dans la parole adulte quotidienne, et si elle devait apparaître, nous traiterions cette production comme une erreur par adjonction.

<sup>3</sup> *Phonologie du Français Contemporain* : <http://www.projet-pfc.net/>.

Tableau 2: Liste des énoncés amorces et des énoncés attendus dans la tâche de production en contexte préverbal

Contextes	Énoncés amorces	Énoncés attendus
(1) Contexte “clitique sujet + verbe”	les garçons jouent au foot les garçons lisent les garçons ferment les filles s’en vont les filles sortent la maman et sa fille font les courses on regarde on coupe on recule	ils arrosent ils écrivent ils ouvrent elles arrivent elles entrent elles achètent on écoute on épluche on avance
(2) Contexte “clitique objet + verbe”	les garçons plantent les fleurs les garçons ferment les portes on coupe les pommes la fille a 5 pommes le garçon a 3 ballons le monsieur a 4 dés il n’y a pas beaucoup de planches il y a peu d’élèves il y a peu de pommes	ils les arrosent ils les ouvrent on les épluche le garçon en a trois la fillette en a deux la dame en a un il y en a beaucoup/plus/plein il y en a beaucoup/plus/plein il y en a beaucoup/plus/plein
(3) Contexte “GN pluriel + verbe”	les filles font du cheval les mamans discutent les filles s’en vont de l’école les garçons jouent dans la piscine les garçons lisent les garçons jardinent la fille a 5 pommes la fille a 4 dés la fille a 3 toupies le garçon a 2 robots le garçon a 2 chiens le garçon a 2 voitures	les garçons écoutent les enfants ouvrent les garçons arrivent les filles attrapent les filles écrivent les filles arrosent les garçons ont deux ballons les garçons ont trois camions les garçons ont deux bateaux les filles ont trois chiens les filles ont trois chats les filles ont trois bateaux

### 2.3. Tâche de dénomination d’images en contexte nominal

Le matériau linguistique utilisé comprend 6 Mots<sub>2</sub> (*arbre, avion, éléphant, escargot, ordinateur, ours*), et 3 Mots<sub>1</sub> (*un, deux, des*). Ces mots étaient représentés par des images en simple, double et multiples exemplaires. On demandait ici à l’enfant de dire ce qu’il voyait sur l’image qui lui était présentée (la présentation des images a été rendue aléatoire pour chaque enfant). Dix-huit productions de liaisons étaient attendues.

### 3. RÉSULTATS

#### 3.1. *Évolution des productions de liaison dans les groupes verbaux*

Les données que nous présentons ici portent sur l'ensemble des productions enfantines correspondant à un contexte préverbal, même si les éléments lexicaux réalisés n'étaient pas ceux ciblés par la tâche. Par exemple, si l'enfant a dit *ils arrosent*, alors que nous attendions *les garçons arrosent*, nous avons conservé la donnée en requalifiant son contexte de production (de GN à clitique sujet). Globalement, 22,8% (1392/6100) des productions enfantines correspondent à la séquence Mot1–Mot2 attendue, mais 42,6% (2600/6100) présentent un contexte de liaison préverbale<sup>4</sup>. Notre échantillon de données s'appuie donc sur 2 600 contextes préverbaux qui se répartissent de façon non homogène dans les trois tranches d'âge : 407 contextes dans la première, 937 dans la deuxième et 1256 dans la dernière.

##### 3.1.1. *Contexte "clitique sujet/clitique objet + verbe"*

Dans un premier temps, et à des fins de comparaison entre les performances dans les deux tâches (cf. page 10), nous analyserons ensemble les contextes "clitique sujet + verbe" et "clitique objet + verbe". En effet, comme dans le contexte déterminant-nom, on attend ici la réalisation d'une liaison, ce qui n'est pas le cas pour le contexte GN pluriel + verbe. Nous avons donc calculé, pour chaque enfant, un pourcentage indiquant la proportion des 3 principaux types de productions sur l'ensemble des contextes de liaison produits. Un seuil minimal de 5 occurrences de contextes de liaison produits a été appliqué pour le calcul du pourcentage.

Le taux de liaisons justes moyen est de 97,4% et se répartit selon les tranches d'âges de la façon suivante (cf. [Figure 1](#)).

Comme l'illustre la [figure 1](#), aucune évolution avec l'âge n'est relevée sur les taux de liaisons justes (Kruskall-Wallis :  $K = 1,991$ ;  $p = 0,3696$ ) ni sur les omissions ( $K = 1,694$ ;  $p = 0,4288$ ). Quant aux erreurs par remplacement, elles sont quasi inexistantes : seuls 4 enfants de la tranche d'âge 2 en ont produit.

##### 3.1.2. *Contexte GN pluriel + verbe*

Nous abordons le contexte "GN pluriel + verbe" séparément pour deux raisons. La première est que ce contexte est particulier puisque la liaison n'y est pas réalisée dans la langue adulte. La seconde est que nous avons recueilli peu d'occurrences pour ce contexte (504). De ce fait, nous procédons ici à une analyse globale de la répartition des différents types de productions sur l'ensemble des occurrences recueillies.

Sur les 504 occurrences recueillies dans ce contexte : 87,5% correspondent à des productions justes et 12,5% à des erreurs par ajout d'une consonne de liaison. Parmi les erreurs ( $n = 63$ ), on trouve une proportion massive d'ajouts de /z/ (93,7%,  $n = 59$ ); les erreurs restantes correspondent à 2 ajouts de /l/, 1 ajout de /n/ et 1

<sup>4</sup> Plus précisément, dans près de 63% des cas où nous attendions la production d'un GN pluriel, l'enfant produit un clitique sujet, soit par une dislocation à gauche (*les enfants ils entrent*), soit par remplacement du GN par un clitique.

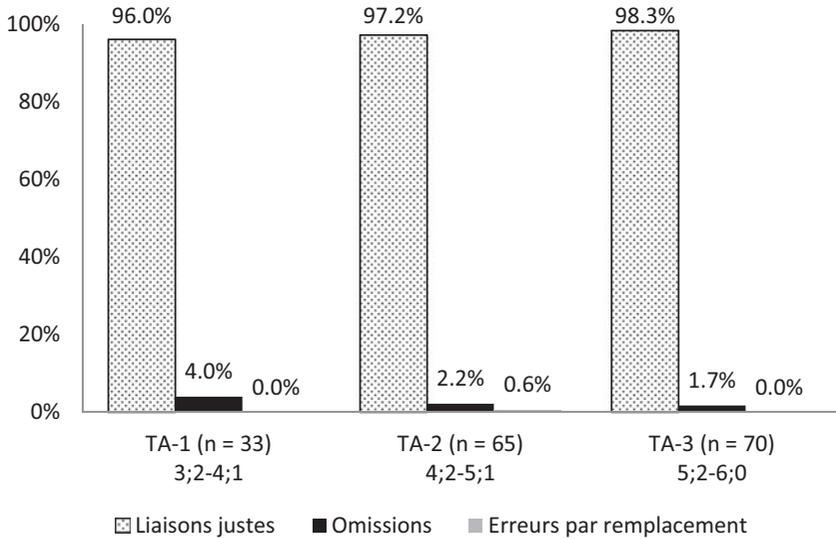


Figure 1: *Types de productions en contexte "clitique (sujet et objet) + verbe" selon les tranches d'âges*

ajout de /t/. Notons par ailleurs que dans tous les cas d'ajouts de /z/, le GN sujet est au pluriel (*les filles, les gars, etc.*). En outre, parmi les 8 Mots2 les plus souvent produits<sup>5</sup> (au moins 4 fois), celui pour lequel nous constatons le plus d'erreurs est *ont* (19,5%<sup>6</sup> des productions sont des erreurs en /z/). En observant les résultats par tranche d'âge, on relève que, dans la première, seulement 2 Mots2 (*a* et *ont*) sont produits plus de 3 fois, précédés ou non de liaisons erronées. Parmi ces Mots2, seul *ont* suscite des erreurs. Dans la deuxième tranche d'âge, 8 Mots2 apparaissent plus de 3 fois et parmi eux, seuls 2 provoquent des erreurs (*ont* et *écoutent*). Enfin dans la dernière tranche d'âge, 8 Mots2 apparaissent plus de 3 fois et 4 entraînent des erreurs (*arrosent, attrapent, écrivent, ont*).

On constate donc que les erreurs impliquant la forme *zont* sont présentes dès la première tranche d'âge et se maintiennent dans les deux suivantes<sup>7</sup>. Par ailleurs, on note que les formes verbales concernées par des erreurs se diversifient au fil des tranches d'âge.

<sup>5</sup> Les 8 mots2 sont les suivants : *a, arrosent, attrapent, écoutent, écri-, écrivent, ont, ouvrent*.

<sup>6</sup> Parmi les 256 contextes GN + *ont* produits, 50 sont des ajouts de /z/ (19,5%), 204 correspondent à des productions sans liaison (productions justes), et on relève un ajout de /l/ et un ajout de /t/.

<sup>7</sup> Notons au passage que des erreurs de liaison avec la forme *zont* sont aussi attestées chez l'adulte. Dans le corpus ESLO, une jeune fille (RN166, tranche d'âge 15-25 ans) lors d'un repas (ESLO2\_REPAS\_1260) : *moi c'est les premiers que je connais qui z-ont jamais vu un seul James Bond quoi*. En conférence de presse filmée, le footballeur F. Ribéry : *γ a eu des choses qui z-ont été dites*. ([http://www.wat.tv/video/ribery-aux-journalistes-vous-3j9c9\\_2exyh\\_.html](http://www.wat.tv/video/ribery-aux-journalistes-vous-3j9c9_2exyh_.html))

3.2. Évolution des productions de liaison dans les groupes nominaux

Les résultats présentés portent sur l'ensemble des réponses qui correspondent à un contexte de liaison même s'il ne s'agit pas du contexte attendu par le protocole.

Pour chaque type de production, nous avons calculé un pourcentage individuel reflétant sa proportion sur l'ensemble des contextes de liaison produits. Comme précédemment, ce pourcentage a été calculé à partir d'un seuil minimal de 5 occurrences de contextes de liaison produits.

La figure 2 fait apparaître leur répartition et leur évolution au fil des trois tranches d'âges.

Toutes tranches d'âges confondues, nous constatons que les liaisons justes sont les productions majoritaires (85%), suivies des erreurs (8,4% d'omissions, 6,6% d'erreurs par remplacement). Nous relevons une augmentation significative des liaisons justes avec l'âge (Kruskal-Wallis :  $K = 23,055$ ;  $p < 0,0001$ ). Toutefois, l'observation de cet effet entre chaque tranche d'âge révèle que seule l'augmentation de 75,5% en TA-1 à 81,5% en TA-2 est significative (Mann-Whitney :  $U = 1267,000$ ;  $p = 0,0015$ ). Les erreurs par remplacement diminuent régulièrement entre les trois tranches d'âge pour ne représenter plus que 1,7% des productions dans la dernière ( $K = 36,655$ ;  $p < 0,0001$ ). Cet effet de l'âge se retrouve entre TA-1 (14,1%) et TA-2 ( $U = 1310,000$ ;  $p = 0,0011$ ) et entre TA-2 (6,5%) et TA-3 ( $U = 2280,500$ ;  $p = 0,0011$ ). Quant aux omissions, leur proportion reste stable entre 3 et 6 ans ( $K = 0,886$ ;  $p = 0,6420$ ).

Ces évolutions – augmentation puis stabilisation des justes, diminution des remplacements, stagnation des omissions – sont totalement cohérentes avec les

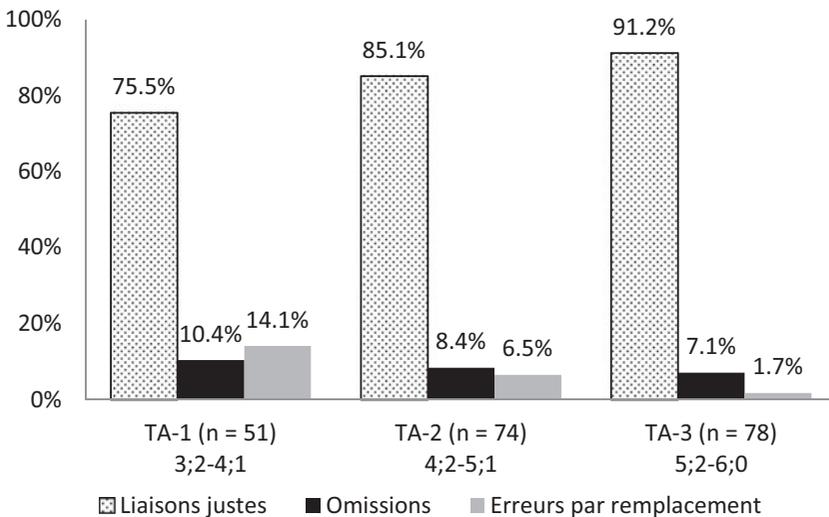


Figure 2: Types de productions en contexte "déterminant + nom" selon les tranches d'âges

études précédentes réalisées à partir de protocoles similaires<sup>8</sup> et de tranches d'âges comparables.

### 3.3. Comparaison des productions de liaisons après clitique préverbal et après déterminant

Dans cette section, nous mettons en perspective les résultats obtenus en contexte "déterminant + nom" à ceux obtenus dans les groupes verbaux là où une liaison réalisée est attendue, c'est-à-dire dans le contexte "clitique (sujet et objet) + verbe".

La figure 3 montre que les performances en liaisons justes sont plus élevées dans les groupes verbaux que dans les groupes nominaux. Toutes tranches d'âges confondues, les performances dans les groupes verbaux sont significativement supérieures à celles relevées dans les groupes nominaux (Wilcoxon :  $Z = -6,166$ ;  $p < 0,0001$ ). Cette différence se manifeste également dans chaque tranche d'âge (TA-1 :  $Z = -3,095$ ;  $p = 0,0020$ ; TA-2 :  $Z = -3,735$ ;  $p = 0,0002$ ; TA-3 :  $Z = -3,895$ ;  $p < 0,0001$ ).

Le même patron de résultat est observé pour les erreurs par remplacement dont la proportion est moins importante en contexte verbal qu'en contexte nominal (toutes tranches d'âges confondues : 0,2% contre 6,6%;  $Z = -4,568$ ;  $p < 0,0001$ ),

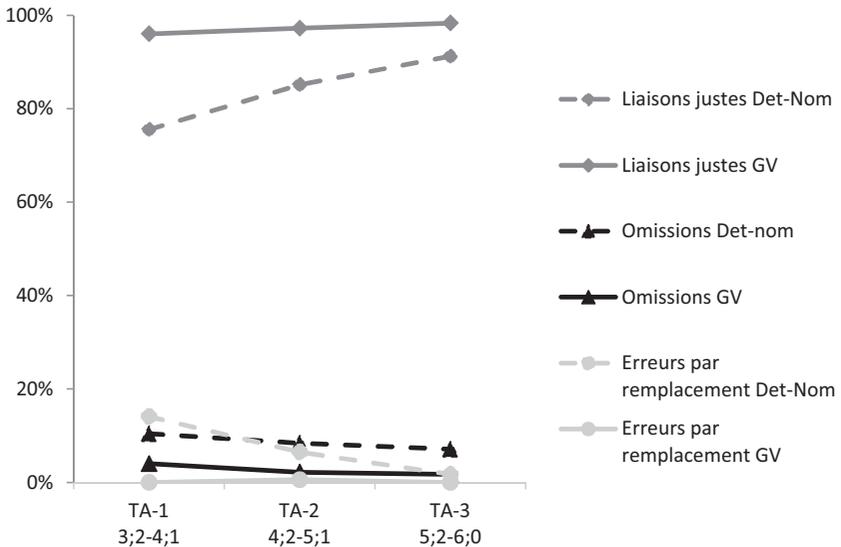


Figure 3: Comparaison des types de productions en contexte nominal et verbal selon les tranches d'âges

<sup>8</sup> Dans les études précédentes (Dugua, 2006; Nardy, 2008; Chevrot et al., 2011), seuls deux Mots1 étaient impliqués (*un* et *deux*), contrairement à ici où on attend aussi le Mot1 *des*.

dans chaque tranche d'âge (TA-1 :  $Z = -3,061$ ;  $p = 0,0022$ ; TA-2 :  $Z = -2,743$ ;  $p = 0,0061$ ; TA-3 :  $Z = -2,023$ ;  $p = 0,0431$ ).

Enfin, bien que les omissions soient également moins présentes en contexte verbal qu'en contexte nominal (toutes tranches d'âges confondues : 2,4% contre 8,4%;  $Z = -4,665$ ;  $p < 0,0001$ ), elles ne se distinguent significativement que dans les deux dernières tranches d'âges (TA-2 :  $Z = -2,713$ ;  $p = 0,0067$ ; TA-3 :  $Z = -3,318$ ;  $p = 0,0009$ ).

#### 4. DISCUSSION/CONCLUSION

La mise en perspective des productions de liaisons préverbaux après clitique et des liaisons prénominales montre que les premières sont acquises plus tôt que les secondes. Cette tendance est également observée dans des données recueillies en situation naturelle d'interaction chez des enfants de 2–5 ans (corpus ALIPE,<sup>9</sup> Liégeois et al., 2014; Liégeois, soumis). En outre, le profil des erreurs suscitées par ces deux contextes est différent. Comment expliquer ces différences entre l'acquisition des liaisons en contexte déterminant-nom et clitique-verbe qui sont deux cas de liaisons catégoriques ?

Globalement, les séquences clitique-verbe impliquant une liaison sont plus fréquentes dans l'input que les séquences déterminant-nom, certaines d'entre elles, telles que *on a*, *on est*, *ils ont*, *en a* comptant même parmi les cadres lexicaux les plus fréquents entourant les liaisons (Laks et al., 2014; Liégeois, 2014). Par ailleurs, dans le discours adressé à l'enfant, les pronoms clitiques liaisonnants sont employés avec un nombre très limité de Mots2 (Liégeois, 2014). Cette stabilité est propice à la mémorisation de constructions figées formant précocement la base pour la production de liaisons justes. Ces caractéristiques distributionnelles pourraient donc expliquer l'avance développementale des liaisons entre clitique et verbe. Inversement, la composition interne des suites déterminant-nom est beaucoup plus variée. Cette variabilité favoriserait la segmentation précoce des unités qui les composent et la formation plus rapide de constructions abstraites à partir des constructions figées préalablement mémorisées (étape 1 du scénario de l'acquisition de la liaison nominale). En raison de ce moindre figement, les unités nominales s'autonomiseraient plus rapidement des séquences déterminant-nom que les unités verbales s'autonomisent des séquences clitique-verbe.

Même si les séquences clitique-verbe sont davantage figées, elles sont toutefois segmentées par l'enfant qui dispose de formes verbales à initiale vocalique (*arrive*), mais aussi parfois de formes verbales commençant par une consonne (*zarrivent*, *zont*). Toutefois, ces deux types de formes, qu'on retrouve également pour les noms, n'ont pas la même genèse, ni le même usage ou la même fréquence dans les deux contextes.

<sup>9</sup> Corpus « Acquisition de la Liaison et Interactions Parents-Enfant »

La présence d'erreurs par adjonction en contexte GN + verbe (*les garçons /z/ ont*) suggère que l'enfant segmente des formes |Consonne de liaison-Verbe| formellement proches des variantes *zours*, *navion* segmentées en contexte nominal. Toutefois, en contexte clitique-verbe, les erreurs par remplacement (telles *on zarrive*) sont inexistantes, alors qu'elles sont très nombreuses jusqu'à 4 ans entre déterminant et nom (*un zours*, *les navions*). Plusieurs raisons peuvent être avancées pour expliquer la moindre prégnance des variantes de type |Consonne de liaison-Verbe| comparativement aux variantes |Consonne de liaison-Nom|. Tout d'abord, les verbes sont rencontrés à la fois après des clitiques (donc parfois avec liaison) et après des GN (donc jamais avec liaison). Dans ce dernier cas, la forme à initiale vocalique est disponible dans l'input (*les enfants arrivent . . .*), ce qui rend sa segmentation plus facile que celle d'une forme nominale telle que *ours*, qui apparaît rarement sans être enchaînée sur une consonne résultant d'une liaison ou de l'élision du déterminant. Les enfants segmenteraient donc précocement des formes verbales à initiale vocalique qu'on retrouve dans les erreurs par omission entre clitique et verbe et dans les séquences GN + verbe. A ces raisons distributionnelles s'ajoute un facteur d'ordre morphologique qui concerne la variation de la terminaison du verbe après une même consonne de liaison. Pour un verbe donné (*arriver*), la consonne de liaison /z/ suivant les clitiques *nous*, *vous*, *ils*, *elles* précède des formes verbales différentes : *arrivons*, *arrivez*, *arrivent*, *arriveront*, etc. La forme phonologique des séquences |consonne de liaison-Verbe| est donc moins stable que celle des séquences |consonne de liaison-Nom|. Il en résulte que leur mémorisation comme variante est moins directe.

Une autre tendance importante de nos résultats, également observée par Liégeois (soumis) à partir de l'analyse de productions enfantines spontanées, est que l'ajout de consonne de liaison entre un GN et un verbe concerne essentiellement la consonne /z/, toujours dans un contexte pluriel. Il est donc probable que s'élabore, en production, un lien entre la pluralité et la consonne /z/, lien qui a d'ailleurs été montré en compréhension par des tâches de pointage et de regard préférentiel dès l'âge de 30 mois (Legendre et al., 2010). Bien que cette valeur sémantique de la liaison soit attestée en contexte déterminant-nom chez l'enfant (Nicoladis et Paradis, 2010), son apparition serait moins précoce puisque les jeunes locuteurs produisent des séquences de type *un zours* pour désigner des objets uniques.

Les processus d'acquisition de la liaison entre clitique et verbe et entre déterminant et nom semblent donc emprunter des voies distinctes. Nous disposons maintenant des premières bases empiriques (à la fois expérimentales et écologiques) pour l'étude de l'acquisition de la liaison préverbale que nous devons compléter afin de proposer un scénario développemental et répondre à plusieurs questions qui restent en suspens. Il s'agira notamment d'examiner les productions d'enfants plus jeunes (entre 2 et 3 ans) et de déterminer plus précisément les processus de segmentation en contexte préverbal pour préciser le statut lexical de la liaison qui suit les clitiques. En effet, s'il semble que la liaison /z/ puisse être encodée précocement à l'initiale d'un petit nombre de formes verbales très fréquentes (par exemple, *zont*), l'absence d'erreurs par remplacement après clitiques suggère que ce

cas n'est pas général et que la consonne de liaison pourrait être épenthétique ou s'attacher précocement aux formes lexicales plurielles des clitiques plutôt qu'au mot<sup>2</sup>. Enfin, il s'agira d'analyser plus précisément l'acquisition de la liaison préverbale en séparant le contexte clitique sujet + verbe du contexte clitique objet + verbe afin d'observer si le premier, du fait de sa plus grande fréquence dans l'environnement est acquis plus tôt que le second. De tels résultats conforteraient le rôle prépondérant de la fréquence dans la construction du système linguistique et apporteraient un argument supplémentaire en faveur des modèles basés sur l'usage.

*Adresse pour correspondance:*

*e-mail:* [celine.dugua@univ-orleans.fr](mailto:celine.dugua@univ-orleans.fr)

#### RÉFÉRENCES

- Bybee, J. (2010). *Language, Usage and Cognition*. Cambridge: Cambridge University Press.
- Chevrot, J.-P., Chabanal, D. et Dugua, C. (2007). Pour un modèle de l'acquisition des liaisons basé sur l'usage: trois études de cas. *Journal of French Language Studies*, 17(1): 103–128.
- Chevrot, J.-P., Dugua, C. et Fayol, M. (2009). Liaison, word segmentation and construction in French: a usage-based account. *Journal of Child Language*, 36(3): 557–596.
- Chevrot, J.-P., Dugua, C., Harnois-Delpiano, M., Siccardi, A. et Spinelli, E. (2013). Liaison acquisition: debates, critical issues, future research. *Language Sciences*, 39: 83–94.
- Chevrot, J.-P. et Fayol, M. (2000). L'acquisition de la liaison: enjeux théoriques, premiers résultats, perspectives. *Lidil*, 22: 11–30.
- Chevrot, J.-P., Nardy, A. et Barbu, S. (2011). Developmental dynamics of SES-related differences in children's production of obligatory and variable phonological alternations. *Language Sciences*, 33: 180–191.
- Content, A., Kearns, R. K. et Frauenfelder, U. H. (2001). Boundaries versus onsets in syllabic segmentation. *Journal of Memory and Language*, 45: 177–199.
- De Jong, D. (1994). La sociophonologie de la liaison orléanaise. In: C. Lyche (dir.), *French Generative Phonology: Retrospectives and Perspectives*. Salford: ESRI, pp. 95–129.
- Delattre, P. (1947). La liaison en français, tendances et classification. *The French Review*, 21(2): 148–157.
- Dugua, C. (2006). Liaison, segmentation lexicale et schémas syntaxiques entre 2 et 6 ans: un modèle développemental basé sur l'usage. Thèse de doctorat, Université Stendhal, Grenoble.
- Dugua, C. et Chevrot, J.-P. (2015). Acquisition des liaisons prénominales: complémentarité des approches transversale et longitudinale. *Lidil*, 51: 35–63.
- Dugua, C., Spinelli, E., Chevrot, J.-P. et Fayol, M. (2009). Usage-based account of the acquisition of liaison: evidence from sensitivity to plural/singular orientation of nouns. *Journal of Experimental Child Psychology*, 102: 342–350.
- Durand, J., Laks, B. et Lyche, C. (2005). PFC: un corpus numérisé pour la phonologie du français. In: G. Williams (dir.), *Actes des Deuxièmes Journées de la Linguistique du Corpus*. Rennes: Presses Universitaires de Rennes, pp. 205–217.

- Durand, J. et Lyché, C. (2008). French liaison in the light of corpus data. *Journal of French Language Studies*, 18(1): 33–66.
- Encrevé, P. (1983). La liaison sans enchainement. *Actes de la Recherche en Sciences Sociales*, 46: 39–46.
- Goldberg, A. E. (2006). *Constructions at Work. The Nature of Generalization in Language*. New York: Oxford University Press.
- Laks, B., Calderone, B. et Celata, C. (2014). French liaison and the lexical repository. In: C. Celata et S. Calamai (dir.), *Advances in Sociophonetics*. Amsterdam / Philadelphia: John Benjamins, pp. 29–54.
- Legendre, G., Barrière, I., Goyet, L. et Nazzi, T. (2010). Comprehension of infrequent Subject–Verb agreement forms: evidence from French-learning children. *Child Development*, 81(6): 1859–1875.
- Liégeois, L. (2013). De l'analyse au partage des données, quel(s) format(s) choisir? L'exemple d'un corpus d'interactions parents–enfant. In: M. Damiani, K. Dolar, C. Florez-Pulido, R. Loth, J. Magnier et A. Pegaz (dir.), *Traitement de corpus* (Actes de Coldoc 2012). Paris: Modyco, pp. 128–142.
- Liégeois, L. (2014). Usage des variables phonologiques dans un corpus d'interactions naturelles parents–enfant: impact du bain linguistique et dispositifs cognitifs d'apprentissage. Thèse de doctorat, Université Blaise Pascal, Clermont-Ferrand.
- Liégeois, L. (soumis). Acquisition des liaisons catégoriques nominales et verbales: de la lexicalisation à l'abstraction des constructions. Article soumis en vue de la publication d'un numéro de revue coordonné par D. Legallois et A. Patard.
- Liégeois, L., Chanier, T. et Chabanal, D. (2014). Corpus globaux ALIPE: Interactions parents–enfant annotées pour l'étude de la liaison. Nancy: Ortolang. [<http://hdl.handle.net/11041/alipe-000853>]
- Mattys, S. L. et Jusczyk, P. W. (2001). Do infants segment words or recurring continuous patterns? *Journal of Experimental Psychology: Human Perception and Performance*, 27: 644–655.
- Nardy, A. (2008). Acquisition des variables sociolinguistiques entre 2 et 6 ans: facteurs sociologiques et influences des interactions au sein du réseau social. Thèse de doctorat, Université Stendhal, Grenoble.
- Nicoladis, E. et Paradis, J. (2011). Learning to liaise and elide comme il faut: evidence from bilingual children. *Journal of Child Language*, 38(4), 701–730.
- Siccardi, A. (2015). Unités et séquences dans le lexique adulte et enfantin. Thèse de doctorat, Université Grenoble Alpes.
- Tomasello, M. (2003). *Constructing a Language: A Usage-based Theory of Language Acquisition*. Cambridge, MA : Harvard University Press.
- Wauquier, S. (2009). Acquisition de la liaison en L1 et L2: stratégies phonologiques ou lexicales? *Aile*, 2: 93–130.