



Production de liaisons dans l'input parental

Damien Chabanal, Loïc Liégeois

► To cite this version:

Damien Chabanal, Loïc Liégeois. Production de liaisons dans l'input parental. Christiane Soum-Favaro, Annelise Coquillon et Jean-Pierre Chevrot. La liaison : approches contemporaines, Peter Lang, pp.263-282, 2014, Sciences pour la commun!cation. <hal-00988817>

HAL Id: hal-00988817

<https://hal.archives-ouvertes.fr/hal-00988817>

Submitted on 9 May 2014

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.

Manuscrit auteurs

Chapitre publié dans :

Soum-Favaro, C., Coquillon, A. et Chevrot, J.-P. (2014). *La liaison : approches contemporaines*. Berne : Peter Lang

Référence du chapitre :

Chabanal, D. et Liégeois, L. (2014). Production de liaisons dans l'input parental. In C. Soum-Favaro, A. Coquillon, & J.-P. Chevrot (dir.) *La Liaison : Approches Contemporaines*. Berne : Peter Lang, pp. 263-282.

1. Introduction

Les études sur la liaison se sont pour la plupart focalisées sur des corpus issus de productions d'adultes (Agren, 1973 ; Mallet, 2008) ou d'enfants (Dugua, 2006 ; Nardy, 2008) sans que ne se croisent, autour d'un même corpus, les interactions parents-enfant. Concernant l'acquisition de ce sandhi, différentes hypothèses et autres scénarios apparaissent à ce jour à partir de corpus plus ou moins importants (Chevrot et al., 2009 ; Wauquier, 2010) sans pour autant que les productions adultes en direction d'enfants ne soient décrites qualitativement et quantitativement. Grâce au projet *Phonlex*, nous avons pu recueillir six corpus denses d'interactions parents-enfant dans le but de décrire l'input parental et de mesurer son influence sur l'acquisition de la liaison. Notre première démarche est descriptive : il s'agit de rendre compte des liaisons que l'enfant entend réellement dans l'input. Pour ce faire, nous décrirons les contextes morphosyntaxiques et la nature lexicale des mots les plus présents. Une autre finalité, afin de tester le scénario basé sur l'usage proposé par Chevrot et al. (2009), concerne le fait de savoir si l'enfant peut se baser sur des structures récurrentes, régulières et limitées l'aidant à inférer un fonctionnement plus général. Enfin, pour répondre aux observations émises par Wauquier (2010) sur le lien non avéré entre input et acquisition de la liaison, nous démontrerons grâce à l'analyse d'interactions parents-enfant que la production de formes erronées par les enfants semble fortement influencée par l'input adulte.

2. L'acquisition de la liaison

Aujourd'hui, un des thèmes majeurs dans les recherches en acquisition du langage concerne l'encodage de la variation chez le jeune sujet et les dispositifs cognitifs réguliers mis en place lors de la production de formes variables. Ces travaux donnent lieu à deux types de débats principaux dans lesquels nous intervenons à travers l'étude de la liaison. Le premier concerne le lieu de la variation. Est-elle ancrée phonologiquement ou lexicalement ? En clair, est-ce qu'il existe des contraintes voire des règles phonologiques générales, intervenant sur la production variable de la liaison catégorique ou variable sans faire référence à la nature même du mot et à sa fréquence en mémoire (Wauquier, 2010) ? Ou bien, est-ce que le lexique lui-même pourrait avoir encodé des constructions lexicales que le locuteur activerait selon des procédures davantage liées à l'usage qu'à des principes abstraits (Bybee et Hopper, 2001 ; Dugua, 2006) ?

Cette question est fortement liée au deuxième débat qui anime nos travaux. Il a pour thème les processus à l'œuvre dans l'acquisition de la liaison. Globalement, deux approches

s'affrontent. L'une, abstractionniste et plutôt inductive, préconise une acquisition de la liaison contrainte par des principes abstraits (Wauquier-Gravelines et Braud, 2005). L'autre, exemplariste (Chevrot et al., 2009 ; Pierrhumbert, 2001), part plutôt du principe selon lequel « *language structure emerges from language uses* » (Tomasello, 2003). Dans ce cadre, c'est ce que l'enfant entend et produit qui le conduit à façonner un fonctionnement langagier régulier. A partir de ces hypothèses, un scénario développemental de l'acquisition de la liaison a été proposé. Il prédit que dans un premier temps l'acquisition repose sur la mémorisation d'une construction large regroupant mot1 et mot2 à partir des constructions les plus présentes dans l'input. Au cours de la deuxième étape, l'enfant va segmenter cette construction large et détacher le mot1 du mot2 sur la base du schéma Consonne-Voyelle, récurrent en français. Ainsi, la consonne de liaison (à partir de maintenant : CL) va rester attachée à l'initiale du mot2, engendrant la mémorisation de plusieurs exemplaires d'un même mot2 ([zuks], [nuks] et [tuks] par exemple). Par la suite, des traitements de type analogique motivés par la fréquence et la nature des collocations mot1-mot2 dans l'input et dans l'output engendreraient la construction de schémas abstraits, sans pour autant que les généralisations effacent cette base empirique. Ainsi, au stade terminal du scénario que nous défendons, l'enfant aurait à disposition des schémas abstraits permettant la production d'énoncés qu'il n'a jamais entendus. Dans cette optique, la variation phonologique serait donc modélisée sur des bases empiriques liées à la fréquence des constructions.

Les critiques récentes faites à l'encontre du scénario basé sur l'usage portent notamment sur le fait que les enfants seraient peu sensibles à l'input. En effet, selon Wauquier (2010), les formes erronées produites par l'enfant ne seraient pas majoritairement des formes produites dans l'input adulte. Selon l'auteure : « ces données [sur les erreurs de liaison] nous montrent que les enfants ne reproduisent pas du tout l'input adulte et qu'ils produisent des formes qu'ils n'ont jamais entendues ». De ce point de vue, les productions des enfants seraient davantage liées à la fréquence de la CL ainsi qu'à l'harmonisation consonantique (Wauquier-Gravelines et Braud, 2005).

De nombreuses études attestent de la reprise de l'input parental par les enfants pour de nombreux énoncés. Behrens (2006) rend compte d'une étude longitudinale réalisée à partir d'un corpus dense composé d'interactions entre un jeune locuteur allemand et ses parents. Léo est enregistré durant trois ans (entre 1;11 ans et 4;11 ans). D'après ces données, le langage précoce est hautement conservateur puisque les énoncés produits par l'enfant lors de sa quatrième année sont dans 63% des cas une reprise exacte d'un énoncé parental, que 27%

diffèrent légèrement (substitution, suppression ou addition d'un mot dans l'énoncé) et que 10% seulement seraient des créations substantielles. Dans le même ordre d'idées, Kidd, Lieven et Tomassello (2006) établissent que chez des enfants anglophones âgés de 2;10 ans et 5;8 ans, la fréquence relative avec laquelle un verbe apparaît dans une construction syntaxique dans l'input prédit la capacité des jeunes enfants à s'en souvenir et à reproduire ces constructions.

Par ailleurs, Cameron Faulkner, Lieven et Tomassello (2003) relèvent que le discours de douze mères anglaises adressé à l'enfant serait relativement restreint en termes de structures, ce qui faciliterait la mémorisation de structures majoritaires et récurrentes. Ils constatent qu'un peu plus de la moitié des énoncés des mères commencent par 52 syntagmes identiques (impératif, copule, question) formés de deux mots ou de deux morphèmes. Parmi ces 52 syntagmes, 45% commencent avec seulement 17 mots différents. Les observations sur les discours enfantins révèlent également une forte corrélation entre les structures utilisées par les mères et celles utilisées par les enfants.

Ces études témoignent de trois faits d'importance concernant l'input parental :

- la reprise d'énoncés par les enfants et leur faculté de mémorisation,
- le rôle prépondérant joué par la fréquence de certaines structures de l'input sur la vitesse de l'acquisition de ces dernières par l'enfant,
- le fait de proposer aux enfants une combinatoire limitée.

Ces éléments argumentent en faveur de la théorie basée sur l'usage (Langacker, 1986 ; Barlow & Kemmer, 2002 ; Tomasello, 2003). Dans le cadre de la liaison, ce modèle prédit que l'enfant mémoriserait plus rapidement des formes fréquentes et les reproduirait avant d'en déduire des principes plus généraux et de construire des schémas abstraits. Ce scénario serait facilité et rendu possible par la présence d'une combinatoire limitée et régulière formée de constructions récurrentes et fréquentes. Nous nous cantonnerons ici à décrire si l'input parental recueilli et analysé présente bien ces tendances.

3. Méthodologie

Les résultats proposés dans ce travail sont issus de l'analyse de corpus denses¹ recueillis en situation naturelle d'interaction entre trois fillettes et leurs deux parents respectifs. Pour

¹ Un corpus dense se compose d'enregistrements quotidiens relativement courts (1h – 1h30) d'interactions entre un enfant et ses parents sur une période relativement longue.

² Ce tableau ne contient pas exactement les mêmes valeurs de contextes de liaisons par rapport au tableau 2 car nous

récolter ces corpus, nous avons mis à disposition des parents un enregistreur numérique équipé d'un microphone omnidirectionnel intégré. Pendant une semaine, ceux-ci ont été chargés d'enregistrer leur enfant une heure par jour au cours de moments propices aux interactions parents-enfant. Cette méthode de recueil des données comporte deux principaux avantages. Premièrement, le volume relativement important des interactions recueillies permet l'analyse de phénomènes linguistiques peu fréquents. De plus, la méthode qui consiste à laisser aux parents le soin de gérer les enregistrements, peu intrusive, minimise les biais causés par la présence d'un observateur au domicile de l'enfant (Tomasello & Stahl, 2004). La majeure partie des temps de recueil a eu pour cadre le bain, les repas et des moments de jeux ou de lecture. Les données récoltées permettent de couvrir une période d'acquisition assez large, allant de 2;4 ans à 3;4 ans (cf. tableau 1).

	Salomé	Lola	Prune
Âge des fillettes	2;4 ans	3 ans	3;4 ans

Tableau 1 : âge des fillettes lors des enregistrements.

Au total, nous avons donc à notre disposition un corpus d'environ 21 heures (7h par enfant en moyenne) d'enregistrement que nous avons transcrit et annoté. Nous différencions ici deux types d'annotation : les annotations sur le contexte et l'annotation sur la liaison. Les annotations sur le contexte rendent compte du locuteur qui produit l'énoncé, des pauses et des chevauchements de discours. De plus, pour chaque énoncé parental, nous avons renseigné l'adresse du discours. Ainsi, trois types d'énoncés parentaux ont été distingués : les énoncés adressés à l'enfant (AE), les énoncés adressés à un autre adulte (AA) et les énoncés adressés à un groupe composé d'un enfant et d'un ou plusieurs adultes (AT). Cependant, dans cette étude, nous ne prendrons pas en compte ces critères d'adresse, les énoncés parentaux étant analysés dans leur globalité. Nous entendons donc par input parental tout ce que l'enfant entend de ses parents.

En ce qui concerne la liaison, l'ensemble des contextes de réalisation potentielle a été annoté en fonction de quatre critères :

- le contexte syntaxique de production (entre déterminant et nom par exemple),
- la consonne de liaison potentiellement réalisable (/n/, /z/ ou /t/ dans la majorité des cas),

- la consonne de liaison effectivement réalisée (ou l’absence de réalisation),
- le caractère variable ou catégorique de la liaison selon la catégorisation de Chevrot et al. (2007).

Au total, nous avons à notre disposition pour cette étude 5932 contextes de liaisons, réalisées ou non, pour l’ensemble des locuteurs enregistrés. A ces contextes de liaison viendront s’ajouter, pour la dernière partie de nos analyses, un relevé d’erreurs dans les productions de Prune notées “au vol”. Cette pratique est utile et facile à mettre en place pour les linguistes, les contraintes méthodologiques demeurant faibles. Cependant, ces erreurs ne peuvent pas constituer la base d’une généralisation théorique mais viendront compléter notre corpus d’erreurs déjà important.

4. Résultats

4.1 Descriptif de l’input parental

Trois types de résultats seront présentés et discutés ici. Dans un premier temps, nous observerons les taux de production des liaisons variables et catégoriques dans les énoncés parentaux. La suite de nos analyses se focalisera sur la nature des mots¹ employés par les parents des fillettes en contexte de liaison. Enfin, nous mettrons en rapport les fréquences d’apparition des mots¹ et des mots² utilisés en input dans le but de vérifier s’il est possible de relever des collocations mot¹-mot² récurrentes.

4.1.1 Taux de production des liaisons

Dans cette section, nous présentons une description quantitative de l’input parental au niveau de la liaison. L’objectif est d’observer si, quantitativement, l’input que reçoit chacune des trois fillettes est similaire ou si l’on peut relever des différences. Le tableau présenté ci-dessous (tableau 2) regroupe, pour chaque couple de parents, le nombre de contextes de liaison (réalisées ou non) ainsi que le taux de liaisons justes. Les liaisons facultatives non réalisées sont ici comptées comme des liaisons justes.

	Parents Salomé	Parents Lola	Parents Prune
Nombre de contextes de liaisons	1685	785	1531
Taux de liaisons justes	99,0% (667/674)	98,8% (318/322)	98,9% (753/761)

Tableau 2 : contextes de liaison dans l’input

Nous constatons un nombre de contextes de liaison plus ou moins important chez les parents des trois fillettes (cf. tableau 2). Alors que les parents de Salomé et les parents de Prune produisent un nombre de contextes assez proche en sept heures d'enregistrement (respectivement 667 et 753 liaisons réalisées), les parents de Lola en réalisent plus de deux fois moins (318 liaisons réalisées). Cette différence peut être reliée à des circonstances sociolinguistiques, les parents de milieux défavorisés s'adressant moins à leur enfant (Nardy, 2008, pour une revue). Cependant, comme nous nous y attendions, l'input perçu par les trois fillettes est qualitativement équivalent et composé presque exclusivement de liaisons justes. Quelques erreurs apparaissent dans les productions parentales. Il s'agit dans la plupart des cas de liaisons abusives, souvent réalisées entre un mot invariable et un nom pluriel ([katzuks] pour "quatre ours" par exemple). Le bain linguistique serait donc formé de liaisons réalisées justes, dont la fréquence d'apparition varie en fonction des locuteurs. Cependant, lorsque l'on observe ces réalisations en fonction du caractère variable ou catégorique de la liaison², des différences plus nettes se dégagent entre les inputs parentaux (cf. tableau 3).

Ces résultats témoignent d'une grande différence entre la réalisation de liaisons catégoriques et de liaisons variables par les parents, expliquant du même coup la très faible proportion de liaisons variables réalisées justes en situation naturelle chez les enfants (Chabanal, 2003). En effet, alors que les liaisons catégoriques sont maîtrisées et systématiquement réalisées par l'ensemble des parents, l'input parental semble offrir à l'enfant un nombre restreint de liaisons variables réalisées. Une nette variabilité dans la réalisation de ce type de liaison se dégage. Alors que les parents de Salomé et de Prune réalisent la liaison variable dans 11,6% et 13,8% des cas, les parents de Lola ne les réalisent que très rarement (2,7% de liaisons variables réalisées). Le test de Chi² de conformité révèle des différences très significatives entre le taux de réalisation de la liaison variable chez les parents de Lola et chez les parents de Prune et Salomé (respectivement : Chi² = 41.2575, p<0.0001 et Chi² = 31.0878, p<0,0001).

² Ce tableau ne contient pas exactement les mêmes valeurs de contextes de liaisons par rapport au tableau 2 car nous avons retiré quelques contextes où le statut de certaines liaisons nous semble indéfini (ni catégorique ni variable).

	Parents Salomé	Parents Lola	Parents Prune
Taux de réalisation de la liaison variable	11,6% (131/1129)	2,7% (13/475)	13,8% (123/889)
Taux de réalisation de la liaison catégorique	100% (536/536)	99,7% (305/306)	100% (630/630)

Tableau 3 : taux de réalisation des liaisons variables et catégoriques

4.1.2 Nature des mots1

Dans cette partie, notre objectif sera d'observer si l'input proposé à l'enfant est constitué d'un nombre restreint de contextes de liaison ou si, au contraire, le discours parental propose une grande variété de constructions différentes. Cette étude a pour objet de vérifier s'il existe pour le phénomène de liaison une certaine récurrence dans les constructions. Pour ce faire, nous avons relevé les 15 mots1 les plus fréquents produits en input en contexte de liaison catégorique et de liaison variable. Pour chacun d'entre eux, un calcul de fréquence a été effectué en divisant le nombre d'occurrences du mot1 par le nombre total de contextes de liaison, en séparant liaisons catégoriques et liaisons variables. De cette manière, si un mot1 est produit 50 fois pour un nombre total de 500 contextes de liaison variable, le taux de fréquence de ce mot1 en contexte de liaison variable sera de 10%. Nous présentons ci-dessous la fréquence cumulée de la présence des mots1 dans l'input de chacune des fillettes en contexte de liaison catégorique puis de liaison variable.

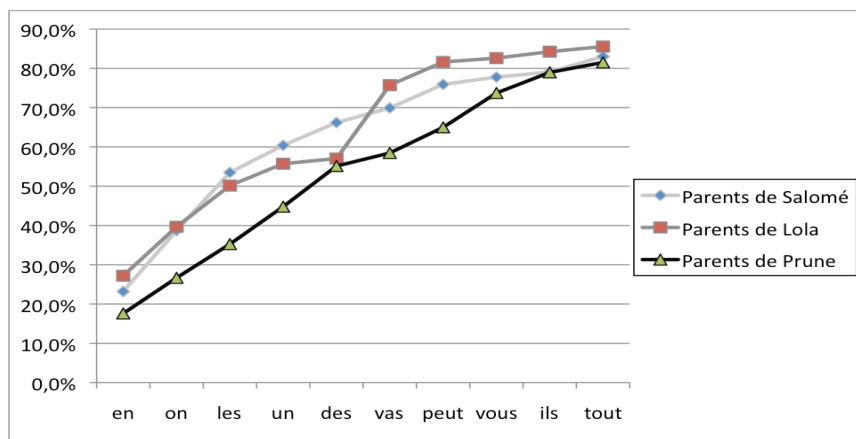


Figure 1 : fréquence cumulée des mots1 produits en contexte de liaison catégorique dans l'input parental des trois fillettes

Ces premières informations indiquent que dix mots1 constituent plus de 80 % des liaisons catégoriques dans le bain linguistique des trois enfants. Par ailleurs, les 5 mots1 les plus

présents constituent entre la moitié et les deux tiers des liaisons catégoriques de l'input. Il y aurait donc peu de régularités à repérer pour l'enfant. Les six locuteurs adultes proposeraient un nombre de mots1 limité à leurs enfants, ce qui pourrait faciliter l'apprentissage du fonctionnement de la liaison et la capacité à générer des schémas récurrents. Par la suite, l'enfant sera en mesure d'introduire des mots nouveaux correspondant à la même fonction grammaticale sur la base des mots souvent rencontrés dans cette position. L'autre information concerne la similarité des mots1 apparaissant dans l'input des trois enfants.

En observant les mots1 produits au moins dix fois par chaque couple parental en contexte de liaison catégorique, il s'avère que cinq mots1 sur sept se retrouvent dans les inputs parentaux des trois fillettes. Nous avons répété les mêmes analyses en contexte de liaison variable dans le but de vérifier si, dans ce contexte, il était possible d'observer les mêmes régularités dans la distribution des mots1 en input parental (cf. figure 2) :

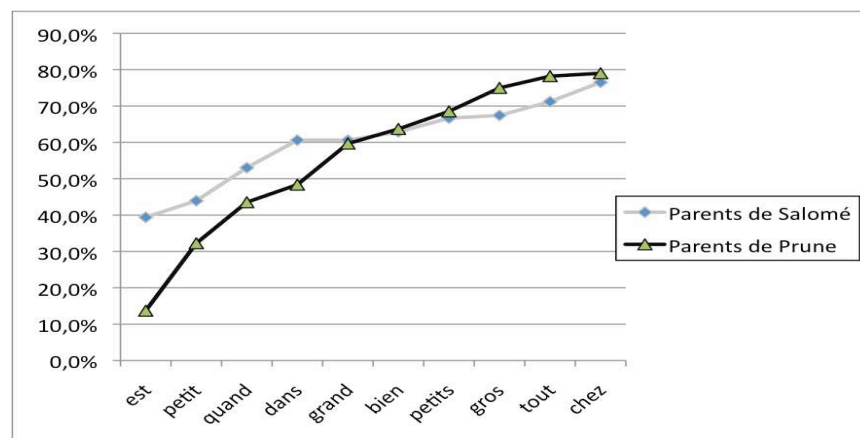


Figure 2 : fréquence cumulée des mots1 produits en contexte de liaison variable dans l'input parental des trois fillettes

Au vue du faible taux de réalisation de la liaison variable dans les productions des parents de Lola (13 liaisons réalisées, cf. tableau 3), nous présentons ici uniquement les données pour les parents de Salomé et de Prune. Les résultats révèlent les mêmes tendances à l'emploi d'un nombre limité de mots1 récurrents en contexte de liaison variable. En effet, les cinq mots1 les plus fréquents représentent environ 60% de l'ensemble des contextes de liaison variable réalisée.

Parents de Salomé		Parents de Lola		Parents de prune	
Mot1	Nombre	Mot1	Nombre	Mot1	Nombre
en	125	en	82	en	111
on	83	vas	57	des	65
les	79	on	38	un	60
est	60	les	32	on	57
un	37	peut	18	vous	55
peut	32	un	17	les	54
des	31	prends	13	peut	41
tout	26	tes	9	ils	33
vas	20	bien	5	petit	23
ton	13	choisis	5	vas	21

Tableau 4 : distribution des mots1 en contexte de liaison catégorique dans l'input parental des trois fillettes

Ces résultats permettent de mettre en exergue trois caractéristiques de l'input au niveau de la liaison. Premièrement, les taux de réalisations de liaisons offerts aux enfants sont différents selon le caractère variable ou catégorique de la liaison. Alors que les parents réalisent correctement la quasi-totalité des liaisons catégoriques, nous notons un faible taux de réalisation de la liaison variable. De plus, ce taux de réalisation varie significativement en fonction des locuteurs. Deuxièmement, notre relevé des mots1 produits dans l'input parental en contextes de liaison catégorique fait apparaître une certaine régularité dans l'apparition de ces mots1. En effet, parmi les dix mots1 les plus utilisés par les parents des fillettes, six se retrouvent dans chaque input parental (cf. tableau 4).

La distribution des mots1 dans l'input semble donc limitée, peu d'entre eux étant à l'origine de la grande majorité des liaisons perçues par l'enfant. Ce dernier semble être exposé à deux types de constructions au niveau des contextes de liaison catégoriques ou variables : des constructions récurrentes très fréquentes (formées de liaisons réalisées ou non), constituant la majorité des liaisons en input et des constructions peu fréquentes (formées de liaisons réalisées ou non) qui ne se retrouvent que ponctuellement dans les productions parentales. A partir de ces informations, nous avons décidé de focaliser notre attention sur ce que recevait l'enfant en termes de constructions mots1-mots2 en contexte de liaison catégorique. Nous avons voulu rechercher si parmi les mots1 des liaisons catégoriques les plus fréquents, l'enfant rencontre des séquences mots1-mots2 variées ou figées. Autrement dit, nous avons cherché à savoir par exemple si l'enfant entend plus souvent "en a" que "en aurait", "en avoir", "en ai"... Cet élément pourrait, au-delà des aspects descriptifs concernant

l'input parental, nous renseigner sur la manière dont l'enfant peut raisonner pour la mise en place de l'alternance phonologique de la liaison.

4.1.3 Production des constructions mot1-mot2 en contexte de liaison catégorique dans l'input parental.

Dans cette section, nous analysons le taux de figement de certaines constructions mot1-mot2 produites par les parents. Ces analyses sont réalisées à partir des quatre mots1 les plus employés dans l'input parental : « en », « des », « les » et « on ». Pour chacun de ces mots1, nous avons relevé les trois collocations mot1/mot2 les plus fréquentes dans les énoncés parentaux. Ainsi, dans les productions des parents de Salomé, les trois constructions les plus fréquentes avec les mots1 « en » sont les suivantes : « en a » (50 occurrences), « en as » (26 occurrences) et « en ai » (10 occurrences). En divisant le nombre total de contextes de liaison dans lesquels apparaît le mot1 « en » (138 occurrences) par le nombre de ces constructions les plus fréquentes (86 occurrences), nous obtenons le taux de figement du contexte de liaison « en + X » dans l'input reçu par Salomé. Dans ce cas précis, le taux de figement du contexte « en + X » chez les parents de Salomé s'élève à 62,3% (ce qui signifie que 62,3 % des liaisons après « en » apparaissent dans trois séquences mot1-mot2). Le tableau ci-dessous regroupe les taux de figement relevés chez les parents des fillettes dans les quatre contextes analysés :

Parents	Mot1	Nombre d'occurrences du mot1	Nombre d'occurrences des trois constructions les plus fréquentes	Taux de figement
Parents de Salomé	En	125	93	74,4%
	Des	31	11	35,5%
	Les	79	25	31,6%
	On	83	43	51,8%
Parents de Lola	En	83	62	74,7%
	Les	32	18	56,3%
	On	38	25	65,8%
Parents de Prune	En	112	88	78,6%
	Des	65	16	24,6%
	Les	54	14	25,9%
	On	57	39	68,4%

Tableau 5 : pourcentage de réalisation des 3 constructions les plus présentes parmi les 4 mots1 étudiés

Nos données laissent apparaître deux caractéristiques au niveau du figement des constructions mot1-mot2 dans l'input parental. Tout d'abord, certains mots1 se trouvent utilisés dans des constructions figées dans des proportions très élevées. Nous notons, par exemple, que les taux de figement des mots1 "en" et "on" sont les plus élevés dans les productions de chacun des couples de parents. Les parents de Prune, par exemple, tendent à utiliser un vocabulaire restreint et récurrent après ces mots1, les taux de figement s'élevant jusqu'à 78,6% (cf. tableau 5). En revanche, les taux de figement des mots1 "des" et "les" sont nettement plus bas, ce qui montre que, dans ces contextes de liaison, les parents tendent à employer un lexique beaucoup plus diversifié. Dans ces contextes, l'enfant perçoit ainsi une combinatoire mot1-mot2 beaucoup plus riche et variée. D'autre part, nous observons une variation de ces taux de figement en fonction du couple de parents. Alors que Lola semble recevoir un input assez faible en termes de variété des constructions, les parents de Prune et de Salomé proposent une combinatoire plus variée à leur fillette. Cette différence est nettement visible au niveau du mot1 "les" (cf. tableau 5).

Le discours proposé à l'enfant, en ce qui concerne la liaison, présente donc deux tendances de combinatoire, entre figement de certaines constructions et diversité. Alors que peu de mots1 constituent la majeure partie de l'input parental (cf. figures 1 et 2), les enfants se trouvent également exposés à un nombre assez important de mots1 produits moins de cinq fois dans nos enregistrements. Au niveau des constructions mot1-mot2, le même constat s'impose : les enfants perçoivent à la fois des séquences mot1-mot2 fréquentes et récurrentes et des séquences beaucoup plus rares. Ces caractéristiques de l'input parental permettraient ainsi aux jeunes locuteurs de pouvoir extraire directement de l'input des séquences figées ou des exemplaires de mot2 qui se trouvent en nombre assez important. Dans un deuxième temps, la richesse et la diversité des combinaisons proposées faciliteraient la construction de schémas abstraits, davantage favorisée par la fréquence des *types* de constructions que par celle des *tokens* (Goldberg, 2003 ; Goldberg, 2006 ; Langacker, 2009).

La dernière partie de notre travail tentera de répondre à l'influence de l'input parental sur les productions erronées des fillettes. Nous observerons si la fréquence des constructions en input peut avoir une incidence sur les productions erronées des jeunes locuteurs.

4.2 Formes erronées dans les productions enfantines : reprise de l'input ou création ?

Nous nous focalisons ici sur les productions erronées de Prune (3;4 ans) en les mettant en relation avec les productions de ses parents. Parallèlement au corpus dense recueilli en

situation naturelle d'interaction, nous avons pu récolter, pour Prune, des formes erronées notées "à la volée" au cours du même mois. Au total, nous avons pu constituer un corpus de 110 erreurs en frontière mot1-mot2 du type une z-abeille [ynzabɛj], un z-œuf [œzœf]. Comme nous l'avons dit précédemment, nous faisons l'hypothèse (Chevrot et al., 2009) que ces erreurs seraient liées à une segmentation syllabique en consonne-voyelle faisant émerger dans le lexique de l'enfant des formes lexicales de mots2 avec le phonème de liaison présent en initial. Cette mauvaise segmentation multiplierait les exemplaires ("z-oiseau", "t-oiseau", "n-oiseau") que l'enfant pourrait mémoriser pour une même entité lexicale ("oiseau"). L'enfant aurait donc à sa disposition, pour un même mot2, différentes formes phonologiques activables après un même mot1.

Les erreurs relevées dans les productions de Prune ont été classées en fonction de la consonne de liaison mise en jeu (/n/, /t/ et /z/). Dans le but d'observer si le choix de l'exemplaire utilisé en erreur était conditionné par les productions justes de Prune ou de ses parents, nous avons relevé pour chaque mot2 l'ensemble des exemplaires produits. Ainsi, nous avons à notre disposition deux séries de données :

- l'ensemble des exemplaires de mots2 produits en contexte d'erreur par Prune (110 occurrences),
- pour ces mêmes mots2, l'ensemble des exemplaires produits par Prune et ses parents dans des contextes de liaison variable ou catégorique réalisée juste.

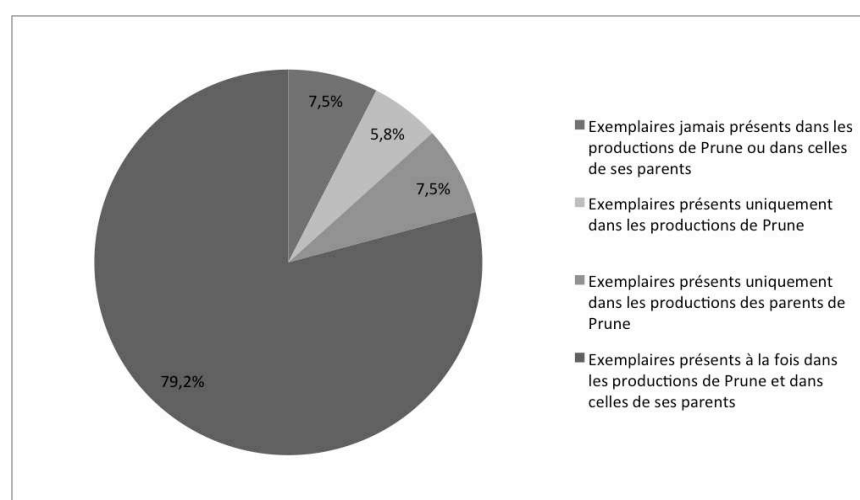


Figure 3 : présence en input et en output des exemplaires de mots2 produits par Prune

A partir de ces données, nous avons souhaité mettre en relation les effets de fréquence et la variation lexicale. Les questions concernent l'effet du bain linguistique sur le choix de l'exemplaire :

- l'enfant utilisera-t-il en contexte d'erreur un exemplaire qu'il a forcément déjà produit ou entendu en contexte juste ?
- l'enfant qui entend et produit plus souvent un exemplaire (par exemple l'exemplaire "n-oiseau") le sollicitera-t-il davantage en contexte d'erreur (par exemple "beau n-oiseau") ?

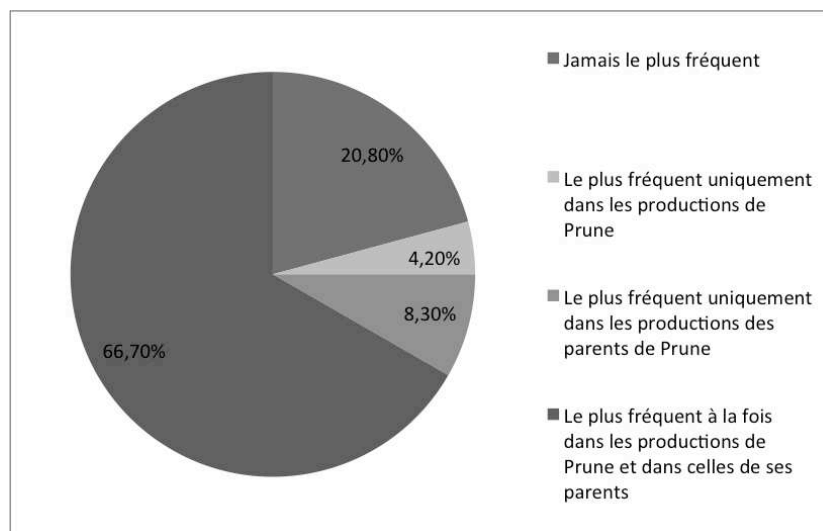


Figure 4 : fréquence des exemplaires produits en contexte d'erreurs

En réponse à la première question, il s'avère que Prune reprend majoritairement des exemplaires qu'elle a déjà produits et entendus dans des contextes justes de liaison (cf. figure 3). Les résultats montrent que 79,2% des exemplaires d'erreurs sont présents à la fois dans les productions justes de Prune et de ses parents. Par ailleurs, 7,5% des exemplaires produits en contexte d'erreur sont présents uniquement dans les productions parentales alors que 5,8% sont présents exclusivement dans les productions de Prune. Nos données suggèrent que l'enfant reproduit essentiellement, dans ses erreurs, des exemplaires de mots² apparaissant dans les événements d'usage (en production ou en réception). En effet, seulement 7,5% des exemplaires employés en contexte d'erreur ne sont jamais présents dans les productions de Prune ou de ses parents. On constate donc une faible part de création dans les contextes d'erreur et une variation des exemplaires qui semble conditionnée par le bain linguistique (dans 84,7% des cas).

Concernant la deuxième question, nous constatons, dans la plupart des contextes d'erreur (2/3), une tendance chez Prune à utiliser l'exemplaire de mot2 le plus produit par elle-même et ses parents en contexte de liaison juste.

Pour ce qui est des exemplaires non en lien avec leur fréquence d'apparition chez Prune ou ses parents (21% des cas), des observations dans les tours de parole précédant les erreurs ont montré qu'il n'y avait pas de lien direct avec la récence de cet exemplaire. Cependant, le peu de mots2 observés (5) ne permet pas de formuler de conclusions définitives. Cette hypothèse sera discutée dans de futures études grâce à d'autres corpus denses de ce type recueillis et analysés dans le cadre du projet ALIPE (*Acquisition de la Liaison et Interactions Parents Enfant*).

5. Discussion

Nos observations de productions parentales extraites de corpus denses nous ont permis de décrire l'input parental au niveau de la liaison, quantitativement et qualitativement. Nos résultats indiquent que l'input diffère quantitativement entre les trois couples de parents étudiés. Les parents de Lola, par exemple, produisent plus de deux fois moins de contextes de liaison que les parents de Salomé pour une méthodologie de recueil des corpus identique. Qualitativement, si les taux de réalisation juste de la liaison sont identiques, les données sur les productions des parents de Lola révèlent des taux de réalisation de la liaison variable significativement différents de celui des autres couples.

Les résultats concernant les collocations mots1-mots2 des parents apportent des éléments importants sur les effets qu'ils pourraient avoir en termes d'acquisition de la liaison. En effet, l'enfant semble être exposé à deux types de constructions au niveau de la liaison : des contextes lexicaux identiques et récurrents où le mot1 est généralement suivi d'un nombre très restreint de mots2 et des contextes plus ouverts où le mot1 est suivi d'une grande diversité de mots2 différents. Ces éléments pourraient l'aider à fixer des liaisons et à percevoir les types de liaison où par exemple les constructions [un + mot2] sélectionnent toujours la consonne /n/. D'autre part, les choix des exemplaires opérés par Prune sont effectués à partir d'une sensibilité à la fréquence des mots2 présents dans l'input parental et non pas en ayant recours à des connaissances phonologiques plus générales, sans lien avec le lexique mémorisé. Dans l'ensemble, les données présentées ici vont dans le sens d'une reconnaissance d'un vrai lien entre phonologie et lexique.

6. Références bibliographiques

- Agren, J. (1973). *Etude sur quelques liaisons facultatives dans le français de conversation radiophonique : fréquences et facteurs*. Uppsala University Press.
- Barlow, M., & Kemmer, S. (Eds.) (2000). *Usage Based Models of Language*. Stanford California : CSLI Publications.
- Behrens, H. (2006). The input-output relationship in first language acquisition. *Language and Cognitive Process*, 21, 2-24.
- Bybee, J. L., & Hopper, P. (Eds.) (2001). *Frequency and the emergence of linguistic structures*. Typological Studies in Language, 45. Amsterdam/Philadelphia : John Benjamins Publishing Company.
- Cameron-Faulkner, T., Lieven, E., & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science*, 27(6), 843-873.
- Chabanal, D. (2003). *Un aspect de l'acquisition du français oral : la variation socio-phonétique chez l'enfant francophone*. Université Paul-Valéry - Montpellier III.
- Chabanal, D., & Liégeois, L. (2010). Description de l'input dans le cadre du recueil de corpus denses parents-enfants. *Phonlex 2010 : Liaison and other sandhi phenomena*.
- Chanier, T., Liégeois, L., Chabanal, D., & Lotin, P. (2012). *Projet Acquisition de la Liaison et Interactions Parents-Enfant*. Laboratoire de Recherche sur le Langage. Clermont Université. <<http://lrl-diffusion.univ-bpclermont.fr/alipe>>
- Chevrot, J.-P., Chabanal, D., & 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(01), 103-128.
- Chevrot, J.-P., Dugua, C., & Fayol, M. (2009). Liaison acquisition, word segmentation and construction in French : a usage based account. *Journal of Child Language*, 36(3), pp. 557-596.
- Dugua, Céline. (2006). *Liaison, segmentation lexicale et schémas syntaxiques entre 2 et 6 ans. Un modèle développemental basé sur l'usage*, doctorat, Université Stendhal - Grenoble III.
- Goldberg, A. E. (2003). Constructions : a new theoretical approach to language. *Trends in Cognitive Sciences*, 7(5), 219-224.
- Goldberg, A. (2006). *Constructions at Work : The Nature of Generalization in Language*. Oxford : Oxford University Press.

- Kidd, E., Lieven, E., & Tomasello, M. (2006). Examining the role of lexical frequency in the acquisition and processing of sentential complements. *Cognitive Development*, 21(2), pp. 93-107.
- Langacker, R. W. (1986). An Introduction to Cognitive Grammar. *Cognitive Science*, 10(1), pp. 1-40.
- Langacker, R. W. (2009). A dynamic view of usage and language acquisition. *Cognitive Linguistics*, 20(3), 627-640.
- Mallet, G. (2008). *La liaison en français : descriptions et analyses dans le corpus PFC*, doctorat, Université Paris Ouest Nanterre La Défense.
- Nardy, A. (2008). *Acquisition des variables sociolinguistiques entre 2 et 6 ans : facteurs sociologiques et influences des interactions au sein du réseau social*, doctorat, Université Stendhal - Grenoble III.
- Pierrehumbert, J. (2001). Exemplar dynamics : Word frequency, lenition and contrast. In J. L. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structures*. Typological Studies in Language, 45. Amsterdam/Philadelphia : John Benjamins. pp. 137-158.
- Tomasello, M. (2003). *Constructing a Language : A Usage-Based Theory of Language Acquisition*. Harvard University Press.
- Tomasello, M., & Stahl, D. (2004). Sampling childrens spontaneous speech : how much is enough ? *Journal of Child Language*, 31(1), 101–121.
- Wauquier, S. (2010). Acquisition de la phonologie “du” français : des usages à la structure. *Langue française*, 168, 127-144.
- Wauquier-Gravelines, S., & Braud, V. (2005). Proto-déterminant et acquisition de la liaison obligatoire en français. *Langages*, 158, 53-65.