



**HAL**  
open science

**STATUT DES REPRESENTATIONS  
PHONOLOGIQUES EN ACQUISITION,  
TRAITEMENT DE LA PAROLE CONTINUE ET  
DYSPHASIE DEVELOPPEMENTALE**

Sophie Wauquier

► **To cite this version:**

Sophie Wauquier. STATUT DES REPRESENTATIONS PHONOLOGIQUES EN ACQUISITION, TRAITEMENT DE LA PAROLE CONTINUE ET DYSPHASIE DEVELOPPEMENTALE. Linguistique. EHESS, 2005. tel-03252302

**HAL Id: tel-03252302**

**<https://hal.science/tel-03252302>**

Submitted on 7 Jun 2021

**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.

Ecole des Hautes Etudes En Sciences Sociales

**« STATUT DES REPRESENTATIONS PHONOLOGIQUES EN  
ACQUISITION, TRAITEMENT DE LA PAROLE CONTINUE ET  
DYSPHASIE DEVELOPPEMENTALE »**

Dossier de candidature au diplôme  
D'Habilitation à Diriger des Recherches  
volume 1 : mémoire de synthèse  
présenté par Sophie Wauquier-Gravelines  
novembre 2005

**MEMBRES DU JURY**

Joaquim Brandão de Carvalho  
Emmanuel Dupoux  
Pierre Encrevé , directeur  
Jean Lowenstamm  
Noël Nguyen  
Marilyn Vihman

*Schließlich brauchen sie uns nicht mehr, die Früheentrückten, man entwöhnt sich des Irdischen sanft, wie man den Brüsten milde der Mutter entwächst.  
Aber wir, die so große Geheimnisse brauchen, denen aus Trauer so oft seliger Fortschritt entspringt- : « könnten » wir sein ohne sie ?*

R.M. Rilke, Duineser Elegien, I Elegie.

*La base empirique de la science objective ne comporte donc rien d'absolu ». La science ne repose pas sur une base rocheuse. La structure audacieuse de ses théories s'édifie en quelque sorte sur un marécage. Elle est comme une construction bâtie sur pilotis. Les pilotis sont enfoncés dans le marécage mais pas jusqu'à la rencontre de quelque base naturelle ou « donnée » et, lorsque nous cessons d'essayer de les enfoncer davantage, ce n'est pas parce que nous avons atteint un terrain ferme. Nous nous arrêtons, tout simplement, parce que nous sommes convaincus qu'ils sont assez solides pour supporter l'édifice, du moins provisoirement.*

Karl R. Popper, La logique de la découverte scientifique, Ch. V : le problème de la base empirique

## Ouverture

Mes travaux se situent à l'interface entre phonologie et psycholinguistique (acquisition et adulte). En linguistique, l'Ecole générative, et N. Chomsky (1966, 1968, 1975, 1990) en particulier, ont joué un rôle fondateur dans l'émergence des « Sciences Cognitives » en reposant les bases d'une réflexion sur le mentalisme linguistique, qui se réclame de la tradition rationaliste issue de Descartes et s'appuie sur l'innéisme des représentations linguistiques ainsi que sur une conception dualiste stricte entre esprit/cognition et substance.

La conception de la cognition proposée par Chomsky, repose explicitement, à partir de 1965, sur le postulat d'une dichotomie entre la « compétence » désignant les représentations (et secondairement les procédures) mentales abstraites et en grande partie innées qui constituent la grammaire des locuteurs-auditeurs ; et la « performance », désignant les opérations et procédures psychologiques de mise en oeuvre de cette faculté de langage, dans des opérations concrètes de perception et de production. Cette dichotomie, qui semblait au départ relever plutôt d'un partage des tâches bien compris, a entraîné, dans les faits un morcellement de la réflexion élaborée autour de la conception mentaliste du langage que Chomsky proposait. Les impératifs méthodologiques de la linguistique, traitant de la compétence d'une part, et de la psycholinguistique et de la phonétique expérimentales, rendant compte de la performance d'autre part, ont creusé respectivement cet écart.

La créativité scientifique dont ont fait preuve ces domaines depuis vingt ans était peut-être à ce prix. Néanmoins, il n'est pas sûr qu'une telle position soit le gage exclusif d'une fertilité renouvelée et continue et qu'elle puisse fonder définitivement une réflexion empiriquement fondée sur les facultés linguistiques humaines, sans mener à des apories inhérentes à une méconnaissance respective des apports de chaque domaine. La convergence de réflexions théoriques et de résultats expérimentaux en provenance de divers domaines en interface avec la phonologie, dans une optique résolument, pluridisciplinaire me paraît être également une manière

fertile d'aborder le problème du réalisme psychologique d'une phonologie interne abstraite et autonome. Cette question est le point central de mes travaux.

En effet, la constitution en cours d'un réseau de disciplines permettant de rendre compte "du cognitif", nécessite non seulement que chacun des domaines apporte sa contribution théorique et ses résultats, mais également que soient pensées les interfaces entre ces différents domaines au sein d'une théorie visant l'unification des champs. A cet égard, la juxtaposition de l'étude psycholinguistique et de la théorie phonologique n'est pas en soi la garantie d'une explication plus satisfaisante des phénomènes phonologiques en tant qu'ils sont cognitifs, si elle n'est pas envisagée comme une troisième voie visant à dépasser les apories respectives des deux champs convoqués.

Par ailleurs la tentative épistémologique qui consisterait à greffer tels quels les modèles actuels de la compétence phonologique dans les modèles psycholinguistiques me paraît également problématique : l'échec des travaux dans le cadre, par exemple, de la théorie de la complexité dérivationnelle en fut la preuve pour la syntaxe.

Il convient par conséquent de construire les articulations entre ces différentes disciplines, -et c'est là assurément que nous manquons encore cruellement de modélisations et de clairvoyance-, dans un modèle cognitivement pertinent rendant compte de ce qui relève respectivement des procédures et des représentations abstraites, lors des opérations de production et de perception du langage simultanément.

C'est dans cet esprit, que j'ai cherché à me donner une formation pluridisciplinaire. Après avoir découvert la linguistique et la phonétique historique à l'Université de Lille où j'ai obtenu une maîtrise de littérature française (1985) et l'avoir pratiquée à l'ENS Fontenay-St Cloud (1985-1990), j'ai passé un DEA de phonétique en 1987 (Paris 7/Paris 3) sous la direction d'Antoine Culioli, un DEA de Sciences Cognitives option psychologie expérimentale en 1991 (EHESS, Paris 6, Polytechnique) sous la direction de Juan Segui, puis j'ai présenté une thèse de doctorat sous la direction de Jean Lowenstamm, soutenue en 1996 à l'Université Paris

7 mettant en relation phonologie théorique et psycholinguistique expérimentale pour rendre compte de l'utilisation de la compétence phonologique lors de la segmentation de la parole continue chez l'adulte.

Le projet de travail dans son ensemble vise donc très précisément une reformulation de la question des interfaces que la phonologie entretient avec la psycholinguistique (adulte et enfant) et porte plus précisément sur le statut des représentations phonologiques dans les modèles de traitement adultes ainsi que dans les modèles d'acquisition.

Mes activités de recherche s'inscrivent dans une épistémologie rationaliste et mentaliste, retenant l'existence d'une phonologie interne abstraite et autonome utilisée en temps réel par les locuteurs-auditeurs et mise en place par les enfants lors de l'acquisition entre 0 et 3ans.

Elles s'articulent donc autour des deux axes suivants

- La place et le rôle des représentations phonologiques dans les phénomènes d'accès au lexique et de traitement de la parole continue chez l'adulte normal
- L'acquisition phonologique et l'émergence des représentations lexicales chez l'enfant normal et dysphasique

Dès mon doctorat, j'ai entrepris d'avancer des propositions théoriques et des résultats expérimentaux visant à esquisser l'interface entre phonologie générative, phonétique et psycholinguistique expérimentales chez l'adulte. Une partie du travail publié rend compte de ces travaux (chapitre I). Des recherches élaborées à partir de nouvelles hypothèses théoriques sont actuellement en cours sur la psycholinguistique de l'adulte (j'en rends compte au chapitre II).

La plupart des recherches présentées dans ce mémoire concernent l'acquisition de la phonologie chez les enfants « normaux » et dysphasiques. L'étude de l'acquisition d'une part et des troubles du langage parlé d'autre part me semblent

constituer des moyens particulièrement efficaces de nourrir ce type de questionnements, d'autant que cet angle d'approche a été assez peu développé jusqu'ici en France. Les données d'acquisition proposent à notre observation des productions et des conduites perceptives s'appuyant sur des représentations abstraites en devenir qui, on peut le supposer, entrent en continuité avec les représentations abouties de l'adulte. L'observation de ces données nous permet donc de falsifier les modèles adultes et d'en tester la cohérence : pour qu'un modèle adulte soit plausible, il faut en effet qu'il soit « apprenable » et au moins partiellement spécifié chez l'enfant qui l'acquiert, à moins d'envisager l'hypothèse d'une discontinuité entre le modèle adulte et le modèle enfant. C'est une hypothèse qu'on ne peut pas exclure *a priori* mais qui reste à interroger et à prouver.

Par ailleurs, outre cet apport des données d'acquisition à la théorie linguistique, l'étude expérimentale de ce matériau offre une résistance à la généralisation tout à fait passionnante et invite à l'observation détaillée. Les données linguistiques enfantines présentent une variété et une richesse qui méritent « en soi » l'intérêt en ce qu'elles permettent (comme c'est le cas en diachronie également) d'observer un système qui change et qui varie. La logique d'évolution, et en particulier la logique structurelle qui se fait jour lors de l'émergence du système phonologique d'un enfant, ouvre directement une fenêtre sur la structure que l'on peut inférer dans la grammaire adulte.

Enfin, la mise en relation des données d'enfants « normaux » et « dysphasiques » pour un même phénomène peut donner lieu à des applications cliniques : seule une connaissance précise des démarches de production et de perception chez le tout-petit, reposant sur des descriptions fines et détaillées et sur des protocoles expérimentaux reproductibles peut nous permettre d'espérer des applications dans le diagnostic des troubles du langage parlé et de leur rééducation. La précocité de ces prises en charge est souvent le gage d'une récupération quasi-totale des facultés linguistiques.

Les recherches sont en grande partie centrées autour d'un phénomène de sandhi du français : la liaison. Je vois s'exprimer dans le choix de cet objet de recherche, profondément lié à la métrique du français, outre la conséquence de mon monolinguisme, de ma formation ainsi que le goût et l'expérience d'abord littéraires et poétiques que j'ai des langues en général et de ma langue maternelle en particulier, le « Cumulative Paradox » labovien mentionné par Encrevé (1988), soutenant que plus une langue est connue, plus on peut faire de découvertes à son sujet. La liaison est sans doute l'un des phénomènes phonologiques les plus étudiés par des linguistes francophones et non francophones. Pourtant, alors que je commençais à me former en phonologie et psycholinguistique, j'ai constaté avec beaucoup d'étonnement que ces phénomènes n'avaient été que très peu étudiés en psycholinguistique de l'adulte et en acquisition. C'est donc ce à quoi je me suis attachée, d'abord en produisant des données ensuite en les confrontant aux modèles théoriques existants. Dans mes travaux, je défends l'idée selon laquelle le traitement chez l'adulte et l'acquisition chez le petit enfant sont conditionnés par une représentation abstraite et hautement structurée de la consonne de liaison et ne reposent pas sur un traitement ou un apprentissage supplétif et soumis au contexte de réalisation.

Sont présentés à côté de ces recherches sur la liaison, mes autres travaux sur l'acquisition des structures syllabiques, les phénomènes de troncations et réduplications chez l'enfant et la modélisation de l'acquisition phonologique à partir d'un gabarit lexical abstrait permettant de formaliser et d'expliquer ces différents phénomènes d'acquisition observables dans les données des enfants francophones

J'expose dans un premier temps (chapitre I) une synthèse des travaux réalisés depuis mon doctorat dans les domaines sus-cités correspondant aux publications placées dans le dossier joint que je soumetts au jury d'habilitation à diriger des recherches

Mais, un tel mémoire n'aurait sans doute que peu de sens s'il n'avait d'autres raisons d'être que d'entériner une pratique et une réflexion arrêtées en l'état. Les travaux publiés, les commentaires et critiques que j'en ai reçus ont ouvert de



nouvelles perspectives, de nouveaux projets que je présente au chapitre II. Sont exposés là, les réflexions, interrogations et sujets sur lesquels je travaille actuellement, qui le plus souvent constituent des prolongations, développements, reformulations, réanalyses parfois des données ou des hypothèses posées au chapitre précédent concernant la question des représentations phonologiques chez l'adulte ainsi que les propositions théoriques pour un modèle gabaritique autosegmental de l'acquisition phonologique.

Enfin, le mémoire d'habilitation à diriger les recherches offre un espace vierge et unique permettant l'exposition d'idées et surtout d'interrogations en gestation, parfois spéculatives, qui trouveraient difficilement, en l'état, leur place ailleurs. J'ai, dans cet esprit, mis en forme (Chapitres III et IV), des interrogations portant sur mes choix théoriques et les pratiques de recherche actuelles qui en découlent. Je pense ainsi témoigner également d'une recherche que j'espère vivante, active et présenter au jury les projets et progrès espérés.

Je remercie Pierre Encrevé, qui est mon garant dans cette démarche d'habilitation, pour l'attention qu'il a portée à mon travail ainsi que pour sa disponibilité. Je remercie également les membres du jury qui me font l'honneur d'accepter de participer à cette évaluation de mes travaux de recherche. Merci également à Joaquim Brandão de Carvalho et Tobias Scheer pour leur lecture des chapitres III et IV ainsi qu'à Hamida Demirdache, Cécile Fougeron, Nathalie Vallée et Annie Riolland.

Merci à Laurence Ordronneau, Natascha Siehoff, pour leur soutien généreux, amical, musical ainsi qu'à Sophie Aslanides et Gaëlle Tywoniuk, pour leur présence, affectueuse et bienveillante à tous égards.

## CHAPITRE I : SYNTHESE DES TRAVAUX

« Pour faire un tube de toilette  
 En chantant sur cet air bête  
 Avec des jeux de mots laids  
 Il faut pondre des couplets  
 Permits que je te réponde  
 C'est sûr, faut que tu les pondes  
 Bon, mais que dois-je pondre ?  
 Que ponds-je, Que ponds-je  
 Pot pot pot potpodet pot  
 Le dernier mot qui t'a servi était : « Ponds-je »  
 Serviette éponge ! parfait !...  
 Allez, vas-y, je vais t'aider  
 J'apprécie quand de toi l'aide-GANT DE TOILETTE  
 Me soutient cela va beau –CE LAVABO  
 Coup plus vite c'est bien la vé – C'EST BIEN LAVE  
 Rité, ça nous le savons A NOUS L'SAVON DE  
 TOILET'  
 [...]  
 CHANSON DE TOILETTE TRES POETIQUE TOC »

B. Lapointe, Le tube de toilette, *Intégrale*, Domens, 1994

## 1. Introduction

### 1.1. Linéarité, temporalité, segmentations et troncations

Chez Saussure, la représentation linéarisée du signifiant est érigée très explicitement comme un principe de la langue.

« Le signifiant étant de nature auditive, se déroule dans le temps seul et a les caractéristiques qu'il emprunte au temps : a) il représente une étendue, et b) cette étendue est mesurable dans une seule dimension : c'est une ligne. [...] Par opposition aux signifiants visuels (signaux maritimes, etc.), qui peuvent offrir des complications simultanées sur plusieurs dimensions, les signifiants acoustiques ne disposent que de la ligne du temps ; leurs éléments se présentent l'un après l'autre ; ils forment une chaîne. Ce caractère apparaît immédiatement dès qu'on les représente par l'écriture et qu'on substitue la ligne spatiale des signes graphiques à la succession dans le temps.»

F. de Saussure, *Cours de Linguistique générale*, 1966, p.103

« la ligne phonique doit être découpée comme aux ciseaux ... c'est ainsi que s'affirment les unités.

Un syntagme comporte un ordre, une suite linéaire, quelle qu'en soit l'ampleur... si nous pouvons découper les mots, c'est une des conséquences de ce principe »

F. de Saussure in Cao Xuan Hao, *Phonologie et linéarité, réflexions critiques sur les postulats de la phonologie contemporaine*, 1985, p.32

A la question de la linéarité du signifiant est donc nécessairement attachée celle de la segmentation. L'analogie musicale est convoquée pour définir la structure du signifiant comme unilinéaire et temporelle et pour considérer au final que les constituants (quel que soit le niveau de constituance concerné) entretiennent d'abord des rapports syntagmatiques formulés en termes de séquentialité.

« Il y a ici un caractère spécifique de la matière phonique, non mis suffisamment en relief, c'est de se présenter à nous comme une chaîne acoustique, ce qui entraîne immédiatement le caractère temporel qui est de n'avoir qu'une seule dimension. On pourrait dire que c'est un caractère linéaire : la chaîne de la parole, forcément se présente à nous comme une ligne. Les différences qualitatives... n'arrivent à se traduire que successivement<sup>1</sup> ... Tout forme une ligne, comme d'ailleurs en musique.

Il n'y a rien de plus uniforme, de plus pauvre, que l'ordre de la langue : la parole (comme la musique sans les accords) est linéaire. »

F. de Saussure in Cao Xuan Hao, *Phonologie et linéarité, réflexions critiques sur les postulats de la phonologie contemporaine*, 1985, p.30

La linéarité étant envisagée comme un principe structurel d'organisation de la langue, la possibilité du découpage des unités sur l'axe unilinéaire du temps est identifiée au caractère discret et catégoriel de ces mêmes unités.

Or ces déclarations saussuriennes postulent, me semble-t-il, une isomorphie de principe ou du moins une correspondance entre linéarité (du signifiant) et temporalité (du signal de parole) qui tendrait à identifier les relations syntagmatiques qu'entretiennent entre elles les unités découpables à la séquentialité purement temporelle de la substance sonore. Ce faisant, la segmentation ne peut alors être effectivement envisagée que comme un « découpage aux ciseaux » dans la matière sonore organisée sur l'axe unilinéaire du temps. Par ailleurs, l'analogie avec l'écriture et la référence explicite qu'en donne Saussure préfigure toutes les apories des analyses psycholinguistiques qui viendront<sup>2</sup> ainsi que des limites d'une phonologie

---

<sup>1</sup> C'est moi qui souligne.

<sup>2</sup> Problème détaillé en 2.1 de ce chapitre I.

unilinéaire (et notamment pour l'analyse orthographique de la liaison, notamment par Schane (Encrevé, 1988)).

Il est intéressant de noter que l'ambiguïté de la formulation saussurienne met immédiatement en lumière « le » problème que soulèvent pour un phonologue les questions de segmentation de la chaîne parlée en temps réel lors de l'accès au lexique. Elle reflète les ambiguïtés qui sont celles de la plupart des modèles psycholinguistiques de l'accès au lexique

- i) comment s'opère la mise en relation entre forme et substance, linéarité et temporalité, relations syntagmatiques et séquentielles
- ii) Doit-on postuler l'existence de procédures intermédiaires permettant cette mise en relation
- iii) A quel niveau exact et quand les unités de perception et de production utilisées par les locuteurs-auditeurs sont-elles calculables et calculées, « découpables » ?

C'est d'abord de cette question du rapport entre linéarité et temporalité dont je me suis occupée dans mes travaux i) sur le traitement de la parole continue chez l'adulte lors de l'accès au lexique ii) sur l'acquisition de la liaison chez le petit enfant, à partir de la question de l'établissement de la frontière gauche de mot pour l'accès au lexique.

Un phénomène phonologique du français comportait des caractéristiques structurelles permettant de poser assez radicalement ces questions de segmentation à la frontière gauche de mot et de mise en relation de la continuité temporelle du signal avec les représentations phonologiques. La liaison, dont la réalisation entraîne une réorganisation syllabique sur la frontière de mot, brouille en effet la conception d'une segmentation par « découpage aux ciseaux » d'unités lexicales entretenant des rapports de séquentialité simples. Enfin elle soulève, avec d'autres phénomènes, la question i) de l'existence du contexte disjonctif {C/#} à la frontière gauche, ii) de la

particularité structurelle de cette position (Lowenstamm, 1999) et iii) de l'exploitation que les auditeurs peuvent en faire en temps réel.

Chez l'adulte, la question qui m'a occupée concerne la formalisation précise des formes lexicales entrant en contexte de liaison, utilisables et utilisées par les auditeurs en temps réel. En d'autres termes, je me suis demandée à partir de quelles représentations phonologiques travaillait le locuteur, découpant le signal de parole en temps réel, dans une langue comme le français comportant ce sandhi. J'ai tenté de montrer que contrairement à ce que suppose la déclaration saussurienne, c'est justement le caractère plurilinéaire des représentations qui permet un traitement efficace de la temporalité du signal, en particulier dans le cas de la liaison.

Chez l'enfant, j'ai formulé des hypothèses concernant les stratégies de segmentation retenues pour reconnaître des frontières ambiguës et j'ai traité de la question de l'établissement des représentations lexicales et de l'émergence des régularités phonologiques en l'absence d'un lexique de référence fini, comparable à celui de l'adulte.

Ces travaux ont été complétés par des travaux sur le traitement de la liaison chez les enfants dysphasiques et les données obtenues ont ouvert un autre champ d'interrogations dont les travaux présentés rendent également compte. Ces données reflètent d'abord un pattern de réponse spécifique en ce qui concerne le traitement de la liaison chez les enfants dysphasiques qui pourrait être envisagé comme un marqueur de la dysphasie. Mais par ailleurs, les enfants testés présentent tous un déficit majeur qui les amène à réduire drastiquement la taille de l'empan de signal qu'ils sont capables de produire, même à un âge assez avancé. Ils opèrent, pour réduire les mots, des troncations des syllabes internes de mot. C'est également ce que font les petits francophones non dysphasiques à des âges moindres sur les mots excédant deux syllabes.

Toute cette partie des recherches concerne donc i) la question de la taille de l'unité initiale chez les enfants francophones ii) la nature des contraintes qui

conditionnent cette unité iii) l'évolution de cette unité (mémoire de travail, structure segmentale du mot, complexité syllabique du mot) iv) les déficits observés lors de l'évolution (déficit de la mémoire de travail, déficit du stockage lexical ou sous-spécification de représentations non-distinctives).

Enfin ces données nous ramènent aux questions du traitement de la séquentialité lexicale en temps réel, du recours à des régularisations et structures phonologiques et du rapport entre linéarité du signifiant et temporalité du signal : quand la totalité de l'item visé ne peut pas être produit (à cause de l'âge ou d'un déficit), l'observation du matériau prosodique et segmental qui sera préservé nous renseigne à la fois sur la disponibilité relative chez les enfants des informations phonologiques et sur les stratégies de préservation retenues. Il se trouve qu'elles ne sont pas du tout aléatoires. Ces données nous montrent bien que les enfants ne découpent pas « aux ciseaux » et que leurs productions articulent linéarisation de l'information phonologique et temporalité du signal en s'appuyant sur l'existence de positions prosodiquement et phonologiquement fortes et non sur le respect d'une séquentialité strictement temporelle.

## **1.2. Données et prises de données**

Tant pour l'acquisition que pour la psycholinguistique de l'adulte, j'ai utilisé et confronté, en fonction des problèmes que j'envisageais, des données de diverses natures (perception / production, données spontanées et induites, orales et écrites) et retenu par conséquent plusieurs techniques de prise et d'analyse des données.

En production, pour les travaux d'acquisition, j'ai recueilli un corpus spontané longitudinal permettant de suivre l'évolution d'un enfant (C.) entre 1,9 et 2,3 ans. La plupart des travaux ont été réalisés sur la base de ce corpus segmenté puis reporté, stocké et analysé dans la base de données Childphon (Rose, 2003). J'ai travaillé à partir de 2001 sur plusieurs versions préliminaires de cette base de données qui a depuis été considérablement développée (Rose et al., 2005).

La plupart des hypothèses d'acquisition ont été élaborées par ailleurs à partir de tests de dénomination d'images (pour les travaux sur la liaison ainsi que pour les travaux sur les troncations / réductions) réalisés à l'Ecole Decroly à St Mandé, à l'école St Elisabeth de Nantes, mais elles ont été ensuite re-testées et réétudiées à partir de ce corpus longitudinal qui a permis d'inscrire les hypothèses dans une continuité développementale.

Elles ont été également examinées à la lumière de données complémentaires de corpus dirigés ou semi-dirigés de production de bébés (à partir de livres et de jouets) enregistrés en crèche à Nantes lors d'études transversales (enregistrement de plusieurs bébés à la même tranche d'âge).

Mes travaux en dysphasie ont testé, les mêmes hypothèses théoriques (sur la liaison et la troncation), en contrepoint des données d'enfants non dysphasiques. Les mêmes dispositifs expérimentaux (dénomination d'images) ont donc été utilisés. Ces enregistrements ont été réalisés dans le cadre du projet de recherche concernant la dysphasie de développement (CNRS/Paris 5 « Axes d'excellence 1995 », GIS, Sciences de la cognition), sous la direction de Célia Jakubowicz. J'avais, dans ce projet, la responsabilité de l'axe « phonologie ». Les enfants ont été enregistrés chez eux, puis à l'hôpital Robert Debré à Paris.

Pour tous ces travaux, j'ai réalisé des évaluations qualitatives des productions d'enfants (jugement de grammaticalité, dénomination d'images), présentées généralement en pourcentage ou histogrammes.

Pour les recherches en psycholinguistique de l'adulte, j'ai travaillé exclusivement sur la base de données de perception en ayant recours à des tests chronométriques de comportements perceptifs de sujets pour des tâches comme la détection de phonèmes ou la reconnaissance de mots. Ces données ont été obtenues au laboratoire de psychologie expérimentale Paris V René Descartes dirigé à cette époque par Juan Segui. Dans ce cas, les données ont été évaluées par des tests statistiques.

### 1.3. Position théorique

Si je me suis appliquée à travailler systématiquement à partir de corpus larges, contrôlés, tant en perception qu'en production, en m'appuyant sur des dispositifs expérimentaux psycholinguistiques et en ne me refusant l'éclairage d'aucun type d'empirie, je ne me reconnais pas pour autant dans une approche qui s'appuierait sur une théorisation exclusivement « data-driven » ou « bottom-up » des phénomènes phonologiques.

J'ai travaillé sur la base de postulats définissant des objets théoriques n'ayant aucune réalité substantielle observable (consonnes flottantes, autosegments, lignes, squelette de positions pures) et servant de fondement à la modélisation réaliste d'une phénoménologie empirique, elle-même substantiellement réelle. De même, pour l'acquisition, je n'écarte pas *a priori* le postulat de l'existence de la « GU », du « LAD », ou de la « compétence phonologique » n'ayant aucune réalité substantielle dans la mesure où leur existence, leur réalité et la scientificité de la démarche qui consiste à les inventer pour rendre compte de données empiriques physiquement réelles ne contreviennent pas aux règles de construction, validation, falsification des modèles.

Le cadre que j'ai retenu est celui de la phonologie autosegmentale. Avec la théorie autosegmentale, Goldsmith (1976) pose en effet le principe formel de l'existence de plusieurs niveaux interprétatifs parallèles possibles, constitués d'autosegments alignés où chaque ligne est définie indépendamment des autres (segments, tons, syllabes, constituants syllabiques, traits, éléments, accents, etc). Ceci permet donc d'envisager que la linéarité du signifiant ne soit pas analysable de manière unique et uniforme, ni alignée automatiquement, comme le suppose Saussure, sur la temporalité du signal.

Chaque ligne est segmentée de manière autonome et indépendamment des autres lignes avec lesquelles elle est articulée conformément à un certain nombre de conventions et de règles. Ceci suppose donc que la mise en relation entre les représentations et le signal lors des opérations de traitement n'impose pas une seule



unité de segmentation, mais que plusieurs organisations linéaires différentes peuvent être calculées et calculables simultanément à partir du signal. Ce formalisme permet donc une mise en relation beaucoup plus subtile entre un signifiant plurilinéaire et le signal dont l'organisation temporelle peut être interprétée simultanément à partir d'unités différentes. Contrairement à ce que dit Saussure, avec un tel modèle des représentations phonologiques, le caractère temporel de la chaîne acoustique n'impose pas qu'il y ait unidimensionnalité de l'information et que les différences ne puissent être exprimées que séquentiellement. Pour reprendre sa métaphore, ce formalisme donne la possibilité d'un découpage « aux ciseaux » par autant de paires qu'on a de lignes parallèles.

Pour mes travaux sur le traitement de la liaison chez l'adulte, la représentation plurilinéaire des unités lexicales et la formulation de l'existence d'un squelette de positions pures comme matrice temporelle des représentations lexicales autosegmentales supposant l'absence de rattachement systématique des segments flottants (Encrevé, 1988, Encrevé & Scheer, 2005) m'ont donné les outils conceptuels qui m'étaient nécessaires pour interroger le rapport entre temporalité du signal et linéarité du signifiant et rendre compte des données psycholinguistiques testant la segmentation de la chaîne parlée en parole continue.

Pour les travaux en acquisition j'ai également travaillé principalement à partir de formalismes autosegmentaux. La logique formelle consistant à autonomiser les niveaux de représentations et à les faire fonctionner dans des plans différents construit une géométrie qui offre une explication extrêmement efficace et élégante de nombreux phénomènes caractérisant les productions enfantines qui ont pu être envisagées comme l'expression de représentations sous-spécifiées : métathèses, harmonies consonantiques, simplifications de l'inventaire phonémique, acquisition des structures syllabiques. Elle permet en particulier de mettre au jour l'existence de logiques structurelles identiques dans des productions en apparence très variables ou sans relation entre elles. En cela, ces modèles remplissent les conditions qu'on est en droit d'attendre d'un modèle digne de ce nom : conditions de simplification et d'abstraction de l'empirie, d'adéquation explicative par rapport aux données considérées dans leur complexité phénoménologique (possibilité d'établir des

généralisations satisfaisantes en particulier), capacité de prédiction des observables possibles à partir des données.

Je pense ensuite que les possibilités formelles qu'offre la théorie autosegmentale même si elles ont été utilisées dans quelques travaux (Menn, 1978 Macken, 1992, 1995), ont été insuffisamment exploitées dans le domaine des recherches en acquisition et en psycholinguistique de l'adulte pour l'accès au lexique. Le déferlement du cadre otiste dans le champ de la psycholinguistique ainsi que le ralliement -grégaire, massif quoiqu'intellectuellement peu convaincant- à la théorie de l'Optimalité (OT par la suite), des phonologues générativistes travaillant en acquisition, a accéléré l'abandon du cadre autosegmental et des possibilités formelles qu'il comportait, avant qu'il ait pu, selon moi, « révéler » autant que nécessaire les données et les hypothèses de traitement et d'acquisition phonologique, notamment pour ce qui concerne des langues à morphologie non concaténative.

J'ai pour ma part délibérément écarté le cadre théorique proposé par OT pourtant quasiment unanimement retenu en acquisition. Je ne l'ai utilisé ni dans mon travail en psycholinguistique de l'adulte, ni dans mes recherches en acquisition pour les raisons suivantes. Il me semble que la variété des propositions possibles qui ont été faites ou qui sont faites dans le cadre de la théorie de l'Optimalité en font d'abord une méta-théorie essentiellement pourvoyeuse d'une terminologie et d'un formalisme homogènes qui permettent de décrire de manière unifiée de très nombreux phénomènes empiriques, mais la question du réalisme psychologique de la logique générale du cadre OT ne me paraît pas pertinente en soi<sup>3</sup>.

Le problème du réalisme cognitif ne peut être abordé dans OT que par une discussion portant sur la nature et la motivation des contraintes. Certaines versions d'OT définissent des contraintes sur une base perceptive (Boersma, 1999), d'autres sur une base articulatoire ou sur une base psychologique spécifiquement linguistique ou non et à partir d'arguments bottom-up (Hayes et al., 2004, Steriade, 2004). Si l'on considère que le fondement substantiel d'une contrainte phonologique est la garantie

---

<sup>3</sup> Même si la genèse du modèle OT s'articule à ces questions de réalisme, notamment par la continuité théorique qui s'établit entre la Phonologie Naturelle et OT (cf. Scheer, 2004b, Lyche, 2005).

absolue de son réalisme psychologique, dans ce cas, on peut effectivement considérer que la question du réalisme linguistique est envisagée dans ce modèle, qui peut tolérer la simultanéité de contraintes variées et diversement motivées<sup>4</sup>. Mais on pourrait arguer tout aussi bien qu'une contrainte visant à maximaliser les attaques syllabiques est réaliste en ce qu'elle exprime une logique de marque universelle, vérifiée dans tous les inventaires phonologiques connus et reflète les régularités cognitives à l'origine de régularités structurelles présentes dans toute langue humaine. Ceci signifie que, fondamentalement, l'architecture du cadre OT en soi ne nous dit rien quant à la logique acquisitionnelle, et ne me semble pas proposer les moyens d'une validité cognitive supérieure. De même, la logique compétitive organisant la hiérarchie de contraintes ne me paraît pas non plus apporter une meilleure modélisation de quelque procédure psychologique que ce soit.

On pourrait avancer que le modèle OT ayant recours à des réorganisations de contraintes permettant de générer l'output attendu, offre un cadre théorique qui peut rendre compte d'une logique développementale et donc de l'évolution des formes progressivement produites par les enfants (Bernhardt & Stemberger, 1998). Mais la puissance du modèle contraint si faiblement la variation et l'évolution des formes que, selon moi, elle les décrit plus qu'elle ne les explique. Chaque stade d'acquisition correspond à une hiérarchie différente, dont l'existence par rapport à la hiérarchie correspondant au stade d'acquisition précédent, n'est au final intrinsèquement motivée que par elle-même. On peut alors craindre que la succession des stades et les stratégies d'acquisition retenues et reflétées par les hiérarchies de contraintes successives ne soient empiriquement infondées et générées par la logique même du modèle OT. Des débats nombreux ont régulièrement lieu à cet égard pour évaluer le fondement et la motivation phonétique, articulatoire, psycholinguistique des contraintes proposées (Lyche, 2005) mais je pense que la valeur explicative et prédictive des propositions faites dans un modèle aussi peu contraint sont rarement satisfaisantes. Du moins ne m'ont-elles jamais convaincue suffisamment pour que j'ai envie de m'y arrêter.

---

<sup>4</sup> On pourrait même envisager, en caricaturant à peine, des contraintes d'origine maturationnelle sans aucune spécificité linguistique comme « poussée dentaire » ou « indépendance du voile ».

## **2. Organisation phonologique et traitement de la parole continue : liaison et accès au lexique**

### **2.1. Le problème : segmentation de la parole continue et accès lexical**

#### 2.1.1. Contraintes de la parole continue

La théorisation des problèmes de segmentation de la chaîne parlée est plutôt récente en psycholinguistique. Les premiers travaux en reconnaissance lexicale (Morton, 1969, Forster, 1976) ont en effet concerné la reconnaissance des mots écrits et ont, dans un premier temps, importé dans les études sur la parole, les hypothèses de segmentation et de traitement formulées pour l'écrit où la définition de l'unité de traitement n'est pas en soi problématique, puisque visuellement, les frontières lexicales sont marquées par des blancs entre les mots. La nécessité qui se fera jour ensuite (Cole & Jakimik, 1980, Marslen-Wilson, 1987, Marslen-Wilson & Tyler, 1980, Tyler & Frauenfelder, 1987) de théoriser le rapport entre linéarité du signifiant et temporalité du signal n'était pas apparue. Les hypothèses retenues pour la parole ont donc légitimement considéré, dans un premier temps que l'on pouvait retrouver les mots dans l'entrée parlée à partir du découpage « aux ciseaux » préconisé par Saussure.

Les travaux consacrés spécifiquement au traitement de l'entrée parlée (Cole & Jakimik, 1980, Marslen-Wilson, 1987, Marslen-Wilson & Tyler, 1980, Tyler & Frauenfelder, 1987) ont néanmoins pointé de manière très perspicace les apories d'une telle conception et montré les limites théoriques d'une hypothèse rabattant les questions de l'accès lexical à l'oral sur celles de l'écrit (Bradley & Forster, 1987).

Pour des raisons anatomiques, la parole consiste en une succession de gestes articulatoires organisés au sein d'une enveloppe prosodique, et produits séquentiellement. L'information acoustique n'est donc physiquement disponible que pendant une durée très courte, séquentiellement délivrée et ne peut plus être récupérée en cas de traitement erroné. Recevant cette information acoustique continue et linéaire, l'auditeur doit en extraire des unités significatives et distinctives, et

contrairement à ce que semble considérer Saussure, ceci ne suppose pas obligatoirement qu'il utilise un signifiant unilinéarisé et temporellement contraint.

L'accès lexical repose sur l'hypothèse de l'existence d'un « lexique mental » stockant les mots connus par le sujet. L'accès lui-même consiste en un appariement entre les représentations stockées dans le lexique interne et les unités dégagées de l'analyse sensorielle du signal. Le phénomène est généralement modélisé en trois étapes (Frauenfelder & Tyler, 1987) : i) la recherche de l'information pertinente dans le signal et la segmentation, ii) l'activation des mots possibles à partir de l'information sensorielle iii) la sélection du mot recherché dans le lexique mental et la résolution finale des ambiguïtés de segmentation. Ces trois étapes sont, bien entendu, respectivement modélisées en fonction des deux autres. Le point principal opposant les modélisations de l'accès lexical, à l'époque où je traitais de ce sujet, concernait la manière dont ils définissaient l'information sensorielle et l'unité ou les unités utiles et utilisables en temps réel par l'auditeur. Deux grands types de réponse étaient proposés opposant modèles lexicaux (où l'accès se fait à partir d'une unité « mot ») et modèles prélexicaux (où l'accès se fait à partir d'une unité infra-lexicale, comme la syllabe, le pied ou la more).

### 2.1.2. Modèles lexicaux et séquentiels : aligner lexique et signal

Les modèles lexicaux de première génération (Cole & Jakimik, 1980) dont le plus célèbre est très certainement « Cohort » (Marslen-Wilson & Welsh, 1978, Warren & Marslen-Wilson, 1987, 1988, Lahiri & Marslen-Wilson, 1991) postulent d'une part que les mots sont reconnus séquentiellement dans l'ordre où ils se présentent, quelles que soient les relations syntagmatiques qui les unissent. Par ailleurs, ils considèrent que la frontière gauche d'un mot est déductible de la fin du mot qui précède. La segmentation est donc foncièrement conditionnée par l'encodage exact de l'information correspondant au début de mot et réalisée à partir d'une représentation phonétique (Cohort I) ou phonologique sous-spécifiée du début de mot (Cohort II), construite à partir de l'information sensorielle disponible dans le signal et incluant les variantes libres et complémentaires. Le traitement envisagé est donc très

strictement séquentiel et linéaire. Il propose un algorithme rapide et non contraint par les caractéristiques de la langue cible. Dès lors, un modèle très strictement aligné sur la frontière gauche de mot postule *a priori* un traitement très strictement linéaire du signal, ne correspondant pas aux effets de coarticulation observables. Par ailleurs il a été montré que le traitement perceptif en soi n'était pas strictement linéaire et que seuls 40% des mots du lexique pourraient en contexte être reconnus sans ambiguïté avant le commencement du mot suivant (Luce, 1986). Enfin, l'énoncé peut comporter des chaînes offrant plusieurs possibilités de segmentation. Une suite comme [krokodil] peut ainsi être segmentée « crocodile », « croque Odile » etc. Pour un auditeur ces suites ne sont pas ambiguës puisque l'information acoustique (et en particulier l'information supra segmentale) permet une compréhension absolument claire de telles suites. Mais un modèle strictement aligné sur l'information segmentale de début de mot ne peut pas en rendre compte.

Pour remédier à ces problèmes une autre génération de modèles lexicaux d'inspiration connexionniste « Cohort II », « Trace » (Mc Clelland & Elman, 1986) postule qu'il n'y a pas d'alignement strict entre le signal et les unités lexicales. Ces modèles permettent le maintien simultané de plusieurs hypothèses de segmentation dont la résolution est obtenue par une dynamique d'activation et d'inhibition d'informations (lexicales, morphologiques, syntaxiques, phonologiques etc) distribuées dans le système. Toutes les segmentations possibles correspondant à l'entrée sensorielle sont susceptibles d'être activées. Ainsi pour une suite comme [krokodil], « croc », « crocodile », « crocodilien » « croque Odile » « croque ode ile » « croque eau d'île » etc etc vont être activées. Seule subsistera l'unité dont le taux d'activation permet l'émergence au détriment des autres. Sa reconnaissance impose les frontières qui le séparent du mot précédent et du mot suivant et qui permettent d'activer ou d'inhiber les candidats activés à la suite. La segmentation est donc en quelque sorte résiduelle.

Ces divers modèles présentent des qualités de simplicité et de rapidité qui s'accordent bien aux temps de réaction mesurés expérimentalement ainsi qu'à la vitesse de la compréhension en temps réel d'un énoncé oral. Ils sont néanmoins problématiques en ce qu'ils assimilent complètement la contradiction lisible dans le

texte saussurien présenté ci-dessus et reproduisent la conception du « mot » préconisée par les premiers modèles par analogie avec le traitement visuel.

Dans tous les cas, la conception linéaire de l'accès semble résulter non pas d'une observation directe des caractéristiques acoustiques et temporelles du signal mais d'une projection dans le signal de représentations linéarisées, d'un signifiant « temporalisé » tel que le préconise Saussure et confondant implicitement forme et substance. Car, si le signal peut se lire comme une succession d'événements acoustiques reflétant une succession de gestes articulatoires, on ne trouve pas dans le signal de corrélats acoustiques clairs permettant de définir une unité telle que le mot ou le phonème. On retrouve pourtant dans tous ces modèles une représentation du mot comme une suite de phonèmes linéarisés dont l'activation isolée ou l'activation / inhibition qu'ils subissent en réseau permettent la reconnaissance lexicale. Ils supposent que, pour le locuteur, le signal est toujours lexicalement interprété et bien qu'il soit envisagé explicitement que les représentations contenues dans le lexique mental sont multidimensionnelles, ces modèles projetent dans le signal une transcription phonétique grossière des mots, avec un seul niveau de représentation phonématique unilinéaire. Ces unités se résument *in fine* à une représentation orthographique plus ou moins bien adaptée à la modalité auditive.

Cette approche définit en outre une unité « mot » qui est en soi linguistiquement controversée et non universelle et dont on ne voit pas clairement à quoi elle pourrait correspondre sinon à une conception alphabétique simplifiée, qui est la moins susceptible de correspondre à une unité universellement utilisée.

Enfin, cette conception omet l'existence de l'information supra-segmentale dont il a pourtant été montré qu'elle est l'un des facteurs majeurs d'organisation temporelle de la parole et qu'elle joue un rôle décisif dans les procédures de segmentation chez l'adulte et le bébé (Jusczyk et al., 1992).

En conclusion, pour ces modèles, la segmentation est une question résiduelle de l'accès lexical, aucune procédure de segmentation n'est clairement formulée avant

l'accès. Le « découpage » du signal résulte de la reconnaissance des mots, eux-mêmes définis *in fine* comme des représentations orthographiques.

### 2.1.3. Modèles prélexicaux : trouver l'unité d'accès au mot dans le signal

Les modèles prélexicaux, au contraire, ne présupposent pas que le signal est perçu immédiatement comme une forme interprétée et formulent explicitement des hypothèses de segmentation. Ils définissent par conséquent l'information utilisable par l'auditeur pour opérer un « découpage » du signal préalable à l'accès et permettant de retrouver optimalement les unités lexicales à partir d'une information suprasegmentale ou d'une unité infralexical. Cette modélisation du traitement postule en effet de manière implicite qu'il existe dans le signal des informations acoustiques, phonétiques correspondant à des frontières lexicales ou des indices permettant à l'auditeur de développer des inférences sur la place éventuelle d'une frontière. De nombreuses recherches ont tenté de mettre en évidence l'existence d'informations spécifiques permettant d'opérer la segmentation préalable à l'accès. Des indices acoustiques tels que la variation et les reset de F0 (Vaissière, 1995), l'allongement (Oller, 1973, Christophe, 1993) ou des informations prosodiques telles que l'accent (Cutler & Carter, 1987) mais aucun de ces indices ne peut être considéré, de manière isolée comme un marqueur de frontière lexicale stable et fiable.

Les pauses physiologiques ont également été étudiées, mais s'il a été montré qu'elles sont utilisées universellement comme marqueur de frontière (Vaissière, 1983), elles permettent de dégager des constituants par ailleurs prosodiquement conditionnés (par l'accent, les allongements etc), correspondant à des syntagmes plutôt qu'à des mots.

Une autre conception de l'accès prélexical a été envisagée, supposant que la segmentation s'appuie sur l'unité métrique de référence de la langue considérée. Cutler et al. (1986, 1992) ont en effet proposé à partir de travaux réalisés sur l'anglais (langue où le mètre est le pied et où il existe un accent lexical) et le français (langue où le mètre est la syllabe et où il existe un accent de groupe) que les auditeurs segmentent en utilisant une routine mise en place lors de l'acquisition du langage et



adaptée à l'unité métrique de leur langue. Ceci les amène à proposer un algorithme de segmentation en syllabes pour les francophones, en pieds trochaïques pour les anglophones, en morae pour les locuteurs de japonais. Cette proposition repose sur une conception en « Principes et paramètres » : les enfants seraient dotés d'un principe universel de segmentation prélexical leur permettant de découvrir les unités de traitement pertinentes dans leur langue. Ce principe serait ensuite paramétrisé par le contact avec l'input délivrant l'unité métrique utile.

Pour le français, ceci suppose donc une routine strictement ascendante qui outre les problèmes qu'elle pose en acquisition relève d'une analogie post-hoc qui reconduit les apories d'une conception linéaire de la segmentation, le signal n'étant plus envisagé comme la projection d'un signifiant linéarisé en unités lexicales mais comme une suite de syllabes que l'on pourrait découper « aux ciseaux ». L'hypothèse reste la même, seules changent la taille et la nature de l'unité segmentée.

Les travaux de Mehler et al. (1981) sur le rôle de la syllabe dans l'organisation des chaînes et la reconnaissance des mots qui sont à l'origine du développement de l'hypothèse métrique, ont montré des effets syllabiques à *l'intérieur* des mots (« pa » ou « pal » dans « palace » par exemple) présentés isolément mais qui n'ont pas été validés pour des énoncés de parole continue. Les auteurs (et toutes les autres publications qui ont suivi sur ce sujet) ont, me semble-t-il, montré que les sujets qu'ils avaient testés étaient sensibles à la structure métrique des mots de leur langue et ont mis en évidence que, pour un francophone, la syllabe est une unité de structuration, d'organisation à l'intérieur du mot. Rien ne permet d'affirmer pour autant que la syllabe est l'unité de segmentation et d'accès lexical qu'utilisent les francophones. Des recherches postérieures ont d'ailleurs montré que si la syllabe pouvait aider la segmentation en français, l'hypothèse trop simple d'une stratégie d'alignement stricte et systématique entre frontières de mots et frontières de syllabes devait être rejetée (Frauenfelder & Content, 1999). Content & al. (2001) ont également mis en évidence plus récemment que les constituants syllabiques (attaque *vs* coda) étaient utilisés différemment par les auditeurs francophones pour la segmentation du signal lors des procédures de traitement. Ceci indique que la syllabe ne peut pas être envisagée

comme un « bloc » compact d'accès au lexique que les locuteurs utiliseraient de manière séquentielle en l'alignant sur les entrées lexicales.

Enfin, outre que ces expériences supposent une conception plate et linéaire de la syllabe qui est phonétiquement et phonologiquement définie de manière peu claire (de quelle syllabe parle-t-on exactement et à quel niveau de représentation de l'objet travaille-t-on : syllabe articulatoire, syllabe perceptive, syllabe phonologique ?), l'hypothèse de segmentation qui en découle n'est pas tenable au regard des phénomènes de liaison enchaînée et d'enchaînement. Si les francophones développaient une stratégie de segmentation alignée sur les frontières syllabiques, ils segmenteraient systématiquement de manière erronée les suites contenant une liaison ou un enchaînement (Wauquier-Gravelines, 1996, 1999). Dans les modèles précités, qu'ils soient prélexicaux ou lexicaux la segmentation de « une égérie » en \*« une neige et rit » aurait un taux d'activation au moins équivalent à celui de la segmentation attendue car « neige » est un mot possible et un début de syllabe possible correspondant à un mot possible.

Or même intuitivement, sur la seule base de la pratique de sa langue, un locuteur-auditeur du français peut soupçonner que ce n'est pas le cas.

## 2.2. La liaison et l'accès au lexique

### 2.2.1. La liaison, phénoménologie

La liaison est un phénomène de sandhi externe se produisant sur la frontière gauche des catégories lexicales majeures en français et qui donne lieu à un double phénomène : quand deux voyelles sont en contact sur une frontière lexicale

i) une consonne latente peut apparaître

[1] "un enfant" : "un" [ɛ̃] + "enfant" [ɑ̃fɑ̃] est prononcé [ɛ̃nɑ̃fɑ̃]

ii) cette consonne latente sera le plus souvent-mais pas obligatoirement (Encrevé 1988)- resyllabée à l'attaque du second mot

[2] [ɛ̃nɑ̃fɑ̃] est toujours syllabé [ɛ̃/nɑ̃/fɑ̃]

[3] [desɔldazɑ̃glɛ] « des soldats anglais »,

peut être syllabé [de/sɔl/daz/ã/glə] ou [de/sɔl/da/zã/glə]

La réalisation du sandhi est cependant conditionnelle puisque dans certains cas la réalisation de la liaison sera

i) obligatoire et toujours enchaînée

[4] [ɛ̃nãfũ] « un enfant », [ilzɔ̃] « ils ont », [sɔ̃til] « sont-ils »

ii) facultative et diversement réalisée (avec enchaînement, non-enchaînement, coups de glotte, schwa, Encrevé, 1988, Encrevé & Scheer, 2005)

[5] [desɔldazãglə] / [desɔldaãglə] « des soldats anglais »,

[6] [vuzetzẽvite] / [vuzetẽvite] / [vuzet ʔẽvite] / [vuzetə zẽvite]

« Vous êtes invités »

iii) impossible

[7] \*[ɛ̃sɔldatãglə] « un soldat anglais »

[8] \*[dezeRo] « des héros »

Les consonnes utilisées sont les suivantes {n, z, t, R, p, g}

### 2.2.2. Trouver les mots en contexte de liaison

Dans les modèles présentés plus haut, on considérera que le locuteur utilise en temps réel une représentation phonologique. Le traitement va nécessairement comporter l'action d'une ou plusieurs composantes phonologiques opérant un filtrage du signal, visant à convertir l'information acoustique extraite du signal, en unités de plus en plus abstraites, et aboutissant ultimement à l'unité lexicale sous-jacente. Mais comme je l'ai montré dans mes travaux, tous les modèles qu'ils soient lexicaux ou prélexicaux font des prédictions qui échouent à rendre compte de la facilité et de la rapidité de traitement de la liaison en parole continue.

Le phénomène de liaison a ceci de particulier qu'il va poser simultanément un problème de segmentation puisqu'il insère un segment sur la frontière gauche du mot et un phénomène d'allomorphie puisque les mots peuvent être réalisés avec une alternance consonantique à l'attaque.

Si l'on considère le mini-corpus suivant :

[9] « un zèbre » : [ẽzɛbr] => [ẽ] + [zɛbr]

[10] « un hérisson » : [ẽerisɔ̃] => [ẽ] + [erisɔ̃]

[11] « un éléphant » : [ẽnelefã] => [ẽ] [nelefã]

[12] [ẽnelefã] vs [dezelefã] vs [ẽzɛbr] vs [dezebr]

les hypothèses de traitement réalisées uniquement à partir des formes de surface, amènent nécessairement à postuler soit l'allomorphie des déterminants : [ẽ] vs [ẽn] ou [de] vs [dez] soit l'allomorphie des noms [nelefã] [zelefã] [elefã] vs [zɛbr] / \*[ɛbr] dont il faut modéliser la dimension psycholinguistique dans le modèle d'accès.

Dans les modèles lexicaux (Trace, Cohort I et II), les représentations phonématiques unilinéaires retenues ne i) permettent jamais de dégager une unité lexicale pourvue d'une consonne flottante ii) n'apparaissant que dans certains contextes, puisque la segmentation est consécutive à la reconnaissance opérée mot par mot sur une base phonématique, et hors de toute contrainte contextuelle.

Si l'on prend par exemple une entrée comme [lɔpətiami], tous les modèles lexicaux activeront /lɔ/ puis /pəti/ « petit » ou /pətit/ « petite » puis /ami/ « ami » ou /tami/ « tamis », les segmentations consécutives à la reconnaissance des mots sont « le petit tamis » ou « \*le petite ami » et ne permettent pas de récupérer « le petit ami » sans qu'il y ait des procédures de traitement rétroactives. Aucune suite contenant une liaison enchaînée ne peut être traitée adéquatement si l'on n'admet pas qu'il faut que le modèle inclue i) une représentation lexicale de l'adjectif masculin<sup>5</sup> /pəti + t/ permettant de savoir que l'adjectif au masculin peut se réaliser soit avec un / t / soit sans / t / donc une formalisation de la consonne flottante ii) la possibilité de savoir quand ce /t/ flottant se réalise ou non.

Pour ce qui concerne la deuxième famille de modèles, reposant sur une segmentation prélexicale, l'alignement supposé des débuts de mots sur les frontières syllabiques (pour le cas de l'hypothèse métrique rapportée au français) crée des

<sup>5</sup> On exclut ici l'hypothèse selon laquelle le « t » de liaison serait le même que le « t » morphologique du féminin. Mais même si on retenait cette hypothèse, on aurait besoin alors d'une forme unique /pəti + t/ et de contraintes explicitant les cas de réalisation du /t/ en tant que « t » de liaison ou « t » du féminin.

segmentations obligatoirement fausses, puisque la CL donne dans le signal une information sensorielle pour les frontières de mot / frontières de syllabe ne correspondant pas aux frontières lexicales et que les mots monosyllabiques constituent souvent des formes enchâssées dans des mots plus longs. Pour des suites comme [ẽnelefã] [dezelefã], toutes les segmentations suivantes seront réalisées « un nez les fend » « \*des haies les fend », « des zélés faons », « \*un nélé faon » avant la récupération de la forme canonique du mot de manière rétroactive. L'auditeur doit non seulement savoir que « un » (ou « éléphant » si l'on considère que la consonne est attachée au deuxième mot) ont des formes allomorphes, mais il doit également avoir appris que la découverte de ces formes allomorphes est perturbée par l'existence d'une resyllabation qui peut entraîner une interprétation erronée du début de mot ou de la syllabe de début de mot.

Tous les modèles d'accès au lexique retenant au final une conception linéaire et alphabétique de la chaîne parlée envisageaient peu ou prou le traitement comme un découpage « aux ciseaux ». Ils prédisaient donc, pour le traitement de la liaison, un coût de traitement consécutif à une ambiguïté d'établissement de frontière et supposaient plus ou moins implicitement l'existence de procédures rétroactives permettant une récupération des mots à initiale vocalique dans un contexte de liaison. Or, ceci ne correspond pas à la facilité avec laquelle les locuteurs-auditeurs de français produisent et traitent les liaisons en parole continue.

### **2.3. Expériences et résultats**

Aucune recherche n'avait été menée en accès au lexique chez des francophones sur des phénomènes phonologiques masquant la frontière gauche de mot comme la liaison, alors qu'il en existe un certain nombre (Lowenstamm, 1999) notamment dans les langues celtiques (phénomène dit de « mutation »). Matter (1986), Dejean de la Bâtie (1993), Bradley & Dejean de la Bâtie (1990) ont, dans des travaux trop rarement cités, été les premiers à soulever cette question en tentant d'expliquer la difficulté tenace de traitement de la liaison (en lecture et en production de parole spontanée) chez des apprenants de français langue seconde, même de

niveau avancé. En observant les difficultés de ces apprenants de français langue seconde, ils avaient ainsi confirmé les prédictions des modèles d'accès lexical existants.

Mais personne n'avait mis en évidence le fait que tous les modèles d'accès au lexique disponibles à cette époque sans exception faisaient des prédictions supposant un coût cognitif supplémentaire pour les francophones de langue maternelle, qui ne manifestent pourtant pas de difficulté de traitement de la liaison. C'est la question à laquelle je me suis attachée dans mes travaux. Deux séries d'expériences ont été réalisées.

### 2.3.1. Coût cognitif à la reconnaissance de mot

Mon premier souci a été d'observer l'existence de cet éventuel coût de traitement lors de la récupération d'un mot à initiale vocalique lors de l'accès au lexique dans un contexte de liaison enchaînée. J'ai donc proposé une expérience de reconnaissance de mot en contexte de liaison vs contexte de non liaison (« un petit éléphant » vs « un joli éléphant ») sur la base d'une tâche simple de reconnaissance avec mesure de temps de réaction. Les résultats (aucune différence significative des temps de réaction pour les deux types de contexte présentés) montrent que, contrairement aux prédictions des modèles lexicaux et prélexicaux, la présence de la consonne sur la frontière ne gêne absolument pas l'accès au lexique et qu'il n'y a donc aucun « coût » cognitif particulier. Ces résultats ont ensuite été répliqués (Gaskell et al. 2002, Spinelli et al. 2003). Tous ces travaux ont montré en ayant recours à des tâches diverses (phoneme-monitoring, word-monitoring, cross-modal-priming) que la liaison n'entravait aucunement l'accès au lexique des francophones, mais au contraire, qu'elle pouvait même être un élément facilitateur de l'accès (Spinelli et al. 2003). La constance des résultats malgré la variété des tâches utilisées permet d'écarter une interprétation liée aux conditions expérimentales et plaide en faveur d'une robustesse d'effet correspondant à ce qui est observable en parole naturelle : la liaison n'empêche pas l'accès lexical chez des locuteurs-auditeurs francophones.

### 2.3.2. La consonne de liaison est-elle une consonne comme les autres ?

Dans une deuxième série d'expériences, j'ai testé la sensibilité des sujets à la consonne de liaison (CL).

J'ai utilisé une tâche de détection de phonème généralisée (les sujets devaient réagir dès qu'ils entendaient le phonème-cible, n'importe où dans une phrase), dans quatre expériences en contexte obligatoire et constaté qu'en l'absence d'une consigne indiquant clairement aux sujets de détecter aussi les CL, elles étaient régulièrement non perçues par les sujets testés.

Pour les expériences 1 et 2, les facteurs manipulés étaient les suivants : nature de la consonne (*/n/ un navire/ un avion* vs */t/ petit éléphant / petit téléphone*), nature du lien syntaxique entre les deux mots entrant en liaison (clitique vs adjectif antéposé) et nature du constituant syntaxico-prosodique où se réalise la liaison. La mesure retenue était le pourcentage de non-détection. Le graphique ci-dessous montre cette difficulté de détection spontanée de la CL.

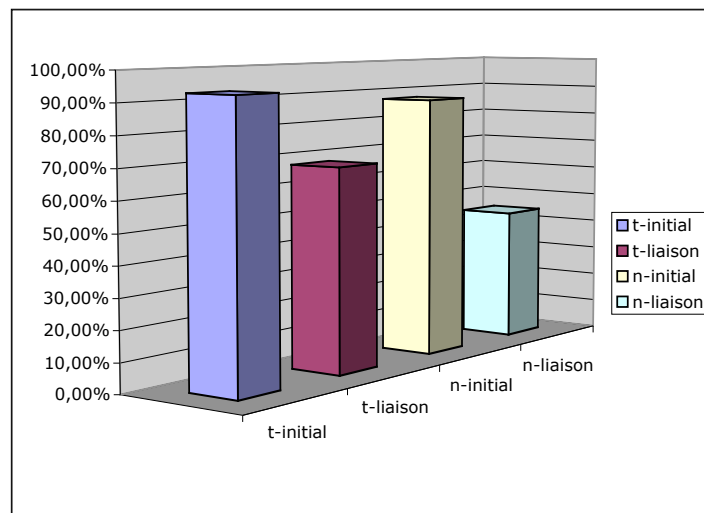


Figure 1 : % de détection spontanée de la CL ([n] [t]) en groupe clitique et en syntagme phonologique

Pour les expériences 3 et 4, les mêmes facteurs étaient manipulés, mais à l'issue d'un entraînement au cours duquel la non-détection de la CL était constatée, on précisait aux sujets que la CL devait aussi être détectée.

La mesure retenue était le pourcentage de non-détection et le temps de réaction. Malgré ce rappel, on constate un maintien quand même important de cette

absence de détection de la CL. Le graphique ci-dessous montre la persistance de cet effet de non-détection malgré la consigne demandant explicitement de détecter la consonne flottante.

Pour t, la différence de % de détection n'est pas significative, pour ces deux expériences, mais les temps de réaction le sont ( $t_{\text{subj}}(1, 23) = 2,89$   $p=0.0009$ ).

	/t/ liais	/t/ init	différence
TR	619 ms	573 ms	46 ms

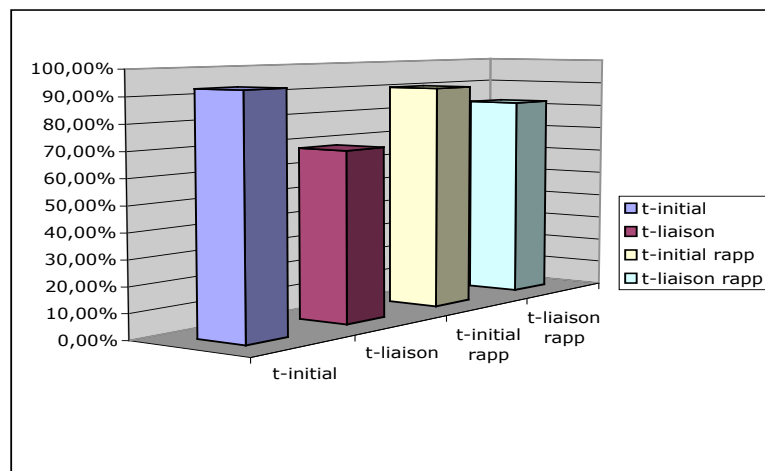


Figure 2 : % détection spontanée et de détection après rappel de la consonne [t] en syntagme phonologique (différence non significative)

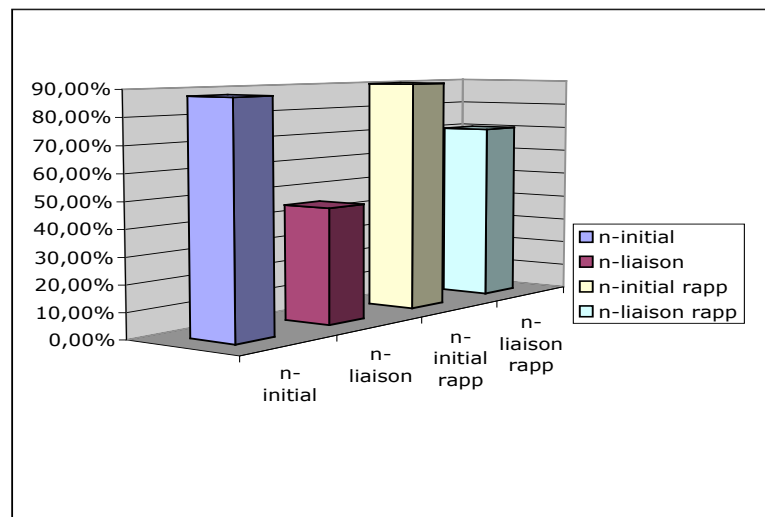


Figure 3 : % de détection spontanée et de détection après rappel de la consonne [n] en groupe clitique



On constate donc un effet de non-détection de la CL pour le /n/ et un TR plus long pour le /t/ (pouvant résulter soit de la nature même de la consonne, soit de la nature du lien syntaxique entre le mot 1 (clitique *vs* adjectif antéposé) et mot 2 (nom)) qui se maintient même si les sujets ont reçu une consigne explicite leur signalant la nécessité de détecter aussi les consonnes de liaison.

Plusieurs explications peuvent être envisagées pour rendre compte de ce phénomène.

## 2.4. Analyses

L'analyse possible de ces résultats montrant simultanément une « surdité » à la CL et une absence totale de difficulté de l'accès lexical en contexte de liaison appelle deux interprétations non nécessairement exclusives l'une de l'autre mettant en cause soit l'éventuelle spécificité acoustique de la CL, soit sa nature phonologique propre.

### 2.4.1. Rôle de l'information montante

Il a été longtemps admis par la tradition phonétique française que les consonnes de liaison (CL pour la suite) n'ont pas de spécificité acoustique. Grammont (1914), Passy (1917), Nyrop (1925) ont tous soutenu l'absence de particularité de la CL. Delattre (1947), à la suite des travaux expérimentaux de l'Abbé Rousselot (1924) et en accord avec les propositions de Grammont (1933) concernant la syllabation en français a néanmoins considéré que la place de la consonne dans une structure syllabique et à l'intérieur d'un cluster (dans les deux cas il oppose attaque *vs* coda) entraîne une réalisation différente. La consonne d'attaque sera produite en mode tendu avec une tension musculaire croissante alors que la consonne de coda suivra un mode décroissant. Il prédit donc que les consonnes d'enchaînement seront réalisées de manière plus relâchée que les CL enchaînées. [t] dans « petite orange » sera plus relâché que [t] resyllabé dans « petit ami », ce qui sera effectivement vérifié par les mesures de Durand (1936).

Fougeron<sup>6</sup> (2004) a également mesuré de manière beaucoup plus sophistiquée la spécificité acoustique des consonnes d'enchaînement les différenciant des consonnes initiales (CI). Elle a également montré l'utilisation de cette différence en perception. Elle prédit en conséquence une absence de coût de traitement occasionné par les consonnes d'enchaînement resyllabées lors de l'accès lexical.

Delattre (1947) ne fait pas de prédictions très précises sur une éventuelle différence entre consonne initiale (CI) et CL. L'enchaînement entraînant une resyllabation de la CL à l'attaque du deuxième mot lui fait dire qu'elles doivent être identiques mais que les CL peuvent être réalisées avec un degré de tension décroissant associé à une augmentation du degré de voisement sans qu'il explique très clairement pourquoi.

Dejean de la Bâtie (1993) a été la première à mesurer directement l'éventuelle différence de réalisation entre CL et CI. Elle a mesuré pour l'occlusive non voisée [t], l'énergie du burst, la durée d'occlusion et la durée du VOT sur des paires de non-mots (« grand irbois vs « grand tirbois ») et a montré que la durée d'occlusion et la durée de VOT sont significativement plus courtes pour la CL.

J'ai, à la suite des résultats de perception et traitement obtenus dans les expériences présentées ci-dessus, réalisé des mesures acoustiques comparant CI et CL à partir de spectrogrammes, enveloppe et courbe de F0 réalisés avec le logiciel « Unice ». J'ai fait les mesures sur les deux types de consonnes utilisées, [n] pour des comparaisons de type « un navire » « un avion » et [t] dans des comparaisons de type « petit téléphone » « petit éléphant ».

Pour l'occlusive [t] mes résultats confirment ceux de Dejean de la Bâtie : la durée d'occlusion de la CL mais surtout la durée totale de la CL (occlusion + burst+ VOT) est significativement plus courte que la durée de la consonne initiale.

---

<sup>6</sup> Je remercie au passage Cécile Fougeron qui m'a transmis la version Powerpoint de son exposé (2004) ainsi que divers documents élaborés pour des présentations orales et comportant des remarques synthétiques sur les mesures de frontières en français.

Pour la nasale [n], mes données sont beaucoup moins claires : aucun des paramètres de durée retenus, aucun paramètre fréquentiel (modification du pattern de reset de F0) ne permettent de montrer que le [n] réalisé dans une liaison enchaînée (type « un navire » vs « un avion ») est significativement différent du [n] réalisé à l'initiale de mot. Ces résultats ont été confirmés ensuite par les mesures réalisées par Gaskell et al. (2002), Spinelli et al. (2003) et uniquement pour les occlusives qui en CL sont significativement plus courtes que les CI.

De tels résultats invitent donc à une certaine prudence quant à l'hypothèse selon laquelle l'information acoustique permettrait aux auditeurs de résoudre les ambiguïtés de frontière suscitées par la liaison. Comme on l'a vu, les différences n'ont été mises en évidence de manière claire que pour les occlusives d'une part, d'autre part les différences acoustiques de durée qui ont été mises en évidence sont de l'ordre de quelques dizaines de millisecondes (entre 15 et 30 au maximum selon les auteurs).

Enfin, je n'ai, pour ma part, trouvé aucune corrélation statistique significative entre les différences de réaction aux CI et CL ([n] et [t]) et leurs différences de réalisation acoustiques et je n'ai retenu ce mode d'explication ni pour rendre compte de la « surdité » aux phonèmes de liaison, ni pour expliquer l'absence de coût cognitif à la reconnaissance de mots.

Les auteurs ayant établi l'existence d'une corrélation entre différence de réalisation et performances de traitement (Dejean de la Bâtie, 1993 ; Gaskell et al. 2002, Spinelli et al. 2003) ne considèrent pas pour autant que cette différence acoustique puisse être le facteur prépondérant permettant aux auditeurs de désambigüiser les frontières et argumentent plutôt en faveur de la convergence entre informations acoustiques et informations lexicales et phonologiques.

Le rôle des représentations lexicales est envisagé ci-dessous.

#### 2.4.2. Rôle de l'information lexicale : quelles représentations ?

Les auteurs ayant étudié les mêmes phénomènes pour les apprenants de français langue seconde (Dejean de la Bâtie, 1993) puis pour les francophones de langue maternelle (Gaskell et al. 2002, Spinelli et al. 2003) ont généralement spéculé sur le rôle de l'information lexicale descendante en comparant le traitement de la frontière créée par CI vs CL, avec des tâches réalisées avec mots (« petit téléphone » / « petit éléphant ») vs non-mots (« grand irbois » / « grand tirbois »). Leur analyse est la suivante : les sujets vont combiner en général les informations montantes ou descendantes et les utiliser alternativement selon leur disponibilité respective (par exemple usage important de l'information phonétique pour traiter « grand irbois » / « grand tirbois » puisque l'information lexicale est inexistante dans ce cas). Mais aucun des auteurs ne s'est interrogé précisément sur la structure phonologique des représentations lexicales permettant de rendre compte des résultats obtenus.

C'est l'angle de travail que j'ai retenu. Compte tenu des questions de linéarisation et de traitement qui m'intéressaient, et étant donné l'absence de travaux existant alors sur ce sujet, je me suis interrogée sur la nécessité d'inclure des représentations phonologiques dans les modèles d'accès lexicaux ainsi que sur la formalisation de la liaison permettant d'expliquer i) la rapidité et l'aisance de traitement lors de l'accès lexical dans des contextes où la liaison est réalisée et enchaînée, ii) la surdité phonologique au phonème de liaison.

Du point de vue phonologique, la liaison constitue un des cas de réalisation de l'alternance consonne/∅ en français que l'on rencontre également dans l'inflexion verbale ("il coud" [ilku] / "ils cousent" [ilkuz] ), adjectivale ("petit" [pəti] / "petite" [pətit]) ou la dérivation nominale ("saut" [so] ⇒ "sauter" [sote] ) ou adjectivale ("grand" [grɑ̃] ⇒ "grandir" [grɑ̃dir]) et qui entraîne une allomorphie des bases lexicales avec réalisation ou absence de réalisation d'une consonne flottante, dont l'apparition est conditionnée par un nombre de contraintes variées.

Pour ce qui concerne précisément la liaison, deux questions se posent qui sont intrinsèquement liées : celle de la nature de la CL (quelle analyse peut-on avoir de

l'objet phonologique « latent » que constitue le phonème de liaison, est-il purement phonologique ou relève-t-il d'une explicitation lexicale, morphologique) et celle du rattachement de la CL (est-elle ou représentée lexicalement, attachée ou non attachée et si c'est le cas sur quel mot, quelle position et comment ?). Ceci renvoie évidemment au problème de savoir comment, et à partir de quelles représentations on peut rendre compte non seulement de sa réalisation enchaînée et non enchaînée (Encrevé, 1988) mais également de sa non-réalisation. De très nombreuses propositions ont été avancées dans divers cadres phonologiques non-génératifs, génératifs linéaires, génératifs multilinéaires et, génératifs par contraintes, OT (pour une synthèse critique cf. Encrevé 1988, Paradis & El Fenne 1995, Côté, 2002).

Les très nombreuses analyses du phénomène peuvent se ramener à quatre propositions théoriques que nous exposons rapidement ci-dessous. Les trois premières ont été conçues dans les cadres génératifs *The Sound Pattern of English*<sup>7</sup> (Chomsky & Halle, 1968) et plurilinéaires. Elles ont tenté de donner avec plus ou moins de bonheur et d'adéquation empirique (Encrevé, 1988) un statut théorique à l'objet phonologique singulier qu'est la liaison.

L'hypothèse supplétive (Bybee, 2001) refait surface dans le cadre des propositions exemplaristes que l'on peut rapprocher, en psycholinguistique, des modèles épisodiques et constitue une hypothèse de supplétion lexicale qui n'envisage pas la liaison précisément comme un phénomène phonologique mais en fait une analyse lexicale.

#### 2.4.2.1. *L'effacement*

Comme le rappelle la synthèse historique et critique d'Encrevé (1988), la liaison fut longtemps envisagée d'un point de vue orthoépique notamment par les phonéticiens français. Schane fut le premier qui formalisa avec toute la fortune qu'on lui connaît la liaison en rendant compte ... de la non-réalisation de la liaison et simultanément de l'élision. Il propose (Schane, 1968, 1974) donc la très fameuse « French Truncation Rule » reprise et reformulée ensuite par divers auteurs (Milner 1967, puis Chomsky & Halle 1968, Dell 1970).

---

<sup>7</sup> Dorénavant SPE.

On peut en résumer l'esprit de la manière la suivante : la CL est formalisée dans la représentation lexicale comme un segment ordinaire appartenant au mot 1 et soumis à un processus d'effacement dans certains contextes.

Cette « French Rule » a suscité de nombreux et fructueux débats dont je ne traiterai pas ici car tel n'est pas mon objet (pour le détail des débats cf. Encrevé 1988, p.87 *et sq.*) Les critiques qu'on peut en faire concernent d'abord l'absence de formalisation de la spécificité phonologique de la CL. Dans son fameux texte traitant de « la réalité psychologique des phonèmes », Sapir (1933) relate comment un informateur de sarcee qu'il interrogeait avait été capable d'attribuer pertinemment des interprétations phonologiques différentes à des formes phonétiques identiques, et avait manifesté une sensibilité aux segments latents, n'ayant pas de réalisation substantielle, sur la base de sa seule intuition de locuteur, et sans connaissance orthographique des formes concernées. Ceci l'amène à soutenir 1- que les unités phonémiques ont une réalité psychologique même chez des locuteurs analphabètes 2- que les locuteurs ont une intuition très sûre des représentations phonologiques de leur langue. Mais ceci laisse également supposer que les consonnes latentes n'ont pas pour les locuteurs –auditeurs le statut de consonne ordinaire. Or, nous l'avons vu, les données perceptives et acoustiques infirment la proposition consistant à envisager la CL comme une consonne « comme les autres ». Bien que je ne considère pas que le problème soit d'abord et surtout phonétique, les résultats mentionnés plus haut indiquent que, déjà sur le plan acoustique, la CL n'est pas un segment ordinaire qui serait détruit dans les contextes non adéquats.

D'autre part, les consonnes de liaison ne peuvent pas être ramenées à des segments ordinaires car leur non-réalisation, contrairement à la non-réalisation des segments ordinaires n'affecte en rien la forme lexicale du mot concerné (Encrevé, 1988).

Au vu des résultats psycholinguistiques que j'avais obtenus, ce formalisme s'avérait peu utile dans la mesure où il propose une représentation lexicale unilinéaire et alphabétique qui me semble être la raison pour laquelle les modèles psycholinguistiques échouent à rendre compte du traitement en temps réel de la liaison enchaînée.

Enfin, les motivations de la règle, -et en particulier la motivation cognitive de la règle de délétion - sont pour le moins opaques et peinent à rendre compte de toute la variété phénoménologique observable en production : liaison non obligatoires, non enchaînée (Encrevé, 1988) français méridional (Sauzet, 1999) ainsi que des résultats obtenus en perception.

#### 2.4.2.2. *L'épenthèse*

En réponse à cette proposition et en réaction à l'arbitraire de la « French Truncation Rule » les tenants d'une phonologie plus « concrète » ont proposé une hypothèse épenthétique (Klausenburger, 1974, Tranel, 1981a) postulant une forme lexicale nue (par ex: /pəti/) correspondant à ce que la grammaire traditionnelle appelle « la forme brève » dont on dériverait « la forme longue ». La consonne apparaissant sur la frontière de mots est considérée comme résultant d'une opération d'insertion d'un matériau segmental et/ou métrique permettant la syllabation de la frontière. Tranel (1981a, 1981b, in Klausenburger, 1984) a donné des versions variées des procédures dérivationnelles permettant de rendre compte de la réalisation de cette consonne et de sa nature phonétique. Les critiques que l'on peut soulever à l'égard de cette hypothèse sont sensiblement les mêmes que celles que l'on pouvait retenir à l'égard de l'hypothèse de délétion mais à l'absence de formalisation du caractère spécifique de la liaison s'ajoute l'impossibilité d'expliquer clairement ce qui justifie la valeur segmentale que prend la consonne. Si la réalisation de /n/ dans [ɛ̃nami] n'est pas justifiée par l'existence de la nasalité réalisée sur la voyelle dans /ɛ̃/, comment expliquer que ce soit /n/ qui se réalise et non pas /t/, coronale non marquée ou aléatoirement, n'importe quel autre segment du français pouvant, soit structurellement occuper cette position, soit refléter l'usage contextuel. On remarquera que c'est ainsi qu'agissent les enfants qui réalisent aléatoirement la liaison avec n'importe quelle consonne pouvant remplir cet office (« le zéléphant », cf 3. ci-dessous).

Par ailleurs, du point de vue de la psycholinguistique de l'adulte, il est difficile d'envisager une hypothèse de traitement qui consisterait en une « épenthèse inversée » donc à une délétion en compréhension. Elle devrait rendre compte du passage d'une

consonne substantielle et réalisée à ... rien dans le lexique, dont l'existence dans le signal devient alors complètement injustifiable. On serait alors amené à justifier une règle qui consiste à générer des segments qui gênent la perception. Dans la mesure où, comme nous l'avons vu, pour tous les modèles psycholinguistiques, la CL gêne l'accès, cette proposition amène à considérer la liaison comme un « bruit » présent dans le signal que l'auditeur doit éliminer en temps réel parce qu'il entrave la reconnaissance des mots, gêne le traitement, entraîne un coût cognitif pour l'accès et dont on ne voit pas, dans ce cas, pourquoi le système l'aurait généré. Ceci paraît effectivement peu compatible avec l'argument selon lequel l'épenthèse constituerait une hypothèse plus « concrète » ou plus « naturelle » donc substantiellement ou cognitivement plus motivée que la troncation, sensée réduire la dimension arbitraire du formalisme et des dérivations. Or, comme nous l'avons rappelé, la liaison n'impose absolument aucun coût cognitif et n'entrave en rien la compréhension des énoncés. Donc au final, les motivations de la règle d'épenthèse ne paraissent pas moins opaques que celles de la règle de délétion et apparaissent même comme plus « anti-naturelles » pour rendre compte des données de compréhension.

Enfin cette hypothèse propose également un modèle unilinéaire et alphabétique des représentations phonologiques dont nous avons montré en quoi ils étaient inadéquats.

#### 2.4.2.3. *Le flottement*

La formulation d'un statut phonologique proprement « différent » pour les consonnes dites « latentes » n'a été véritablement proposée qu'avec l'émergence du cadre autosegmental permettant une définition de ces objets dissociant la position prosodique occupée par la consonne et le contenu segmental qui vient s'y associer et répondant à la conception ou à l'intuition la plus ancienne (Sapir, 1933) de ce qu'est un segment « latent » dont l'existence phonologique est intuitivement connue même quand il n'a pas de réalisation phonétique.

Clements & Keyser (1981) les premiers appliquent la notion de « flottement » aux consonnes de liaison en les assimilant aux « tons flottants ». Ceci les amène, dans un modèle trilinéaire à proposer que les consonnes de liaison ont une existence dans la



représentation lexicale, qu'elles sont rattachées à une position segmentale mais non associées, dans la tire syllabique, à une structure  $\sigma$ .

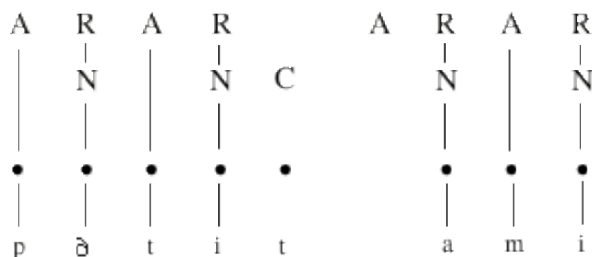
Encrevé (1988) retrace de manière détaillée les diverses propositions ayant exploité les possibilités formelles de la phonologie autosegmentale pour formaliser, à partir de cette première proposition, les différents avatars de la CL jusqu'au point qui nous intéresse directement. L'interprétation de la CL comme résultant par principe de la dissociation entre la position prosodique et l'interprétation segmentale étant posée, les discussions porteront sur la question i) du rattachement de la CL ii) de son degré de flottement. Dans la plupart des modélisations, il est considéré que la CL est attachée à une position squelettale et que par conséquent elle n'est que partiellement flottante.

Ce qui m'intéresse ici est de montrer comment la formulation la plus radicale du flottement, qui théorise la notion « d'autosegment » (Encrevé 1988) et exploite toutes les ressources formelles de la logique autosegmentale offrait non seulement les moyens de décrire mais également d'expliquer les données psycholinguistiques que j'avais obtenues.

La conception du flottement proposée par Encrevé [13] suppose

- i) un flottement de la CL sur la ligne segmentale, donc l'absence de rattachement à une position segmentale
- ii) un flottement de la CL sur la ligne syllabique, donc l'absence de rattachement à une attaque ou à une coda
- iii) une position squelettale disponible permettant l'ancrage de la CL à la fois au plan syllabique (donc en attaque ou en coda) et au plan segmental.

[13]



- iv) La réalisation de CL n'est pas la conséquence d'une dérivation par règles mais le résultat de conventions de bonne formation telles qu'elles ont été paramétrisées pour le français.

Cette conception repose donc sur l'idée que la consonne est syllabiquement et segmentalement flottante, donc « doublement » flottante. Ceci suppose donc qu'aucun rattachement n'est postulé a priori, que la CL est un objet présent dans la représentation lexicale mais qui n'a rien d'ordinaire et qui, par conséquent, ne doit ni être ajoutée ni, à l'inverse être retranchée lors des opérations de traitement. La représentation retenue par Encrevé (1988) conçoit la CL comme « purement flottante », sans aucune association. Si l'on retient ce formalisme, la CL appartient bien à l'unité lexicale (ce qui explique ses caractéristiques segmentales) mais hors de toute ligne, dans une espèce de non linéarité ou de linéarité suspensive (comme on parle en musique de « suspension », lorsqu'on utilise le procédé d'écriture qui consiste à retenir une ou plusieurs notes d'un accord en les ajoutant à l'accord suivant puis à les résoudre dans le deuxième accord). Elle est donc « suspendue » en attente d'une position et non soumise à la segmentation qui s'applique aux lignes syllabiques et segmentales et qui permettent de dégager respectivement les consonnes « ordinaires » et les syllabes bien formées. Ce faisant, elle échappe d'une certaine manière à la séquentialité temporelle exprimée dans le signal qui lui imposerait une position *a priori*. Elle n'est donc pas traitée en temps réel de la même manière qu'une consonne ordinaire même si elle en a quasiment les caractéristiques acoustiques (ce qui explique qu'elle ne soit pas détectée). Je propose qu'elle soit traitée à un niveau de segmentation propre : sa nature phonologique permet un traitement spécifique ce qui

permet d'expliquer aussi pourquoi le non alignement qu'elle crée entre frontière syllabique et frontière lexicale ne gêne pas l'accès lexical.

Ayant connaissance de la structure phonologique de sa langue, l'auditeur reconnaît en temps réel le sandhi comme un sandhi, et « sait » que la CL est un phonème de transition dont les caractéristiques segmentales sont conditionnées par le mot 1 même si ce phonème est réalisé à l'attaque du mot 2. Entendant [ɛ̃n...], l'auditeur « sait » que [ɛ̃] génère une consonne nasale quand il est suivi d'un mot commençant par une voyelle et il a beaucoup plus de chance d'être dans un contexte « un + mot commençant par une voyelle » que « un + mot commençant par « n » », puisque dans le lexique du français, le nombre total de mots commençant par n'importe quelle voyelle est très nettement supérieur au nombre de mots commençant par « n ». Donc la consonne sera préférentiellement reconnue comme une CL et donc extraite du traitement temporel, non calculée à la frontière comme contexte gauche possible. Ceci suppose qu'en français dans une suite comme [ɛ̃nami], la segmentation avec liaison sera la segmentation par défaut et que le sujet segmentera spontanément [ɛ̃(n) + ami] et non [ɛ̃ + nami]. Ceci explique pourquoi la liaison n'impose aucun coût cognitif particulier pour l'accès au lexique et pourquoi la segmentation et la reconnaissance des mots peuvent même être facilitées en contexte de liaison.

Je pense que ce sont les mêmes raisons qui entraînent une non-détection du segment de liaison. Il semblerait qu'on assiste là à quelque chose qui serait de l'ordre de la surdit  phonologique. Objet   contenu segmental ayant pour r le d'exprimer phonologiquement un lien formel entre deux mots, la CL n'est pas   proprement parler un phon me au sens o  une consonne lexicale (ancr e, appartenant pleinement   l'unit  lexicale, distinctive au sein d'un syst me phonologique) l'est. Lors de la t che de d tection, il semblerait qu'elle ne soit pas non plus trait e comme telle, qu'elle ne soit ni entendue, ni reconnue parce qu'elle n'est ancr e nulle part mais que c'est un objet « flottant », « virtuel » et qui malgr  sa r alit  phon tique substantielle est bien trait e comme un objet phonologique « virtuel » par l'auditeur. Cette analyse rejoint l'hypoth se sapirienne d'une sensibilit  aux segments latents comme des objets phonologiques sp cifiques.

#### 2.4.2.4. *La supplétion*

L'hypothèse supplétive consiste à défendre l'idée que chaque mot est stocké dans le lexique de manière concrète et détaillée, quasi « phonétique » comme un exemplaire, un « token », intégrant les variations de réalisation de l'unité lexicale, mémorisées comme telles dans le lexique. Ceci consisterait par exemple à proposer que [pəti], [pətit] et [pətiz] sont également stockés comme mots dans le lexique et ne constituent pas trois réalisations différentes d'une même représentation abstraite. Cette conception radicalement incompatible avec les postulats épistémologiques sur lesquels repose la grammaire générative, et formulée au nom d'une « naturalité » phonologique s'est récemment ré-imposée plus franchement dans le cadre de phonologies mettant en cause le primat des représentations abstraites sur la substance phonique (Coleman, 2002) dans le cadre de la phonologie « des exemplaires » privilégiant des effets de fréquence et d'usage dans l'émergence des généralisations phonologiques ainsi qu'une conception procédurale de la phonologie. Bybee (2001), en particulier, rend compte de la phonologie de la liaison par l'hypothèse d'un stockage exhaustif de toutes les variantes possibles d'une unité lexicale entrant en contexte de liaison avec les diverses contraintes contextuelles pouvant entraîner la liaison. Elle défend l'idée que la liaison pourrait être conditionnée non seulement par des contraintes morphosyntaxiques, prosodiques et lexicales mais surtout par des effets de fréquence. « enfant » ne sera donc pas stocké comme une unité lexicale isolée mais dans tous les contextes où il apparaît « un enfant », « un petit enfant », « les enfants », « mes enfants », « tes enfants », « un grand enfant ». L'analyse qui sera faite ici est que, dans tous les cas, les déterminants et les adjectifs antéposés à valeur déterminative sont stockés avec le mot dans le lexique, et l'expression « un petit enfant » est envisagée comme une unité lexicale et non comme un syntagme nominal. La réalisation (vs non réalisation) des liaisons est donc conditionnée par la fréquence de co-occurrence des unités ou la fréquence d'usage de l'expression avec ou sans liaison. Ceci prédit donc que i) dans « un + enfant » la liaison est toujours réalisée, alors que « un + anachorète » (co-occurrence peu fréquente) la liaison pourrait ne pas être réalisée ii) « des soldats

anglais » est une liaison facultative car la co-occurrence de soldats + anglais est rare et que, quand elle est produite par les locuteurs, il est rare qu'ils fassent la liaison.

Ceci tendrait à prouver, selon elle, que le degré de disponibilité d'un contexte de liaison en perception et production est directement lié à sa trace en mémoire.

« The chapter concludes with the argument that phonological alternations viewed as having syntactic conditioning are due to the frequency of certain combinations of words and provides evidence for chunks, or unit of processing and storage that are used in speech »

J.Bybee, *Phonology and Language Use*, 2001, chap. 7

On voit bien comment cette hypothèse supplétive proposée par Bybee (2001) suppose un stockage exhaustif d'alternances utilisables en « mapping » direct entre le signal et le lexique lors des procédures de traitement, et rejoint les propositions des modèles « épisodiques » où les représentations lexicales sont élaborées indépendamment de tout formalisme *a priori* et envisagées comme des phénomènes émergents à partir d'un réseau ou d'un treillis (de spectres acoustiques pour LAFS ou de traces mnésiques auto-organisées pour Goldinger & Azuma (2003)). Le traitement psycholinguistique est donc envisagé d'un point de vue essentiellement procédural et s'appuie sur l'hypothèse d'une très puissante capacité mémorielle Nguyen (2005). Ce type de modèle postule *a priori* que l'utilisation, par l'auditeur, de représentations phonologiques abstraites lors des procédures d'accès n'est pas un problème pertinent et que le traitement de la liaison en temps réel ressortit d'une modélisation générale du traitement de la variation lexicale. Ces propositions n'étaient pas véritablement disponibles à l'époque où j'ai travaillé sur cette question et c'est la raison pour laquelle je n'ai pas testé les hypothèses qu'elles proposaient pour le traitement de la liaison en temps réel. Cette question est l'objet d'un travail en cours qui est exposé au chapitre II.

## **2.5. Hypothèse de traitement de la parole continue en temps réel.**

A l'issue de cette première série de travaux concernant le traitement de la liaison en parole continue, le modèle de traitement que j'ai présenté exploite l'utilisation des possibilités de segmentation simultanées que permet la phonologie

autosegmentale et propose un traitement parallèle de l'information segmentale et prosodique.

L'un des problèmes que soulèvent les modèles d'accès au lexique lexicaux et infra-lexicaux que nous avons présentés est qu'ils envisagent toujours que l'accès lexical soit réalisé à partir d'une unité de traitement inférieure ou égale au mot à reconnaître. Or, dans une langue comme le français, la plupart des indices prosodiques délimitent des constituants excédant l'unité lexicale (Ferreira, 1993, Gee & Grosjean, 1983, Grosjean et Dommergues, 1983, Grosjean & Gee, 1987, Delais, 1994, 1995, Padeloup, 1989, 1990). Les indices retenus sont le patron intonatif démarcatif, la distribution des accents, les indices de durée qui délimitent des structures prosodiques à partir du signal produit (et non à partir de la projection d'une structure syntaxique de type X-barre dans le signal, comme le propose la phonologie prosodique). Diverses versions en ont été données (Grosjean & Gee, 1987, Delais, 1994, 1995, Padeloup, 1989, 1990) et toutes les unités proposées s'inscrivent dans une structure hiérarchisée dont les caractéristiques peuvent se résumer ainsi : les niveaux d'organisation prosodique vont de la syllabe à la phrase avec une deux ou unités intermédiaires appelées « groupe prosodique », « groupe de souffle ». Ces unités intermédiaires de longueur moyenne oscillant entre 4/5 et 6/7 syllabes répondent au principe d'équilibrage (Delais, 1994, 1995) qui consiste à créer une organisation prosodique en essayant de maintenir une symétrie entre les constituants à un niveau donné, ceci suppose que ces groupes comportent plus ou moins le même nombre de syllabes et que les regroupements s'opèrent parfois en accommodant la structure syntaxique. J'illustre par les quatre exemples ci-dessous.

[14]  
(l'hippopotame)<sub>4</sub> (nage dans le Nil)<sub>4</sub>

[15]  
(le gros hippopotame)<sub>6</sub> (de la vallée du Nil)<sub>6</sub> (ne nage pas vite)<sub>4</sub>

[16]  
(le grand chien nage)<sub>4</sub>

[17]  
(le grand chien nage dans le Nil)<sub>7</sub> ou (le grand chien)<sub>3</sub> (nage dans le Nil)<sub>4</sub>

J'avais proposé que les informations utilisées lors de la production de ces groupes prosodiques (réajustements en vue de l'équilibre, durée correspondant à la programmation dans la boucle phonologique, congruence avec la structure syntaxique ou argumentale) soient utilisées aussi en perception pour délimiter un domaine de traitement à l'intérieur duquel se ferait le traitement phonologique et que segmentation segmentale et segmentation prosodique soient menées en parallèle (figure 4).

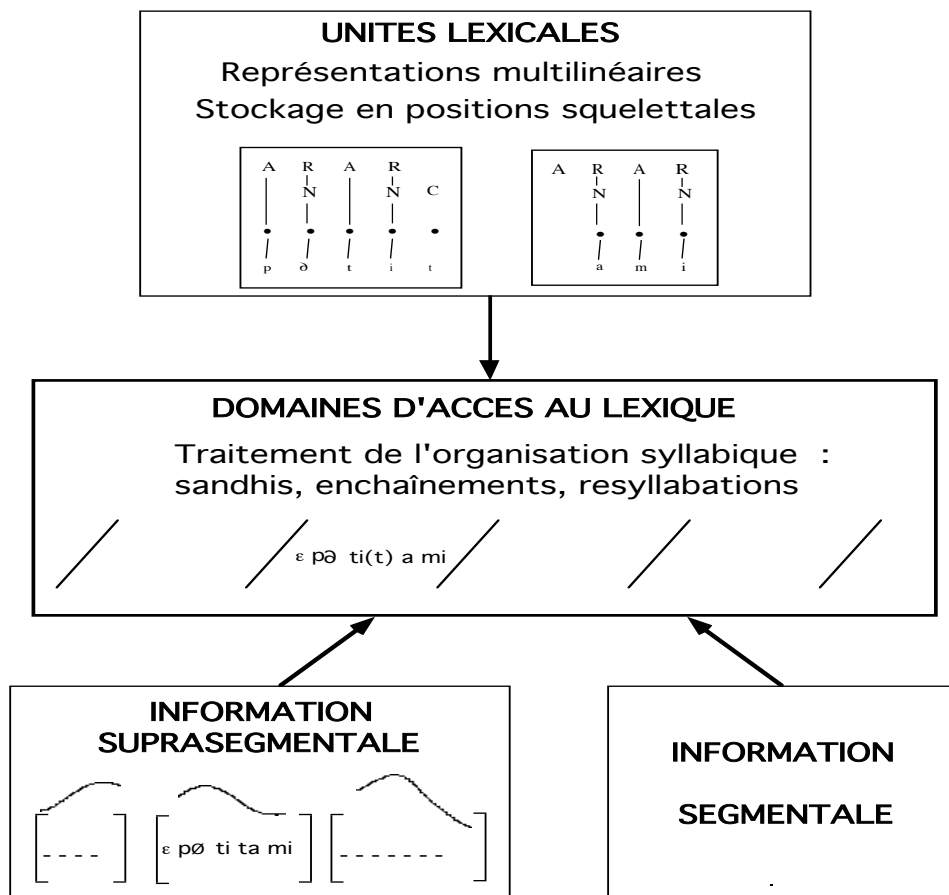


Figure 4 : Proposition de modèle d'accès au lexique en parole continue

A côté de la segmentation infra lexicale est opérée une segmentation en « domaines perceptifs » définis par les informations acoustiques présentes dans le signal (accents, allongements, patrons intonatifs). C'est à ce niveau qu'en français serait déterminée l'organisation syllabique de la chaîne. L'accès est réalisé par

convergence de cette double information montante avec l'information multilinéaire descendante vers un niveau constitué de domaines phonologiques à partir desquels sont reconnus les mots. Cette double dynamique permet que la sélection des mots s'opère au niveau de constituants abstraits après traitement des divers phénomènes phonologiques gênant la reconnaissance lexicale (sandhis, enchaînements et resyllabations) sur leur domaine de réalisation. L'existence d'un niveau de domaines abstraits où se réalise l'accès lexical contraint drastiquement le nombre de candidats lexicaux possibles qu'il est nécessaire d'activer ou d'hypothèses de segmentation syllabiques nécessaires pour retrouver le mot recherché. Si l'on reprend l'exemple des mots enchâssés qui pourraient être contenus dans [ləkrəkodilkrəkodil], un auditeur réalisant la segmentation à partir des deux domaines abstraits /ləkrəkodil/ /krəkodil/ va éliminer automatiquement dans un domaine comme /ləkrəkodil/ « croque, ode, île, Odile » et éviter automatiquement certaines segmentations ambiguës.

Dans le cas de la liaison, qui est comme nous l'avons vu un phénomène dont la réalisation est contrainte par un contexte syntaxico-sémantique, c'est à ce niveau que pourraient se résoudre les décalages entre les segmentations parallèles correspondant aux différentes unités interprétables à partir du signal. Le domaine prosodique apporte par exemple une information suprasegmentale définissant le contexte où la liaison peut ou ne peut pas être réalisée. L'auditeur disposera d'une information supplémentaire lui permettant de déterminer si le contexte autorise ou non la liaison et donc si les suites peuvent ou non inclure des segments interprétables comme des consonnes flottantes. Dans ce cas là, elles peuvent être exclues du calcul linéaire sur la ligne segmentale et sur la ligne syllabique simultanément et ne donnent pas lieu à des erreurs de segmentation.



## Travaux traitant de ces questions

### Publications

**Wauquier-Gravelines, S.** (1992), Les interférences phonétiques comme approche des confusions perceptives, *Actes des XIXèmes journées d'étude sur la parole*, Bruxelles, 19-22mai 1992

**Wauquier-Gravelines, S.** (1994), Segmentation lexicale en français parlé : la liaison enchaînée, *Actes des XX èmes Journées d'Etudes sur la Parole*, juin 1994, Lannion.

**Wauquier-Gravelines, S.** (1995), Detecting Ghost Phonemes : The 'liaison enchaînée' in French, *Proceedings of XIIIth International Congress of Phonetic Sciences*, Stockholm, 13-19 août 1995.

Segui, J & **Wauquier-Gravelines, S.** & (1995), La perception du langage parlé, in *Rééducation orthophonique.*, **182**,171-182.

**Wauquier-Gravelines, S.** (1995), Interferences between phonemes : Evidence for « perceptual-domains » in continuous Speech Perception, *Proceedings of Eurospeech Conference*, Madrid, 18-21 septembre 1995.

**Wauquier-Gravelines, S.** (1996), Organisation phonologique et traitement de la parole continue, Thèse de doctorat en linguistique théorique et formelle, Université Paris 7, Jussieu, ms.

**Wauquier-Gravelines, S.** (1999), Segmentation lexicale et parole continue : la linéarité en question. *Recherches Linguistiques de Vincennes*, **28**, Presses Universitaires de Vincennes, Université Paris VIII

**Wauquier-Gravelines S.** (2004), Les 'liaisons dangereuses' Psycholinguistique et phonologie : une interface complexe, in Ferrand, L., & Grainger, J. (Eds.) *Psycholinguistique cognitive, Essais en l'honneur de Juan Segui*, Bruxelles : De Boeck Université, Collection Neurosciences et Cognition.

**Wauquiers-Gravelines S.** (à paraître), Phonology and psycholinguistics, time to stike the balance ? in Ferrand, L., & Grainger, J. (Eds.), Mouton de Gruyter, Berlin<sup>8</sup>.

### Communications

**Wauquier-Gravelines, S.** (1992), Les interférences phonétiques comme approche des confusions perceptives, Communication orale XIXèmes Journées d'étude sur la Parole, Bruxelles, 19-22mai 1992

**Wauquier-Gravelines, S.** (1994), Segmentation lexicale en français parlé : la liaison enchaînée, Communication orale XX èmes Journées d'Etudes sur la Parole, Lannion.3 juin 1994,

**Wauquier-Gravelines, S.** & Segui J.(1994), Segmentation de la parole continue : la détection des phonèmes fantômes, *Troisième Journée scientifique du Réseau Cogniseine*, 6 juin 1994, Paris

**Wauquier-Gravelines, S.** (1994), La perception du langage parlé, *Sensorialité et Langage : 13èmes journées de Neuropsychologie et de Rééducation organisées par le Groupe de Recherche sur les*

---

<sup>8</sup> Ce manuscrit a également été proposé à MIT press.

*Apprentissages et les Altérations du Langage ( GRAAL)*, Centre hospitalier Emile Roux, Limeil-Brévannes, 25 novembre 1994,

**Wauquier-Gravelines, S** (1995), Interferences between phonemes : Evidence for « perceptual-domains » in continuous speech Perception, *Eurospeech Conference*, Madrid, 18-21 septembre 1995.

**Wauquier-Gravelines, S.** (1995), Detecting Ghost Phonemes : The ‘liaison enchaînée’ in French, *XIIIth International Congress of Phonetic Sciences*, Stockholm, 13-19 aout 1995.

**Wauquier-Gravelines, S.** (1996), Recognition of words in fluent French, Problems, data and model, Conférence invitée, Département de Phonétique de l'Université d'Umea, Suède, 12 avril 1996

**Wauquier-Gravelines S.** (2004), Les liaisons dangereuses. Psycholinguistique et phonologie, Conférence invitée *Journée Internationale de Psycholinguistique cognitive, en l'honneur de Juan Segui*, LPE-CNRS, Université R. Descartes-Paris 5, 6 novembre 2004.

### 3. Acquisition de la liaison

#### 3.1. Introduction

Alors que je travaillais sur le traitement de la liaison chez l'adulte, j'avais constaté que les enfants autour de 3 ans commettaient des erreurs systématiques de liaison ( "le nane" pour l'âne, "le zwazeau" pour l'oiseau) et semblaient mettre en place provisoirement des représentations lexicales erronées. Aussi bizarre que cela paraisse, ces erreurs n'étaient ni décrites ni répertoriées ni analysées dans la littérature scientifique, pédagogique ou orthoépique.

Or les questions que j'avais soulevées via le traitement par l'adulte me semblaient se poser de manière encore plus pertinente pour l'acquisition. Car, en ce qui concerne l'enfant, le problème de la segmentation se pose de manière beaucoup plus complexe. Le lexique mental ne pouvant être reçu de manière innée, l'enfant qui apprend à parler une langue doit constituer un répertoire de formes lexicales de référence, sur la base de ses expériences linguistiques et de sa connaissance du monde extra-linguistique. Pour retrouver les mots dans le signal, contrairement à l'adulte qui peut se servir de ses connaissances lexicales et phonologiques, les nouveau-nés puis les petits enfants ne peuvent s'appuyer ni sur une information lexicale finie pré-stockée, ni sur un système phonologique achevé et doivent opérer un tri et une segmentation efficace et organisée du matériau sonore qui les environne sur la seule base de leur expérience sensorielle. La liaison, crée des ambiguïtés de frontière lexicale qui ne pourront pas chez l'enfant être résolues comme je l'ai proposé pour l'adulte par le recours à une représentation phonologique abstraite et multilinéaire des unités lexicales.

A partir de 1994, j'ai donc d'abord recueilli un corpus spontané d'environ cinquante occurrences chez une quinzaine d'enfants âgés de 2 à 4 ans. Les observations semblaient poser également l'existence d'une période "critique" où les enfants commettent régulièrement des erreurs sur les mots en contexte de liaison. J'en donne quelques exemples ci-dessous

[18]

\*[lenan] pour [lezan] *les ânes* (Aurélien ; 3,1)

[19]

\*[lenelefã] pour [lezelefã] *les éléphants* (Marie; 3) ;

[20]

\*[ëzwazo] pour [ënwazo] ou \*[lenwazo] pour [lezwazo] *les oiseaux* (Aurélien, 2,8 ; Lélia, 2,7)

[21]

\*[zəvəamanekɔl] pour [zəvəamɔnekɔl] *je vais à mon école* (Joseph ; 3)

[22]

\*[elamalasenaft] pour [elamalasezft] *elle a mal à ses aphtes* (Léa ; 3,1)

[23]

\*[blãfnezelesetë] pour [blãfnezelesetnë] *Blanche-Neige et les sept nains* (Lélia, 3,5)

[24]

\*[zəmãzedezem] pour [zəmãzedenem] *j'ai mangé des nems* (Lélia, 3,7)

Dans la plupart des cas ([18] à [22]), on observe une réalisation erronée du mot dans un contexte de liaison impliquant la réalisation d'une consonne qui n'est pas la consonne attendue sur l'attaque de mot. Par exemple, [lenan] pour [lezan] puisque après l'article défini "les" libérant le segment flottant [z] le mot [an] est réalisé [zan]). Ceci signifierait que les enfants ont apparemment une représentation lexicale erronée du mot "âne" qui inclut le "n" initial à l'attaque du mot comme une consonne lui appartenant pleinement d'où les productions telles que "les nanes", "le nane", "le petit nane". Ils n'ont donc pas à ce stade "compris" que le "n" entendu dans "un âne" provient du sandhi et n'appartient pas au mot lui-même.

Un autre type d'erreurs se fait jour dans ce corpus, ([23] et [24]), erreurs plus rares mais également très intéressantes et témoignant sans doute d'une hésitation de l'enfant quant au statut de la frontière gauche du mot envisagé. Dans l'énoncé [blãfnezelesetë], on constate que le mot "nain" est amputé de sa consonne initiale dans un contexte où il est précédé d'un mot se terminant par une consonne pouvant donner lieu à un enchaînement. Ceci signifie que l'enfant fait alors l'hypothèse que le mot désigné est "ain" et que la forme "nain" a résulté d'une liaison avec le déterminant "un" libérant un [n] flottant. Très logiquement cette enfant produit donc

[ɛ̃nɛ̃, døzɛ̃, sɛtɛ̃]. Je pense qu'on peut envisager qu'il s'agit là d'une surgénéralisation inappropriée. On voit que ces erreurs ont été produites tardivement par Lélia (3,5 et 3,7 ans), ce qui invite à penser qu'une fois qu'elle a eu compris que les consonnes réalisées en début de mot pouvaient être des consonnes flottantes, elle a interprété toutes les consonnes initiales comme des consonnes flottantes.

Dans tous les cas, quoi qu'il en soit, les erreurs sont associées à une représentation lexicale erronée de l'item produit et liées à une question de segmentation de la chaîne.

A la suite de ces premières constatations, j'ai mené, à partir de 1997, des travaux plus systématiques à partir d'énoncés produits en situation expérimentale (dénomination d'images) afin d'apporter des réponses aux questions suivantes :

- i) Est-ce que cette période apparaît chez tous les enfants ?
- ii) Dans quelles circonstances (et quels contextes liaisons obligatoires *vs* facultatives, enchaînées *vs* non-enchaînées) ?
- iii) A quel âge ces erreurs apparaissent puis se résolvent-elles ?
- iv) Peut-on y voir les manifestations de la mise en place de la phonologie adulte ?

### **3.2. Premières données, premières analyses**

Ces questions ont été abordées dans une série d'expériences menées d'abord seule, puis avec Virginie Braud<sup>9</sup> dans le cadre d'une maîtrise avec des enfants monolingues et Sandrine Ferré<sup>10</sup> dans le cadre d'un DEA avec des enfants bilingues.

---

<sup>9</sup> BRAUD, V. (1998) "Acquisition de l'information phonologique : exemple de la liaison", Mémoire de maîtrise, Université de Nantes, ms

<sup>10</sup> FERRÉ, S. (1998) "Les effets du bilinguisme sur la paramétrisation des segments flottants : exemple de la liaison en français", Mémoire de DEA sous la direction de J-P Angoujard, Université de Paris VIII.

### 3.2.1. Données et résultats : enfants monolingues

#### 3.2.1.1. Expériences

Deux groupes d'enfants ont participé à cette étude. Les deux groupes d'enfants étaient constitués d'enfants monolingues de langue maternelle française. Tous ces enfants étaient scolarisés en école maternelle à Paris et à Nantes et ont été testés dans le cadre de l'école.

Le premier groupe (E3) était composé de 12 enfants (6 garçons et 6 filles) âgés de 3 à 3,8 ans. Le deuxième groupe (E5) était composé de 12 enfants (6 garçons et 6 filles) âgés de 4,10 à 5,7 ans. La tâche consistait en une dénomination d'images présentées accompagnées de la question suivante : "Qu'est-ce que tu vois sur cette image", induisant une réponse incluant l'article indéfini qui fait obligatoirement la liaison quand il est suivi d'un mot commençant par une voyelle. Les enfants devaient produire des items lexicaux appariés en fréquence dans les contextes suivants (le tableau 1 donne des exemples de contextes singuliers. Les mêmes mots étaient présentés aussi en contexte pluriel). Les enfants devaient donc produire au total 24 items au singulier et 24 items au pluriel avec liaison obligatoire et enchaînée.

Comme le montre le tableau 1, nous avons contrôlé le facteur « longueur de mots » dans la mesure où la liaison entraîne une réorganisation prosodique de l'unité phonologique et prosodique que les enfants doivent produire. J'avais intuitivement supposé que les mots longs subiraient des transformations spécifiques.

	Liaison	Non-liaison
monosyll	"un + ours" [ɛ̃nurs]	"un + loup" « des loups » [ɛ̃lu]
bisyll	"un avion" [ɛ̃navjɔ̃]	"un château" [ɛ̃ʃato]
trisyll	"un écureuil" [ɛ̃nekyrœj]	"un perroquet" [ɛ̃peroke]
quadrisyll	"un hippopotame" [ɛ̃nipopotam]	"un rhinoceros" [ɛ̃rinoserɔs]

Tableau 1 : Exemple de matériel utilisé

### 3.2.1.2. Résultats globaux

La figure 5 présente les résultats globaux des enfants E5, la figure 6 présente les résultats globaux des enfants E3. BL signifie “ bonne liaison ” (réponse attendue en bleu), FL signifie “ fausse liaison ” (autre réponse que réponse attendue, en rouge), NR signifie “ non-réponse ” (en jaune). Les réponses sont rangées par ordre décroissant de performance de gauche à droite.

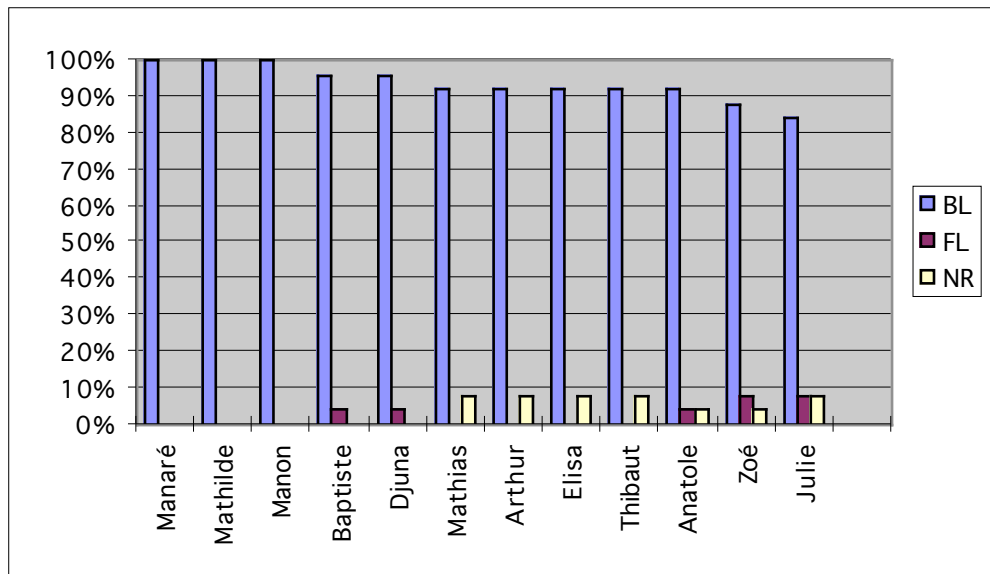


Figure 5 : Résultats globaux pour les enfants de 5 ans

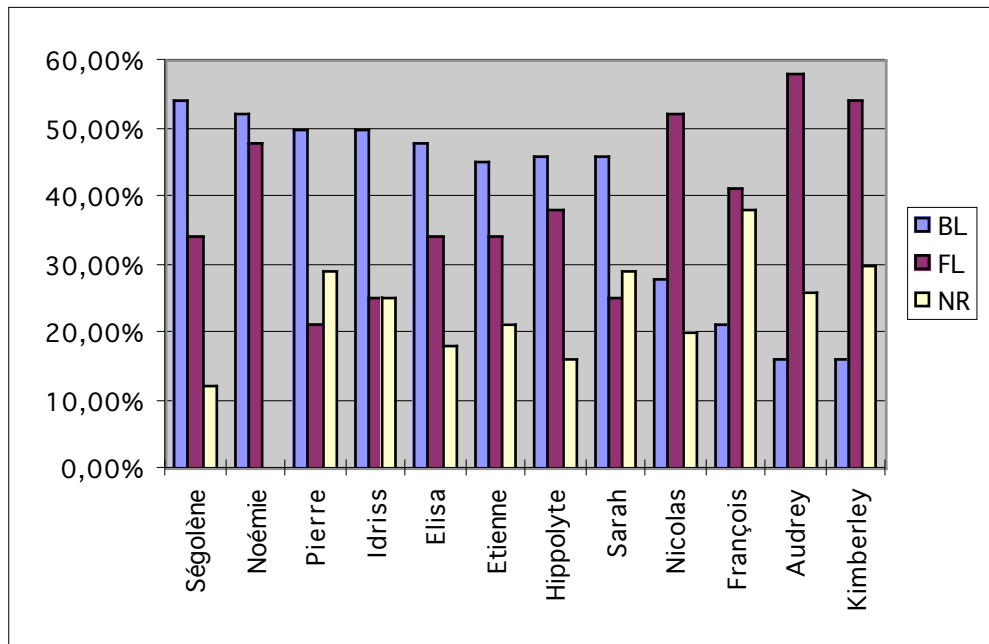


Figure 6 : Résultats globaux pour les enfants de 3 ans

On voit très clairement sur les figures 5 et 6, d'une part que les enfants de 5 ans n'ont aucune difficulté avec la tâche de dénomination d'image (le taux de non-réponse est extrêmement bas) et d'autre part qu'ils ont complètement acquis la liaison obligatoire entre déterminant et nom à 5 ans.

Les performances des enfants de 3 ans, sont beaucoup plus hétérogènes. Je les commente ci-dessous.

### 3.2.1.3. Résultats détaillés des enfants de 3 ans

Comme nous pouvons le constater, le groupe des enfants de 3 ans manifeste des difficultés pour les mots présentés en contexte de liaison. Mais malgré l'hétérogénéité des performances, les fautes ne sont pas aléatoires. L'observation des données nous a permis d'opérer une classification systématique en deux grands types d'erreurs. Des erreurs portant sur la CL elle-même (ECL) et des erreurs portant sur la forme du mot proposé (ELEX). Le tableau 2 présente les types d'erreurs rencontrés.



Erreurs sur CL (ECL)	Erreurs lexicales (ELEX)
<b>Suppression de la consonne</b> ex : "un enfant" → [ɛ̃/ɑ̃fɑ̃] au lieu de [ɛ̃nɑ̃fɑ̃]	<b>troncations (1 syllable ou plus)</b> ex. : un hélicoptère / → [ɛ̃kɔ̃ptɛʁ] au lieu de [ɛ̃nelikɔ̃ptɛʁ]
<b>Mauvaise consonne</b> ex. : "un enfant" → [ɛ̃zɑ̃fɑ̃] au lieu de [ɛ̃nɑ̃fɑ̃]	<b>Modifications phonémiques</b> ex. : un aspirateur → [ɛ̃ʁɑ̃ʁastatœʁ] au lieu de [ɛ̃naspiʁatœʁ]

Tableau 2 : Catégories d'erreurs pour l'analyse des données des enfants de 3 ans

Comme on peut le voir (figures 7 et 8) le nombre de fautes de consonnes et le nombre de fautes lexicales sont à peu près équivalentes mais les enfants associent très rarement les deux types d'erreurs. Pourtant, la distribution des erreurs n'est pas équivalente pour tous les items présentés : l'analyse plus détaillée des données indique en effet très clairement une répartition des erreurs en fonction de la longueur syllabique (figure 8) : les mots de 1 et 2 syllabes présentent majoritairement des fautes concernant la consonne elle-même alors que les mots de 3 et 4 syllabes sont principalement source d'erreurs lexicales, consistant en modifications phonémiques et surtout troncations quasi systématiques que je ne commente pas ici mais dont je traite aux points 4 et 5 de ce chapitre.

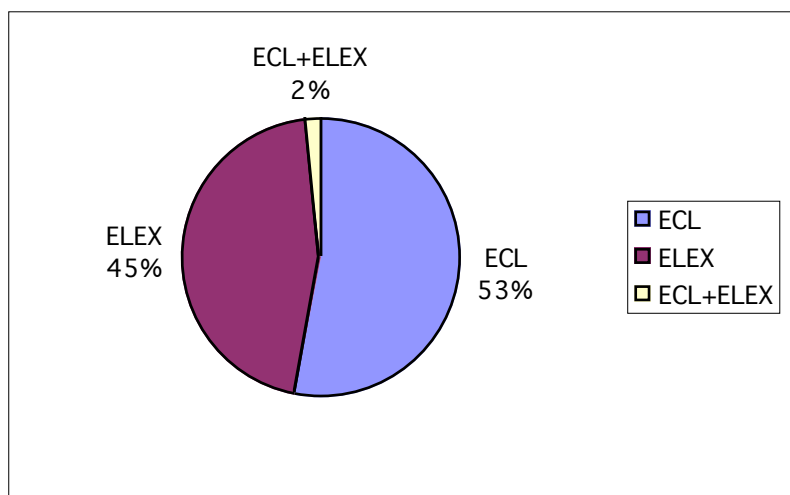


Figure 7 : Types d'erreurs pour les enfants E3

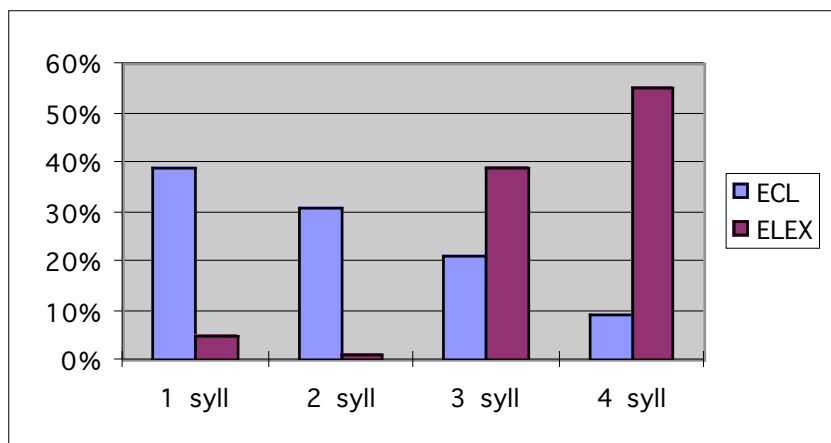


Figure 8 : Types d'erreurs en relation avec la longueur syllabique des items pour les enfants E3

- *ECL1 : Suppression de la consonne*

L'absence de CL est une erreur réalisée de manière marginale par les enfants. Cette erreur semble attribuable à un biais expérimental : les “ suppressions ” se sont produites dans le cas où le mot n’était pas connu. Dans ce cas, l'expérimentateur donnait alors le mot à l'enfant sous une forme “ canonique ” excluant la liaison (ex. : ça s'appelle “ ours ”) puis lui reposait la question : “ Qu'est ce que tu vois sur l'image ”. Dans ce cas-là un certain nombre d'enfants reprenaient le mot tel quel en y accolant ou non le déterminant et en marquant une pause entre le déterminant et le nom ([urs] ou [ɛ/urs]). Mais de manière spontanée ce type d'erreur n'apparaît jamais.

- *ECL2 : Mauvaise consonne*

La récapitulation des erreurs de “ mauvaise consonne ” (MCE) se trouve dans le tableau 3 tous enfants confondus (n'ont pas été inclus les items tri et quadrisyllabiques qui ne présentent pratiquement pas d'erreurs de ce type car ils sont régulièrement tronqués).

Items	MCE	Items	MCE
<b>monosyll</b>		<b>bisyll</b>	
• oeuf	[ɛzø]	• avion	[ɛlavjɔ̃]
• arbre	[ɛzarbr] [ɛlarbr]	• oiseau	[ɛzwazo]
• âne	[ɛzan]	• indien	[ɛzɛdjɛ] [ɛlɛdjɛ]
• oeil	[ɛzjø]	• enfant	[ɛzɑ̃fɑ̃]
• ours	[ɛlurs]	• éclair	
• aigle		• oignon	[ɛzoɲɔ̃]

Tableau 3 : Erreurs de mauvaise consonne pour les enfants E3

Toutes les erreurs produites résultent d'un remplacement du [n] par un [z] ou un [l] à l'exclusion de toute autre consonne disponible en français et du [z] par un [n] ou un [l]. La consonne proposée est dans ce contexte soit le [n] issu du déterminant singulier (95% des fautes sur le pluriel) soit le [z] issu du déterminant pluriel (72% des fautes au singulier), soit [l] issu de l'article défini élide.

### 3.2.2. Données et résultats : enfants bilingues

Les mêmes expériences sur du matériel identique ont été réalisées avec 13 enfants bilingues français-anglais âgés de trois à quatre ans au moment de la passation.

La démarche était ici prospective et visait à établir l'existence d'éventuels patterns différents dûs au bilinguisme. Les moyennes de réponse globales ne permettent pas d'établir des différences avec les enfants monolingues dans leur ensemble, on constate seulement un éventuel délai de résolution des erreurs dans les productions bilingues.

Mais l'analyse plus détaillée des données montre un pattern différent pour les enfants bilingues de langue maternelle française et les enfants de langue maternelle anglaise.

Les enfants de langue maternelle française ont un taux d'erreur comparable aux enfants monolingues alors que les enfants bilingues

- i) montrent un retard marqué dans la maîtrise du phénomène de liaison enchaînée

- ii) font des erreurs de non-liaison en ne réalisant pas la consonne sur la frontière. Erreurs dont nous avons vu qu'elles sont rares chez les enfants monolingues.

Ces données plaident en faveur d'une paramétrisation différenciée des segments flottants à partir de l'input selon l'importance de la langue cible reçue par l'enfant. Elles montrent en particulier que l'acquisition du « segment flottant » est réalisé différemment pour l'acquisition du français L1 et L2, ce qui a été montré par Dejean de la Bâtie (1993). Ceci semble être vrai, même dans le cas où les deux input sont reçus précocement et simultanément.

### 3.2.3. Premières analyses

Ces premières données expérimentales associées au mini-corpus recueilli auparavant permettaient d'apporter une première série de réponses aux questions posées initialement.

#### 3.2.3.1. Généralisation du phénomène : âge et type d'erreurs

L'acquisition du français entraîne de manière plus ou moins longue et prononcée et apparemment pour tous les enfants apprenant cette langue, la réalisation d'erreurs systématiques sur les liaisons. Les erreurs se manifestent en général à partir de 2 ans et ont complètement disparu à 5 ans. Elles sont donc contemporaines des acquisitions lexicales et morphosyntaxiques. On peut envisager l'existence d'une période "critique" entre 2 et 4 ans, où les enfants réalisent des fautes de liaison systématiques sur les contextes obligatoires. Les enfants bilingues montrent un retard à l'acquisition de la liaison enchaînée.

En ce qui concerne le type d'erreur, l'expérience n'a testé qu'un seul contexte obligatoire et les erreurs sont très peu variables. La non-réalisation de la liaison obligatoire semble marginale ([[ǝ/urs]). Presque toutes les erreurs sont des erreurs de

remplacement d'une consonne par une autre ([ɛzurs] [ɛlurs] pour [ɛnurs]). La consonne par défaut, prioritaire semble être le [n] du singulier qui sert alternativement avec le [z] pour remplir l'attaque à la frontière gauche de mot.

Toutes les liaisons réalisées, qu'elles soient justes ou fautives sont réalisées enchaînées.

Les données spontanées sont un peu plus variables et, comme nous l'avons vu, ajoutent aux données expérimentales des erreurs de surgénéralisation consistant à considérer la consonne de début de mot ([n] dans « nains » comme une CL).

Dans les deux cas, les fautes des enfants témoignent donc de la difficulté pour eux de différencier « consonne lexicale » et « consonne flottante » ainsi que de leur impossibilité de leur attribuer par conséquent leur place et leur distribution respectives. Dans la plupart des cas, la CL est interprétée comme une consonne initiale et l'enfant pendant un certain temps semble avoir stocké le mot sous une forme erronée comportant une attaque sur la frontière gauche (« le nane » pour « l'âne »).

A ce stade de ma recherche, j'ai donné une première interprétation de mes données à partir du modèle déclaratif d'Angoujard (1997).

### 3.2.3.2. *Première interprétation : l'absence de sous-spécification*

Angoujard (1997) développe un système à base de contraintes d'association n'utilisant qu'un seul niveau de représentation. Le but de la représentation consiste à "mettre en évidence toutes les contraintes portant sur une représentation donnée" (Angoujard 1997).

Les séquences sonores bien formées sont définies comme des chaînes de segments recevant une triple définition. Chaque segment est d'abord décrit par sa composition en éléments (Kaye, Lowenstamm et Vergnaud 1985, Harris 1994, Angoujard 1997).

La justification des positions squelettales est ensuite légitimée par l'existence d'une structure rythmique. Ainsi toute séquence sonore peut être décomposée en une

série de modèles rythmiques de base. Les segments sont également définis par leur position rythmique (pos)

	x	
x	x	(x)
1	2	3

Les positions 1 et 2 sont universelles : 1 correspondant à un creux de sonorité, 2 un sommet, équivalent au maximum de sonorité. La position 3 est optionnelle et paramétrisée selon les langues et correspond à une baisse de sonorité par rapport au sommet

Enfin, tout segment se voit associé un trait pc (position de courbe) résultant de l'association à tout modèle rythmique d'une courbe prosodique.

Tout séquence phonique est donc représentée par une succession de triplets <seg, pos, pc>. Ainsi [a] dans "ami" sera représenté comme ci-dessous. "Seg" désignant la composition en élément, "pos" désignant la position dans la grille métrique et pc désignant la position prosodique (ici s = sommet de sonorité).

seg = tete =A

op = @

pos = 2

pc = s

Dans ce cadre, un attribut seg, pos ou pc n'ayant pas de valeur associée peut rester sous spécifié. Ainsi l'attaque de mot vide, correspondant à la position précédant le [a] de [ami] est spécifiée pour pos (on sait que c'est obligatoirement une position 1) mais elle reste sous-spécifiée pour seg et pc (on ne connaît ni sa composante en éléments, ni son degré exacte de sonorité). De la même manière la CL dite "flottante" est présente dans la représentation lexicale mais si elle est spécifiée pour seg (le segment flottant que va dégager par exemple le déterminant [ẽ] est obligatoirement un [n]), elle reste sous-spécifiée pour les attributs pos et pc tant qu'une position libre dans la chaîne ne lui permet pas d'être associée à une valeur pos et pc.

La liaison résulte alors de l'unification entre deux représentations lexicales et de la mise en place d'une contrainte de liaison unifiant les représentations (Angoujard,1997). Les erreurs systématiques constatées par le recours à une tâche de dénomination d'images confirment l'hypothèse que nous avons émise en observant les erreurs de production spontanées. Tous les mots produits sont automatiquement pourvus d'une consonne initiale. La CL est donc systématiquement interprétée comme une consonne lexicale. Ceci signifie donc que ces enfants ont des représentations lexicales ne comportant pas d'attaque vide et par voie de conséquence pas non plus des segments flottants. L'existence des segments flottants s'avère en effet non nécessaire puisque si tous les mots de la langue ont une attaque pleine, les segments flottants n'ont jamais de site leur permettant de se réaliser. L'absence d'attaque vide et l'absence de segments flottants chez ces enfants résulte en fait d'un même phénomène : l'absence de sous-spécification des positions. Leurs productions laissent penser qu'ils vont avoir spontanément tendance à tout spécifier et donc qu'ils vont au début de l'acquisition lexicale spécifier simultanément seg, pos et pc : tout segment ayant une spécification en éléments se verra spontanément attribuer également une position rythmique et une position de courbe et inversement. Or comme nous l'avons dit (cf supra), l'attaque vide et le segment flottant suppose une sous spécification partielle. L'explication que nous proposons est donc la suivante. Lors de l'acquisition lexicale, les enfants commencent par spécifier les positions en tout ou rien : chaque position reçoit une spécification simultanée pour seg, pos et pc d'où la production des énoncés erronés en contexte de liaison. Puis les enfants raffinent leurs paramètres phonologiques en constatant que les 3 types de spécifications ne sont pas nécessairement concomitantes et simultanées et mettent en place la possibilité de spécifications partielles ou sous-spécifications dans certaines positions. La sous-spécification de ces cas particuliers que sont les segments flottants et les attaques vides relèveraient de la paramétrisation propre au français. La maîtrise de la liaison serait alors l'indice d'un raffinement des représentations lexicales.

Cette hypothèse propose donc i) un premier état des représentations lexicales chez l'enfant où tous les mots de catégorie lexicale ont une attaque syllabique pleine, ii) une acquisition postérieure des représentations acceptant une syllabe initiale à attaque vide. Elle a le double avantage de proposer de ce fait un parallélisme entre

l'ordre d'apparition des structures syllabiques et leur degré de marque (Kaye et Lowenstamm, 1984), supposant que les enfants commencent tous avec des syllabes CV et que la première unité lexicale produite suit ce schème non marqué (unité CVCV).

Par ailleurs, elle confirme les expériences psycholinguistiques proposant chez l'enfant l'existence d'une routine de segmentation reposant sur les structures les plus fréquentes (Peters 1985).

Pourtant, Rialland (à paraître) souligne que, indépendamment du cadre théorique retenu, cette interprétation ne retenant qu'une forme de référence est problématique en ce qu'elle contredit les données empiriques existantes, laissant apparaître très précocement chez l'enfant francophone des mots commençant par une voyelle et cite plusieurs exemples (cf.[25]) publiés par Boysson-Bardies (1996).

[25]

[ɔ:ljo] *allo*, Léo (10 mois)

[haelo] *encore*, Léo (10 mois)

[awa] *au revoir*, Charles (1).

[ættæ] *attends*, Marie (1,3)

[a] “*assis*” Charles (1,3),

[ata] ou [atã] *attends*, Charles (1,8)

De même, Freitas (1996, 2003) a reporté l'existence précoce de mots à attaque vide en portugais (cf. chII.5) et proposé une analyse du phénomène.

J'en ai donc conclu que ma manière d'appréhender l'acquisition de la liaison à partir d'un seul type de données était problématique à ce stade, notamment parce que la prise de données a neutralisé la variation effectivement observable dans les corpus spontanés. Mais le cadre déclaratif retenu n'est pas ici en cause, il permettrait tout aussi bien de rendre compte des différentes réalisations du phonème de liaison par les enfants.



### 3.2.3.3. Deuxième interprétation : l'épenthèse

Afin d'expliquer la possibilité d'attaques vides précoces, j'ai donc proposé une interprétation alternative de mes données, sur la base des remarques de Rialland (à paraître) et introduit l'hypothèse d'une épenthèse consonantique, en retenant les 3 stades suivants.

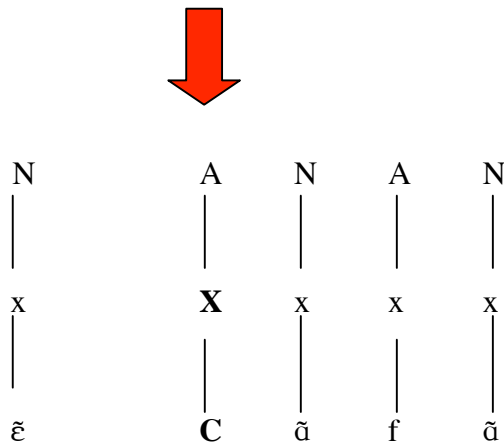
**Stade 1** : avant 2 ans et avant l'explosion lexicale, les mots sont produits de manière isolée et peuvent être réalisés avec et sans attaque vide

**Stade 2** : Autour de 2 ans et au-delà, l'unité de programmation augmente et commence à être resegmentée pour en extraire le déterminant. Les contraintes de syllabation vont donc s'appliquer en accord avec l'input reçu. Le résultat va produire une syllabation visant à satisfaire le principe du « maximum onset principle » engendrant les structures les moins marquées

***Contrainte de réorganisation syllabique : créer autant que possible des structures CV complètes pour syllaber l'unité prosodique lors de la construction du NP.***

Ceci amène donc l'enfant à remplir les attaques de mots lexicaux. Or, il se trouve confronté à une contradiction : une contrainte de syllabation implique le remplissage des attaques de mots alors que certaines entrées lexicales ont une attaque vide (éléphant, au-revoir, arrête, assis, attends) et qu'il les a mémorisées comme telles. Les fautes de liaison seraient alors l'indice d'une stratégie visant à s'accommoder de cette contradiction. Pour répondre à cette difficulté, l'enfant réalise alors l'épenthèse d'une consonne non spécifiée segmentalement [26] permettant la syllabation, et dont la nature est conditionnée par les occurrences disponibles dans sa langue. La nature de la consonne est, à ce moment de l'acquisition, logiquement déterminée par des inférences statistiques en fonction des contextes (dont les contextes de liaison) dans lesquels le mot a été rencontré.

[26]



A ce stade, l'enfant sait qu'il faut une consonne mais ne sait pas laquelle, il essaie par mécanisme d'épenthèse toutes celles qu'il a rencontrées dans ce contexte. Ceci expliquerait le pourcentage plus élevé de “ n ” proposés en attaque.

### **Stade 3 :** Autour de 3 ans 3,5 ans : disparition des erreurs de liaison

Alors que les unités lexicales commencent à se raffiner, on assiste à un phénomène de bootstrapping morphologique : l'enfant découvre que le déterminant « un » relâche une consonne nasale au féminin, que les adjectifs comme « petit », « grand », « gros » relâchent une consonne lors d'opérations d'inflexion (petite, grande, grosse) ou de dérivation nominale (petitesse, grandeur, grandir, grosseur, grossir). On peut penser qu'à partir du moment où il encode les consonnes flottantes dans les représentations lexicales sous-jacentes, il n'a plus besoin d'une consonne épenthétique pour syllaber le groupe clitique. A ce stade, les erreurs disparaissent logiquement. L'enfant sait qu'il faut une consonne et récupère celle qui est contenue dans la représentation sous-jacente du mot qui précède : la contrainte de remplissage des attaques vides résulte alors d'une stratégie transitoire de réparation avant que ne s'élabore la généralisation phonologique consécutive à l'encodage des segments flottants qui sera ensuite opératoire chez l'adulte.

### 3.3. Acquisition de la liaison : acquisition phonologique ou lexicale ?

Entre temps, le domaine s'était considérablement enrichi des données et propositions de Chevrot & Fayol (2000, 2001) qui m'ont amenée à chercher des données complémentaires. Il m'est apparu que la méthode expérimentale que j'avais retenue pour la prise de données gommait la variété des productions enfantines observable en productions spontanées.

J'ai recueilli un corpus longitudinal et des données transversales me permettant d'observer le tout premier lexique émergeant avant les fautes de liaison. Par ailleurs, j'ai commencé à cette époque des travaux i) sur les troncations et ii) sur l'acquisition des attaques syllabiques iii) sur l'acquisition des déterminants et ceci m'a obligée à réfléchir à la forme de la première unité lexicale produite par l'enfant. C'est également pour cette raison que j'ai adopté, à partir de là le cadre autosegmental et repris l'hypothèse gabaritique de Macken (1995) et Vihman (1993) pour l'acquisition. J'ai alors essayé de proposer une modélisation des représentations émergentes de l'enfant au sein d'un modèle unifié rendant compte simultanément de tous les phénomènes empiriques observables (je présente ces recherches dans le chapitre II, travaux en cours) dont l'acquisition de la liaison.

Ces nouvelles données, mes interrogations autour de la première unité produite par les enfants ainsi que mon désaccord avec les propositions faites par Chevrot & Fayol (2000, 2001) puis Dugua, (2002), Chevrot, (2003), Chevrot & Fayol, (soumis), Chevrot et al. (2005) ont ouvert un dialogue régulier avec ces auteurs sur cette question où nos propositions se discutent en temps réel parfois même dans le même volume (Chevrot et al. (2005), Wauquier-Gravelines & Braud, (2005)) et m'ont permis d'enrichir les analyses de l'acquisition de la liaison en français que j'avais proposées.

### 3.3.1. Données et analyse nouvelles

Dans leurs deux premières publications traitant de l'acquisition de la liaison, Chevrot & Fayol (2000, 2001) mentionnent, bien qu'ils ne les présentent pas de manière détaillée, des données d'acquisition de liaison d'une autre nature que celles que j'avais présentées et qui ne corroboraient que partiellement les miennes.

Contextes	3ans	7ans	11ans
<b>obligatoires</b>			
a) après déterminant (les ours)	77 %	97 %	100 %
b) clitique préverbal (j'en ai, ils ont)	91 %	100 %	99 %
c) expression figée (tout à coup)	100 %	100 %	100 %
<b>Total liaisons obligatoires</b>	<b>84%</b>	<b>99 %</b>	<b>99 %</b>
<b>facultatives</b>			
a) après adjectif épithète (un petit ours)	48%	-	-
b) après un nom pluriel (des enfants idiots)	2%	0%	0%
c) après auxiliaire avoir (ils ont un)	-	0%	0%
d) après auxiliaire être (ils sont ici)	0%	12%	20%
e) après verbe (il vient aussi)	1%	0%	5%
f) après mot invariable (quand un, en Espagne)	3%	24%	12%
<b>Total liaisons facultatives</b>	<b>5%</b>	<b>12%</b>	<b>12%</b>

Tableau 4 : Liaisons chez trois enfants (3, 7, 11 ans, d'après Basset (2000) in Chevrot & Fayol (2000)

Ils présentent les données de Basset (2000) (cf. tableau 4), ayant observé 2363 contextes de liaison chez des enfants de 3,7 et 11 ans et mentionnent l'étude longitudinale d'une petite Sophie entre 25 mois et 42 mois. L'analyse des données longitudinales (non publiées) apporte des arguments très consistants montrant qu'un même enfant utilise, pendant la période critique, plusieurs formes concurrentes simultanément ([ɛ̃lurs] [ɛ̃nurs]), dans une même séance d'enregistrement, et n'établit pas, comme je l'avais soutenu, sa représentation sur la base d'un seul input de référence. Ceci signifie que ma prise de données par un test de dénomination d'images a induit un type de réponse masquant la variété des réponses possibles pour un même enfant et entre les enfants.

De plus, les résultats d'une expérience psycholinguistique de syllabation induite leur permettent de montrer que les enfants, contrairement à ce qui a été montré

(Peters,1985), ne semblent pas toujours privilégier une routine de segmentation CVCV.

Chevrot & Fayol ont, sur cette base, proposé que l'acquisition de la liaison relève d'un apprentissage lexical progressant item par item et indifférent aux contextes prosodiques et morpho-syntaxiques de réalisation du sandhi. Ceci suppose que les enfants doivent apprendre au cas par cas chaque contexte où la liaison se réalise. Ils ont considéré que l'acquisition de la liaison

« progresse à travers des items lexicaux, ou des séquences d'items comme un apprentissage lexical plutôt que par la généralisation à l'intérieur de domaines syntaxiques [...] processus long progressant item par item dirigé par les données rencontrées dans l'environnement linguistique »

J-P Chevrot & M. Fayol : *L'acquisition de la liaison, enjeux théoriques*, LIDIL 22, 2000, p.26

Cette hypothèse lexicale a ensuite été reformulée à la lumière de la phonologie des exemplaires (Bybee 2001), de l'acquisition par induction et mémorisation (Tomasello, 2000). Pour la liaison ont été retenues les hypothèses de Morin (2003) et Côté (2005) et un « scénario développemental » a été envisagé Dugua (2002), Chevrot (2003), Chevrot et al. (2005) en trois étapes dont plusieurs versions ont été proposées.

Dans tous les cas, les auteurs maintiennent l'idée générale que l'acquisition se fait par des routines reposant sur des effets de fréquence et de répartition statistique des items dans la langue entraînant une mémorisation progressive des formes et des contextes d'association des deux mots entre lesquels se produit le sandhi (mot 1 et mot 2).

Le scénario développemental proposé dans sa dernière version (Chevrot et al. 2005) se compose de 3 étapes dont les deux premières sont des reformulations des propositions faites par Morin (2003). La troisième étape défend l'hypothèse d'une CL comme une épenthèse réalisée sur le mot 2.

### **Etape 1**

Les enfants vont procéder à une segmentation erronée des unités lexicales, consécutive à l'ambiguïté de frontière gauche créée par la liaison : ils vont donc stocker plusieurs variantes distinctes d'un même mot. Alors que l'adulte disposera d'une seule représentation du mot « ours » /urs/, l'enfant stockera pendant un certain moment des variantes /nurs/, /zurs/, /turs/ l'amenant à produire « le nous », « le tours », « le zours ».

### **Etape 2**

Alors qu'il dispose d'un certain nombre d'exemplaires erronés du mots « ours » comportant une consonne à l'attaque, l'enfant doit apprendre à apparier correctement les suites mot 1 + mot 2 (soit par exemple « un » + « nous », « deux » + « zours » etc). Selon les auteurs, cet apprentissage est conditionné par l'exposition aux séquences mot1-mot2 bien formées rencontrées dans l'environnement langagier. C'est donc l'usage du langage en réception qui contribuerait principalement à structurer ce réseau d'associations.

### **Etape 3**

La troisième étape est caractérisée par l'apprentissage de la non-application du processus d'épenthèse. Les auteurs considèrent en effet qu'une fois que l'enfant a stabilisé des formes alternantes et qu'il sait les convoquer en contexte, il découvre que la consonne n'appartient ni au mot 1 ni au mot 2 mais qu'elle est insérée par épenthèse, donc qu'elle peut ou non être placée sur l'attaque du mot 2. La résolution des erreurs de liaison serait consécutive à cette maîtrise du processus d'épenthèse permettant à l'enfant d'éviter l'insertion quand elle n'est pas pertinente.

Je soutiens au contraire que les erreurs d'acquisition de la liaison relèvent d'une analyse phonologique. Je postule que les enfants se servent d'une grammaire structurée en voie de paramétrisation et utilisent des généralisations fondées sur ces connaissances des contraintes phonologiques universelles gouvernant le fonctionnement des langues pour réaliser la segmentation et construire une représentation phonologique abstraite du sandhi. Je considère par conséquent que les erreurs constatées plus haut sont les traces de la mise en place progressive de cette généralisation morpho-phonologique.

### 3.3.2. Critique de l'approche lexicaliste de la liaison

Comme je l'ai dit plus haut, les données de Basset (2000), les remarques de Chevrot & Fayol (2001) sur l'étude longitudinale de Sophie, ainsi que le scénario développemental proposé par Chevrot et al. (2005) permettent d'affiner la chronologie d'acquisition de la liaison.

Pour la liaison obligatoire, comme nous l'avons montré, des erreurs apparaissent à partir de deux ans. A trois ans, les erreurs sont encore significativement présentes, à 5 ans elles ont pratiquement disparu, mais Chevrot et al. (2005) montrent également que des erreurs résiduelles relevant de généralisations peuvent également se produire « un nèbre » pour « un zèbre » jusqu'à 6 ans. Pour la liaison obligatoire, nous observons donc un phénomène assez régulier s'inscrivant dans une chronologie peu variable d'un enfant à l'autre. Cette régularité plaide par conséquent en faveur d'une acquisition de type phonologique non contrainte par l'usage. Car si l'acquisition des liaisons, y compris des liaisons obligatoires, était exclusivement contrainte par l'usage, on observerait une variabilité beaucoup plus importante des erreurs et une chronologie beaucoup plus floue.

L'observation des données (Braud, 1998, Ferré 1997, Basset, 2000) montre très clairement que la mise en place de la liaison obligatoire et de la liaison facultative ne suivent pas la même chronologie : l'acquisition de la liaison obligatoire est très

systematique et complètement maîtrisée quand s'achève l'acquisition phonologique, vers 3/4 ans alors que l'acquisition des contextes de liaisons facultatives est très variable et idiosyncrasique.

Or, l'approche lexicaliste ne permet pas de distinguer l'acquisition des liaisons obligatoires et l'acquisition des liaisons facultatives, et le raisonnement consistant à arguer de la variété des productions des enfants sur les contextes obligatoires pour défendre un apprentissage lexicaliste au cas par cas, me semble reposer sur un raisonnement biaisé consistant à rendre compte du général en s'appuyant sur l'observation du spécifique.

Si l'on peut admettre que l'enfant apprenne item par item les conditions de réalisation des liaisons facultatives, ceci ne prouve pas qu'il n'a pas opéré auparavant, sur les contextes obligatoires, une généralisation phonologique qui s'appuie sur une représentation abstraite. Ce qu'il apprend ensuite, c'est à appliquer cette généralisation item par item sur les contextes facultatifs. La chronologie proposée par Basset (2000) décrit nettement cette tendance. Elle amène clairement à conclure qu'à 5 ans, les liaisons obligatoires sont complètement acquises, alors que les liaisons facultatives sont encore problématiques à 11 ans.

Chevrot et al. (2005) rendent justice à cet argument en supposant, qu'à un moment, les enfants vont faire des généralisations et qu'ils ne peuvent pas attendre d'avoir entendu au moins une fois tous les contextes de liaison obligatoire pour les produire. Ils proposent donc qu'après un stade de mémorisation des séquences mot 1 + mot 2, considéré dans une version antérieure comme le processus explicatif de la résolution des erreurs de liaison, les enfants formalisent un processus épenthétique dont la maîtrise expliquerait la disparition des fautes. Cependant le lien exact qui amène les enfants à supposer à partir de leur mémorisation des contextes adéquats i) que la CL est placée sur mot 2 et épenthétique ii) qu'il existe des contextes où elle est



réalisée et d'autres où elle ne l'est pas, reste mystérieux. On ne comprend pas en particulier pourquoi ni comment le stockage mémorisé des séquences mot 1 + mot 2 entraîne une généralisation de type phonologique. On ne voit pas non plus pourquoi cette généralisation est une épenthèse, ni pourquoi c'est une épenthèse à droite sur l'attaque du mot 2 ce qui présuppose que la liaison non-enchaînée n'est pas possible.

Le seul avantage qu'on peut trouver à cette hypothèse est qu'elle permet de justifier le contenu segmental de l'épenthèse et d'en limiter apparemment l'arbitraire que nous avons souligné quand nous avons présenté cette conception de la liaison pour le traitement chez l'adulte (cf. 2.4.2.2). Si les enfants ont mémorisé systématiquement « un + néléphant » et qu'ils resegmentent en « un + (n) éléphant » ils donnent effectivement un contenu segmental à l'épenthèse.

Mais deux objections peuvent être faites. Tout d'abord, la résolution du contenu segmental que prend l'épenthèse ne retire rien à l'arbitraire du processus épenthétique en soi, et si l'on doit supposer que les auteurs le retiennent au final comme ce qui aboutit au modèle de l'adulte, pourquoi insérer et pourquoi là ?

Ensuite cette hypothèse prédit, quoi qu'en disent les auteurs, que les locuteurs ne pourront pas traiter les contextes, même les contextes obligatoires, qu'ils n'auraient pas rencontrés dans l'enfance. Si la resegmentation de « un + nanachorète » en « un + (n) anachorète » qui permet au sujet de produire la suite sans faute résulte de la mémorisation du « bon contexte » « un + anachorète », le locuteur-auditeur rencontrant pour la première fois ce contexte à 35 ans, bien après la phase d'acquisition, ne devrait pas pouvoir traiter cette information, puisque, dans ce scénario développemental, le contenu segmental de la consonne résulte du stockage lors des périodes d'acquisition et n'est pas justifié par sa nature phonologique propre (appartenance au mot 1, « n » étant un segment flottant). Dès lors qu'elle n'a pas été stockée précocement, la consonne ne devrait pas être disponible pour un adulte, or

tout adulte sait traiter en temps réel (perception /production) un contexte de liaison obligatoire qu'il n'a jamais rencontré (je pense que le phénomène est productif et que jamais un adulte francophone de langue maternelle ne produira « un zanachorète »). Ceci oblige à envisager que l'acquisition des liaisons obligatoires dure toute la vie, alors que les données montrent qu'à 4/5 ans le problème est réglé pour ce contexte, et que les seuls contextes qui sont sujets à variation et peuvent être l'objet d'erreurs y compris chez les adultes (cuirs, velours, pataquès) sont les contextes facultatifs (réalisation/ non réalisation, enchaînement/ non-enchaînements).

On voit par ailleurs comment ce « scénario développemental » apparaît assez compliqué et peu logique puisque l'hypothèse retenue considère que les enfants opèrent une première segmentation séparant les mots, puis les réassemblent puis les resegmentent. La logique phonologique ou linguistique qui serait à l'œuvre dans ce « scénario développemental » me semble très opaque. L'existence de ces trois étapes ainsi que leur enchaînement chronologique me paraissent arbitraires, descriptifs, non explicatifs et jamais fondés sur l'existence d'acquisition linguistique (bootstrappings) ou psychologique parallèles (mémoire de travail, faculté perceptives, capacités de traitement). On ne voit pas dès lors en quoi ce scénario « développemental » est motivé, puisque ce terme suggère implicitement un modèle psychologique maturationnel dont on ne voit absolument pas trace dans les analyses proposées.

Enfin, phonologiquement parlant, la logique retenue me paraît fondamentalement en contradiction avec les prédictions que font les modèles dont elle se réclame. Les modèles exemplaristes (Bybee, 2001, Pierrehumbert, 2002) supposent chez l'adulte un stockage supplétif des représentations phonologiques (types) fondé sur la fréquence en contexte des réalisations de surface (des « tokens »). Sur cette base on peut supposer que les prédictions de ces modèles pour l'acquisition considèrent

que la segmentation et le stockage des enfants sont conditionnés par la fréquence d'usage des mots en contexte de liaison. Or les contextes de réalisation les plus fréquents sont tous des contextes considérés comme « obligatoires » (Boula de Mareuil et al., 2003) où la liaison est toujours réalisée et toujours enchaînée et jamais fautive. Ceci implique donc que les enfants entendent toujours [ɛ̃nurs / lezurs] rarement [urs] [ɛ̃n # urs] jamais [nurs / zurs / ɛ̃zurs / lenurs] et que si leurs productions étaient directement conditionnées par l'input, ils ne devraient faire aucune faute sur les liaisons obligatoires à aucun moment de l'acquisition. Les seuls contextes où on devrait voir apparaître des erreurs seraient les contextes variables, peu fréquents, non enchaînés, c'est à dire tous les emplois qui manifestent un choix des locuteurs et reflètent effectivement l'usage. Or les données montrent le contraire, le contexte où l'on constate le plus régulièrement et le plus fréquemment des erreurs est le contexte le « plus obligatoire » qui soit et le moins variable, le contexte déterminant-nom. Si, comme le retiennent Chevrot et al. (2005), à partir du modèle de Tomasello (2000)

« Les énoncés du jeune enfant ne seraient pas organisés en catégories et en structures abstraites, mais agencés concrètement autour de séquences et de mots particuliers. Dans le modèle d'acquisition qui en découle, les changements développementaux sont guidés par l'usage. »

M. Tomasello (2000), The item-based nature of early syntactic development, *Trends in Cognitive Science*, **44**

pourquoi la première segmentation (étape 1 du scénario développemental) retiendrait-elle des formes qui ne sont quasiment jamais disponibles dans l'input ? Il faut que cette première segmentation repose alors sur une analyse qui amène les enfants à détacher le déterminant de manière précoce et on voit mal comment ils pourraient le faire sur la base de l'input seul où le déterminant n'est jamais détaché, et sans l'appui de ressources morpho-syntaxiques précoces. Les routines de segmentation précoce appuyées sur l'input vont nécessairement amener l'enfant à segmenter le déterminant pro-clitique avec le nom, même s'il retient une routine CVCV (et pourquoi celle-ci

d'ailleurs, car dans l'input en français il entendra beaucoup d'attaques vides, y compris en début de mot).

Au final, en retenant une solution qui assimile stockage des exemplaires à généralisation phonologique et où la mémorisation de l'information devient un principe explicatif, Chevrot et al. (2005) appuient leur analyse sur un modèle dont les prédictions contredisent l'existence des phénomènes qu'ils prétendent expliquer.

Ceci montre, enfin, que les modèles exemplaristes font de fausses prédictions pour l'acquisition de la liaison (cf. ch III).

### **3.4. Nouvelle analyse : la liaison comme un cas de variation**

Afin d'évaluer plus précisément la part de variation dans le processus d'acquisition de la liaison obligatoire et l'explication formelle qu'on pouvait en donner, j'ai recueilli des données longitudinales qui m'ont permis d'inscrire l'observation des erreurs dans une perspective développementale (apparition, pleine période et régression).

#### 3.4.1. Données longitudinales.

Les données longitudinales font apparaître effectivement une variation (synchronique) par tranche d'âge mais également une évolution diachronique des formes produites pour un même mot.

On peut trouver des formes comportant des attaques vides [27], des formes nominales isolées à attaque vide [28], des formes nominales en contexte de liaison avant deux ans sans consonne [29], des formes nominales en contexte de liaison avant deux ans avec consonne [30].

[27]

[aəwa] *au-revoir*, Claire (1,4)

[aɛ] *arrête*, Claire (1.5)

[28]

[əma] *fromage*, Claire (1,7)[itɔ:] *Victor*, Claire (1,8)[oije] *Olivier*, Claire (1,9)[an] *âne*, Claire (1,8)[eã] *éléphant*, Claire (1,11)[aty ] *voiture*, Claire (1,11)

[29]

[epətɪaty ] *les petites voitures*, Claire (1,11)[æpətɪəma] / [æɣ<sup>R</sup>oəma] *un petit / un gros fromage*, Claire (1,11)[æpətɪija] *un petit coquillage*, Claire (2)[pətiuəjɛ̃] *Petit Ours brun*, Claire (2)

[30]

[lepətɪaty ] *les petites voitures*, (Claire 2,1)[æpətɪija] *un petit coquillage*, (Claire 2)

[31]

[grotut, grojut, pətɪjut, pətɪtut, lɔtus, letus, ~~lɔtɔ~~ <sup>R</sup>ɔtɔjus] *gros ours, petit ours, l'ours, les ours, les trois ours* Claire (2,2)

[lepətɪtan, lətan, letan] *l'âne, les ânes*, Claire (2,4)

[32] des formes qui n'ont pas été « apprises » puisqu'elle n'ont pas été données par l'input.

[lejajo, lezjajo, ɛzjajo] *les oiseaux, un oiseau*, Claire (2,4)[ɛfɛfɛfã] [ɛlefã] [defɛfɛfã] *un éléphant, des éléphants* Claire (2,5)

input adulte

[ɛnaɾkãsjɛl] *un arc-en-ciel*[laɾkãsjɛl] *l'arc-en-ciel***output Claire (2,5)**[vœvwaləmakãsjɛl] *je veux voir l'arc-en-ciel*[ləpakãsjɛlamwa] *l'arc-en-ciel à moi*

input adulte :

[lɔpapaʃəval setɛnetalɔ] *Le papa cheval c'est un étalon*

**output Claire (2,5)**

[ɛletalɔ]

Ces données permettent effectivement d'observer une variété importante dans la production d'un même enfant. On voit d'abord

i) que certains mots sont produits avec attaque vide, sans déterminant [27] [28], [29] de manière précoce (entre 16 et 20 mois), avant l'explosion lexicale. On remarque que ces productions sont des mots à référent limité ou unique (noms propres, « au revoir ») dont on peut penser qu'ils sont mémorisés en contexte. Ils comportent des attaques vides en lieu et place des attaques branchantes ou des fricatives (non encore acquises à cette période) dans la cible adulte et qui peuvent être tronqués ([ija] pour coquillage). Ils sont constitués quasi exclusivement de voyelles [eã].

ii) que, à cette même période, certains de ces mots tronqués [29] en contexte de liaison « un petit + mot tronqué » voient s'ajouter une consonne [æpətijja], *un petit coquillage*.

iii) que, les erreurs « de liaison » à proprement parler consistant à produire une consonne fautive en lieu et place de la consonne flottante attendue, apparaissent massivement, mais vers 24-25 mois, [31] [32], à l'époque de l'explosion lexicale et sur des NP (det + nom). Cette consonne peut être une consonne possible dans l'input ([ɛlurs] ou [ɛzurs] pour [ɛnurs]), soit une consonne qui n'est pas dans l'input. Pour les données de Claire, yod est employé très régulièrement (pour l'analyse de ce phénomène, cf ch. II), mais également des consonnes harmonisées ([ɛfɛfɛã], pour « un éléphant »).

Il semblerait donc que les erreurs de CL apparaissent massivement quand la productivité lexicale des enfants augmente, qu'elles apparaissent sur le NP [det + nom], qu'elles sont consécutives à l'apparition du contexte prosodico-syntaxique constituant le domaine de réalisation de la liaison obligatoire et qu'elles ont un lien

avec l'acquisition du déterminant proclitique et la détermination des contextes gauches pour les mots commençant par une voyelle.

### 3.4.2. Acquisition de la liaison : détacher deux mots

Les analyses que j'avais produites jusqu'alors m'opposaient à Chevrot sur les points suivants. J'avais retenu une analyse phonologique dans un modèle autosegmental s'opposant à une analyse lexicale supplétive, et j'avais rattaché la consonne au mot 1 alors qu'il considère qu'elle est stockée sur mot 2. Mais nos analyses partageaient pourtant la caractéristique suivante : dans les deux cas, l'acquisition de la liaison avait été envisagée comme l'apprentissage des contraintes régissant la frontière résultant de l'ajout de mot 1 et mot 2. Lui comme moi avons postulé implicitement que les enfants manipulent librement des unités déjà segmentées (mot 1 et mot 2) et que les erreurs reflètent des hésitations quant à la forme du sandhi qui se réalise entre les deux.

Je me suis alors demandé si nous n'avions pas posé le problème à l'envers et si les erreurs de liaison ne pouvaient pas résulter du détachement de mot 1 et mot 2 et de la création de la frontière entre le déterminant et le nom. J'ai donc cherché des données transversales permettant d'observer la forme des premières unités produites au moment du début de l'explosion lexicale.

#### 3.4.2.1. Premières formes produites

J'ai observé sous cet angle 1061 items qui ont été recueillis au cours de conversations spontanées et semi-induites en milieu pré-scolaire (crèche) avec 6 enfants entre 16 mois et 2,5 ans. Les mots se répartissent en 3 catégories (tableau 5).

i) La catégorie substantifs : substantifs présentés seuls (*tortue / une tortue / la tortue*) ou avec un présentatif<sup>11</sup> (*c'est une tortue, voilà la tortue, / c'est tortue, voilà tortue*).

---

<sup>11</sup> Bien qu'on puisse considérer que dans ce cas le substantif est en situation prédicative, nous avons présenté par commodité ici ces productions dans la catégorie des substantifs.

ii) les items qui comportent un verbe conjugué ou non avec une structure argumentale même minimale (*j'en ai à la maison, moi, oui, ils sont beaux, il est tout doux le lézard, arrête*).

iii) La catégorie « autres » : les adverbes (ex : *oui, non, encore*) ainsi que les quelques adjectifs qualificatifs produits isolément ou avec un présentatif *bleu, c'est bleu*.

<i>substantifs</i>	<i>IP</i>	<i>autres</i>
64 %	21%	15%

Tableau 5: répartition des items

Pour les substantifs, 4 catégories ont été distinguées

NCS= nom commun seul,

NCD = nom commun avec déterminant,

PRS = présentatifs seuls,

PRD= présentatifs avec déterminants

<i>NCS</i>	<i>NCD</i>	<i>PRS</i>	<i>PRD</i>
28%	72%	41%	59%

Tableau 6: répartition des substantifs

On voit que les enfants produisent prioritairement des substantifs (64%) en leur associant la plupart du temps un déterminant (NCD=72% PRD=59%<sup>12</sup>)

Si l'on rapproche ces données de données longitudinales de Claire [33], on observe également la présence d'un "proto-déterminant" sur les noms communs avant deux ans.

[33]

ÂGE	CIBLE	FORME PRODUITE
01;10 ;03	<i>le chien</i>	[lótɛ̃]
01;10,04	<i>la vache</i>	[əja ]
01;10,04	<i>la vache, l'âne</i>	[laja/lətan]
01;10,06	<i>le pot</i>	[ləpo]
01;10 ;03	<i>le bébé</i>	[ləbebe]
01;10,04	<i>le ballon</i>	[abalɔ̃]

<sup>12</sup> La différence entre ces deux situations peut être attribuée à la situation prédicative du nom dans des constructions comme « c'est + subst ». Nous ne souhaitons pas entrer dans le détail de cette analyse ici



01;10 ;03	<i>un éléphant</i>	[ɔɛjã]
01;11,06	<i>le hérisson</i>	[ləijɔ̃]
01;11,05	<i>la brosse à dent</i>	[anadã]
01;10,06	un médicament	[apamã]

Braud (2003) a également produit des données transversales et elle montre que sur les mots longs (3 et 4 syllabes) dont les syllabes initiales ont été tronquées, le déterminant est néanmoins maintenu.

Donc si l'on observe des données d'acquisition précoces (Braud, 2003, Wauquier-Gravelines & Braud, 2005), on constate que les premiers items produits par les enfants sont en priorité des substantifs et qu'ils sont produits très souvent avec un proto-déterminant complètement intégré à l'unité lexicale. Ce déterminant peut prendre diverses formes, voyelle ou syllabe épenthétique ([lə] [ə] [le] [ɔ] [a] [y]), et il est réalisé même quand les premières syllabes du mot ont été tronquées ([ɛ̃potam], *un potam* pour *un hippopotame* Braud, 2003). Il semblerait donc que mot 1 et mot 2 ne forment qu'une unité, n'ont pas d'existence autonome pour l'enfant apprenant le français et que le proto-déterminant ne soit pas initialement segmenté (Wauquier-Gravelines 2003, Wauquier-Gravelines & Braud, 2005). La « forme de référence » au premier stade d'acquisition semble inclure la plupart du temps le proto-déterminant.

#### 3.4.2.2. Acquisition de la liaison : nouvelle hypothèse

La mise en évidence de la présence précoce d'un déterminant non segmenté m'a amenée à réinscrire l'hypothèse d'acquisition de la liaison dans une chronologie plus large en 4 stades.

##### **Stade 1**

La plupart des substantifs sont produits avec un déterminant non segmenté. A ce stade, l'enfant dispose d'un lexique très restreint (stade des 50 mots, Vihman, 1996). Il travaille sur la base d'une forme lexicale où les attaques consonantiques sur la frontière gauche et en interne de mot peuvent être indifféremment vides ou remplies. La représentation de l'unité est certainement une forme « globale » non

analysée. Les données de perception confirment ce fait (Hirsch-Pasek & al., 1987, Jusczyk & al., 1992). On ne constate pas à ce stade, « d'erreurs de liaison » à proprement parler.

## Stade 2

Les erreurs de liaison semblent bien apparaître et se systématiser au moment de l'explosion lexicale (autour de deux ans). L'augmentation rapide et la mise en place du lexique amènent l'enfant à opérer une reconnaissance plus fine des unités en vue d'établir les formes canoniques des mots. A ce stade, un choix doit être fait quant à la segmentation de suites comme par exemple [sedezeleḞā], [uiletōneleḞā] et la détermination de la forme de l'unité lexicale et du déterminant.

Les données que nous avons présentées suggèrent que lors de ce découpage, les enfants vont segmenter d'abord en s'appuyant sur les unités phonologiques les moins marquées (Peters 1985), et qu'ils vont autant que possible remplir les attaques vides pour dégager des schémas CV à l'attaque de mot. L'existence de telles resegmentations lexicalisées ("Nounours") en atteste également. Leur découpage est réalisé en accord avec le « Maximal Onset principle » (MOP).

Je propose que les erreurs de liaison soient l'une des manifestations la plus précoce de l'indépendance des positions prosodiques et des contenus segmentaux qui y sont rattachés. Lors de la segmentation du déterminant, les enfants vont, en accord, avec le MOP privilégier le rattachement de la position prosodique à l'attaque de mot 2 pour maximaliser l'attaque au détriment de la coda précédente.

La nature segmentale de la consonne n'est pour autant pas définitivement fixée à ce moment de l'acquisition et semble effectivement déterminée le plus souvent par des inférences statistiques en fonction des contextes (dont les contextes de liaison) dans lesquels le mot a été rencontré, mais yod [j] peut être également utilisé par défaut (Claire 24 mois), ou le choix peut résulter d'une harmonie consonantique ([f] dans le cas de éléphant [ēfefeḞā] pour Claire 27 mois) ou être strictement idiosyncrasique, comme dans le cas de [m] ou [p] pour « arc-en-ciel » [ləmakūsjeḞ], [ləpakūsjelamwa].

L'enfant a recours à l'utilisation d'une consonne qu'il a entendue en contexte. Ceci expliquerait le pourcentage plus élevé de "n" proposés en attaque (Braud & Wauquier-Gravelines, 1999) ou  $C = \{n,t,z,l\}$

[34]		↓				
	N	A	N	A	N	
	x	<b>X</b>	x	x	x	
	ẽ	<b>C</b>	ã	f	ã	

Un yod [j] peut être également utilisé par défaut (Claire, 2)

[35]		↓				
	N	A	N	A	N	
	x	<b>X</b>	x	x	x	
	ẽ	~	j	a	j	o

L'harmonie consonantique ([f] dans le cas de éléphant [ẽfɛfɛfã] pour (Claire 2,3)

[36]		↓					
	N	A	N	A	N	A	N
	x	<b>X</b>	x	x	x	x	x
	ẽ	<b>f</b>	e	f	e	f	ã

Cette consonne peut être idiosyncrasique, comme dans le cas de [m] ou [p] pour « arc-en-ciel » [lãmakũsjɛl], *le marc-en-ciel*, [lãpakũsjɛlamwã], *le parc-en-ciel à moi*

[37]		↓						
	N	A	N	A	N	A	N	A/C
	x	<b>X</b>	x	x	x	x	x	x
	ẽ	<b>p</b>	a	k	ã	sj	ɛ	l

On pourra dès lors concevoir la variation exprimée par des *learning paths* spécifiques comme ne reflétant que des variations de surface. Dans tous les cas, l'attaque syllabique est remplie et la variation ne touche que l'interprétation segmentale de la consonne. Elle est la manifestation variable d'un principe général de marque (maximalisation des attaques) lors de la résolution d'une ambiguïté de segmentation.

Dans une perspective développementale, la variation peut donc s'interpréter comme l'expression de stratégies différentes et spécifiques pour arriver au même but. Le listage exhaustif de toutes les occurrences rencontrées n'est une condition ni nécessaire ni suffisante pour expliquer l'émergence d'une grammaire fondée sur l'usage. La phonologisation et la stabilisation des formes lexicales ne supposent pas un stockage des formes lexicales erronées mais la production de ces formes erronées reflète plutôt, selon nous, une heuristique visant à résoudre un problème morpho-phonologique particulier (ici la construction de la structure syllabique et la segmentation déterminant-nom, la cliticisation du déterminant).

### **Stade 3**

Ce stade est identique au stade 3 proposé en 3.2.3.2.. La disparition des erreurs de liaison est consécutive à un phénomène de bootstrapping morphologique qui permet d'encoder la consonne flottante dans les représentations lexicales sous-jacentes.

### **Stade 4**

Malgré cette régularisation des liaisons obligatoires, on peut voir subsister longtemps des erreurs portant sur les liaisons facultatives et non enchaînées (Basset 2000). Pour ce qui concerne ce type de liaison, j'avais effectivement considéré à ce stade de mes analyses, puisque les réalisations ne sont pas systématiques, qu'il s'agit

d'un apprentissage lexical au cas par cas (Chevrot & Fayol, 2001) et que la fréquence d'occurrence va déterminer la disponibilité de certaines réalisations (Bybee, 2001).

### 3.4.3. Conclusion

A ce stade du travail, la production de données et les diverses analyses et contre-analyses que j'en avais faites dans un dialogue constant avec une hypothèse lexicaliste me permettaient d'établir assez clairement un certain nombre de faits et de répondre aux questions posées en exergue

- i) Est-ce que cette période apparaît chez tous les enfants ?
- ii) A quel âge ces erreurs apparaissent puis se résolvent-elles ?
- iii) Dans quelles circonstances (et quels contextes liaisons obligatoires vs facultatives, enchaînées vs non-enchaînées) ?

Oui, on peut constater l'existence d'une période critique entre 2 et 5 ans pour l'acquisition des liaisons obligatoires présentes chez la plupart des enfants francophones et s'inscrivant dans une chronologie et une logique maturationnelle assez stricte. Ce qui plaide selon moi en faveur d'une acquisition conditionnée par des structures innées ou très précocement disponibles et non conditionnées par l'input. Les erreurs sur les liaisons facultatives apparaissent plus tard et se résolvent également plus tard. Elles s'inscrivent dans une chronologie beaucoup plus lâche, ce qui m'avait amenée à concéder qu'elles pouvaient effectivement comme le propose Chevrot & Fayol (2000, 2001) être apprises au cas par cas par mémorisation.

Les travaux actuels, à la lumière de la reformulation du flottement et de la non-association proposée par Encrevé & Scheer (2005) m'ont amenée à reconsidérer cette proposition qui envisage la variation comme un phénomène lié aux stratégies d'acquisition mais qui en néglige la dimension inhérente (un même enfant produit des formes alternantes au cours d'un même enregistrement de prise de données).

Rendre compte de l'acquisition de la liaison consiste en effet à analyser comment les enfants développent des stratégies variables sur un phénomène lui-même variable. La

question de la variation doit donc être posée en perspective, comme « au carré ». Je m'y emploie au chapitre II, 2.

Je défends également l'hypothèse selon laquelle l'acquisition de la liaison obligatoire se fait dans le contexte (NP : det + nom) qui est le premier et seul contexte obligatoire où les erreurs apparaissent au moment de l'explosion lexicale. Ceci plaide en faveur d'une acquisition liée à l'acquisition du déterminant pro-clitique conditionnée par un bootstrapping morphologique, même si ces phénomènes reflètent un lien étroit entre acquisition phonologique et lexicale (je développe cette réflexion de manière détaillée au chapitre IV).

Des données longitudinales et transversales plus précoces (10-16 mois) et plus nombreuses sont à ce stade nécessaires pour affiner plus précisément la compréhension du lien entre acquisition du déterminant et acquisition de la liaison.

## **Travaux traitant de ces questions**

### **Publications**

Ferré S., & **Wauquier-Gravelines, S.** The processing of phonological information by bilingual subjects : liaison in French ». *Proceedings of XIVth International Congress of Phonetic Sciences*, San Francisco, août 1999.

**Wauquier-Gravelines S.** (2003), La question du réalisme des formalisations phonologiques contemporaines : que nous apprennent les données d'acquisition" in Wauquier-Gravelines S. & Angoujard, J.P. (eds), *Phonologie : Champs et perspectives*, ENS / éditions, Lyon

**Wauquier-Gravelines, S.** & Braud, V. (2005), Proto-déterminant et acquisition de la liaison obligatoire en français », *Revue Langage, Nouvelles approches de la liaison*, n° coordonné par Laks B., Chevrot J.P. & Fayol M.

**Wauquier-Gravelines, S.** (soumis), L'acquisition comme un cas de variation, l'exemple de la liaison en français, *Langue française*, n° coordonné par J. Brandao de Carvalho & Marc Klein

### **Communications**

**Wauquier-Gravelines, S.** & Ferré S. (1999), Effets du bilinguisme sur la paramétrisation des segments flottants : exemple du français. *VIIIth International Congress for the study of Child Language*, San Sebastian, juillet 1999

- Ferré S. & **Wauquier-Gravelines, S.** (1999), The processing of phonological information by bilingual subjects : liaison in French. *XIVth International Congress of Phonetic Sciences*, San Francisco, aout 1999
- Wauquier-Gravelines, S.** & Braud, V. (1999), Acquisition phonologique en français, *VIIIth International Congress for the study of Child Language*, San Sebastian, juillet 1999.
- Wauquier-Gravelines, S.** (2000), Réalisme et validité cognitive des modèles phonologiques : que nous apprennent les données d'acquisition ?" Conférence invitée, *Journées du GDR de Phonologie*, 26 juin, 2000, EHESS, Paris.
- Wauquier-Gravelines, S.** (2001), Contraintes d'acquisition sur le traitement des segments flottants en français : le cas de la liaison", *Conférence au colloque « Représentations, contrainte et stratégie de réparation »*, organisé par le GDR 1954, Toulouse 15/16 novembre 2001
- Wauquier-Gravelines, S.** (2002), Statistical learning or phonological generalisation The case of acquisition of liaison in French, *Invited Conference, Special Session « Acquisition » of Tenth Phonology Meeting of Manchester*, 23 mai 2002
- Wauquier-Gravelines, S.** (2002), Segmentation de la chaîne parlée et détermination des frontières gauches dans l'acquisition du lexique précoce, *Colloque du GDR de Phonologie*, Grenoble, 6-8 juin 2002
- Wauquier-Gravelines, S.** (2002), Realism of Constraints in the Acquisition of Liaison in French, *Ind NaPhC*, Concordia University, Montréal, 28 avril 2002
- Wauquier-Gravelines S.** (2003), L'acquisition comme un cas de variation, exemple de la liaison en français, *Colloque PFC, Phonologie et phonétique du français, données et théories*, Paris, 11-13 décembre 2003

#### **4. Troncations et réductions.**

Lors des premières expériences sur la liaison chez les enfants de 3 ans et 5 ans, j'avais introduit un facteur "longueur de mots" de manière un peu prospective. La liaison étant contrainte par un facteur prosodique résultant de l'ajout et de la redistribution de positions, j'avais envisagé que la liaison puisse être conditionnée par la longueur du mot 2.

Les données ont effectivement confirmé cette intuition, mais en la détachant de l'hypothèse initiale et au final, la problématique s'est ensuite élaborée en marge des travaux sur la liaison. La manière dont les enfants gèrent la longueur des mots, coupent, modifient, redistribuent le matériel segmental, dans les premières unités offre une fenêtre intéressante sur l'émergence des représentations phonologiques.

Je présente ici les analyses des coupes et redistributions segmentales (réductions et harmonies consonantiques) opérées dans les mots trisyllabiques et quadrisyllabiques testés en contexte de liaison et dans les mots contrôles de cette même expérience (présentée en 3.2.1) puis les analyses des phénomènes concomitants de troncation et réduction que j'ai réalisées sur des données longitudinales complémentaires.

##### **4.1. Troncation des mots longs en dénomination d'images (données obtenues sur la liaison)**

Je présente en [38] les items proposés et les productions des enfants obtenus dans l'expérience de production de mots tri et quadrisyllabiques en contexte de liaison.



[38]

Items	ELEX	Items	ELEX
<b>trisyll</b>		<b>quadrisyll</b>	
• escargot	[ĕkargo] [ĕsargo] [ĕsago]	• aspirateur	[ĕpasatæR] [ĕraatæR] [ĕrastatæR]
• ananas	[ĕnana] [ĕnana] [ĕnas]		[ĕpiratæR][ĕpisratæR] [ĕspiratæR]
• artichaut	[ĕtiʃo]	• épouvantail	[ĕpuvĕtaj] [ĕpupuvĕtaj] [ĕvĕtaj]
• éléphant	[ĕnefĕ] [ĕlefĕ]	•	[ĕpipopotam] [ĕpopotam] [ĕpotam]
• arrosoir	[ĕROZWAR] [ĕRAROZWAR]	hippopotame	[ĕkORdeĕ] [ĕkORneĕ] [ĕRARajĕ]
• écureuil	[ĕWAR]	• accordéon	[ĕreabl] [ĕpermeabl]
	[ĕkyroej]	• imperméable	[ĕnikotER][ĕnikoptER]
		• hélicoptère	[ĕkoptER][ĕkotER]

A partir des erreurs relevées, nous pouvons faire les remarques suivantes :

On voit d'abord que tous ces mots sans exception ont été produits avec une attaque pleine alors que ce sont tous des mots à initiale vocalique, comme si la liaison avait été faite malgré la troncation, comme pour les mots courts. Les enfants les ont en effet produits, soit en redoublant une consonne interne du mot (ex. : [ĕRAROZWAR], [ĕpipopotam]), une syllabe interne du mot (ex. : [ĕpupuvĕtaj]) ou en opérant une troncation sur une frontière syllabique (ex. : [ĕROZWAR], [ĕpopotam], [ĕpotam], [ĕpuvĕtaj], [ĕvĕtaj]).

En ce qui concerne les troncations, nous observons que ce phénomène est massif pour les mots tri et quadrisyllabiques. Dans la plupart des cas, les mots sont réduits à des unités tri et bisyllabiques voire monosyllabiques (1 seul cas). Il semblerait donc qu'une limite prosodique s'impose lors de la production d'une chaîne sonore et que ces enfants de 3 ans ne sont pas capables, en général, de produire un groupe excédant 3 à 4 syllabes. On constate d'ailleurs que ces troncations se produisent également sur les mots tri et quadrisyllabiques à initiale consonantique : [ĕniseros] pour [ĕrinoseros], [ĕparpi] pour [ĕparaplɥi], [ĕpejĕ] pour [ĕpapijĕ] et sont donc indépendantes du phénomène de liaison en soi. La réduction syllabique n'entraîne pas nécessairement de modification du substantif qui reste reconnaissable.

On observe également que la troncation ne touche jamais les syllabes finales. L'initiale est quasi systématiquement retirée, certaines syllabes internes également,

mais la syllabe finale est préservée dans tous les cas. On peut avancer deux explications qui ne sont pas nécessairement exclusives. La chute de la syllabe initiale du mot attendu peut-être attribuable à son faible poids métrique : dans un syntagme nominal comme [ɛ̃pʊvɑ̃taj], l'accent de groupe est porté par la syllabe finale, l'initiale de mot, placé ici en deuxième position dans le syntagme occupe une position de prétonique interne qui la fragilise.

On remarque également que la chute de la syllabe initiale de mot élimine automatiquement la syllabe à attaque vide qui pose problème déjà dans les mots courts. On peut penser que les enfants gèrent ainsi efficacement une double difficulté : ils réduisent un groupe syllabique trop long en supprimant d'abord la syllabe la plus problématique.

Enfin on remarque que la chute de la syllabe initiale n'entraîne pas pour autant la non-réalisation du déterminant qui est présent sur tous les items.

A partir de ces premières observations obtenues lors d'expériences sur la liaison, je me suis donc intéressée plus systématiquement aux phénomènes de troncation en français<sup>13</sup>.

#### **4.2. Troncations et reduplications chez les enfants francophones**

J'ai travaillé sur cette question à partir de données longitudinales de Claire [40 a, b, c, d] et de quelques données de Clara [39 a, b] (français québécois, 16 / 26 mois qui m'ont été communiquées par mon collègue Yvan Rose (Rose, 2000).

---

<sup>13</sup> D'abord de manière isolée puis dans le cadre d'une collaboration avec V. Braud qui a fait une thèse sur ce sujet BRAUD, V. (2003), *Acquisition de la prosodie chez les enfants francophones. Les phénomènes de troncations*, Thèse de doctorat, Université de Nantes, ms.

**[39a] : mots monosyllabiques**

ÂGE	CIBLE	FORME CIBLE	FORME PRODUITE
01;04.07	<i>oiseau</i>	[wazo]	[əwəzú:]
01;04.07	<i>lion</i>	[ljɔ]	[əlá:]
01;04.07	<i>lit</i>	[li]	[əjí:]
01;04.14	<i>puppy</i>	[pɔpi]	[əp <sup>h</sup> ɛp <sup>h</sup> í]
01;04.15	<i>sandale</i>	[sãdal]	[ədadæ]

**[39b] : mots trisyllabiques et quadrisyllabiques**

ÂGE	CIBLE	FORME CIBLE	FORME PRODUITE
01;09.29	<i>pantalon</i>	[pãtalɔ]	[ətæloé]
01;11.18	<i>magnétophone</i>	[máetɔfɔn]	[əfɔ <sup>"</sup> fɔn]
02;00.02	<i>coquelicot</i>	[kɔkliko]	[əpi:kó:]
02;02.07	<i>espadrille</i>	[ɛspadɛij]	[əpæ <sup>"</sup> dɛij]

**[40a] : mots monosyllabiques**

ÂGE	CIBLE	FORME PRODUITE
01;10 ;03	<i>le chien</i>	[lɔtɛ]
01;10,04	<i>la vache</i>	[ɔja ]
01;10,04	<i>la vache, l'âne</i>	[laja/lɔtan]
01;10,06	<i>le pot</i>	[lɔpo]
01;10,14	<i>Claire</i>	[le]

**[40b] : mots bisyllabiques**

ÂGE	CIBLE	FORME PRODUITE
01;10 ;03	<i>le bébé</i>	[lɔbebe]
01;10,04	<i>le ballon</i>	[abalɔ]
01;11,05	<i>Didou</i>	[didu]
01;10,22	<i>Maman</i>	[mamã]

**[40c] : mots trisyllabiques et quadrisyllabiques**

ÂGE	CIBLE	FORME PRODUITE
01;10 ;03	<i>un éléphant</i>	[ɔɛjã]
01;10,04	<i>éléphant</i>	[eã]
01;10,05	<i>deux éléphants</i>	[døeã]
01;11,06	<i>le hérisson</i>	[løijɔ̃]
01;10 ;03	<i>Aurélien</i>	[jɛ̃jɛ̃]
01;11,04	<i>Olivier</i>	[oje]
01;11,05	<i>la brosse à dent</i>	[anadã]
01;10,06	<i>un médicament</i>	[apamã]

**[40d] : mots trisyllabiques et quadrisyllabiques 2ème type**

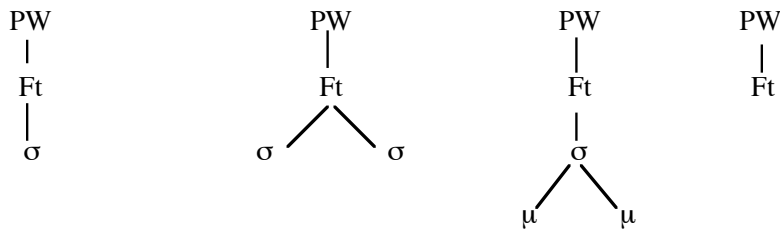
ÂGE	CIBLE	FORME PRODUITE
01;10 ;03	<i>le chocolat</i>	[ekola]
01;11,04	<i>le chocolat</i>	[løkokola]
01;11,16	<i>le chocolat</i>	[lekokola]
01;10 ;08	<i>un crocodile</i>	[ɛ̃koti]
01;10,23	<i>un crocodile</i>	[œkukudi]
01;11 ;02	<i>un crocodile</i>	[ekokodi]

Les phénomènes de troncations, très fréquents dans les données enfantines, et ce quelle que soit la langue cible, ont été abordés en général dans le cadre de la phonologie prosodique (Nespor & Vogel, 1986). Fikkert (1994), Demuth & Fee (1995), proposent que le niveau PW (phonological word) constitue l'unité prosodique retenue pour la mise en place des premières généralisations phonologiques et divers phénomènes observés lors de l'acquisition (harmonies consonantiques, fronting, gliding...). L'acquisition se ferait en quatre étapes détaillées ci-dessous ([41] et [42])

**[41]**

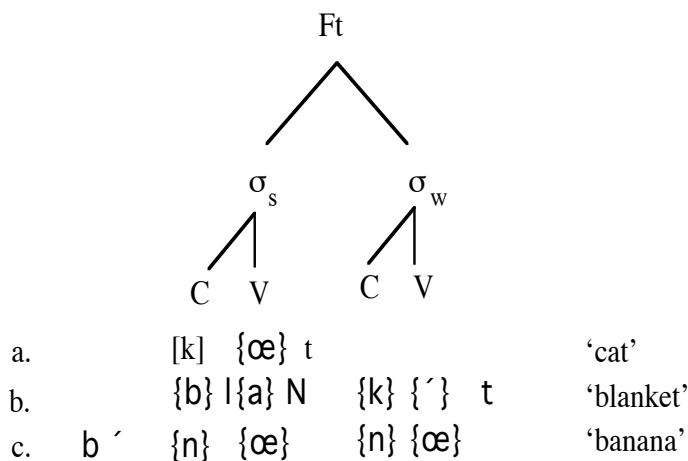
- I : Syllabes CV : (ex : ba ou na)
- II Mot minimal : CVCV (bana ou nana)
- III pied binaire accentué : ('nana, 'nanas')
- IV pieds accentués : ('baba'nana)

[42]



Et les enfants opéreraient des réductions systématiques pour intégrer la cible adulte dans le gabarit formé par un pied binaire, formé d'une syllabe accentuée suivie d'une syllabe non-accentuée [43] (Fikkert, 1994).

[43]



Cette proposition postulant la complexification progressive de l'unité à partir de la syllabe accentuée sur le domaine du pied, suppose d'abord un ordre d'acquisition strictement linéaire et invariant, progressant de bas en haut dans l'arbre prosodique et qui ne rend pas compte de la variété des items produits par les enfants. Cette conception, en particulier ne permet pas d'expliquer la différence établie par Waterson (1987) entre « enfants segmentaux » privilégiant des unités courtes où les cibles segmentales sont très bien réalisées et « enfants prosodiques » produisant des suites longues très bien intonées mais sous articulées.

Par ailleurs cette analyse reposant sur un stade III, composé d'un pied métrique suppose que les enfants vont passer nécessairement par un stade qui intègre le calcul d'un accent lexical. Les travaux sur les troncations ont jusqu'à Rose (2000) porté quasi-exclusivement sur les langues germaniques et ont amené la formulation de l'hypothèse dite du « biais trochaïque » (Allen & Hawkins, 1978, 1980) Gerken (1994) pour l'anglais et Fikkert (1994) pour le néerlandais, supposant que les enfants vont construire les premières unités lexicales en privilégiant la structure métrique « strong-week » et opérer des troncations en conséquence. Toutes les langues n'étant pas, loin s'en faut, des langues à accent lexical reposant sur une métrique trochaïque, l'hypothèse a ensuite été reformulée dans le cadre de « principes et paramètres ». Hayes (1995) propose que les enfants utilisent tous un pied binaire dont la tête serait paramétrisée au contact de l'input (ce qui revient à dire qu'on observera quelle que soit la langue cible des patterns initiaux bisyllabiques trochaïques « strong-week » ou iambiques « week-strong ») et que le paramètre par défaut est le pied trochaïque. Rose (2000) a proposé que le français soit acquis en miroir de l'anglais et que les enfants passent par un pied iambique privilégiant les deux dernières syllabes des unités qu'ils produisent selon un schéma « week-strong ».

Or, comme Braud (2003) le discute longuement, cette hypothèse est hautement problématique pour les deux raisons suivantes : le français n'est pas une langue à accent lexical où l'accent est toujours réalisé sur la dernière syllabe pleine (autre que schwa) de l'unité lexicale, mais un accent de groupe à valeur démarcative placé sur la dernière syllabe d'un groupe prosodique (Di Christo, 1999).

Par ailleurs, l'existence de pieds binaires a été discutée pour une langue comme le français où l'unité métrique est la syllabe.

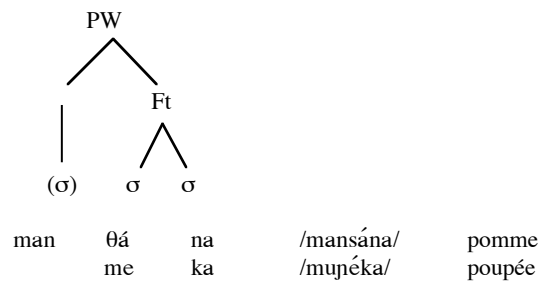
Enfin les données de Claire et les données de Braud (2003) infirment cette analyse. Les enfants francophones produisent, comme nous l'avons déjà montré pour la liaison des unités précoces bisyllabiques tronquées quand la cible excède deux syllabes, mais cette unité lexicale est toujours très précocement associée à un site phonologique correspondant au déterminant. Malgré la troncation des syllabes initiales pour les

unités tri et quadrisyllabiques, le déterminant est toujours préservé (Braud & Wauquier-Gravelines, 2005).

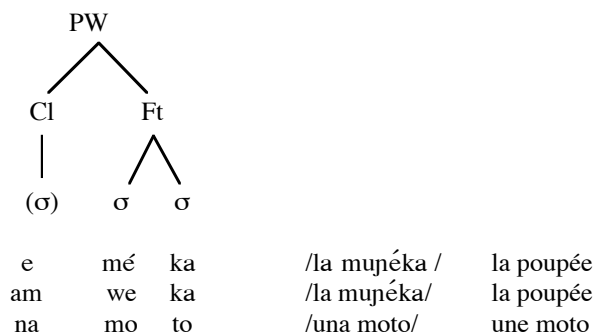
Des travaux sur l'espagnol-argentin, ont amené un assouplissement du modèle prosodique (Demuth, 2001) permettant de rendre compte de la variété des données produites ne se limitant pas à des pieds trochaïques.

A partir des données de Sofia (18 /19 mois), Demuth propose les premières unités suivantes

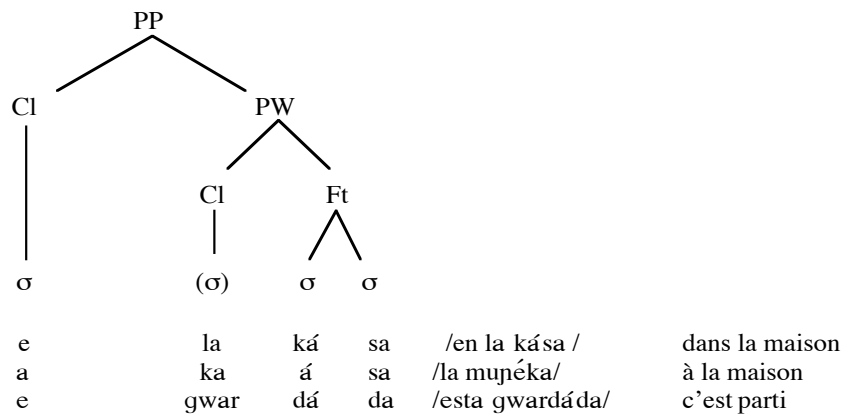
[44]



[45]



[46]



où l'on observe d'une part que plusieurs unités assez hautes dans la hiérarchie prosodique peuvent être simultanément disponibles (Phonological Word (PW) et Phonological Phrase (PP)) même à un stade très précoce, et que la variation observable dans les données est plus aisément envisageable à partir de cette proposition.

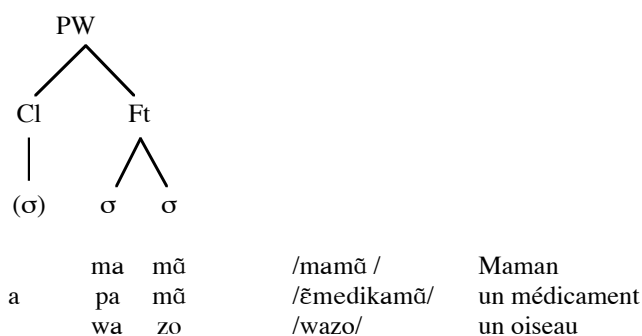
Par ailleurs, on observe que l'unité première PW inclut la possibilité de représentation du clitique (au moins de son site phonologique) également à un stade précoce d'acquisition. Enfin, cette modélisation minimise l'influence de l'accent lexical et le recours à la notion de pied accentué lors de l'acquisition phonologique. Elle paraît donc plus susceptible de rendre compte de l'acquisition de la première unité lexicale en français où il n'existe pas d'accent lexical et où les mots sont très rarement produits en isolation.

Néanmoins, elle pose une limite de taille qui permet de faire des prédictions concernant l'unité produite : on constate en effet que l'unité produite par l'enfant ne peut excéder 3 syllabes qu'à condition que les syllabes supplémentaires soient prosodifiées à un niveau supérieur [45]. Ceci suppose donc que toutes les séquences quadrisyllabiques sont des PP et que toute unité lexicale excédant trois syllabes (ex un rhinocéros) sera tronqué à ce premier stade. Comme je l'ai montré (Wauquier-Gravelines, 2003), ces propositions permettent de rendre compte des données de Clara et de la plupart des données de Claire.



Elle rend compte comme on peut le voir en [45] de la présence des déterminants sur tous les noms communs qui ne sont, en français, pratiquement jamais produits en isolation, de la préservation de toutes les unités n'excédant pas trois syllabes (det + mot bisyll) et du maintien du déterminant et de la troncation simultanée des syllabes initiales pour les unités excédant trois syllabes (det + mot bisyll) et dépassant la structure prosodique disponible.

[47]



Pourtant les données de Claire en [40d] ainsi que les données de Braud (1998) [47] montrent l'existence en français de phénomènes de reduplication associés à des troncations sur les mots longs qui infirment cette analyse

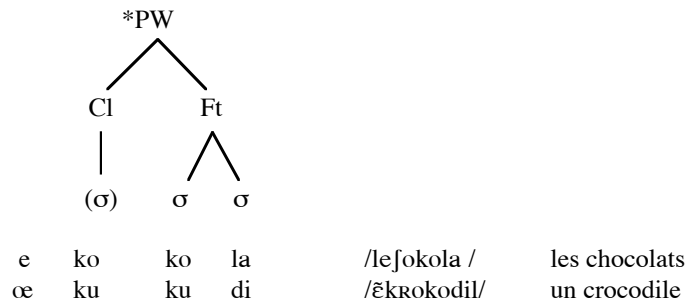
[48]

Mots monosyllabiques	Mots trisyllabiques	Mots quadrisyllabiques
« ours » [ẽnunuks]	« arrosoir » [ẽRORɔZwǎʃ] [ẽReaRɔZwǎʃ]	« aspirateur » [ẽRaRastatœʃ] [ẽRaatatœʃ]
« os » [ẽnonos]	« coccinelle » [ynkokosineɛ]	« accordéon » [ẽRaRaʒɔ̃]
		« épouvantail » [ẽpupuvãtaj]
		« hélicoptère » [ẽnenikɔtœʃ] [ẽniniɔtœʃ]

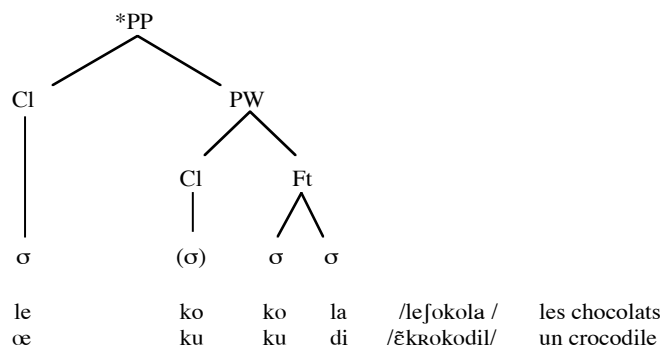
En effet, les mots produits en [40d], comme les mots redupliqués dans les données de Braud (2003) débordent la structure prosodique permise, soit en excédant le PW, soit

en nécessitant que le déterminant clitique soit prosodifié au niveau PP. Ces deux propositions sont également irrecevables comme nous le montrons ci-dessous.

[49]



[50]



### 4.3. Synthèse et conclusion

Les données de français que nous avons analysées, eu égard aux phénomènes de troncation et de reduplication nous permettent de dégager les régularités suivantes.

i) au-delà de 3 syllabes, les enfants réalisent systématiquement des coupes syllabiques ainsi que des redistributions du matériel segmental ("un potam" pour "un hippopotame", "un pisrateur" "un rastateur" pour "un aspirateur"). S'ils réalisent toujours un groupe prosodique cohérent, la taille de cette unité semble être drastiquement limitée.

ii) Les coupes effectuées ne sont pas aléatoires. La première et la deuxième syllabes sont quasi systématiquement omises, alors que la syllabe finale est toujours préservée, ainsi que la pénultième. Le caractère accentué de cette syllabe en français lui donne un poids prosodique qui tend à la maintenir prioritairement lorsqu'une troncation s'opère (Wauquier-Gravelines, 2003a, Braud, 2003, Braud & Wauquier-Gravelines, 2004). Les résultats suggèrent plutôt que la mise en place et l'utilisation des représentations lexicales ne se fait pas de manière strictement linéaire et que la production est guidée par des mécanismes impliquant une projection du matériel syllabique et segmental dans des unités prosodiques prédéterminées.

On remarque également que bien que la ou les premières syllabes de l'unité lexicale soient tronquées, la première syllabe de l'unité produite (futur déterminant) est maintenue de manière précoce. Ceci confirme ce que nous avons pu observer pour l'acquisition de la liaison.

iii) la réorganisation du matériel segmental dans l'unité prosodique n'est pas non plus aléatoire et semble également obéir à une logique non linéaire. Il semblerait que les enfants distribuent sur les positions syllabiques disponibles, un nombre réduit de consonnes et de voyelles reconnues dans les mots des adultes en accord avec des phénomènes d'harmonie vocalique et consonantique.

A la suite de Macken (1995), on peut considérer que les représentations lexicales enfantines pourraient être envisagées comme des gabarits prosodiques : "rarastateur" serait le résultat d'une redistribution de 3 consonnes (r,s,t) et de deux voyelles [a], [œ]) sur un gabarit de type CVCCVCVC selon leur ordre d'apparition.

L'acquisition lexicale consisterait alors en une adéquation entre gabarits prosodiques et matériau segmental progressivement élaborée jusqu'à la forme finale du mot.

Ces trois aspects des données (limite de l'unité prosodique, troncation de syllabes et redistribution non aléatoire d'un matériel lexical réduit) suggèrent effectivement l'existence de gabarits prosodiques et lexicaux. Ces analyses ont contribué à la formulation d'une hypothèse gabaritique (Braud & Wauquier-

Gravelines, 2004) sur laquelle je travaille actuellement et que je présente au chapitre II. 5 (projets en cours).

## **Travaux traitant de ces questions**

### **Publications**

**Wauquier-Gravelines S.** (2003), La question du réalisme des formalisations phonologiques contemporaines : que nous apprennent les données d'acquisition" in Wauquier-Gravelines S. & Angoujard, J.P. (eds), *Phonologie : Champs et perspectives*, ENS / éditions, Lyon

**Wauquier-Gravelines, S.** (2003), Troncation et reduplication. Peut-on parler de gabarits morphologiques dans le lexique précoce ?, in Fradin B., Dal G., Hathout N., Kerleroux F., Roché m., Plénat M.(eds) *Les unités morphologiques. Silexicales 3*, Université de Lille III.

Braud, V. & **Wauquier-Gravelines, S.** (2005), Approche gabaritique des phénomènes de troncation du français», *Actes des Journées d'Etudes sur la Parole*, avril 2004, Fez.

### **Communications**

**Wauquier-Gravelines, S.** (2002), Troncation et reduplication. Peut-on parler de gabarits morphologiques dans le lexique précoce ?, *Troisième forum international de Morphologie*, Université de Lille III, 21 septembre 2002.

Braud, V & **Wauquier-Gravelines, S.** (2005), Approche gabaritique des phénomènes de troncation du français, *Actes des Journées d'Etudes sur la Parole*, avril 2004, Fez.

## 5. Travaux en dysphasie développementale

Il a été proposé que la dysphasie développementale constitue un trouble qui touche la compétence morphologique (Gopnik, 1994). Par ailleurs de nombreux travaux ont été réalisés sur les troubles syntaxiques dans la dysphasie de développement (Léonard, 1997). Afin d'évaluer le type d'atteinte dont peut être affectée la compétence phonologique chez les enfants dysphasiques, je me suis proposé, en prenant la responsabilité de l'axe « Représentations phonologiques et segmentation en mots » au sein des projets coordonnés par Célia Jakubowicz, d'étudier les erreurs de production des liaisons enchainées chez un groupe d'enfants dysphasiques.

Comme je l'ai exposé ci-dessus (3. Acquisition de la liaison), à partir de 5 ans environ, les enfants francophones normaux ne réalisent plus le type de segmentations erronées qui peuvent s'observer avant cet âge (« le zoiseau » pour « l'oiseau »). Ceci suggère qu'à cet âge, ils ont acquis les représentations phonologiques leur permettant de traiter ce type de frontières. Si la dysphasie développementale invalide la capacité d'utiliser la compétence phonologique, on devrait voir subsister des segmentations erronées au-delà de 5 ans.

Par ailleurs, une des thématiques de ce projet consistait i) à déterminer la nature exacte des déficits affectant spécifiquement la composante phonologique de la grammaire des enfants atteints de troubles dysphasiques, ii) à établir, autant que possible, la distinction entre ce qui relevait respectivement de troubles de la compétence phonologique ou de troubles de la performance. Le but de ce travail était d'observer si les déficits constatés portaient sur les représentations configurationnelles ou sur les savoirs procéduraux. Les deux points de vue ont été également défendus par de nombreux chercheurs (pour synthèse, cf. Leonard 1997), mais leurs débats ont porté essentiellement sur la composante syntaxique de la grammaire.

Pour répondre à ces questions, j'ai travaillé sur la liaison enchaînée en français. Comme je l'ai montré, ce phénomène de sandhi nécessite effectivement une représentation appropriée des « segments flottants » et relève donc de la mise en place de la compétence phonologique, dont l'émergence est contrainte, comme nous l'avons montré par des caractéristiques structurelles. Autour de 3 ans, les enfants entrent dans une période critique, correspondant à la mise en place des paramètres propres de la liaison en français. Cette période critique se caractérise par des productions systématiquement erronées qui disparaissent à partir de 5 ans. Les enfants SLI présentant assez souvent sur le plan syntaxique et morphologique les mêmes erreurs de production que les enfants normaux, avec un décalage dans le temps, notre hypothèse était donc la suivante : si l'on constate chez des enfants SLI, ayant dépassé la période critique, la persistance de représentations fautives de la liaison constatées chez les enfants de 3 ans, on pourra donc supposer que ces enfants ont des déficits spécifiques de la composante phonologique dont il conviendra de préciser l'origine exacte.

## **5.1. Données et analyses**

### **5.1.1. Expérience**

Les enfants dysphasiques (à partir de maintenant D) que j'ai testés constituent un groupe de 18 sujets âgés de de 5,7 à 13 ans, retenus par C. Jakubowicz pour l'ensemble du projet et testés en phonologie et en syntaxe.

La tâche de dénomination d'images était identique à celle utilisée avec les enfants normaux de 3 ans et 5 ans. La plupart des enfants dysphasiques ont été testés également dans une pièce calme dans le cadre de l'école spécialisée qui les accueille. Quelques uns appartenant à des familles où plusieurs enfants avaient reçu un diagnostic de dysphasie ont été testés chez eux, dans des conditions tout à fait comparables à celles des autres enfants (famille 1: Nathalie, Bénédicte, Béatrice ; famille 2 : Elodie, Jennifer ; famille 3 : Johan, Frédéric, famille 4 : Vanessa, Julien).

Les données ont été intégralement transcrites en API. Les résultats ont été calculés à partir de la transcription API.

### 5.1.2. Résultats

Les données des enfants D ont été analysées en comparaison avec les résultats des enfants normaux présentés ci-dessus en partie 3.2.1. Le même histogramme a été réalisé pour les enfants dysphasiques.

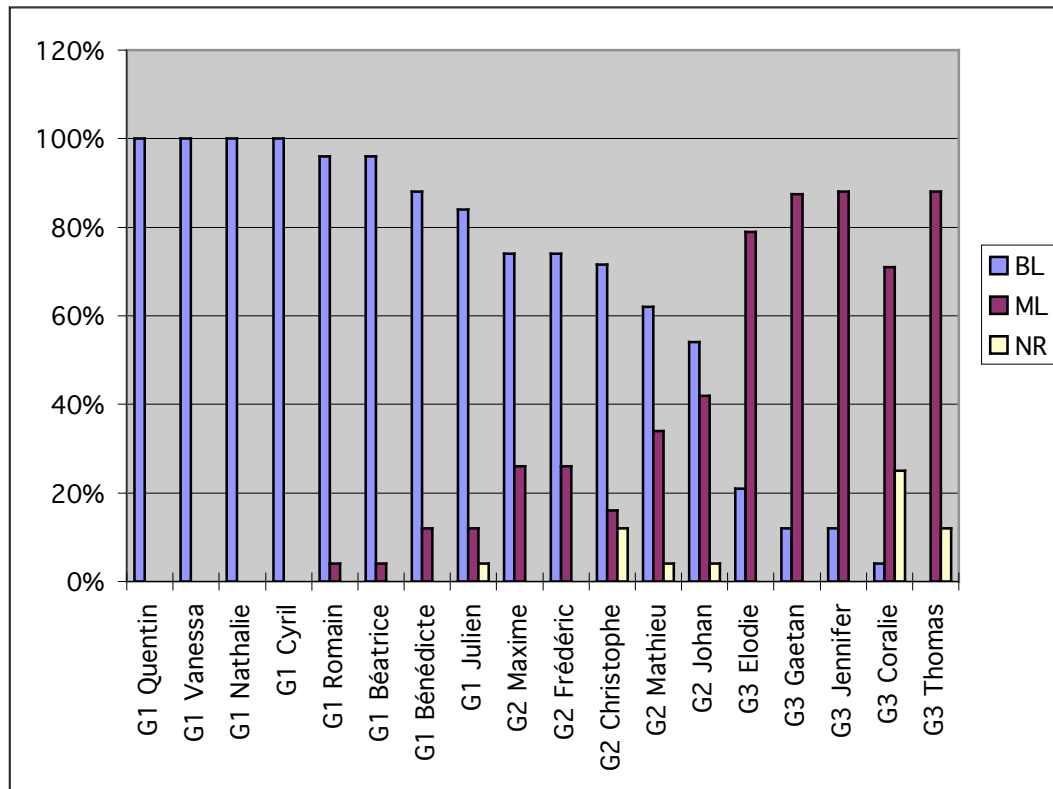


Figure 9: Résultats globaux pour les enfants D

Les performances globales des enfants D se situent entre celles des enfants de 3 et de 5 ans, mais quelques différences apparaissent pourtant : on constate un très faible pourcentage de non-réponses (3%) : la tâche semble donc avoir été bien comprise par les enfants qui se sont montrés très coopératifs même quand ils réalisent des scores d'erreurs très élevés. Les enfants de trois ans qui ont manifesté des taux d'erreurs comparables à ceux des enfants D obtiennent des taux de NR beaucoup plus élevés. On remarque ensuite une très grande hétérogénéité des résultats correspondant

à divers degrés de sévérité du déficit, ce qui m'a amenée à dégager 3 sous-groupes d'enfants dont les données ont été analysées séparément<sup>14</sup>.

#### *5.1.2.1. Résultats des enfants D 1*

Les enfants du groupe D 1 ont des résultats comparables à ceux des enfants de 5 ans (95,5% de BL, 4% de FL et 0,5% de NR) avec un taux de non-réponse moins élevé. La différence est sans doute attribuable à leur âge (ils ont tous largement dépassé 5 ans). La tâche peut leur sembler plus facile. On peut également penser que leur lexique est plus important. Les enfants D 1, comme les enfants de 5 ans ont parfaitement mis en place les contraintes propres à la liaison en français et ne présentent manifestement aucun trouble phonologique lié à ce phénomène de sandhi.

#### *5.1.2.2. Résultats des enfants D 2*

Les enfants du groupe D 2 ont des résultats meilleurs que ceux du groupe des enfants de 3 ans avec 67,2% de BL, 28,8% de FL et 4% de NR. Concernant le faible taux de NR on peut l'attribuer aux mêmes raisons que celles présentées pour les enfants D 1. On remarque également que leur taux de BL est beaucoup plus élevé que celui des enfants de trois ans.

Les erreurs portant sur la liaison elle-même sont peu nombreuses mais sont différentes des erreurs réalisées par les enfants de 3 ans. Elles sont constituées essentiellement d'omission de la CL (type [ẽ / elefã]) alors que les enfants de 3 ans produisent majoritairement des fautes consistant à remplacer la CL par une consonne erronée (type [ẽzelefã] ou [ẽlelefã]). Ceci plaide en faveur d'une discontinuité développementale chez les enfants dysphasiques laissant envisager une atteinte spécifique de la compétence phonologique.

Mais on remarque également que l'absence de réalisation des liaisons varie avec la durée syllabique des mots, les erreurs de liaison consistant à omettre la consonne se produisent essentiellement sur les mots trisyllabiques et quadrisyllabiques, un peu moins pour les mots bisyllabiques et pratiquement pas pour

---

<sup>14</sup> L'appartenance des enfants aux groupes I, II ou III est reporté dans la figure 10.



les mots monosyllabiques, La corrélation entre les omissions de la CL et la longueur de mot plaide donc plutôt en faveur d'une interprétation prosodique du trouble. En effet si les erreurs reflétaient un déficit de la compétence phonologique chez ces enfants, elles devraient se manifester également sur les mots courts.

Il semblerait que pour ces enfants, une limite prosodique s'impose lors de la production de la chaîne et qu'ils montrent des difficultés à produire un groupe excédant 3 à 4 syllabes (la majorité des mots quadrisyllabiques sont tronqués et comme pour les enfants de trois ans, la première syllabe est régulièrement coupée). La non réalisation de la liaison peut alors s'envisager comme l'incapacité de lier le déterminant et le nom dans un groupe prosodique comptant alors  $1+3 = 4$  syllabes et une manière de produire l'item sous la forme de deux unités prosodiquement distinctes. Cette faute portant sur la liaison elle-même pourrait relever en fait d'un déficit prosodique et non d'un trouble spécifique de la compétence phonologique.

#### Cas particulier de Johan

Seul Johan, qui est le plus jeune des enfants D, manifeste des erreurs tout à fait comparables à celles que produisent les enfants de 3 ans en remplaçant la CL par le [z] issu du déterminant pluriel ou par [l] (par exemple : "un ours" [ɛ̃lurs] pour [ɛ̃nurs]). On peut interpréter les erreurs de Johan en terme de retard par rapport au groupe des enfants de trois ans puisqu'à son âge (5,7), les erreurs de liaison devraient être résolues. Il constituerait un cas d'enfant tel qu'il a été décrit par Léonard (1997) où le déficit manifeste un retard sur un continuum normal d'acquisition plutôt qu'un déficit spécifique.

#### *5.1.2.3. Résultats des enfants D 3*

Les enfants du groupe D 3 présentent un déficit très sérieux et obtiennent des scores beaucoup plus bas que les enfants de trois ans (12,2% de BL, 82,8% de FL et 5% de NR). La difficulté de la tâche ne peut pas être ici incriminée. La moyenne de 5% de NR résulte des performances d'une seule enfant (Coralie, 5,10) l'une des enfants les plus jeunes du groupe D et qui a montré beaucoup de difficultés à terminer

l'expérience. Tous les autres enfants ont des taux de NR très bas : ils ont participé volontiers à l'expérience et ont manifestement très bien compris la tâche proposée.

L'explication du taux de FL particulièrement élevé est assez délicate à dégager. On constate une grande variabilité d'erreurs chez les enfants D 3 ainsi que la production d'énoncés parfois phonologiquement si éloignés de l'item attendu qu'il est difficile de les reconnaître et d'établir alors une catégorisation des fautes à partir des critères précédemment retenus (par exemple [ẽɔɾtɛl] pour [ẽnelikɔptɛɾ] *un hélicoptère* ou [saatœ] pour [ẽnaspiratœɾ] *un aspirateur*). Deux groupes plus homogènes peuvent néanmoins être dégagés

#### Elodie et Jennifer

Elodie et Jennifer sont deux sœurs. Elles produisent des énoncés très peu déformés très aisément reconnaissables. Leurs performances sont très homogènes, quasi identiques et se caractérisent de la manière suivante :

i) elles produisent plus de 80% des items mono, bi et trisyllabiques en supprimant la liaison, quand le mot passe à 4 syllabes, elle opèrent systématiquement des troncations.

ii) elles ont tendance à produire les réponses en épelant une à une chaque syllabe (ie [ẽ/e/li/kɔp/tɛɾ]), comme si le groupe voire le mot lui-même ne constituait pas une unité prosodique en soi. On peut donc penser que leurs performances manifestent la systématisation de ce que nous avons pu observer pour les enfants D 2 : l'unité qu'elles sont capables de produire se limite plus drastiquement à 2 syllabes, à partir du moment où elle doivent produire un groupe de type déterminant + mot >2 syllabes, elles opèrent des coupes ou des segmentations internes au groupe prosodique et épellent syllabe par syllabe.

#### Gaëtan, Thomas et Coralie

Les performances de ces trois enfants ont été regroupées malgré leur variabilité. Les données sont parfois tellement déviantes par rapport à la cible attendue qu'il est difficile de les analyser. On peut résumer leurs erreurs de la manière suivante

i) Ces trois enfants font très peu de fautes de consonnes isolées et ces fautes se produisent essentiellement sur les mots monosyllabiques.

ii) Sur les mots tri et quadri syllabiques on constate des erreurs lexicales massives sans lien nécessaire avec la liaison. Ces enfants ne semblent pas avoir spécifiquement de problèmes avec la liaison : malgré leurs performances très faibles ils réussissent à produire au moins 1 ou 2 items monosyllabiques en réalisant correctement la liaison. Leurs résultats nous invitent à considérer qu'ils ont des difficultés phonologiques lexicales et prosodiques plus massives qui incluent éventuellement des difficultés à réaliser la liaison. Leurs résultats confirment les descriptions des troubles phonologiques sérieux chez les enfants D synthétisées par Léonard (1997). Chez tous ces enfants on retrouve les erreurs décrites, parfois cumulées pour un même mot : réduction de clusters, « gliding » des liquides [R] et [L], délétion des consonnes finales et initiales, dévoisement des consonnes finales. Ingram (1987) propose également que pour les enfants D le degré de déficit phonologique soit l'inverse des compétences lexicales de l'enfant et du nombre de mots qu'il possède. Il semble que cette remarque s'applique à ce dernier sous-groupe d'enfants D 3 pour qui on constate un manque du mot et des temps de réponse très longs.

iii) On retrouve chez ces trois enfants, les mêmes difficultés prosodiques constatées pour Elodie et Jennifer. Ils éprouvent apparemment de grandes difficultés à produire une unité prosodique excédant 2 ou 3 syllabes. Les troncations et simplifications de la structure syllabique sont très systématiques dès qu'on leur présente des mots de deux syllabes. Comme chez les enfants N3, les items sont tronqués pour la ou les premières syllabes alors que la syllabe finale est toujours préservée. Par ailleurs, Gaetan et Thomas ont tendance à réduire le groupe en omettant le déterminant, Coralie réalise systématiquement cette tendance. Ces enfants sont donc apparemment incapables d'inclure le mot attendu dans une unité prosodique structurée.

## 5.2. Synthèse et discussion des résultats

### 5.2.1 Synthèse des résultats

Pour les enfants D, on constate d'abord une importante variabilité des résultats correspondant à des degrés variés de déficits. Certains enfants D comme les enfants de 5 ans, ont intégralement maîtrisé les contraintes de mise en place de la liaison. Pour d'autres au contraire, on constate un déficit important, voire massif mais qui diffère de celui des enfants normaux de 3 ans. Les enfants D ne semblent pas avoir réellement de problèmes avec la liaison en soi. Leurs difficultés portent essentiellement sur la maîtrise de contraintes prosodiques qui gouvernent la réalisation des consonnes de liaison. Si les enfants de trois ans et D 2 et 3 montrent tous des difficultés à produire des items en contexte de liaison, leurs difficultés ne sont apparemment pas de même nature. Ce qui est manifestement commun à tous ces enfants est le manque de maîtrise d'unités prosodiques dépassant, selon le cas, deux à trois syllabes. Les enfants de 3 ans, comme les enfants D 2 et 3 effectuent sur les mots trisyllabiques et quadrisyllabiques des troncations systématiques ainsi que des omissions de déterminant réduisant la longueur prosodique des groupes dans lesquels ces mots apparaissent. Ces difficultés sont-elles systématiquement associables au phénomène de liaison ? Il semblerait que ce ne soit pas le cas. Si l'on regroupe les réponses des enfants D2 et D 3 sur les items contrôles commençant par une consonne et ne comportant aucune difficulté liée à la liaison, on constate un taux d'erreurs s'élevant brutalement à partir des mots tri et quadrisyllabiques. Comme pour les items en liaison, ces erreurs sont essentiellement des erreurs de troncations portant sur la ou les première(s) syllabe(s), des omissions du déterminant, des dictions épelant syllabe par syllabe. Ceci invite donc à considérer que le problème prosodique commun aux enfants D 2 et 3 est d'une part très consistant, d'autre part indépendant du phénomène de liaison lui-même.

### 5.2.2. Limite des constituants prosodiques pour les enfants de 3ans, D 2 et D 3

On constate l'existence, apparemment indépendante de la liaison elle-même, de limites portant sur la taille des groupes prosodiques que ces enfants sont capables de

produire. Les enfants de 3 ans ne peuvent en moyenne pas dépasser des groupes de 4 syllabes. Pour les enfants D, une limite plus sévère s'impose à 3 syllabes voire 2 syllabes chez les enfants D 3 les plus atteints. Cette contrainte s'exprime, nous l'avons vu, par des réductions des groupes prosodiques consistant en troncations d'une ou plusieurs syllabes initiales, omission du déterminant, épellation des mots. Toutes ces procédures sont selon nous interprétables comme une absence de maîtrise de la procédure de production qui suppose la projection de l'information sonore dans une structure. En d'autres termes, ces enfants semblent incapables de programmer une unité de production dépassant une certaine taille. Ces résultats recourent les propositions diverses envisageant les déficits des enfants D comme relevant d'une capacité de « processing » limitée en production. Leonard (1997) présente en effet de nombreuses données permettant d'envisager chez les enfants D la persistance parfois dramatique d'une capacité de « processing » limitée qui se manifeste par ailleurs plus précocement chez les enfants normaux. Léonard (1987) a donc proposé que les difficultés des enfants D ne soient pas envisagées en terme de rupture par rapport à un développement normal mais que les enfants D manifestent un faible niveau de compétence retardé sur un continuum général d'acquisition. Cette conception semble particulièrement compatible avec les données obtenues : les enfants D2 et D3 ne manifestent pas de difficultés massives en relation avec leur compétence phonologique, tous réussissent tant bien que mal à réaliser des liaisons sur les items courts, alors que tous les enfants D 2 et 3 qui sont âgés de 6 à 13 ans, manifestent des difficultés de « processing » comparables à celles d'enfants de 3 ans. Les troubles semblent bien donc être d'abord d'origine procédurale.

Deux explications peuvent être envisagées. Soit ces déficits sont la manifestation d'une difficulté à programmer une unité de traitement en production. Ces enfants ne manipuleraient pas aisément tous les constituants prosodiques disponibles dans leur langue. Leurs difficultés devraient alors être envisagées comme un déficit de l'encodage phonologique dans un modèle de production.

L'autre interprétation que l'on peut retenir serait une difficulté liée non pas à la structure du constituant mais à la quantité de matière sonore (dont la mesure est à déterminer : nombre de syllabes, durée pure ?) que les enfants peuvent projeter dans

une structure. Cette hypothèse amène alors à poser le problème en termes de limite de la mémoire de travail et déficit ou retard dans l'utilisation de la "boucle phonologique".

#### *5.2.1.1 Déficit de l'encodage phonologique dans un modèle de production*

J'ai repris cette notion au modèle de Levelt (1989, 1992) compatible avec la phonologie prosodique exposée plus haut. C'est également le modèle de production retenu par Léonard (1997).

Levelt propose qu'à l'issue de la formulation de l'unité lexicale, celle-ci soit encodée et éventuellement resyllabée pour être produite. Selon lui, afin de réaliser l'encodage de l'unité lexicale programmée, le locuteur sélectionne la représentation phonologique adéquate (au sein d'une structure arborisée du type de celles que propose la phonologie prosodique) en fonction du pattern prosodique dans lequel le mot produit va s'insérer.

La structure prosodique est délivrée par un composant parallèle au formulateur. Ce générateur produit des unités prosodiques à l'intérieur desquelles viennent s'insérer les représentations phonologiques sous-jacentes, obtenues à la sortie de l'encodage phonologiques et stockées lexicalement. Cette insertion dans une unité prosodique permet au locuteur de planifier son énoncé et impose la nécessité d'un ajustement entre représentations phonologiques des mots isolés et caractéristiques des unités prosodiques. Cet ajustement peut nécessiter un recours aux représentations optionnelles (Levelt donne l'exemple du choix de l'allomorphe "ve" dans le cas de "I've"). L'utilisation de ces représentations optionnelles, permet d'ajuster les mots planifiés dans cette structure, et entraîne des modifications phonologiques de surface telles que la resyllabation.

Les difficultés constatées chez les enfants D2 et D3 pourraient alors consister en une limite de l'encodage phonologique. On peut penser que tous les constituants prosodiques universellement représentés ne sont pas simultanément disponibles et que les enfants n'y ont que progressivement accès, de même qu'ils ne produisent pas spontanément et immédiatement toutes les phrases bien formées possibles dans leurs langues mais qu'ils ont recours à des structures syntaxiques qui se complexifient

progressivement. Les capacités de « processing » seraient donc limitées par le constituant qu'ils sont capables de sélectionner lors de l'encodage phonologique. Chez les enfants normaux, on assisterait à une appropriation progressive des structures disponibles (Gerken, 1994) correspondant à une remontée dans l'arbre prosodique et parallèle aux acquisitions syntaxiques et conditionnant les capacités de « processing ». A trois ans, seules les unités de rang inférieur au groupe clitique pourraient être aisément manipulées. Chez les enfants dysphasiques beaucoup plus avancé en âge, se manifesterait un « blocage » à un niveau de l'arbre prosodique qui serait selon le cas le groupe clitique ou le syntagme phonologique (« phonological phrase »). Cette interprétation confirme l'interprétation de déficits morphologiques et syntaxiques (omissions de déterminants, pronoms et inflexions à l'intérieur de groupe clitiques et phonological phrases) constatées par Mc Gregor & Leonard (1994 ), Bortolini & Léonard (1996).

#### 5.2.1.2. *Déficit de la boucle phonologique*

Cette interprétation n'est pourtant que partiellement satisfaisante puisque les enfants ont tous été testés sur des groupes clitiques et qu'ils ne manifestent des difficultés qu'à partir des groupes clitiques supérieurs ou égaux à 3 syllabes. Il semblerait donc que ce qui leur pose problème, lors de l'encodage phonologique, n'est pas tant la sélection d'un niveau de représentation prosodique que la mise en adéquation de la structure sélectionnée avec une quantité donnée de matière sonore à insérer dans cette structure. L'interprétation nécessite alors le recours aux travaux concernant la mémoire de travail portant sur utilisation des ressources mémorielles en général ensuite appliquée aux tâches langagières. La mémoire de travail est un système stockant et traitant une quantité limitées d'informations, nécessaires à la réalisation d'une activité cognitive en cours. Elle est constituée d'un composant central qui gère le flux et la redistribution des informations ainsi que l'interface avec la mémoire à long terme. Lui sont adjoints deux composants annexes : la *boucle phonologique* et l'*ardoise visuo-spatiale*.

La *boucle phonologique* gère toute l'information codée verbalement. Les recherches concernant l'architecture et le fonctionnement de la boucle phonologique ont été très

fécondes depuis une vingtaine d'années (pour une synthèse, cf. Gathercole & Baddeley, 1993). Il se dégage de ces travaux des résultats suivants pouvant éclairer les difficultés des enfants D et des enfants normaux.

Ces recherches ont mis au jour les faits suivants : la capacité d'utilisation de la mémoire phonologique engrangée est directement influencée par la longueur des mots. Selon les auteurs, cet effet est dû à la durée articulatoire des items et non au nombre de syllabes contenues. Le système de mémoire ne stockerait donc pas un nombre de syllabes potentielles mais la durée pure du mot, et ce, en temps réel. Plus un mot est long, plus sa trace en mémoire est difficile à rappeler, plus sa récupération en vue d'une restitution est difficile. Gathercole & Baddeley (1990, 1993) ont établi l'existence chez les enfants D d'un déficit de stockage de l'information phonologique dans la mémoire de travail. Ce déficit peut avoir diverses origines. Parmi les interprétations avancées, les auteurs suggèrent l'imprécision de l'analyse segmentale lors du stockage, qui entraîne une représentation phonologique floue, peu discriminative et par conséquent difficile à activer en production. Plus généralement, les auteurs proposent qu'il y ait, chez les enfants D une limite de stockage (dont l'origine n'est pas précisée) dans la mémoire de travail tant en termes de quantité (le nombre et la taille des items stockables sont limités) qu'en terme de qualité (les items stockés sont déformés). Cette limite entraînerait chez ces enfants une pauvreté des représentations lexicales effectivement confirmée par des données impliquant des tâches lexicales et montrant que les enfants D ont beaucoup plus de difficultés à produire, retrouver des items lexicaux que les enfants normaux du même âge (Kail et al. 1984, Kirchner & Klatsky, 1985).

Cette explication serait tout à fait compatible avec les résultats obtenus. Les enfants D 2 sur les mots tri et quadrisyllabiques, et les enfants D 3, dès les items monosyllabiques, réalisent, à des degrés divers, des déformations phonologiques des items proposés en contexte de liaison comme en contexte contrôle. Dans le cas du deuxième groupe d'enfants D 3 (Gaetan, Coralie et Thomas), certains items n'étaient même pas reconnaissables. Par ailleurs, certains d'entre eux (Romain, Christopher) étaient affectés d'un "manque du mot" se manifestant par un temps de réponse très



ralenti et certaines non réponses étaient commentées par les enfants comme une incapacité à restituer une information lexicale pourtant connue ("je sais pas le dire").

Reste le problème que posent les données d'Elodie et Jennifer qui découpent les items syllabe par syllabe mais en ne déformant pas du tout les items. Le codage mémoriel des items ne semble donc pas être qualitativement déficitaire mais quantitativement restreint à une syllabe. Pour elles, l'encodage phonologique ne peut être réalisé que sur la base de l'unité syllabique.

Chez ces enfants D au contraire, il semblerait que l'on ne puisse observer aucun déficit de représentation phonologique et que les erreurs soient uniquement attribuables à un déficit plus ou moins grave de processing.

### **5.3. Conclusion**

Ces recherches amènent donc à conclure de la manière suivante : je n'ai pas mis en évidence, chez les enfants D l'existence d'atteintes spécifiques portant sur des représentations de la compétence phonologique. Ces enfants sont capables de produire des liaisons. Les enfants D 2 et 3 manifestent des difficultés de production, déforment des mots ou refusent de verbaliser à partir de 3/4, syllabes chez les enfants du 2ème groupe D 3, les déficits affectant apparemment la mémoire de travail sont massifs dès la deuxième syllabe, c'est à dire sur des mots monosyllabiques qui sont difficilement restitués et très déformés.

Leurs difficultés apparaissent plus manifestement comme des difficultés de processing : la projection d'une certaine quantité de matière sonore à l'intérieur d'une unité de production est fortement handicapée chez certains de ces enfants D, et ce indépendamment de leur âge. Nos résultats confirment la synthèse des données présentées par Leonard (1997), puisque les déficits constatés semblent interprétables comme des déficits de la mémoire de travail entraînant des atteintes lexicales qui peuvent s'avérer très sérieuses.

## Travaux traitant de ces questions

### Publications

**Wauquier-Gravelines S.**, (à paraître) "'Liaison enchaînée' in French: Representational or processing deficit ?" in, *Essays on Syntax, Morphology and Phonology in SLI* in Celia Jakubowicz, Léa Nash and Kenneth Wexler (Eds.), MIT Press.

### Communications

**Wauquier-Gravelines, S.** (2002) : « Acquisition de la Liaison et problèmes de traitement chez les enfants dysphasiques » conférence invitée, Concordia University, Montréal, 24 avril 2002

**Wauquier-Gravelines, S.** (2001) : « Troubles prosodiques et phonologiques dans la dysphasie de développement », Conférence invitée de l'Université de Genève, 27 mars 2001.

**Wauquier-Gravelines, S.** (2000). "Troubles prosodiques dans la dysphasie de développement : déficit structurel / déficit mémoriel ?" *XXVIIème Symposium de l'APSLF* : "La maîtrise du langage", 14-16 septembre 2000, Nantes.

**Wauquier-Gravelines, S. & S.**; Jakubowicz, C.; Sauzet, P. ; Durand, C.; Franc, S. (1997a) Liaison enchaînée and Derivation in French : Representational or processing deficit ?, *Workshop on SLI*, organisé par Celia Jakubowicz, CNRS URA 316, septembre 1997, Paris

**Wauquier-Gravelines, S.** (1997b) : Troubles phonologiques dans la dysphasie de développement, *Journée d'Information de l'Equipe de Recherche AAI*, Université de Nantes, 11 avril 1997.

**Wauquier-Gravelines, S.**, Jakubowicz, C., Sauzet, P., Durand, C., Franc, S. (1997c) : Phonological knowledge in Developmental Language Disorders : "Liaison enchaînée" and Derivations in French, *G ALA*, Edimbourg, 4 au 6 avril

## CHAPITRE II : TRAVAUX EN COURS

### 1. Introduction

Ce chapitre rend compte des travaux en cours, plus ou moins avancés selon le cas. Certains d'entre eux ont donné lieu à des communications orales dans des colloques internationaux ou des séminaires, ou sont en voie de publication ou mis en forme dans des textes soumis ou manuscrits. Ces communications et publications en cours seront présentées à l'issue de chaque partie.

Les recherches concernant la liaison en français sont en continuité directe (2. acquisition de la liaison, nouvelles perspectives et 3. traitement de la liaison enchaînée chez l'adulte) avec des travaux et hypothèses présentées dans le chapitre I et proposent de nouvelles analyses à partir de données identiques pour la partie 2. Pour la partie 3, des données issues de la thèse ont été réexploitées mais de nouvelles données de perception retestant de manière plus systématique les hypothèses proposées sont en voie de réalisation. Le travail présenté là est réalisé en collaboration avec Noël Nguyen et Leonardo Lancia (Université d'Aix en Provence), Maïtine Bergounioux (Université d'Orléans) et Betty Tuller (Florida Atlantic University, USA) au sein d'une ACI Systèmes complexes en SHS «Vers un modèle dynamique de l'émergence des représentations symboliques dans le traitement de la parole » dirigée par Noël Nguyen.

Pour ce qui concerne les travaux sur la liaison, si les données ne sont que partiellement nouvelles, les analyses le sont totalement et visent

- i) à unifier l'explication de la variation observable dans les données d'adultes (production et perception) et les données d'acquisition.
- ii) à proposer une modélisation cognitive et psychologiquement plausible de la notion de « segment flottant » au sein d'un modèle

psycholinguistique émergent et dynamique, rendant compte des représentations permettant d'expliquer les productions et traitements adultes de la liaison ainsi que de l'émergence des représentations chez l'enfant.

La partie 4. expose un travail nouveau sur l'acquisition des structures syllabiques dans une perspective typologique que j'ai commencé à mettre en œuvre à partir de 2003. J'ai réalisé ce travail pour l'acquisition des attaques syllabiques en français mais cette question est étudiée plus largement au sein d'échanges européens mis en place à partir de 2003 avec Paula Fikkert (Nimègue), Clara Levelt (Leiden) et M.-J. Freitas (Lisbonne). Les données utilisées pour cette recherche sur l'acquisition des structures syllabiques sont celles de l'enfant C utilisées pour les recherches présentées au chapitre I. A celles-ci s'ajoutent les données longitudinales d'un autre enfant L. de 1,9 à 2,6 ans.

La partie 5. présente le travail actuel où je tente, à partir des propositions faites par Macken (1992), Menn (1978) et Vihman (2001), d'unifier l'analyse des données d'acquisition du français dont je dispose au sein d'un modèle gabaritique autosegmental. Ce travail vise

- i) à proposer une hypothèse de formalisation de la première unité utilisée en production par les enfants francophones
- ii) à en établir les régularités structurelles (phonologiques et prosodiques).

Enfin, la dernière partie du chapitre (6.) présente rapidement les perspectives de recherche qui peuvent être envisagées pour des travaux futurs.

## 2. Acquisition de la liaison, nouvelle proposition

### 2.1. Acquisition de la liaison et variation

En conclusion de la partie 3.4 du chapitre I, j'avais terminé en faisant les propositions que je rappelle ci-dessous

- i) Existence d'une période critique entre 2 et 5 ans pour l'acquisition des liaisons obligatoires présentes chez la plupart des enfants francophones
- ii) L'acquisition de la liaison obligatoire se fait dans le contexte (NP : det + nom) qui est le premier et seul contexte obligatoire où les erreurs apparaissent au moment de l'explosion lexicale.
- iii) L'acquisition de la liaison obligatoire est liée à l'acquisition du déterminant pro-clitique conditionnée par un bootstrapping morphologique
- iv) Les erreurs sur les liaisons facultatives apparaissent plus tard et se résolvent également plus tard. Elles s'inscrivent dans une chronologie beaucoup plus lâche, ce qui m'a amenée à considérer qu'elles pouvaient effectivement comme le proposent Chevrot & Fayol (2000, 2001) être apprises au cas par cas par mémorisation.

J'avais donc envisagé que l'acquisition des liaisons obligatoires et facultatives puissent être des processus radicalement différents. J'avais considéré que l'acquisition de la liaison obligatoire est phonologiquement conditionnée et que, par conséquent les erreurs de liaison obligatoires produites par les enfants reflètent la mise en place des représentations phonologiques abstraites.

Mais j'avais en revanche concédé que l'acquisition de la liaison facultative puisse relever d'un apprentissage lexical qui consiste à appliquer, au cas par cas, en fonction de normes d'usage, sur des contextes spécifiques sociolinguistiquement conditionnés, les régularités phonologiques acquises dans le contexte det + nom.

A ce stade de la réflexion, j'avais donc considéré que les types d'erreurs observables sur la liaison obligatoire et sur la liaison facultative étaient différentes. Je n'avais pas envisagé de rendre compte des erreurs sur la liaison obligatoire et sur la liaison facultative à partir d'un formalisme unique. J'avais par conséquent exclu de ma problématique la question de la variation en acquisition. Du moins avais-je formulé la question de la variation chez l'enfant, uniquement sur la liaison obligatoire comme un principe d'apprentissage et donc indépendamment de critères sociolinguistiques.

J'avais proposé que la variation des formes produites pour les liaisons obligatoires (par exemple « un féféfan », « un zéléfan », « un léléfan » pour *un éléphant*), ne constitue que l'expression de *learning paths* spécifiques.

J'avais envisagé la variation comme un phénomène ne touchant que l'interprétation segmentale de la consonne et l'avais interprétée comme l'expression de stratégies différentes et spécifiques pour arriver au même but (remplissage de l'attaque du mot et application du principe universel de maximalisation des attaques).

Mais cette explication est problématique en ce que, si l'on retient par ailleurs comme je l'ai fait pour tous mes travaux en acquisition, l'hypothèse d'une continuité entre les processus observables en acquisition et les grammaires adultes, on aboutit logiquement à la conclusion que la liaison obligatoire et la liaison facultative chez l'adulte ne relèvent pas forcément des mêmes principes explicatifs.

Cette proposition de processus d'acquisition différenciés pour la liaison obligatoire et facultative m'amènent donc à envisager la liaison comme un phénomène hétérogène dont les conditionnements multiples échappent à un principe explicatif unique. Or cette conception est radicalement en contradiction avec le formalisme autosegmental (Encrevé, 1988) que j'ai utilisé dans mes travaux sur le traitement de la liaison chez l'adulte (cf. Ch I).

## 2.2. Vers une explication unifiée de la variation

### 2.2.1 Non-association des segments flottants

La formalisation de la liaison que j'ai retenue pour le traitement chez l'adulte (Encrevé, 1988) propose en effet un principe explicatif unique pour la liaison obligatoire et facultative qu'elle soit ou non enchaînée.

Rappelons la conception du flottement retenue par Encrevé qui a été reformulée par Encrevé et Scheer (2005).

- i) un flottement de la CL sur la ligne segmentale, donc l'absence de rattachement à une position segmentale
- ii) un flottement de la CL sur la ligne syllabique, donc l'absence de rattachement à une attaque ou à une coda
- iii) une position squelettale disponible lexicalement permettant l'ancrage de la CL à la fois au plan syllabique (donc en attaque ou en coda) et au plan segmental et leur faisant face. Le squelette a une réalité cognitive, il matérialise le gabarit temporel de stockage des unités lexicales

Ceci suppose donc que la CL est un objet disponible dans la représentation lexicale. La CL est conçue comme « doublement » flottante. Selon ces auteurs, la logique autosegmentale n'impose pas, en effet, que tout matériel mélodique s'associe aux constituants syllabiques de façon automatique. L'association est conditionnée par la grammaire (compétence) du sujet et en accord avec des conventions de bonne formation. Elle résulte « d'une instruction explicite » dont le conditionnement peut-être morphologique, phonologique ou stylistique.

Pour la liaison facultative l'association est stylistiquement conditionnée, quand le contexte le permet, le locuteur peut i) choisir d'associer ou non la consonne flottante ii) choisir de l'associer soit à l'attaque du mot suivant, soit à son constituant

lexical propre. Les locuteurs-auditeurs connaissent cette variation qui fait partie de leur compétence passive. Les contraintes d'association, dans tous les cas, font donc également partie de la compétence du sujet.

## 2.2.2. Propositions pour l'acquisition de la liaison

### 2.2.2.1. *Liaisons obligatoires vs facultatives : deux conditionnements différents*

Cette formulation autosegmentale radicale de la CL, qui l'envisage comme un « locus variationnis » intégrant toutes les formes de variation dans la compétence, offre un dispositif formel permettant d'unifier l'interprétation des faits d'acquisition pour la liaison facultative et la liaison obligatoire.

Comme je l'ai montré au chapitre I et comme en attestent les données diverses (Basset, 2000, Wauquier-Gravelines, 2003b), l'acquisition de la liaison obligatoire se produit pendant une période allant de 2 à 5 ans et elle est associée aux acquisitions morpho-syntaxiques et plus particulièrement à l'acquisition du déterminant proclitique par bootstrapping morphologique. La conception du double flottement proposée par Encrevé (1988) reformulée dans Encrevé & Scheer (2005) propose que l'association n'est pas automatique et qu'elle est diversement conditionnée. On peut arguer sur cette base que l'association de la consonne lors de la première étape d'acquisition plaide en faveur d'un conditionnement morpho-syntaxique exclusif avec une association strictement automatique. L'acquisition va consister 1) à détacher le segment pour le rendre doublement flottant 2) à inférer les conditionnements possibles de l'association (obligatoires ou non).

Boula de Mareuil et al. (2003) ont réalisé une analyse automatique d'un corpus de 66500 phrases extraites du Monde et lues par 120 locuteurs (100 heures de parole) et pour les 8 contextes obligatoires établis à partir de Delattre (1966) et Encrevé (1988), ils ont obtenu les résultats suivants :



1. Déterminant + nom (*les amis*) : réalisées à **95,3 %**, représente **35,6 %** du corpus
2. adj + nom (*un gros arbre*) : réalisées à **72,5 %**, représente **3,1 %** du corpus
3. adverbe monosyll sauf pas (*tant en ville*) : réalisées à **70,7 %**, représente **4,4%** du corpus
4. verbe + pronom (*sort-il*) : réalisées à **99,2 %**, représente **2,2 %** du corpus
5. pronom atone + verbe (*ils ont*) : réalisées à **83,3 %**, représente **9,2 %** du corpus
6. verbe aux + attribut (*il est évident*) : réalisées à **81, %**, représente **13,6 %** du corpus
7. prep monosyll + nom (*en avance*) : réalisées à **90,5 %**, représente **9,4 %** du corpus
8. quand + (*quand il arrive*) : réalisées à **94,6 %**, représente **0,4 %** du corpus

ce qui signifie que, sur un corpus d'une telle importance, les liaisons obligatoires sont réalisées à plus de 85% et que parmi celles-ci, celles qui représentent de très loin le contexte obligatoire le plus fréquemment produit (contexte déterminant-nom plus d'un tiers de tout le corpus) sont réalisées à plus de 95 %.

On pourra, néanmoins, objecter que ce classement de Boula de Mareuil et al. introduit dans la catégorie des contextes obligatoires, des cas (comme 6. ou 3. voire 8.) où la liaison peut être réalisée sans enchaînement, avec coups de glotte et qu'Encrevé (1988) ne classe pas d'ailleurs dans la catégorie des liaisons obligatoires.

Si l'on s'en tient à une définition plus restrictive des contextes véritablement obligatoires et n'entraînant aucune variation, telle qu'ils ont été définis par Encrevé (1988), on retiendra dans les contextes de Boula de Mareuil les cas 1, 4, 5 auxquels s'ajoutent des formes figées (*Les Etats-Unis, tout à coup*) et selon moi<sup>15</sup> le contexte ouvert par la proposition « en » quand elle n'est pas suivie d'article (ex : « en avance », « en Italie », « en attendant », « en allant ») que Boula de Mareuil répertorie en 7.

On voit donc que les enfants mettent en place leur généralisation sur un seul contexte obligatoire et que ce contexte est, si je peux m'exprimer ainsi, le « plus obligatoire », c'est à dire celui qui ne donnera lieu à quasiment aucune variation

---

<sup>15</sup> Sur la base de mon intuition de locuteur francophone et sans vérification dans un corpus de données.

(liaison toujours obligatoire et toujours enchaînée), si ce n'est l'interprétation segmentale de la consonne définie par le déterminant (un / des ).

#### 2.2.2.2. *Acquisition de la liaison, nouvelle proposition*

J'ai à partir de là envisagé que pour les enfants, l'acquisition de la liaison obligatoire sur le contexte article + nom est un phénomène exclusivement conditionné par des contraintes morpho-syntaxiques<sup>16</sup> et qui serait déclenché par bootstrapping morphologique. On peut dès lors arguer que pour eux, à ce stade, l'acquisition de la liaison obligatoire sur ce contexte constitue la mise en place d'une représentation phonologique de la CL (doublement flottante) qu'ils vont pouvoir ensuite mettre en œuvre librement dans des contextes variés dont un certain nombre sont non obligatoires.

Cette conception de la CL, comme étant un objet doublement flottant, nous permet d'expliquer, à partir d'un même formalisme, les différences d'acquisition pour la liaison obligatoire et facultative et de mettre en continuité les données d'acquisition et les productions adultes. Dans les deux cas, on peut envisager que les fautes de liaison que font les enfants ne montrent pas seulement qu'ils sont en train d'apprendre qu'il existe dans leur langue des consonnes flottantes. Ces fautes reflètent également la mise en place de contraintes d'association différenciées pour les liaisons obligatoires et les liaisons facultatives.

Je propose donc sur cette base de réexpliquer les 4 stades retenus au ch1. 3. 4. 2. 2 de la manière suivante

#### **Stade 1 Tout est associé par défaut**

La plupart des substantifs sont produits avec un déterminant non segmenté. A ce stade, l'enfant dispose d'un lexique très restreint (stade des 50 mots, Vihman, 1996). Il travaille sur la base d'une forme lexicale où les attaques consonantiques sur la frontière gauche et en interne de mot peuvent être indifféremment vides ou remplies. La représentation de l'unité est certainement une forme « globale » non

---

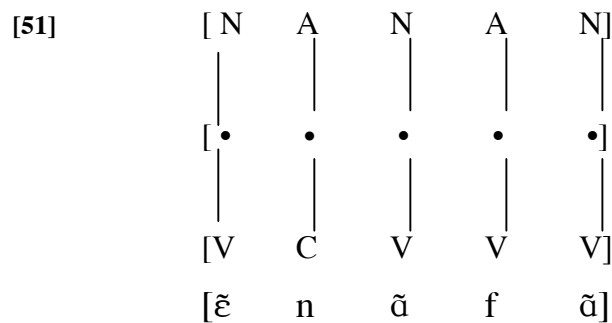
<sup>16</sup> Je n'ai pas dans mes données d'erreurs de liaison portant sur le contexte pronom + verbe (ils ont). Chevrot ne mentionne pas non plus d'occurrences de ce type.

analysée. Les données de perception confirment ce fait (Hirsch-Pasek & al., 1987, Jusczyk & al., 1992). On ne constate pas à ce stade, « d'erreurs de liaison » à proprement parler.

Dans le modèle autosegmental retenu, on peut considérer qu'à ce premier stade d'acquisition, les associations sont réalisées par défaut, et en l'absence d'instructions phonologiques spécifiques non encore acquises, en 1 pour 1. Les tires segmentales et syllabiques sont mises en relation de manière alignée avec la position squelettale leur faisant face [51], tout au long de la chaîne parlée, indépendamment des frontières lexicales. Aucune tire n'est autonome, tout segment est associé.

Ceci correspond aux premières productions<sup>17</sup> qui vont se caractériser en général par

- l'absence de structures syllabiques branchantes
- l'absence de noyaux et d'attaques vides en interne de mot
- des segmentations de la chaîne parlée systématiquement CVCV (Peters 1985),
- une majorité de mots formés sur le schéma CVCV ou (v)CVCV

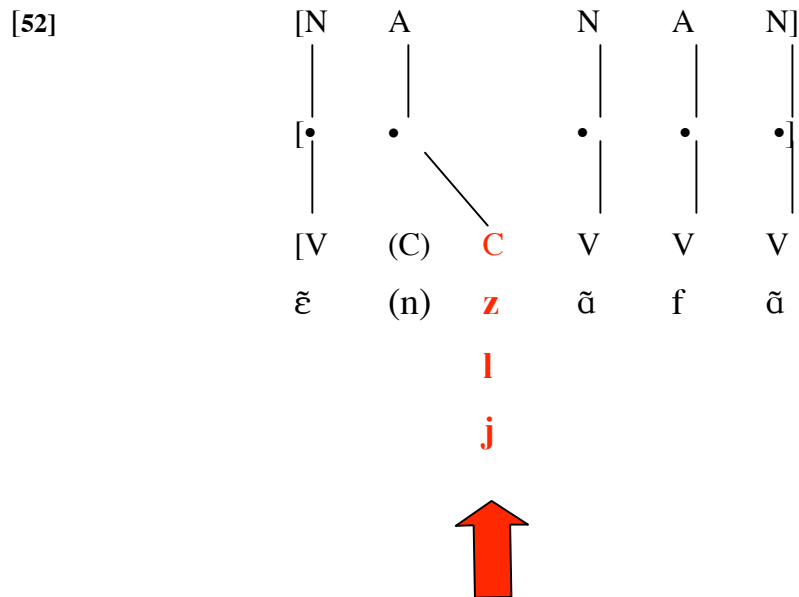


<sup>17</sup> Comme c'est d'ailleurs le cas pour les données d'aphasie également qui vont avoir tendance à être simplifiées pour sauvegarder des schémas CVCV (Valdois & Nespoulous, 1994)

## Stade 2 désassociation-diversification du contenu segmental

Les erreurs de liaison apparaissent au moment du découpage de la chaîne quand un choix doit être fait quant à la segmentation de suites comme par exemple [sedezelefã], [uiletõnelefã] et la détermination de la forme de l'unité lexicale et du déterminant.

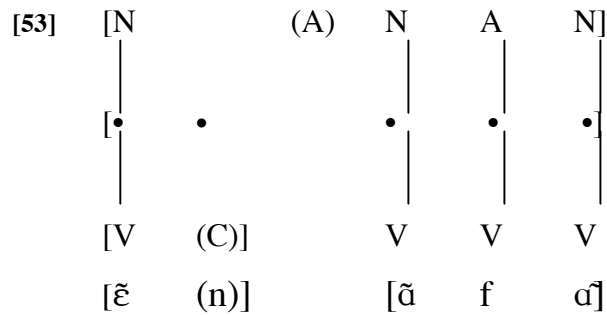
Comme je l'ai soutenu au ch1. 3. 4. 2. 2, les erreurs de liaison sont une des manifestations précoces (avec l'acquisition des structures syllabiques branchantes) de l'indépendance des positions prosodiques et des contenus segmentaux qui y sont rattachés. On peut penser que les erreurs de liaison manifestent le début d'une logique d'association qui ne fonctionne plus en 1 pour 1 en alignant comme à l'étape précédente la position syllabique et la consonne avec la position squelettale leur faisant face. A ce stade la position syllabique est associée au point squelettal (en accord avec le MOP les enfants vont privilégier le rattachement de le remplissage de l'attaque de mot 2 pour maximaliser l'attaque au détriment de la coda précédente) mais le contenu segmental est détaché et devient « flottant » [52] même si la position syllabique ne l'est pas encore. Les enfants associent assez variablement avec divers contenus segmentaux déterminés le plus souvent par des inférences statistiques en fonction des contextes (dont les contextes de liaison) dans lesquels le mot a été rencontré. Ils utilisent donc préférentiellement [l], [n], et [z] [ẽlelefã], [denelefã], [ẽzelefã]. Mais yod [j] peut être également utilisé par défaut (Claire 24 mois), ou le choix peut résulter d'une harmonie consonantique ([f] dans le cas de éléphant [ẽfefefã] pour Claire 27 mois) ou être strictement idiosyncrasique, comme dans le cas de [m] ou [p] pour *arc-en-ciel* [lãmakãsjɛl], [lɔpakãsjɛlamwa].



### Stade 3 : acquisition du double flottement

Ce stade est identique au stade 3 proposé en 3.2.3.2. La disparition des erreurs de liaison est consécutive à un phénomène de bootstrapping morphologique qui permet d'encoder la consonne flottante dans les représentations lexicales sous-jacentes. Ceci suppose deux étapes [53]

- 1) La position syllabique se désassocie également du squelette et devient comme le contenu segmental « flottante ». La consonne est alors pleinement un autosegment au sens où le définit Encrevé (1988). Elle est « doublement » flottante.
- 2) L'association est « reconnue » comme un morphème acquis par bootstrapping quand le déterminant est lui même « reconnu » comme un morphème au moment de l'acquisition morphologique. A ce stade, la contrainte d'association ne peut avoir qu'une origine morphologique.



Morphème => Associer la consonne et l'attaque à la position squelettale

**Stade 4 : les facultatives sont des choix possibles quand le morphème d'association est absent.**

A partir du moment où l'autosegmentalité est reconnue et installée dans la grammaire, elle peut, dans d'autres contextes sociologiquement conditionnés être utilisée librement par choix du locuteur. La compétence intègre cette possibilité. Ceci expliquerait pourquoi on voit subsister plus longtemps des erreurs portant sur les liaisons facultatives et non enchaînées (Basset 2000). Pour ce qui concerne ce type de liaison, on peut considérer que les contextes où l'enfant locuteur sait qu'il a le choix sont tous ceux qui ne sont pas conditionnés morphologiquement. L'absence dans sa grammaire, pour ces contextes, d'un morphème imposant explicitement le rattachement dans l'attaque du mot 2 ouvre un choix et lui laisse la possibilité i) de choisir de rattacher (réaliser la liaison ou non) ii) de choisir où il rattache (dans la coda du mot 1 ou dans l'attaque du mot 2).

Le caractère facultatif des liaisons non obligatoires résulte donc du fait qu'elles ne sont pas associées morphologiquement, mais par choix. Elles apparaissent en quelque sorte dans les contextes résiduels de la liaison obligatoire.

Ceci permet d'expliquer pourquoi les obligatoires sont toujours enchaînées et n'intègrent jamais de coups de glotte ou schwa intermédiaires comme c'est le cas pour les facultatives, car l'association, quand elle est morphologiquement conditionnée ne le permet pas. On trouvera également là l'explication de la différence chronologique entre l'acquisition des obligatoires et des facultatives. Les enfants acquièrent d'abord

et exclusivement la liaison comme un phénomène morpho-phonologique sur le contexte le plus obligatoire c'est à dire dans des contextes où le lien morpho-syntaxique entre les deux mots est le plus étroit (det + nom) au moment de l'acquisition morpho-syntaxique et alors qu'on voit apparaître les erreurs sur les contextes facultatifs beaucoup plus tardivement.

### **Travaux en cours**

#### **Communications**

**Wauquier-Gravelines, S.,** Encrevé, P. & Scheer T., Liaison in French, towards an unified explanation of variation, *colloque PFC, Phonological Variation, the case of French*, CASTL, Tromso, 25-27 août 2005

### **3. Traitement de la liaison chez l'adulte, nouvelles perspectives**

#### **3.1. Introduction**

Ces travaux sont entrepris dans la continuité des recherches sur le traitement chez l'adulte présentées au chapitre I et visent à cet égard une explicitation, une confirmation ou une réanalyse des résultats obtenus montrant une difficulté de perception du phonème de liaison que j'avais attribuée à une « surdit  phonologique ».

Mais ils entrent  galement dans un projet plus large au sein d'une ACI Syst mes complexes en SHS «Vers un mod le dynamique de l' mergence des repr sentations symboliques dans le traitement de la parole » dirig e par No l Nguyen en collaboration avec No l Nguyen et Leonardo Lancia (Universit  d'Aix en Provence), Ma tine Bergounioux (Universit  d'Orl ans) et Betty Tuller (Florida Atlantic University, USA). La liaison est utilis e ici comme un ph nom ne phonologique variable et le but de ce projet est de tenter d'observer l' ventuelle sp cificit  cognitive de la "consonne flottante" et d' tablir par quels m canismes de traitement pr cis elle est per ue et reconnue en temps r el par les locuteurs sans g ner l'acc s au lexique.

Les exp riences sont actuellement en cours, je vais donc pr senter ici rapidement les hypoth ses que nous avons pos es, les m thodes retenues et les r sultats attendus.

Comme je l'ai explicit  au chapitre I 2.4.2., diverses analyses sont disponibles afin de rendre compte de la complexit  ph nom nologique que propose la liaison. Tant en perception qu'en production, le locuteur s'appuie sur une phonologie abstraite qui formalise sp cifiquement le ph nom ne de liaison, soit comme un segment flottant appartenant au mot 1 (Encrev , 1988, Tranel, 1990, 1992, 1995, Paradis El Fenne, 1995), soit comme un morph me latent appartenant au mot 2 (Morin & Kaye,



1982) soit comme une épenthèse (Tranel, 1981a, 1981b), s'appliquant à une forme lexicale nue, soit comme une délétion (Dell, 1970) s'appliquant à une forme lexicale portant le phonème. Dans tous les cas, quel que soit le formalisme retenu, la réalisation de la liaison est d'abord conditionnée par l'existence de contraintes morpho-syntaxiques et prosodiques. Cette conception prédit donc que le phénomène est productif, que n'importe quel mot ou non-mot de la langue commençant par une voyelle et placé, par exemple, dans un contexte de liaison obligatoire, verra se réaliser la liaison.

Mais j'avais également souligné que la conception supplétive de la liaison s'est récemment ré-imposée plus franchement par rapport à l'époque où j'ai réalisé mes premiers travaux sur le traitement de la liaison chez l'adulte dans le cadre de la phonologie « des exemplaires » privilégiant des effets de fréquence et d'usage dans l'émergence des généralisations phonologiques ainsi qu'une conception quasi-exclusivement procédurale de la phonologie.

Les travaux menés dans ce projet visent à interroger de manière plus fine les effets de fréquence et à dépasser les apories respectives de modèles de traitement « abstractionnistes » et « exemplaristes » (Nguyen, 2005).

### **3.2. Modèles abstractionnistes vs exemplaristes**

D'un côté du continuum, les hypothèses abstractionnistes (Lahiri & Marslen-Wilson, 1991 ; FUL Model, Lahiri & Reetz, 2002, Pallier & al, 2001, Levelt, 1989) retiennent la conception dualiste de la cognition humaine et proposent un modèle de la performance doté d'une composante phonologique abstraite. Dans tous les cas prévalent les postulats suivants

- i) il y a d'une part des représentations linguistiques abstraites, d'autre part des procédures de traitement
- ii) elles interagissent avec une certaine logique et dans un certain ordre (toujours le même).

Le traitement consiste alors en une simplification progressive de l'information visant à convertir le signal en représentations abstraites (Nguyen, 2005), le plus souvent sous-spécifiées (Lahiri & Reetz, 2002), ramenées à l'état le plus abstrait possible et

hiérarchiquement structurées (Fitzpatrick & Wheeldom, 2000). La production consiste en une programmation des représentations abstraites aboutissant au signal physique. Les modèles de perception (Lahiri & Marslen-Wilson, 1991 ; Lahiri & Reetz, 2002) et de production (Levelt, 1989, Levelt et al. 1999) sont différents et non-unifiés pour l'instant. Ils ne sont pas réversibles, les phénomènes d'encodage et de décodage n'étant pas réalisés dans les mêmes modalités, mais la composante phonologique (et notamment le formalisme retenu) est *a priori* la même en perception et en production.

Les modèles exemplaristes, comme les modèles épisodiques en psycholinguistique au contraire vont retenir

i) une conception exclusivement procédurale de la phonologie

« It thus appears that lexical meaning may be propositional (though I think that in many cases, the choice of lexical items may also be procedural), but that the rest of language, surely all of the grammar and phonology is procedural knowledge »

J. Bybee, *Phonology and Language Use*, 2001, p. 40

« Thinking of grammatical constructions as procedural units has profound consequences for our view of phonology. Phonology then becomes a part of the procedure for producing and decoding constructions, rather than a purely abstract, psychological system »

J. Bybee, *Phonology and Language Use*, 2001, p. 8

ii) des représentations lexicales minimales voire inexistantes, élaborées indépendamment de tout formalisme *a priori* et envisagées comme des phénomènes émergents (à partir d'un réseau ou d'un treillis (de spectres acoustiques pour LAFS ou de traces mnésiques auto-organisées pour Goldinger & Azuma, 2003), ou un stockage d'exemplaires ou tokens intégrant la variabilité phonologique et phonétique des occurrences (Bybee, 2001)

« Units such as morpheme, segment, or syllable are emergent in the sense that they arise from the relations of identity and similarity that organize representations »

J. Bybee, *Phonology and Language Use*, 2001, p. 7

iii) l'idée que l'expérience et en particulier la fréquence d'usage affectent les représentations

« Units such as morpheme, segment, or syllable are emergent in the sense that they arise from the relations of identity and similarity that organize representations »

J. Bybee, *Phonology and Language Use*, 2001, p. 7

De telles propositions supposent une absence de représentation abstraite unique de la liaison et envisagent les opérations cognitives non plus comme un calcul abstrait sur des objets symboliques ou subsymboliques mais comme un traitement direct de l'objet physique par un encodage permanent des informations sous forme de traces, exemplaires dans une mémoire auto-organisée et dynamique, s'adaptant au contexte. Elles s'appuient sur l'hypothèse d'une très puissante capacité mémorielle.

Comme le souligne Nguyen (2005), la métaphore informatique utilisée à partir des années 50 pour rendre compte de la cognition humaine, et la limitation en mémoire qu'elle a longtemps supposée, ont profondément influencé la conception du traitement de la parole qui en découlait. Le traitement devait « s'apparenter à un processus de réduction de l'information » (Nguyen 2005) et de conversion du signal en représentations abstraites réduisant le coût mémoriel de traitement. Or, pour les modèles épisodiques comme celui que proposent Goldinger & Azuma (2003), la capacité mémorielle n'est plus seulement envisagée d'un point de vue quantitatif, mais également d'un point de vue qualitatif et dynamique. Dès lors, la question de l'économie des ressources en mémoire devient caduque. Cette logique poussée à terme amène donc à envisager qu'il n'y ait plus ni grammaire ni généralisations abstraites. La cognition est alors envisagée comme une mémoire dynamique surpuissante dotée d'un système de calcul où production et perception résultent d'un matching direct entre input-mémoire et mémoire-output dans une dynamique assez proche du rapport stimulus-réponse. La formulation d'hypothèses sophistiquées concernant les représentations linguistiques devient une question secondaire puisque les représentations mentales sont remodelées en permanence par l'usage – et tout particulièrement la fréquence d'usage.

### **3.2 Surdit  phonologique, fr quence, repr sentation phonologique et substance acoustique**

Comme je l'ai expos  au chapitre I, pour des contextes structurellement identiques (NP : det + nom, et adj +nom), la CL est d tect e beaucoup plus difficilement que la consonne initiale de mot, alors qu'il n'y a pas de corr lation statistique entre la diff rence de r alisation acoustique de la CL et la difficult    la d tecter. Je n'ai pas cependant, m me si j'ai contr l  la fr quence des mots exp rimentaux  tabli   cette  poque d'hypoth se particuli re interrogeant le r le particulier de la fr quence d'usage. Les contextes sont identiques pour le mot 1 (un navire / un avion) et les « mot » 2, sont de fr quence quasi- quivalente, ce qui laisserait supposer que ce facteur n'influence pas le traitement.

Des  tudes psycholinguistiques ont  t  r alis es pour tester plus syst matiquement l'influence de la fr quence sur le taux de r alisation des liaisons.

Elles ne permettent pas non plus d'affirmer de mani re d finitive que la fr quence joue un r le d terminant dans la production et le traitement des liaisons. Fougeron & al. (2001a, 2001b), Fougeron & Delais-Roussarie (2004) montrent en effet que la fr quence de mot 2, la co-occurrence de mot 1 et mot 2, le taux de probabilit  d'occurrence de mot 2 apr s mot 1, ainsi que la comp titivit  lexicale (« petit  l phant » et « petit t l phone » vont  tre jusqu'  la production de la troisi me voyelle de tr s forts comp titeurs) ne sont pas des facteurs statistiquement d terminants conditionnant la r alisation de la liaison en fran ais. Selon ces auteurs, seule la fr quence du mot 1 peut conditionner le taux de r alisation de la liaison, sans qu'on puisse savoir clairement si la fr quence du mot 1 (par exemple « un » dans « un enfant ») refl te une pure fr quence d'occurrence lexicale ou le signe de l'existence d'une structure syntaxique contraignant la liaison (un NP, form  d'un clitique et d'un nom).

Nos recherches visent donc à réétudier très systématiquement ce phénomène de « surdit  phonologique » au segment de liaison, d'absence de d tection du phon me en contexte de liaison. Il nous para t que, conform ment   l'hypoth se que j'ai d fendue, cette difficult  de d tection ne peut pas  tre exclusivement et prioritairement guid e par la fr quence d'usage des mots stock s avec leur contexte de liaison (Bybee, 2001). Si c' tait le cas, les unit s « un navire » et « un avion », un « petit  l phant » et « un petit t l phone » seraient stock es comme des unit s lexicales avec le d terminant voire l'adjectif. Dans les deux cas, si ces NP sont de fr quence  quivalente, le « n » serait consid r  comme une consonne interne de mot et la CL n'ayant aucun statut phonologique particulier ne devrait pas  tre plus difficile   d tecter. Les effets de fr quence  tant n anmoins tr s robustes en perception, on ne peut pas exclure non plus une influence au moins partielle de l'usage, lors du traitement des liaisons en temps r el. Les exp riences ont donc pour but d' tablir plus clairement le r le respectif des informations phonologiques, fr quentielles et acoustiques dans le traitement de la liaison en parole continue.

Dans une premi re exp rience, nous avons proc d    une t che de d tection de phon mes dans les contextes i) de liaison, ii) d'encha nement iii) initiale de mots afin de confirmer ce ph nom ne qui a  t   galement r pliqu  par Hall  (2003, cp.). Le but est de r pliquer dans des conditions syst matiquement contr l es l'effet de surdit  phonologique et d'en expliciter les conditionnements.

Nous r alisons un contr le simultan  de la diff rence de r alisation acoustique des consonnes initiales, consonnes de liaison, et consonnes d'encha nement ainsi qu'un contr le de la sensibilit  des sujets   une  ventuelle diff rence de r alisation acoustique (exp rience de cross-splicing). Enfin, nous manipulons la fr quence des mots en liaison ainsi que la fr quence de co-occurrence des mots en liaison.

Dans la phase suivante du projet, le travail consiste ultimement   rendre compte de la notion de « segment flottant » au sein d'un mod le dynamique non

linéaire (Tuller, 2003) où la CL est envisagée comme une transition entre deux possibilités de catégorisation perceptive (i.e deux points d'attraction de l'espace perceptif) respectivement coda du mot 1 ou attaque du mot 2. Dans un tel modèle, nous supposons donc qu'en temps réel l'auditeur opère une modification dynamique de ces hypothèses de segmentation passant d'une analyse VC # V à V # CV. Cette possibilité, pour la CL de permettre un traitement dynamique non linéaire est reflétée par sa nature flottante (au fait qu'elle ne soit associée à rien), qui la met perceptivement « en suspens » (cf. mon analyse au chapitre I 2.4.2.3), en attente d'une position et non soumise à la segmentation qui s'applique aux lignes syllabiques et segmentales. Ceci pourrait expliquer qu'elle ne soit pas traitée en temps réel de la même manière qu'une consonne ordinaire même si elle en a quasiment les caractéristiques acoustiques (ce qui explique qu'elle ne soit pas détectée).

## **Travaux en cours**

### **Communications**

Nguyen, N. & **Wauquier-Gravelines, S.**, Lancia, L. & Tuller B. (2005), Processing liaison consonants in French : new acoustic and perceptual data, *PAPi Conference*, Barcelone, 20-21 juin 2005

Nguyen, N., Lancia, L., Bergounioux, M., **Wauquier-Gravelines, S.**, & Tuller, B. (2005). Role of training and short-term context effects in the identification of /s/ and /st/ in French, *ISCA Workshop on Plasticity in Speech Perception (PSP2005)*, Londres, 15-17 juin 2005

**Wauquier-Gravelines, S.** & Nguyen, N. (2005), Réalisme psychologique des segments latents : intuition phonologique ou effet de fréquence ? *Colloque du réseau français de Phonologie*, Aix-en-Provence, 3-5 juin 2005

**Wauquier-Gravelines, S.** & Nguyen, N. (2005), Is phonological deafness frequency-dependent ? the case of liaison consonant in French, *13th Phonology meeting*, Manchester, 23-25 mai 2005

#### **4. Acquisition syllabique, *learning paths* et marque**

Ce projet de recherche, est, comme je l'ai précisé en introduction, réalisé au sein d'un groupe européen informel mis en place à partir de 2003 avec Paula Fikkert (Nimègue), Clara Levelt (Leiden) et M-J Freitas (Lisbonne). Une communication a été présentée et un article est en préparation.

Ce projet repose sur une approche typologique de l'acquisition syllabique tentant d'établir l'existence, lors de l'acquisition des structures syllabiques, de *learning paths* conditionnés par la structure phonologique de la langue cible. Une partie du projet a été défini en commun (critères d'analyses et choix théoriques). Les hypothèses et analyses portant plus spécifiquement sur le français reflètent ma part de participation à ce projet. L'extension et la mise en place officialisée de ce projet typologique européen en vue de permettre des échanges de données et d'affiner les analyses sont en cours et leur concrétisation attendue pour 2006.

La démarche générale repose sur le constat suivant : les contraintes de marque syllabiques universelles ne peuvent pas à elles seules expliquer la variété des patterns observés dans les langues germaniques (anglais, allemand, néerlandais) et romanes (français, portugais). L'hypothèse est donc que la marque syllabique interfère avec d'autres facteurs linguistiques et psycholinguistiques dont on cherche à établir l'existence et la nature.

##### **4.1. Introduction**

Bien que l'acquisition syllabique ait donné lieu à beaucoup moins de travaux que l'acquisition segmentale, il est généralement admis que, conformément aux propositions de Jakobson (1941/1969), l'ordre d'apparition des structures syllabiques comme l'ordre d'apparition des segments est contraint par la marque. Pour ce qui concerne la syllabe, la structure universelle de marque 0 étant la structure CV (Kaye & Lowenstamm, 1979), on peut donc faire la double prédiction suivante : on s'attend à ce que, quelle que soit la langue cible, les premières structures syllabiques produites par les enfants soient les syllabes CV (Fikkert, 1994, Fikkert et al., 2004, Moskowitz,

1970), et qu'il n'y ait aucun constituant branchant précoce (Fikkert, 1994) et aucun constituant vide précoce (Fikkert, 1994). On s'attend donc à ce que les enfants détruisent les clusters consonantiques et les consonnes finales mais également qu'ils évitent les attaques vides.

Conformément à ces prédictions, Fikkert (1994) a donc proposé que l'acquisition syllabique suivent les étapes suivantes :

<b>Stades d'acquisition</b>	<b>Types syllabiques</b>
Stade 1	CV = O V ☞ <i>attaques = toujours occlusives</i>
Stade 2	CV or V = OV or xV ☞ <i>attaques = occlusives ou attaques vides</i>
Stade 3	CV or V = OV ou xV ou CV ☞ <i>toutes les attaques simples = occlusives, nasales, fricatives ou liquides</i>
Stade 4	CV or V or CCV or CVC ☞ <i>toutes les attaques simples = occlusives, nasales, fricatives ou liquides</i> ☞ <i>les attaques branchantes</i> ☞ <i>les codas</i>

Tableau 6 : Stades universels d'acquisition des structures syllabiques (d'après Fikkert, 1994).

Or les faits empiriques, notamment en français et portugais contredisent ces prédictions.

#### **4.2. Attaques vides en portugais et en français**

On constate en effet l'existence d'attaques vides précoces chez les enfants lusophones (Freitas, 1996, 2003, [54], [55]) et francophones (citées in Boysson-Bardies (1996)), CORPAC Nantes) au stade CV et en concurrence avec les structures CV.



**[54] Joao**

ÂGE	CIBLE	FORME CIBLE	FORME PRODUITE
00; 11.06	<i>eau</i>	[ʼagwa]	[ʼaβa]
01;00.12	<i>est</i>	[ʼɛ]	[ʼɛ]
01;02.01	<i>Il n'y a pas</i>	[nɛw'a]	[a:]

**[55] Ines**

ÂGE	CIBLE	FORME CIBLE	FORME PRODUITE
01;00.25	<i>eau</i>	[ʼagwa]	[a]
01;00.25	<i>est</i>	[ʼɛ]	[ʼɛ]
01; 01.30	<i>Il n'y a pas</i>	[nɛw'a]	[a:]

Pour le français, les exemples présentés pour les données [56] sont extraites de données précoces de Claire. En [57] et [58] sont présentées les “échelles de Guttman” pour les syllabes internes et pour les syllabes finales construites à partir de l’analyse longitudinale de deux enfants Claire (de 1.9 à 2.5 ans) et Louis (de 1.10 à 2.6 ans) (CORPAC, Nantes).

**[56] Claire**

[akɔain] *encore une mandarine* 1, 6

[øøki] *j'en veux une petite* 1,4

[øago] *j'en veux un gros* 1, 3

[lalæjɛ] *Lalla, elle vient* 1, 5

## [57] Echelle de Guttman, syllabe finale (Fikkert et al. 2004 &lt;CORPAC, Nantes)

Claire	CV	V	CVC	CCV	VC	CCVC
1-4	+	+	-	-	-	-
1-6	+	+	(+)	+	-	-
1-7	+	+	+	+	-	-
1-8	+	+	+	+	+	-
1-9	+	+	+	+	+	+
Louis	CV	V	CVC	CCV	VC	CCVC
1-2	+	-	-	-	-	-
1-3	+	+	-	-	-	-
1-11	+	+	+	-	-	-
1-13	+	+	+	+	-	-
1-16	+	+	+	+	(+)	(+)

## [58] Echelle de Guttman, syllabes internes et initiales (Fikkert et al. 2004 &lt; CORPAC, Nantes)

Claire	CV	V	CCV	CVC	VC	CCVC
1-3	+	+	-	-	-	-
1-8	+	+	+	-	-	-
1-10	+	+	+	+	-	-
1-16	+	+	+	+	+	-
1-17	+	+	+	+	+	+
Louis	CV	V	CCV	CVC	VC	CCVC
1-4	+	+	-	-	-	-
1-12	+	+	+	-	-	-
1-16	+	+	+	+	-	-

On constate donc que de manière très précoce, Claire (1.9) comme Louis proposent (1.10), en concurrence de syllabes CV et de syllabes xV à attaque vide.

Par ailleurs, si les syllabes CV non marquées apparaissaient de manière exclusive quelle que soit la langue cible, les premières unités produites seraient constituées exclusivement de structures CV pleines (mots de type [babo] [tato]). Or on constate que, dans les productions précoces de certains enfants, certaines unités lexicales sont produites quasiment sans consonnes (cas présentés par Boysson-Bardies, 1996, Vihman, 1996, Vihman et al. 1994).

C'est le cas de Claire (CORPAC, Nantes).

- [59] [aəwa] *au-revoir* 1,4  
 [aε] *arrête* 1,5  
 [an] *âne* 1,8  
 [eã] *éléphant* 1,11

Ceci montre donc que le modèle d'entrée dans le langage où la première structure syllabique disponible serait CV= OV, n'est pas suffisant pour rendre compte des données observables dans toutes les langues.

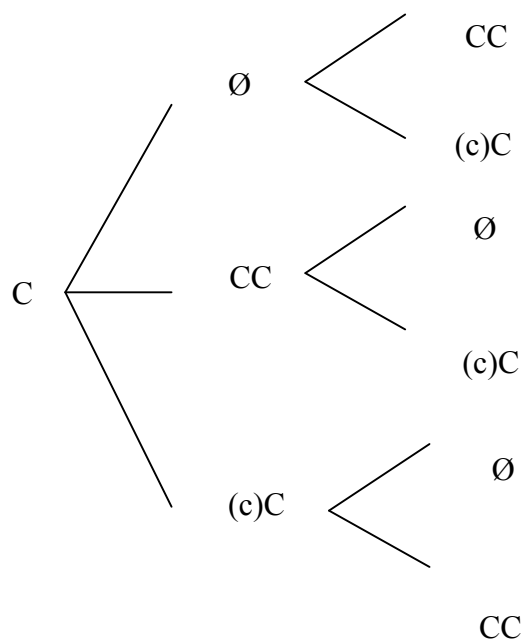
CV n'est donc pas un pattern d'acquisition universel : les effets de marque syllabique isolés ne peuvent rendre compte ni des données d'acquisition dans toutes les langues, ni d'éventuelles stratégies individuelles retenues par les enfants. On doit envisager, comme le propose Moskowitz, qu'au-delà de ce premier stade CV, les enfants vont utiliser des *learning paths* différents pour acquérir les attaques syllabiques (Fikkert et al. 2004, [60])

“Beyond the primary acquisition of the (CV), there is no unique and specific pattern, or order of acquisition of syllable types, which could reasonably be predicted as that which all children follow”

Moskowitz (1970), *Language*, 46, p. 429

Cependant, contrairement à ce que semble envisager Moskowitz, la théorie de la marque prédit également que les *learning paths* ne sont pas en nombre illimité et que leur existence n'est pas purement idiosyncrasique et aléatoire.

[60]



Et nous avons en effet constaté les *learning paths* suivants (pour les syllabes accentuées<sup>18</sup>) qui étaient systématiques et semblaient être conditionnés par la langue cible que l'enfant était en train d'acquérir.

### Langues germaniques

- Néerlandais : CV > CVC > V > CCVC or CVCC
- Anglais : CV > CVC > V
- Allemand : CV > CVC or VC? > V

### Langues romanes

- Portugais : CV > V > CVC > VC > CCV
- Français : CV > V > CVC > CCV > VC > CCVC

Nous nous sommes donc interrogé sur les facteurs susceptibles de contraindre la marque syllabique ou d'interagir avec elle. Nous avons considéré que ces stratégies reflètent une logique développementale structurelle qui pourrait être conditionnée de manière plus complexe par la convergence de la marque syllabique et d'autres facteurs propres à la langue cible comme le rythme (Fikkert et al., 2004).

## 4.3. Facteurs possibles contrariant ou complétant l'effet de marque

### 4.3.1. Fréquence de distribution syllabique dans l'input adulte

Nous avons d'abord comparé les fréquences de distribution dans les différentes langues concernées afin d'observer une éventuelle différence de distribution des structures syllabiques dans l'input adulte permettant d'expliquer les *learning paths* infantiles. Le tableau 7 ci-dessous reprend pour le portugais et le

---

<sup>18</sup> En français, j'ai comptabilisé comme syllabe accentuée les syllabes de fin d'énoncé produit par les enfants. Soit la dernière syllabe d'un groupe prosodique.

néerlandais les valeurs présentées dans Fikkert & al. (2004). Pour le français, les pourcentages de distribution mentionnés sont extraits de Vallée (2004)<sup>19</sup>

	Néerlandais	Français	Portugais
<b>CV</b>	<b>42%</b>	<b>54%</b>	<b>65%</b>
<b>CVC</b>	30%	18%	16%
<b>VC</b>	<b>11%</b>	<b>2%</b>	<b>3%</b>
<b>V</b>	<b>4%</b>	<b>8%</b>	<b>11%</b>
<b>CCV</b>	-	11%	5%
<b>CVCC</b>	4%	4%	-
<b>CCVC</b>	3%	3%	-

Tableau 7 : fréquence de distribution des types syllabiques  
(d'après Fikkert et al. 2004, Vallée, 2004)

Il apparaît que les valeurs de distribution des structures CV *vs* VC ou V montrent que la fréquence de distribution des structures syllabiques dans l'input adulte ne peut pas être considérée comme un facteur déterminant de l'acquisition. Ces données n'offrent absolument aucune explication à l'existence d'attaques vides précoces en portugais et français (*vs* absence dans les langues germaniques). Les fréquences pures prédisent les mêmes *learning paths*, quelle que soit la langue cible.

Par ailleurs, l'affirmation selon laquelle la proportion de syllabes x ou y (p. ex. CV ou CVC ou VC) dans l'input d'entrée conditionne la production de ces structures syllabiques par les enfants est un peu spéculative. On peut en effet se demander, si effet de fréquence il y a, de quelle fréquence on parle (cf à ce sujet chIV): à quelle (s) syllabes l'enfant est-il sensible dans l'input (syllabe accentuée, syllabe initiale) et quelles différences de structure seront pour lui précocement pertinentes (CV *vs* CVC ou CV *vs* VC ?)

<sup>19</sup> Je remercie Nathalie Vallée qui m'a transmis une copie de sa présentation Powerpoint 2004, qui m'a permis de réactualiser les valeurs pour le français par rapport aux données proposées dans Fikkert et al. 2004.

#### 4.3.2. Structure du lexique

Trois spécificités nous sont apparues au moment de l'analyse des données quand nous avons opéré les segmentations lexicales des énoncés et le comptage des syllabes.

Nous avons d'abord constaté que les données des enfants portugais et français comportaient beaucoup plus de voyelles épenthétiques que les données des enfants néerlandais produisant des clusters consonantiques plus fréquents. Pour les données de langues romanes, on constate donc une tendance à éviter les syllabes fermées et les noyaux vides au profit des voyelles (CVCVC vs CVCVCV) et donc à ajouter ou couper des syllabes en vue d'une simplification des structures par évitement de coda ou clusters et préservation des noyaux. Ceci explique également la présence des voyelles épenthétiques en attaque de mot. On constate, en français, plus précocement et en plus grand nombre, l'existence de "mots" commençant par une voyelle de type ([ə] [le] [ɔ] [a] [y]) dont on a de fortes raisons de penser qu'il peut s'agir d'un proto-déterminant qui subsiste même quand les premières syllabes du mot sont tronquées (pour les arguments et l'analyse, cf Ch I 3.4.2). Or, en portugais comme en français, les déterminants sont cliticisés sur le mot qui suit, alors qu'en néerlandais ils se rattachent au mot qui précède (Fikkert et al. 2004). La syllabation conséquente va donc entraîner, pour le français un taux de "mots" (au sens unité lexicale prosodifiée) commençant par une voyelle, très élevé.

Nous avons également établi une différence entre ces langues en terme de longueur des mots dans les stades précoces d'acquisition (les différences sont portées dans les tableaux 8 et 9 Fikkert et al. 2004). Ces données montrent que les enfants francophones vont avoir tendance à produire plus de "mots" ou "d'unités" longues que les enfants néerlandais au stade précoce et au stade plus tardif

Stade précoce	Néerlandais	Français	Portugais
Monosyll.	65 - 90%	< 35 %	< 35
Polysyll.	10 - 35%	> 65%	> 65%

Tableau 8 : différences de longueur de mots stade précoce  
(d'après Fikkert et al. 2004)

Stade plus tardif	Néerlandais	Français	Portugais
Monosyll.	diminuent de 40%	diminuent de <10%	diminuent de 10%
Polysyll.	augmentent de de 60%	augmentent de > 90%	augmentent de > 95%

Tableau 9 : différences de longueur de mots stade précoce  
(d'après Fikkert et al. 2004)

Enfin, un dernier écueil s'est dressé quand il a fallu que nous déterminions respectivement pour les langues concernées ce qu'étaient les syllabes "accentuées". Pour les langues germaniques concernées comportant un accent lexical, la définition était assez claire et les prédictions (Fikkert, 1994) quant à la réalisation différente des syllabes accentuées et non-accentuées, complètement vérifiables. Pour le français les critères permettant de retenir des syllabes accentuées étaient plus difficiles à définir puisque par définition la syllabe accentuée est la dernière du groupe prosodique produit. Ceci suppose donc que les syllabes accentuées dans les chaînes produites et entendues par les enfants francophones ne sont pas "intrinsèquement" fortes ou faibles mais que leur valeur accentuelle dépend de la manière dont elles sont prosodifiées. Comme nous l'avons souligné au ch. I.4.2, le fonctionnement du français ne peut donc pas être rabattu sur celui des langues à accent lexical et le statut de l'accent de fin de groupe ne permet pas que l'on considère *a priori* qu'il soit traité par les enfants francophones au même titre que l'accent lexical dans les langues germaniques. Il semblerait donc que les différences entre les structures accentuelles s'appliquant aux représentations lexicales dans ces langues puissent avoir un effet sur l'acquisition syllabique et la disponibilité respective des structures chez les enfants et puisse être un des facteurs contraignant les *learning paths* décrits plus haut.

A partir de là, l'autre piste que nous avons également envisagée est celle de l'influence du rythme et de la prosodification des unités lexicales dans la langue cible.

#### 4.3.3. Influence du rythme

Comme je l'ai montré plus haut, on voit chez certains enfants francophones (Boysson-Bardies, 1996) et tout particulièrement dans les données de Claire une relative indépendance entre les voyelles et les consonnes. Les mots produits précocement peuvent consister en une suite de noyaux vocaliques sans (ou avec très peu de) consonnes et en particulier pour les séquences longues

[61]

[aeεeme] *ça, c'est fermé, fermé*, Claire (1,11)

[eøiugaje] *Je veux vous coiffer*, Claire (2,1)

[ejejðijeaεt] *Le biberon, il est tombé sur la tête*, Claire (2,1)

[eitɔ:] *C'est Victor*, Claire (2,4)

Et l'on peut suggérer que dans les situations où l'inventaire phonologique consonantique est lacunaire (quand les consonnes cibles ne sont pas des occlusives), Claire adopte une stratégie en deux temps : il semblerait qu'elle remplit préférentiellement les noyaux afin de respecter la longueur prosodique de l'unité cible en posant les voyelles. Le mot semble être produit comme un nombre de sommets vocaliques, plutôt que comme un nombre de syllabes pleines. Ces remarques peuvent d'ailleurs être rapprochées de ce qui est généralement confirmé par les données développementales et les bases de données : l'inventaire vocalique est disponible avant l'inventaire consonantique (Vallée & Boe, 2000).

Ces faits recourent partiellement les observations de Freitas (2003) montrant également une préférence pour la mélodie vocalique. Les données de portugais l'amènent à considérer que de manière assez similaire que pour le français, lors de



leur acquisition lexicale et de la production des premières unités reconnaissables, les enfants portugais vont manifester une préférence pour la mélodie vocalique.

Freitas (2003) a suggéré une explication rythmique à partir des propositions de Ramus et al. (1999). Cette explication nous a semblé une piste susceptible d'éclairer plus généralement les différences constatées entre les données d'acquisition romanes (français et portugais) et germaniques (néerlandais, anglais, allemand).

Du point de vue rythmique, les langues du monde sont traditionnellement classées à partir des unités métriques pertinentes qu'elles sélectionnent (ped métrique ou syllabe). Sont distinguées classiquement les *stress-timed languages* et les *syllable-timed languages* (Abercrombie, 1967, Dauer, 1983)<sup>20</sup>, le français étant considéré traditionnellement comme un bon exemple de *syllable-timed language*.

L'appartenance à une de ces catégories peut être corrélée à un certain nombre de patterns syllabiques et de phénomènes phonologiques récurrents : réduction vocalique, contrastes de quantité et gémiation et nombre important de structures syllabiques complexes vont plutôt caractériser les *stress-timed languages*, alors que les *syllable-timed languages* vont se caractériser par la resyllabation des chaînes, l'introduction de noyaux dans les clusters consonantique (comme le schwa en français) et un inventaire syllabique plus restreint privilégiant les syllabes ouvertes (Dauer, 1983).

Pourtant les travaux les plus récents ont montré que l'isochronie physique caractérisant les différents types de langue est assez difficile à établir clairement et que la dichotomie entre ces deux types de langue reposerait davantage sur la perception qu'en ont les auditeurs (Dauer, 1983; Dauer, 1987; Ramus et al., 1999, Roach, 1982).

Ramus et al. (1999) ont proposé un travail permettant d'établir un lien entre la saillance vocalique en perception et les structures syllabiques éventuellement perceptibles et perçues par les nouveaux-nés. Leur hypothèse repose sur une série

---

<sup>20</sup> Cette analyse classant les langues en deux catégories *stress-timed* vs *syllable-timed languages* a pu être contestée. Wenk et Wioland (1982) apportent des arguments pour discuter de la pertinence d'une telle analyse pour le français. A ces deux catégories s'est en outre ajoutée la catégorie des *mora-timed languages* dégagée de l'analyse métrique du japonais en particulier.

d'expériences montrant que les nouveaux-nés sont précocement sensibles aux structures rythmiques des langues naturelles et sont capables de les discriminer sur ce critère. L'idée est que les enfants ne sont pas capables pourtant d'opérer une analyse détaillée du signal et d'appuyer leur distinction sur des corrélats acoustiques fins des catégories *syllable-timed* et *stress-timed languages*. Ils vont par conséquent appuyer leur perception sur les parties du signal qui sont les plus saillantes et les plus susceptibles d'être discriminantes : les voyelles.

Ramus et al. font en effet l'hypothèse que les enfants vont percevoir le signal comme une suite de voyelles clairement discriminantes séparées par des consonnes perçues comme des intervalles de bruits non discriminants et non analysés par les bébés. Ils ont mesuré sur cette base les intervalles séparant vocaliques et consonantiques et ont dégagé trois variables

$\%V$ : qui est la proportion des intervalles vocaliques, soit la somme des intervalles vocaliques divisée par la durée totale de l'énoncé

$\Delta C$ : qui est la déviation standard de la durée des intervalles consonantiques

$\Delta V$ : qui est la déviation standard de la durée des intervalles vocaliques

et ont établi que, sur la base de  $\%V$ , les enfants nouveaux-nés sont capables d'établir des discriminations rythmique précoces. Les structures rythmiques d'une langue étant corrélés avec des patterns syllabiques spécifiques, l'hypothèse de Ramus et al. est que  $\%V$  est un indice stable et fiable permettant au nouveau-né d'inférer les structures syllabiques pertinentes dans la langue cible qu'il est en train d'acquérir.

Cette hypothèse a été remise en forme par Vigario, Frota et Freitas (2003, in Fikkert et al. 2004) qui ont construit un tableau à double entrée ( $\%V$  en abscisse et  $\Delta C$  en ordonnée, figure 10) et y ont distribué les langues suivantes : anglais, polonais, néerlandais, français, espagnol, italien, catalan, japonais et portugais.

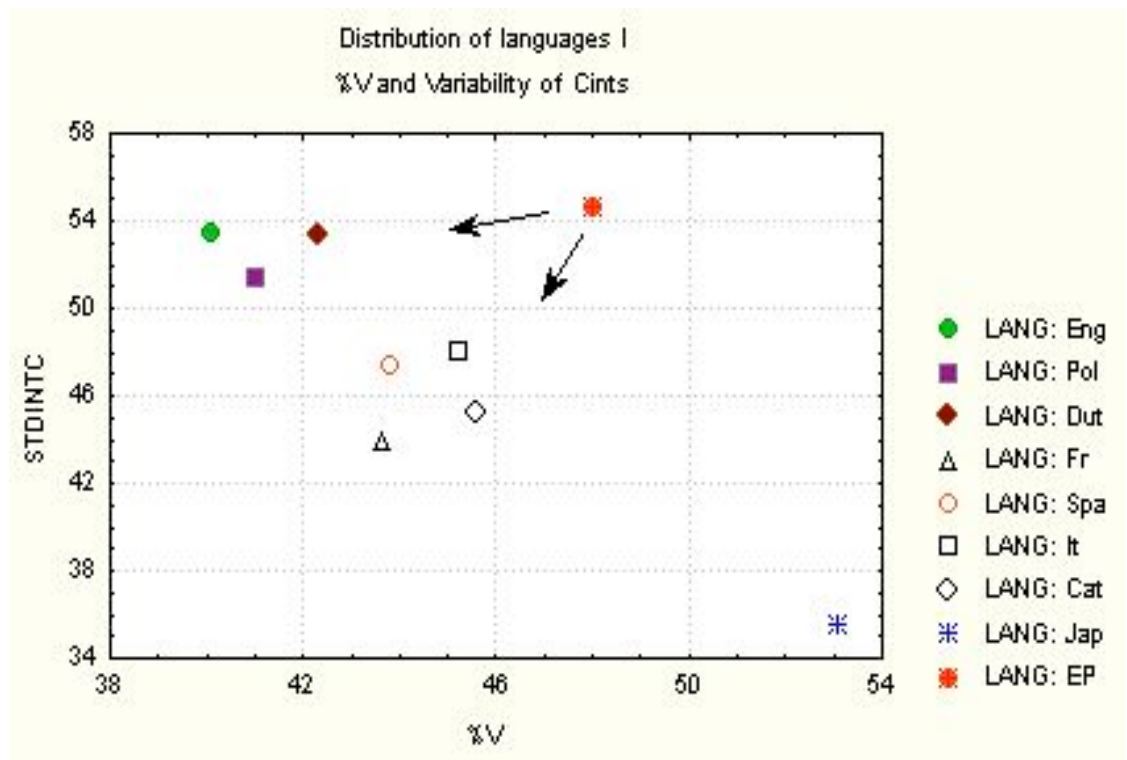


Figure 10 : dispersion des langues selon les critères %V et  $\Delta C$

L'hypothèse fait les prédictions suivantes : une langue *stress-timed* aura tendance à avoir un %V bas et un  $\Delta C$  élevé. A l'inverse, une langue *syllable-timed* aura un %V plus élevé et un  $\Delta C$  bas. Les flèches permettent de montrer comment le portugais qui est considéré comme *syllable-timed language* mais qui néanmoins connaît des phénomènes de réduction vocalique se situe bien, sur la base de ces critères entre le français et le néerlandais.

Sur cette base, nous pouvons retenir %V comme un indice susceptible d'orienter le bébé nouveau-né vers les structures syllabiques pertinentes dans sa langue : un enfant acquérant une langue *stress-timed* (avec %V bas et  $\Delta C$  élevé) aura tendance à dégager des syllabes où la proportion de consonnes par rapport aux voyelles sera plus importante que pour une langue «*syllable-timed*» où à l'inverse (avec %V élevé et  $\Delta C$  bas), la proportion de voyelles par rapport aux consonnes sera plus importante. L'enfant acquérant une langue *stressed-timed* aura tendance à privilégier les consonnes et produira donc plus de codas, d'attaques branchantes là où l'enfant acquérant une langue *syllable-timed* privilégiera les noyaux vocaliques (absence de codas et attaques vides). Cette explication est compatible avec les analyses de données

que nous avons réalisées et qui dégagent les *learning-paths* que je rappelle ci-dessous. Elle permet d'expliquer la présence d'attaques vides précoces en portugais et en français.

### Langues germaniques

- Néerlandais : CV > CVC > V > CCVC or CVCC
- Anglais : CV > CVC > V
- Allemand : CV > CVC or VC? > V

### Langues romanes

- Portugais : CV > V > CVC > VC > CCV
- Français : CV > V > CVC > CCV > VC > CCVC

Elle permet également d'expliquer les différences lexicales que nous avons constatées entre le néerlandais et le français. Il semblerait que pour étendre leur lexique à partir d'une structure CV, les enfants néerlandais vont avoir tendance à ajouter des consonnes autour de la voyelle (CV > CVC ou CCV), là où les enfants francophones vont ajouter un noyau ou une autre syllabe (CV > CV.V ou CVCV). Enfin on peut établir une corrélation entre %V et la longueur des mots. Plus une langue aura des mots polysyllabiques (comportant plusieurs noyaux vocaliques), plus %V sera élevé. Ceci est également conforme aux données du français.

Pour conclure on peut envisager que %V est un indice phonétique (parmi d'autres) qui aide à la phonologisation des structures syllabiques et qu'il interagit avec la marque. Ceci invite à envisager la marque comme un phénomène complexe faisant intervenir marque segmentale, marque syllabique, sensibilité innée de traitement à des indices acoustiques spécifiques comme les %V et capacité à les utiliser pour construire une phonologie (Ramus et al. 2000).

Enfin ces recherches invitent à envisager la marque comme un phénomène relatif et multi factoriel conditionné par les caractéristiques structurelles de la langue cible et la manière dont elles interagissent.

Par rapport à CV, V et CVC sont toutes deux marquées et se différencient toutes deux de CV par un seul degré de marque. Il n'en reste pas moins vrai que pour un enfant francophone CVC sera relativement plus marquée que V par rapport à CV, et inversement pour un enfant néerlandais V sera relativement plus marquée que CVC par rapport à CV. La théorie de la marque syllabique seule limite drastiquement le nombre de *learning-paths* qu'elle peut définir par une progression ajoutant à chaque stade un degré de marque, mais elle ne prédit pas clairement quels sont les *learning paths* attendus en fonction de la langue-cible.

L'interaction entre la marque syllabique et %V permet d'en rendre compte.

#### 4.4. Existence d'un *learning path* en yod

Sur les données de Claire, un autre travail a été réalisé, sur la base de l'analyse du corpus de Claire entre 23 et 26 mois par D. Suet-Bouret (2002) qui a permis de mettre en évidence l'existence de très nombreux yods.

J'ai tenté d'établir sur cette base si l'on pouvait considérer qu'il y avait chez cette enfant un *learning path* en yod et si c'était le cas, comment on pouvait l'interpréter.

##### 4.4.1. Etude distributionnelle

L'étude des données de Claire montre que yod peut être présent dans l'attaque syllabique d'un mot cible adulte où il est normalement attendu ([6]). Il peut être aussi la conséquence d'une métathèse ([7]).

[62] *crayon* : [krejɔ̃] > [kejɔ̃]

[63] *oreille* : [orej] > [oje]

On voit également un comportement très régulier de yod lors de la simplification des attaques occl+ yod [64a] [64b] ou la simplification de clusters consonantiques de surface [65].

[64] yod dans une attaque occl + yod

occlusive tombe

[64a] *pompier* : [pɔ̃pje] > [pɔje]

yod tombe

[64b] *les pièces* : [lepjes] > [εpe/lepε]

[65] *pardon* : [paɾdɔ̃] > [paɟɔ̃]

Mais comme le montrent également les données en [66] et [67], ci-dessous, yod sert également à remplacer systématiquement les fricatives, absentes à ce stade de l'inventaire consonantique de l'enfant et propage une harmonie consonantique palatale, y compris sur les occlusives que cette enfant utilise pourtant facilement dans d'autres contextes.

[66]

	Cible adulte orthog	Cible adulte	Product. enfant
f	<i>éléphant</i>	[elefã ]	[ejã]
v	<i>vache</i>	[vaʃ]	[ja]
s	<i>poisson</i>	[pwasɔ̃]	[pwaɟɔ̃]
z	<i>la maison</i>	[lamezɔ̃]	[lameɟɔ̃]
ʃ	<i>cochon</i>	[koʃɔ̃]	[koɟɔ̃]
ʒ	<i>manger</i>	[mãʒe]	[maɟe]

[67]

	Cible adulte orthog	Cible adulte	Product. enfant
R	<i>rouge</i>	[RUʒ ]	[ju]

[68] *coucou, caché* : [kukukaʃe] > [kukujaje]

*encore une chanson* : [ãkɔ̃vynʃãsɔ̃] > [ãkoɟɔ̃jɔ̃]

*attention* : [atãsɟɔ̃] > [aɟɔ̃jɔ̃]

*devant* : [dãvã] > [jãjã]

*les camions* : [lekamɟɔ̃] > [jaɟajɔ̃]

#### 4.4.2. Analyse

Ces données montrent que Claire utilise yod pour remplacer une fricative, en réduction des clusters et en simplification d'attaque, en alternance avec une occlusive alors que toutes les occlusives sont acquises. Mais on constate simultanément que ce remplacement est systématique et positionnel mais non conditionné par l'environnement segmental, puisque yod n'est pas produit exclusivement dans des contextes palatals, mais également entre voyelles d'arrière et arrondies : *cochon* : [koʃɔ̃] > [kojɔ̃]

Il n'est donc apparemment pas le résultat d'une palatalisation de la fricative.

On constate également que yod est toujours produit en attaque syllabique simple, sur une position qui le désigne automatiquement comme une « vraie » consonne et jamais comme une semi-voyelle. Cette utilisation du glide [j] en français, comme objet purement consonantique, correspond d'ailleurs à l'analyse qui en a été faite pour les données d'adultes en synchronie (Klein, 1991) et en diachronie (Pourin, 2003).

Pour conclure, il apparaît clairement qu'entre 23 et 26 mois, [j] semble constituer, pour Claire, une consonne d'attaque par défaut dont se sert cette enfant, soit pour combler les lacunes d'un inventaire phonémique insuffisant (dans le cas où l'attaque syllabique du mot cible adulte est une fricative ou une liquide, cf. tableau 10) soit parce qu'elle doit produire une structure branchante dont elle ne dispose pas encore. Il semblerait donc que lors de l'acquisition des consonnes elle passe par un "stade yod" qui consiste à remplir les attaques des autres syllabes que OV avec yod, quel que soit le contexte vocalique, puis à propager éventuellement le yod sur d'autres structures et au-delà du mot (Suet-Bouret, 2002).

**Stade 1: OV ou xV > Stade 2 : jV > Stade 3 : VCV > Stade 4 : CCV**

Stades d'acquisition	Types syllabiques
Stade 1	CV or V = OV or xV ☞ <i>attaques = occlusives ou attaques vides</i>
Stade 2	CV or V or jV = OV or xV ☞ <i>attaques = occlusives ou attaques vides</i> <i>remplissage des attaques vides par yod</i>
Stade 3	CV or V = OV ou xV ou CV ☞ <i>toutes les attaques simples = occlusives, nasales, fricatives ou liquides</i>
Stade 4	CV or V or CCV or CVC ☞ <i>toutes les attaques simples = occlusives, nasales, fricatives ou liquides</i> ☞ <i>les attaques branchantes</i> ☞ <i>les codas</i>

Tableau 10 : Stades d'acquisition des structures syllabiques pour Claire

Comment peut-on expliquer l'existence de ce « stade yod » et quelle est sa fonction dans la logique maturationalnelle de la grammaire de cette enfant ?

#### 4.4.3. Pourquoi yod ?

##### 4.4.3.1. Comparaison avec les données d'Alice

Un autre cas d'enfants ayant systématiquement développé des *palatal patterns* est répertorié dans la littérature (Vihman et al., 1994). Les auteurs présentent les conclusions issues de l'étude longitudinale d'Alice, enfant anglophone et décrivent les phénomènes suivants.

De manière précoce (entre 9 et 14 mois), donc au stade de transition entre le babillage varié et la phase des 50 mots, Alice développe 3 types de patterns impliquant des structures ou des segments palatalisés

- i) de 9 à 11 mois, elle produit des suites <jV> (yod voyelle) de type [ja, jo]
- ii) à partir de 10 mois elle commence à produire des diphtongues dont la deuxième cible est la voyelle palatale, [I] relâché dans des monosyllabes comme [æɪ]



- iii) à partir de 13 mois, elle commence à produire des mots clairement reconnaissables qui tous sont construits sur le modèle suivant  $\langle C_0VCi \rangle$  (*baby*, *mommy*) et à 14 mois va même jusqu'à combiner les palatales sur des positions consonantiques et vocaliques [dæɪji] : *daddy*.
- iv) Enfin elle met en place un pattern systématique où la première syllabe (accentuée) est pleinement réalisée et préservée et où elle palatalise consonnes et / ou voyelles sur les autres syllabes qui sont des syllabes faibles ne recevant pas d'accent.

Vihman et al. (1994) analysent les données de cette enfant comme le reflet d'un développement idiosyncrasique, observable d'abord dans la préférence précoce d'Alice pour des yods en période de babillage

“At 8 months, only babble was produced ; there were no identifiable words. Babble vocalisations tended to include yod to an uncommonly great extent. [...]”  
M. Vihman, et al., *First and second language Phonology*, 1994, p.22

et considèrent qu'elle a mis en place un pattern articulatoire palatal lors du babillage, qui l'amène ensuite à choisir yod comme une mélodie consonantique par défaut. Selon elles, Alice “phonologise” progressivement les consonnes par la mélodie palatale qui s'appliquerait de manière préférentielle sur une tire consonantique. L'utilisation du yod manifeste une mélodie palatale autosegmentale se propageant à défaut d'autres structures disponibles (en particulier sur les syllabes faibles, non accentuées).

“Alice's palatalisation pattern has appeared to exist in some sense separate from it's manifestation in any particular word, because it is imposed inconsistently on various portions of different words and it's effect vary from one token to the next. Palatalization appears to have a status of an autosegmental melody.  
We propose that the vocal motor scheme which is first manifested as [j] is gradually shaped into the more extensive and flexible palatal melody, [...] in a phonologically consistent way”  
M. Vihman, et al., *First and second language Phonology*, 1994, p.22

#### 4.4.3.2. 'Palatal pattern' et chemin apophonique

On remarque en effet que ces deux enfants proposant un *palatal pattern* se caractérisent également par l'existence de mélodies consonantiques et vocaliques séparées où le schème CV (occlusive / voyelle centrale) n'est pas toujours prioritaire.

On voit qu'au stade où elles disposent des occlusives et des voyelles, elles s'en servent pour produire les structures les moins marquées du système de leur langue respective (Claire produit exclusivement des syllabes OV et Alice exclusivement des syllabes OV accentuées). Dans les autres contextes, (attaques branchantes, consonnes fricatives, syllabes faibles non accentuées), le matériau segmental disponible est quasiment uniquement constitué de voyelles et les productions se caractérisent par l'existence de constituants vides. Dans ce contexte, le yod et /ou les voyelles palatales semblent donc constituer des segments "par défaut" qui comblent de manière transitoire les lacunes du système phonologique en développement.

Guerssel & Lowenstamm (1996), pour les langues sémitiques, puis Ségéral & Scheer (1998) pour les langues germaniques et sur la base d'une formulation des primitives phonologiques en éléments (Kaye et al., 1985), ont proposé que l'apophonie soit guidée par une logique proprement phonologique qui suit le chemin suivant.

$$0 > I > A > U$$

ce qui signifie donc que, si l'on part de 0 ou d'un constituant vide, on s'attend à ce que ce soit l'élément I (donc la voyelle [i] ou la voyelle [e]) qui émerge.

Si l'on rapproche cette proposition des données d'Alice et de Claire, on se souvient que ces enfants n'ont que des occlusives et des voyelles (pas de fricatives et peu de liquides), des mélodies consonantiques et vocaliques indépendantes et que leurs productions se caractérisent par l'existence de constituants vides.

L'émergence de leur système phonologique semble bien suivre une logique apophonique exprimée par le passage de 0 à I. Toutes les formes de 'palatal patterns' exprimées par ces deux enfants peuvent être expliquées par le passage de l'étape 1 du chemin apophonique (0) à l'étape 2 (I).

Le passage de 0 à I, au niveau de la mélodie vocalique permet d'expliquer la présence massive des diphtongues dont la deuxième cible est la voyelle palatale, [I] relâché dans des monosyllabes comme [æɪ] (pattern 2 Alice jusqu'à 10-13 mois) ainsi que la forme des premières unités lexicales produites par Alice à partir de 13 mois <C<sub>o</sub>V*C*i> (*baby, mommy*) (pattern 3).

Le passage de 0 à I au niveau de la mélodie consonantique explique la préférence marquée pour le yod en position d'attaque (pattern 1 d'Alice, pattern général de Claire), puisque l'élément I surfaçant sur une position consonantique vide produit automatiquement un yod. Ces yods peuvent être considérés comme l'expression de l'élément I, premier pas du chemin apophonique à partir de 0 et dont l'apparition sur une position consonantique est conditionnée par une logique apophonique. Il va donc constituer « par défaut » l'élément qui remplit les positions consonantiques et vocaliques vides. Cette proposition permet une explication unifiée des données produites par Alice et Claire et permet de rendre compte de l'existence des yods chez Claire et Alice mais aussi des deux autres 'palatal patterns' d'Alice exprimés sur la mélodie vocalique.

Par ailleurs la palatalisation, comme l'ont souligné Vihman et al. (1994)<sup>21</sup> — aussi bien pour Alice que pour Claire fonctionne effectivement comme une mélodie qui se propage sur les autres syllabes du mot et parfois même au-delà du mot sur la totalité de la séquence produite. Elle semble par conséquent au moins partiellement indépendante de l'acquisition lexicale et de la mémorisation de l'input environnant.

La propagation plaide assurément en faveur d'un phénomène d'abord phonologiquement conditionné qui échappe au moins partiellement à l'acquisition lexicale. Le pattern de babillage initial proposé par Vihman et al. est donc « phonologisé en soi » et utilisé comme une stratégie d'acquisition indépendamment de supports lexicaux spécifiques et mémorisés.

Ces *palatal patterns* semblent bien constituer un *learning path* pour l'acquisition de l'inventaire phonologique et confirme nos conclusions précédentes : il semblerait que la marque syllabique universelle ne suffit pas à rendre compte, à elle

---

<sup>21</sup> « Palatalization appears to have a status of an autosegmental melody » p. 24

seule, de la variété des productions observables dans les données des enfants et en particulier de la présence de ces yods et de ces segments palataux.

Ici les contraintes de marque se voient en effet contrariées ou complétées par une logique de marque segmentale. Leur entrée dans le système phonologique de leur langue et plus particulièrement l'acquisition de la syllabe semblent donc s'appuyer davantage sur les mélodies segmentales dont ils disposent précocement que sur la structure interne de la syllabe et l'alternance de C et V.

Enfin cette explication nous permet de faire une prédiction que nous souhaiterions vérifier dans l'avenir. On peut s'attendre à ce que les enfants « autosegmentaux » ayant tendance à produire les premiers mots reconnus en s'appuyant préférentiellement sur la mélodie vocalique et donc en produisant précocement beaucoup de constituants consonantiques vides utilisent préférentiellement le yod comme consonne par défaut et ce quelle que soit la langue cible.

#### **4.5. Conclusion**

On ne peut pas considérer CV comme un pattern d'acquisition universel et les effets de marque syllabique isolés ne peuvent pas rendre compte des données d'acquisition dans toutes les langues. Doivent être associés aux contraintes de marque syllabique des effets de marque segmentale, des contraintes rythmiques, prosodiques, lexicales dont les interactions et le rôle exact sont mal connus. On ne peut qu'espérer le développement d'études typologiques de l'acquisition syllabique observant des langues variées. Ceci permettrait sans doute un éclairage radical de certaines questions dont la formulation est fortement conditionnée par le fait que la quasi-exclusivité des recherches en acquisition ont été réalisées sur des langues indo-européennes et parmi elles surtout sur des langues germaniques.

## Travaux en cours

### Publications

**Wauquier-Gravelines S.** & Suet-Bouret D., (2005), Acquisition des attaques syllabiques et ‘palatal patterns’, pourquoi yod ?, ms.<sup>22</sup>

Fikkert, P. Levelt, C. & **Wauquier-Gravelines, S.** (2006)<sup>23</sup>, Syllabic markedness, segmental markedness, rhythm and acquisition, *Recherches Linguistiques de Vincennes*, **34**, Presses Universitaires de Vincennes, Université Paris VIII

### Communications

Fikkert, P., Freitas, M-J., Grijzenhout, J. & Levelt, C., **Wauquier-Gravelines, S.** (2004), Syllabic markedness, segmental markedness, rhythm, and acquisition, *GLOW Conference*, Thessalonique, 18 avril 2004

**Wauquier-Gravelines, S.** & Suet-Bouret D. (2004), Acquisition des attaques syllabiques et ‘palatal patterns’, pourquoi yod ?, *Colloque du réseau français de Phonologie*, Orléans, 2-5 juin 2004

---

<sup>22</sup> J’écris en ce moment une version en anglais de ce manuscrit en vue d’une soumission à « Journal of Child Language ». Je ne mentionne pas cette version dans la liste des publications puisque cet article n’a pas encore été soumis.

<sup>23</sup> Cet article est en préparation mais il appartiendra à un volume accepté par les Presses Universitaires de Vincennes. C’est à ce titre que je le mentionne.

## **5. Vers un modèle gabaritique de l'acquisition en français**

### **5.1. Introduction**

A l'issue de cette présentation de mes travaux, je voudrais montrer comment les résultats que j'ai obtenus sur des données de français m'ont amenée à proposer une modélisation des premières unités produites par les enfants par un gabarit lexical qui serait le domaine des généralisations phonologiques et prosodiques.

Mon but est de formuler une représentation qui puisse offrir une explication unifiée des données dans le cadre d'un même formalisme auto-segmental.

La littérature en acquisition phonologique regorge de nombreuses spéculations autour de la question de la taille de la toute première unité véritablement linguistique produite par les enfants et oppose classiquement en particulier les enfants « prosodiques » et « segmentaux » (Waterson, 1987).

Le premier problème que soulèvent ces spéculations est tout à fait central mais difficile à circonvenir. Il consiste à établir plus précisément le moment exact dans l'acquisition où l'on peut parler de productions « linguistiques ». On peut en effet considérer que les expériences précoces de communication du tout petit bébé à 4 ou 5 mois (babillages, sourires, modulations prosodiques) sont déjà des manifestations linguistiques relevant de la mise en place d'une grammaire ou au contraire envisager qu'on ne puisse véritablement parler de productions linguistiques qu'à partir du moment où l'environnement comprend minimalement le sens exact et hors contexte de ce que profère l'enfant.

Pour la phonologie plus précisément, la question qui va se poser consiste à établir aussi précisément que possible les conditionnements exacts d'une première unité où l'on voit se manifester de manière évidente et systématique des régularités phonologiques (ex : métathèses, harmonies etc ). Le problème est alors d'établir si cette unité est conditionnée lexicalement et morphologiquement et dépend de l'acquisition lexicale ou si au contraire elle émerge sous l'influence de contraintes

essentiellement phonologiques (organisation distributionnelle des sons et des structures indépendante des mots) ou prosodiques sur des unités supérieures au mot.

Selon l'hypothèse retenue les principales unités qui ont pu être envisagées sont

- i) la syllabe (perceptive Dupoux, 2005, Ramus et al. 1999) et /ou articulatoire (Mac Neilage & Davis 1993)
- ii) « entire lexical unit » ou « whole-word » (dont la nature exacte reste à définir mais dont la caractéristique est qu'elle est une unité lexicale ayant un référent (Francescato, 1968, Ferguson & Farwell 1975, Menn, 1978, 1983, Menn & Matthei, 1992, Macken 1992, Vihman, 1996, 2001)
- iii) Unité(s) prosodique (s) supérieure(s) au « mot ». En général le « mot prosodique » ou le « pied » métrique (Fikkert, 1994, Demuth & Fee, 1995).

Je ne rediscute pas ici de manière détaillée la pertinence des trois hypothèses. J'ai traité au ch. I des limites de l'explication fournie par la phonologie prosodique pour ce qui concerne l'acquisition du français, langue sans accent lexical. Cette question est également amplement traitée dans Braud (2003). Par ailleurs je discute également de la possibilité de l'émergence d'une phonologie qui serait indépendante du lexique au ch. IV. J'argumente longuement pour défendre au contraire l'idée que l'acquisition réellement « phonologique » ne peut se mettre en place qu'avec l'acquisition morphologique et lexicale.

J'ai donc retenu la « whole-word » hypothesis supposant que la motivation essentielle qui va amener l'enfant à « découper » des portions de signal et à les stocker est leur distinctivité.

Or l'unité « mot » (et la catégorie des noms en particulier) matérialise la relation signifiant / signifié / référent de manière particulièrement efficace (du moins pour le français et autres langues actuellement observées en acquisition).

Je considère donc bien volontiers avec Francescato que :

« Children never learn sounds, they only learn words and the sounds are learn through words »

G. Francescato, », *Lingua*, **21**, p. 148

Et je propose de formaliser cette unité dans un modèle gabaritique dans la continuité des travaux de Vihman.

## 5.2. L' hypothèse gabaritique

L'hypothèse selon laquelle l'acquisition phonologique se met en place sur un gabarit lexical a été défendue par Menn (1978) et développée par Macken (1992), dans le cadre d'une réflexion plus ancienne sur la *whole-word hypothesis* (cf. Vihman & Croft, à paraître). De manière générale, le terme de « template » ou « gabarit » est assez couramment utilisé en acquisition sans pour cela qu'il désigne nécessairement les unités morpho-phonologiques définies en particulier à partir des travaux sur les langues sémitiques (Pour une synthèse sur la notion de gabarit en phonologie cf. Lowenstamm, 2003).

Mais l'emploi de ce terme retient les caractéristiques formelles essentielles de ce qui caractérise un gabarit dans la définition que Lowenstamm place en exergue du volume qu'il consacre à cet objet

« Dans les langues sémitiques, les mots sont construits selon un nombre fini de modèles. Tout mot construit selon l'un de ses modèles présente une taille fixe, une allure et des caractéristiques d'organisation sonore immédiatement repérables à partir desquelles il est possible d'inférer un nombre important d'autres propriétés grammaticales »

J. Lowenstamm, *Introduction*, RLV, **32**, p.5

A ceci s'ajoute, comme le souligne également l'auteur, l'autosegmentalité des représentations phonologiques (p. 8) ainsi que la possibilité de construire l'interface avec la morpho-syntaxe (p.27) qu'offre l'objet formel appelé « gabarit »



En acquisition, l'emploi de ce terme suppose effectivement un objet fini dont la variabilité est limitée, mettant en évidence la nécessité pour les enfants d'opérer un « mapping » ou « matching » ou « fitting ». Cet objet est par ailleurs une unité lexicale, morphologiquement et prosodiquement contrainte permettant de rendre compte de phénomènes qui témoignent d'une relative indépendance des mélodies (en particulier consonantiques et vocaliques) chez les enfants.

L'hypothèse consiste à proposer qu'ils se servent d'un gabarit spécifique, modèle lexical, qui serait le domaine dans lequel se font les généralisations phonologiques. Ce gabarit évoluerait progressivement sous l'influence des acquisitions.

Un certain nombre « d'erreurs » enfantines relèveraient des difficultés de matching.

#### 5.2.1. Gabarit et acquisition, le modèle de Vihman

Vihman (2001, Vihman & Velleman, 2000, Croft & Vihman, à paraître) a développé à partir de cette idée et sur la base de données variées dans des langues variées (anglais, français, estonien, finnois) une théorisation complète de ce que pourrait être le gabarit d'acquisition. Elle le définit comme un objet idiosyncrasique que retient chaque enfant sur la base de

i) ses expériences précoces de babillage. L'enfant utilisera dans les premiers gabarits un mini inventaire phonologique restreint correspondant à ses premières expériences de production (mots, jeux vocaux, babillages) et fera émerger les autres sons à partir de ce premier inventaire.

ii) ses expériences perceptives. L'intérêt qu'il porte à son environnement et aux interactions verbales qui l'entourent

iii) la fréquence des sons et des structures dans l'input ainsi que la fréquence lexicale.

Vihman envisage donc les gabarits comme des unités émergentes qui ne sont conditionnée ni par la marque, ni par des universaux structurels ou grands principes gouvernant les langues naturelles. Elle établit une continuité entre les expériences neuro-motrices du babillage et la première phonologie tout en les distinguant pourtant. Les schèmes neuro-moteurs qui sont utilisés volontairement reflétant le contrôle neuro-musculaire émergent du tractus vocal ainsi que l'apprentissage implicite des caractéristiques de l'input vont peu à peu faire place à des patterns réguliers de reproduction de mots adultes systématiquement déformés dans le gabarit propre de l'enfant.

Selon elle, la reproduction de ces patterns adultes déformés assure une double fonction : elle soutient la représentation des mots dans la mémoire et aide à la mémorisation de patterns similaires (paires minimales émergentes). Mais elle sert également à limiter l'espace d'apprentissage et de généralisation puisque l'enfant travaille sur un mini-corpus d'objets finis dont la variation structurelle est restreinte.

Dans ce modèle, la production de gabarits relève d'une généralisation à partir de patterns sélectionnés dans l'input et qui sont adaptés pour correspondre aux formats propres de l'enfant. Le gabarit fonctionne comme une « routine » qui automatise le processus de production de mots. La phonologie émerge progressivement par extension et amélioration des ressources articulatoires et phonétiques en production, l'expansion du lexique et l'évolution des représentations. Le phénomène est donc exponentiel, la généralisation phonologique apportant la possibilité d'apprendre de nouveaux mots et vice-versa.

Pour Vihman qui retient une épistémologie constructiviste de type piagétienne, l'acquisition de la phonologie par fixation de routines lexicales et généralisation a en outre le mérite de partager des caractéristiques invariantes que l'on retrouve pour d'autres acquisitions cognitives et /ou motrices chez les enfants comme le fait que la haute variabilité et l'idiosyncrasie précèdent immédiatement la stabilisation. Par ailleurs, cette conception de l'acquisition phonologique peut rendre compte des phénomènes de régression suivis d'un progrès immédiat (phénomène dit de « la

courbe en U ») témoignant de cycles d'organisation / réorganisation des représentations et de l'évolution de leurs représentations. Ce modèle permet donc d'intégrer au mieux la dimension maturationnelle des processus d'acquisition phonologique.

On peut néanmoins objecter que les gabarits proposés sont sans doute plus formellement conditionnés que Vihman ne le suppose.

Il me semble d'abord que, eu égard à la définition de ce qu'est un gabarit, le recours à cette notion suppose implicitement un objet formel structurellement contraint par ses caractéristiques propres et qui ne peut donc pas résulter uniquement d'un choix idiosyncrasique. Il me semble que par définition, un gabarit est un objet phonologique abstrait, ayant éventuellement une correspondance psychologique mais pas un objet physique.

Et si l'on observe les formats des objets proposés par Vihman on voit

i) qu'ils sont contraints par la structure syllabique de la langue cible. Elle montre en particulier qu'il existe très peu de structures CVC dans les gabarits des enfants français alors qu'on en trouve en anglais, allemand et estonien (Vihman & Velleman, 1989) et chez des enfants bilingues anglais/français (Brulard & Carr, 2001). Ces remarques recourent totalement les données que nous avons obtenues et les observations que nous avons pu faire dans le cadre de notre projet d'acquisition des structures syllabiques (Fikkert et al.2004).

On objectera qu'il existe en français, des suites CVC dans des mots de haute fréquence (poule, vache, robe, jambe...). Pourquoi les enfants ne les sélectionnent-ils pas ?

ii) la variété des structures qu'exhibent les données est restreinte (conformément à la définition citée plus haut)

Tous les gabarits proposés dans les données :

CV  
VC  
CVC  
VCV  
CVCV

peuvent être ramenés à 2 structures CV complètes.

Sur la base des propositions de Vihman, je présente donc une version plus structurellement conditionnée du gabarit d'acquisition.

### 5.3. Hypothèse d'un gabarit prosodique pour l'unité primitive en français.

A partir des données et des analyses que j'ai présentées jusqu'à maintenant concernant l'acquisition de la liaison, les troncations et réductions, l'acquisition des structures syllabiques et l'existence d'une indépendance des mélodies vocaliques et consonantiques repérable dans les données, je propose donc le gabarit suivant

[70]  $[, \sigma (\sigma)_n ' \sigma ]$

défini au départ par des contraintes prosodiques puisqu'il est borné par deux accents démarcatifs.

La dernière syllabe porte l'accent de groupe démarcatif qui délimite la frontière droite de l'unité.

La première syllabe est également accentuée. Elle porte un contre-accent en français établissant la frontière gauche de l'unité. L'existence de cet accent de début de mot est proposée par Fonagy (1980) et confirmée par Di Cristo (1999).

Celui-ci décrit en effet

“ une tendance à accentuer la syllabe initiale des mots, ce qui donne naissance à la formation de schèmes barytoniques et d'arcs accentuels dans lesquels seules les syllabes initiales et finales d'un groupe de mot reçoivent un accent.

Il est possible que l'émergence de l'accent initial en français que l'on situe vers la fin du XIX<sup>ème</sup> siècle, soit bien plus ancienne encore, comme semblent le montrer les travaux de Pensom (1993) sur la métrique de textes en ancien français”

A Di Cristo, Le cadre accentuel du français contemporain: essai de modélisation, *Langues*, vol.2,3, p.185.

Selon Di Cristo l'existence de cet accent en français contemporain est accepté par la majorité des prosodistes quelle que soit leur conception du phénomène. Les divergences portent sur l'interprétation exacte de cet accent (emphatique, écho d'accent ou accent secondaire).

Les syllabes initiales et finales sont donc des positions prosodiquement *fortes*. Elles sont posées prioritairement par les enfants. Conformément aux données et aux analyses que nous avons exposées plus haut et en accord avec les données de Braud (2003), elles ne sont jamais tronquées et subissent très peu de déformations. (Braud 2003). Cette première syllabe est le site susceptible de recevoir le proto-déterminant.

Entre ces deux bornes prosodiquement fortes, il y a un n-number de syllabe (s) intermédiaire(s) -la valeur de 'n' est théoriquement comprises entre 0 et l'infini. Braud (2003) a montré que 'n' est toujours  $\leq 2$  dans les productions précoces avant deux ans. A l'inverse de la première et de la deuxième syllabes, ce 'n-site' est instable, les syllabes internes étant celles qui sont systématiquement tronquées, réorganisées. La structure et le contenu segmental de ces syllabes intermédiaires sont le site à partir duquel le gabarit évolue (allongements, ajout de positions, diversification de l'inventaire phonémique)

Je postule que cette structure va être le domaine des généralisations morpho-phonologique et que c'est une unité référentielle (un mot). Je propose l'hypothèse maturationnelle suivante

**Stade I**

$$[71] \quad [ ,cV(cV)_n 'cV ]$$

Le gabarit est d'abord rempli préférentiellement par les voyelles. Les voyelles et les consonnes n'ont pas la même importance pour les enfants. La cible adulte est constitué d'abord par la mélodie vocalique. Les consonnes peuvent être accessoirement produites.

**• Stade II**

$$[72] \quad [ ,CV(CV)_n 'CV ]$$

La structure du gabarit est définie par les noyaux. Les premières consonnes non marquées (essentiellement occlusives) peuvent remplir les positions d'attaque? C'est à ce stade qu'on constate de nombreux phénomènes phonologiques tels que les harmonies consonantiques, les réductions de clusters.

**Stade III**

$$[73] \quad [ ,CVc (CVc)_n 'CVc ]$$

Le gabarit approxime l'output adulte. Les voyelles sont complètement stables et les attaques encore variables avec phénomènes d'harmonies encore attestés. On suppose que c'est à ce stade que les enfants acquièrent les codas ... ou les noyaux vides (selon l'hypothèse syllabique que l'on retient). Je ne veux pas me prononcer sur cette question puisque je n'ai pas travaillé directement dessus, mais l'existence du schwa en français, le %V mentionné plus haut ainsi que, pour ce que j'en ai vu, la différence d'acquisition des structures CVC en finale et en interne de mots, laissent penser que la plupart du temps, ce que les enfants acquièrent, c'est plutôt une position vide intermédiaire.

#### 5.4. Avantages de l'hypothèse gabaritique

Cette proposition formulée ici à une étape liminaire doit être confirmée par des données complémentaires.

En l'état, elle permet d'abord d'exprimer la dimension autosegmentale de nombreuses productions en acquisition (dissociations consonnes / voyelles, reduplications, dissociations attaques / noyaux...) et de la plupart de nos données. Elle nous permet en outre de rendre compte de différents problèmes que nous avons soulevés précédemment.

##### 5.4.1. Proto-déterminant et resegmentation

Ce gabarit permet d'expliquer la présence d'une position forte dans l'unité susceptible de recevoir le proto-déterminant. Cette position n'est pas arbitrairement posée puisqu'elle est déterminée par la présence d'un accent démarcatif.

Cette conception est en outre compatible avec l'input francophone, qui sauf pour les noms propres voit très rarement les noms se réaliser sans déterminant. On peut donc arguer que le proto-déterminant n'est pas analysé comme tel : il est semblait-il interprété comme une position accentuée de début de mot et découpé avec le mot cible. La forme des unités produites par les enfants plaide en faveur de l'arc accentuel défini par Di Cristo (1999). Il semblerait bien effectivement que les enfants établissent clairement les limites droites et gauches de l'unité pour en fixer les bornes et qu'entre ces deux extrémités, les réalisations soient approximatives et floues. La reconnaissance et l'établissement de la forme lexicale se réaliserait par ses extrémités.

Ceci permettrait d'expliquer pourquoi, bien qu'il soit un clitique donc par définition faible, le site du futur déterminant est rempli précocement et maintenu, quand les premières syllabes du mot tombent. On peut considérer qu'il n'est interprété morphologiquement et syntaxiquement qu'après segmentation et détachement du gabarit initial. A cet égard, cette hypothèse est compatible avec la proposition de l'existence d'un site CV initial permettant de rendre compte d'un contexte disjonctif sur la frontière gauche où se réalisent des phénomènes morpho-

phonologiques variées. On pourrait arguer, sur la base des ces données et de ce formalisme qu'en français, dans un premier état, ce site est prosodiquement déterminé et occupé par une syllabe ou une voyelle épenthétique purement phonologiques qui seront ensuite réanalysées morphologiquement et extraites du site.

Ce gabarit est également un NP, premier domaine de réalisation de la liaison obligatoire dont j'ai longuement montré qu'elle est, sur ce contexte le moyen pour les enfants francophones d'acquérir la liaison par bootstrapping morphologique et détachement du déterminant.

#### 5.4.2. Stratégies d'allongement et de diversification par syllabes internes

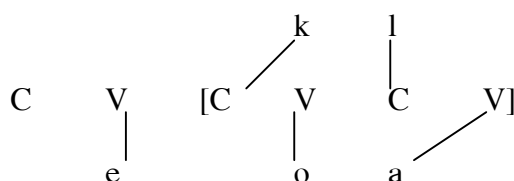
Ce gabarit permet aussi d'offrir une explication aux phénomènes de troncation et reduplication que j'ai évoqués au ch.I.

On voit comment, conformément à ce que j'avais annoncé, les troncations et reduplications témoignent de la nécessité pour l'enfant d'ajuster l'input adulte qui est la cible visée dans le format du gabarit à partir de l'inventaire phonologique dont il dispose. Si l'on reprend les données présentées au chapitre I.4.

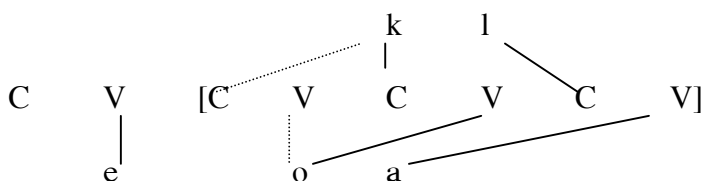
Les mots mono et bi syllabiques ne posent aucunement problème et ne sont jamais ni tronqués ni dupliqués car ils entrent dans le gabarit

Pour les mots excédant 2 syllabes, on peut envisager cette évolution

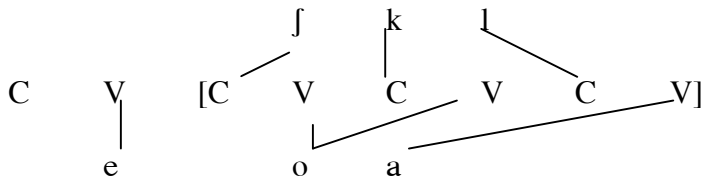
[74] **Stade 1 (Clara) : troncation pour correspondre au format du gabarit n=1**



[75] **Stade 2 : allongement en ajoutant des positions**





[76] **Stade 3 : propagation segmentale**

Reste néanmoins à établir plus clairement sous quelles contraintes s'opèrent la réduplication et l'allongement. On peut penser que l'augmentation de la valeur de *n* peut refléter une maturation de la mémoire de travail permettant à l'enfant d'augmenter en production et en perception la taille de l'unité traitée.

Par ailleurs, je n'ai pas à ce stade d'explication ni de données permettant d'expliquer si l'enfant ajoute une syllabe à la fois, s'il ajoute des syllabes ou des noyaux (ex [raatatœk]) ni comment s'opère la diversification de l'inventaire phonémique à partir du gabarit. On peut supposer même qu'à un même stade plusieurs valeurs identiques de '*n*' sont utilisées, puisque j'ai dans les données observé des phénomènes de copies stéréotypées (ex: [didukaje/ didukaje])

## 5. 5. Conclusion

Ce gabarit peut constituer un premier patron lexical à partir duquel se font les acquisitions phonologiques. C'est une unité prosodique disponible en perception, bornée par accent et contre-accent et donc possiblement segmentable dans l'input.

Par ailleurs c'est une unité fonctionnelle et référentielle complètement motivée pour l'enfant.

Cet objet est formellement contraint par un nombre de positions, types de positions (attaques, noyaux, relations entre ces deux types d'objets) et repose sur une conception autosegmentale des représentations dont les associations sont gouvernées par des principes structurels universels.

Ceci permet de l'envisager comme une unité émergente et motivée mais structurellement contrainte par des principes universels, une portion de signal non contingente à partir de laquelle émergent les généralisations phonologiques.

Comme je l'ai dit, cette première formulation de l'hypothèse gabaritique doit être confrontée à des données plus nombreuses et plus variées.

Par ailleurs, elle définit sans doute une unité lexicale type rendant compte de la structure du français. Mes souvenirs en phonétique historique me font intuitivement penser que ce type de gabarit ne serait peut-être pas non pertinent en diachronie. Mais, je n'affirmerai pas, ce stade, qu'il s'agit d'une structure universelle.

## **6. Propositions de travail**

Dans cette dernière partie, j'indique quelques pistes de travail qui pourraient constituer des prolongements des travaux actuels

### **6.1. Etude de la liaison encore**

Les données dont je dispose me semblent encore insuffisantes pour pouvoir me permettre une généralisation véritablement satisfaisante et donner une vue d'ensemble cohérente sur l'acquisition de ce phénomène phonologique. Trois questions me semblent devoir donner lieu à approfondissement.

i) Seule la liaison obligatoire entre déterminant et nom au sein du groupe nominal a été observée. Or il existe en français d'autres contextes qui génèrent des liaisons obligatoires (Encrevé, 1988).

- au sein du groupe nominal entre l'adjectif antéposé et le nom
- (" un petit ami ")
- au sein du groupe verbal entre la copule verbale et le participe ("est arrivé")

- au sein du groupe prépositionnel entre une préposition courte et le nom (" en avion" (liaison) qu'on opposera à « pendant une promenade » (préposition longue pas de liaison)
- - entre un pronom clitique sujet et un verbe (" ils arrivent ")

Ces contextes mettent tous en relation deux termes ayant des relations syntaxiques étroites. On peut également remarquer que dans tous les cas le mot libérant la consonne flottante est un mot court entrant dans le domaine prosodique du mot avec lequel se fait la liaison. Cette contrainte syntaxico-prosodique a été repérée et expliquée dans le cadre de la phonologie prosodique (Nespor & Vogel 1986). Il me paraît déterminant d'observer systématiquement ces contextes afin d'établir s'ils donnent également lieu à des erreurs, d'établir si les erreurs sont alors du même type

ii) J'ai pour l'instant centré mes recherches sur l'acquisition de la liaison obligatoire en supposant que l'acquisition de la liaison facultative était un phénomène lexical reposant sur la mémorisation des contextes. Si l'hypothèse d'Encrevé & Scheer (2005) peut être retenue pour unifier l'analyse de la liaison, je manque de données précises permettant actuellement de valider cette hypothèse.

Je n'ai pour l'heure pas de données me permettant d'observer ni l'acquisition et les réalisations dans les contextes facultatifs, ni le lien avec la fréquence de réalisation de ces contextes. De telles données permettraient de développer et d'approfondir une réflexion sur l'usage des représentations phonologiques et l'inscription de l'usage dans la grammaire des locuteurs (Encrevé, 1988, Encrevé & Scheer, 2005).

L'une des questions qui devrait en particulier être abordée concerne l'influence des représentations orthographiques dans les erreurs sur contextes facultatifs.

iii) Etude des représentations de la CL chez les enfants normaux bilingues.

Les travaux réalisées dans ce domaine avec mes étudiants dans le cadre de mémoires de maîtrise n'ont pas pu aboutir à des conclusions encore bien affirmées, là encore

faute de données suffisantes. Le recueil de données spontanées chez quelques enfants bilingues ainsi que le travail mené par Ferré (1999) laissent penser que le phénomène d'acquisition de la liaison est plus complexe chez les enfants bilingues et que l'on voit apparaître des phénomènes de conflits de contraintes phonologiques conditionnés par les structures syllabiques permises dans les langues (présence ou absence d'attaque syllabiques vides).

Je souhaiterais étoffer cet axe de recherche.

- en répliquant les expériences faites avec 18 enfants bilingues français / anglais. Je souhaiterais observer d'autres bilinguismes impliquant des langues dont la morphologie diffère plus radicalement de celle du français que celle de l'anglais. Je pense en particulier au bilinguisme français / arabe ou français / allemand puisque ces langues interdisent les attaques vides sur les contextes gauches des mots lexicaux et imposent un coup de glotte systématique à l'attaque. Je souhaiterais également travailler sur le bilinguisme français / breton puisque les langues celtiques présentent un phénomène d'alternance vocalique sur la frontière gauche de mot (les « mutations » ) conditionné par le contexte morpho-syntaxique.
- en réalisant des études longitudinales d'enfants bilingues. La comparaison de ces corpus obtenus dans des conditions expérimentales très semblables me paraît pouvoir fournir des données précieuses permettant d'avancer des hypothèses sur la liaison mais également sur les représentations lexicales précoces des enfants parlant exclusivement ou partiellement le français

## **6.2. Acquisition des déterminants en français : articulation phonologie / syntaxe et sémantique**

J'ai fait l'hypothèse d'une spécification totale des attaques de mots ne permettant d'envisager une attaque vide remplie par une consonne flottante qu'à partir de 3/4 ans. J'ai aussi proposé que les erreurs sont la conséquence d'un bootstrapping morphologique détachant le déterminant du nom.

Ceci laisse supposer que les enfants ne règlent les questions phonologiques que pose le contexte disjonctif en frontière gauche pour le français à proprement parler qu'à partir de 2ans et demi / 3 ans, moment où ils segmentent les déterminants. Et par l'acquisition de la liaison obligatoire. Les autres travaux sur les troncations, l'acquisition des attaques syllabiques semblent confirmer cette présence précoce d'un proto-déterminant en français qui ne tombe jamais même quand les syllabes initiales sont tronquées. Le problème consiste à établir, à partir de quand et sur quelles bases (morphologiques, syntaxiques, sémantiques, prosodiques) ce début de mot, souvent constitué uniquement d'un noyau vocalique est potentiellement analysé comme un proto-déterminant puis un déterminant.

Mon but serait alors d'observer le rôle de la disponibilité prosodique et acoustique des catégories fonctionnelles dans l'acquisition et d'autre part les contraintes phonologiques qui s'appliquent à l'acquisition des déterminants (définis, indéfinis, partitifs du /de le et des / de les). Ce travail pourrait (et devrait) s'inscrire dans un projet articulant acquisition phonologique et acquisition syntaxico-sémantique.

## **6.3. Projet typologique sur l'acquisition de la phonologie**

Le projet dans sa dimension européenne et typologique a été mené jusqu'à maintenant de manière informelle sur la base de rencontres entre Paula Fikkert, Maria-Joao Freitas, Clara Levelt et moi-même autour de travaux portant plus spécifiquement sur l'acquisition des structures syllabiques.

Il est prévu de donner une forme plus officielle à ce projet afin de i) permettre des échanges de données et d'analyse beaucoup plus systématiques, ii) avoir les moyens de réaliser un corpus européen consultable à partir de la base de données Child Phon

(Rose et al. 2005) iii) étendre la problématique initialement centrée sur l'acquisition des structures syllabiques.

Les thèmes de recherche porteraient sur les points mentionnés plus haut (rôle et nature de la marque et influence sur les *learning paths* observables). Mais nous souhaiterions y associer d'autres problématiques comme le lien entre acquisition phonologique et acquisition du lexique, rôle respectifs de la perception et de la production dans l'émergence des représentations phonologiques. Ce projet actuellement à l'étude et en cours de préparation supposerait des collaborations européennes multiples.

## **6.4. Travaux sur la dysphasie**

### 6.4.1. Troubles prosodiques dans la dysphasie de développement

Les travaux sur l'existence d'une éventuelle atteinte de la composante phonologique de la grammaire des enfants dysphasiques ont mis en évidence l'existence d'un trouble prosodique récurrent se manifestant par l'incapacité à construire des groupes prosodiques et à projeter l'information segmentale au sein d'une unité structurée plus large que la syllabe.

Je souhaiterais développer ces recherches dans les deux directions suivantes

i) Ces données ont été obtenues secondairement lors d'expériences sur la liaison. Je souhaiterais avoir recours à des tâches de production et de compréhension permettant de juger plus directement de la fluence des enfants dysphasiques et des déficits prosodiques dont ils semblent atteints.

ii) L'introduction des hiatus au sein de groupes comportant des liaisons apparaît à ce jour comme symptomatique du comportement des enfants dysphasiques (Rialland, à paraître). De nouvelles données confirmant cette observation permettraient d'envisager qu'elle constitue un indice dans le diagnostic précoce des troubles dysphasiques.

### 6.4.2. Les enfants sans consonnes

Ce projet de recherche vient de débiter en collaboration avec Annie Rialland (CNRS, UMR 7018) et Marie-Thérèse Le Normand (Inserm).

Certains enfants dits « dyspraxiques » présentent des troubles particuliers de structuration de la parole les amenant à omettre les consonnes lors de la production de la chaîne parlée (ex. « a a u / a oeu a a a » est produit pour « la voiture, dans le garage » Le Normand & Chevrie-Muller (1991), Le Normand et al. (1993).

Sur la base de l'étude longitudinale de productions spontanées et induites de 4 enfants entre 3 et 8 ans nous avons observé que ces enfants disposent pourtant de capacités praxiques de faire des bruits (maîtrise motrice de l'articulation Tch, t, ch) qui émergent dans le non-linguistique (onomatopées) alors qu'elles sont absentes quand ils produisent des mots. Cet état de fait correspond à un des stades de développement du babillage tels qu'ils sont décrits dans Oller (2002). Les données montrent également une forte dissociation entre les productions « linguistiques » des enfants caractérisées par des suites de voyelles et les productions spontanées mimant des bruits ou des productions chantées où les consonnes sont beaucoup plus présentes.

Enfin ces enfants réalisent de manière optimale les patterns prosodiques appropriés à leurs énonciations (respect des durées, des mélodies, du rythme) même en l'absence quasi-totale de consonnes (comme si la « place » phonologique des consonnes étaient programmée, même si elles ne sont pas réalisées).

Le projet vise à établir l'existence de troubles du babillage pouvant être mis en relation avec les déficits observés et de troubles de la continuité babillage-langage lors de la phonologisation des structures syllabiques. Notre but est d'établir plus précisément l'origine du déficit qui entraîne une incapacité à intégrer l'information consonantique lors de la syllabation de la chaîne.

Une communication et une publication sont en préparation.

## **Travaux en cours**

### **Communications**

**Wauquier-Gravelines, S.**, Le Normand M.-T. & Rialland A. (2005), Syllabic planification and phonological working memory disorder : the case of Consonantless-SLI children, *Xth International Congress for the study of Child Language*, Berlin, 25-29 juillet 2005

## CHAPITRE III : PHONOLOGIE ET COGNITION,

### PROBLEMES ET QUESTIONS

« Chaque chose doit être considérée d'une autre façon selon que l'on se réfère à l'ordre de notre connaissance ou que l'on parle d'elle selon l'existence réelle. »

R.Descartes, Règles pour la direction de l'Esprit, Règle XII

« Je dis « une fleur » et, musicalement s'élève, l'absente de tout bouquet »

S.Mallarmé, Un coup de dé, jamais n'abolira le hasard

### **1. Introduction**

La phonologie et la psycholinguistique élaborent leurs relations au sein de ce qu'il est convenu d'appeler le champ des « sciences cognitives » regroupant les diverses disciplines rendant compte des fonctionnements intelligents humains et artificiels. Cette terminologie est depuis un certain temps d'usage courant. Elle désigne un champ de connaissances dont la légitimité scientifique et institutionnelle est notamment attestée par l'attribution de crédits, la création de formations et un nombre de publications scientifiques conséquent, alors que la définition exacte de l'objet ou des objets de recherche ainsi que la structuration disciplinaire des domaines de connaissance qui organisent ce champ sont divergentes et opaques, pourrait-on dire. Laks (1996) a montré très clairement comment les définitions proposées sont variables et dessinent, au-delà d'un simple débat terminologique, les lignes de fracture qui structurent le domaine et définissent diverses conceptions de la cognition, diverses théories des fonctionnements mentaux et de nombreux débats autour de l'ontogenèse des conduites intelligentes chez l'être humain.



Mais que l'on conçoive le champ à partir des objets, des méthodes, des paradigmes ou de manière plus sociologique à partir des champs disciplinaires concernés, la linguistique générative occupe, quoi qu'il en soit, une place centrale et fondatrice dans le champ des « sciences cognitives », aussi bien en tant qu'objet de recherche qu'en tant que source d'arguments théoriques, de méthodes et d'enjeux épistémologiques ayant permis d'asseoir le « modèle cognitiviste classique » (Piatelli-Palmarini, 2000) au sein duquel se sont construites la dichotomie théorique entre « compétence » et « performance » ainsi que la distribution institutionnelle, scientifique et disciplinaire des tâches incombant respectivement à la linguistique et à la psycholinguistique cognitive.

C'est en se réclamant d'une tradition linguistique rationaliste directement héritée de Descartes que Noam Chomsky entre, dès 1957, dans les débats suscités par ce qu'on a ensuite appelé le « tournant cognitif ». Je ne discuterai ici, ni de la validité historique, ni de la validité épistémologique de la paternité cartésienne revendiquée par Chomsky, ni de la présentation du débat rationalisme / empirisme qui est la sienne et qui est très fermement contestée par historiens et philosophes (Pariente, 1975, Aarsleff, 1982), car tel n'est pas mon propos.

Quoi qu'il en soit, on peut légitimement considérer avec Laks (2005) que la pensée chomskyenne consacre le retour « du sujet cartésien » et s'élabore donc délibérément et explicitement à l'ombre de sa lecture de Descartes, de la grammaire de Port-Royal (« La linguistique cartésienne », 1966) et des philosophes rationalistes du XVIIIème et du XIXème siècle. Le postulat fondamental du rationalisme considère qu'il n'existe aucune forme d'organisation intrinsèque du monde externe à l'entendement et que les structures résultant de la perception (visuelle, auditive, etc) sont imposées à l'environnement par la pensée et non pas extraites de celui-ci pour conditionner les structures de la pensée. Les contenus mentaux sont donc universels, non substantiels (ou relevant d'une autre substantialité que la matière), opaques à l'environnement, invariants dans le temps et dans l'espace. Chomsky a emprunté et adapté un certain

nombre de postulats théoriques qu'il a appliqués à la linguistique et la psycholinguistique cognitives : le mentalisme<sup>24</sup>, le dualisme<sup>25</sup> et l'innéisme<sup>26</sup>.

Ils sont réaffirmés dans tous les textes chomskyens et ce depuis *Aspects<sup>27</sup> of the theory of Syntax* (1965) (ci-après : *Aspects*) :

Mais, alors que les arguments concernant le dualisme, le réalisme psychologique des grammaires, la nature des représentations, le behaviorisme (Laks, 2005), le développement de l'enfant, l'instinct de langage chez le petit humain etc, sont disponibles et certains depuis fort longtemps, Chomsky réalise ce qu'on pourrait appeler un coup de force épistémologique qui, en faisant place nette tant du côté linguistique (Harris, 1993, Encrevé, 1999) que du côté psychologique (Seuren, 1998 ; Harris, 1993) a contribué de fait à l'extraordinaire mais très réelle cohérence du paradigme génératif, au risque d'en caricaturer voire d'en refouler certains enjeux.

<sup>24</sup> La linguistique générative et la psycholinguistique cognitive visent à rendre compte de (« expliquer ») l'activité mentale que constitue la production et la perception linguistique par tout locuteur-auditeur. Les comportements linguistiques observables sont rapportés à un système intériorisé de connaissances abstraites symboliques et de processus de calculs algorithmiques appliqués systématiquement à ces symboles. Idéalement ces symboles et processus sont en nombre fini et doivent être formalisés selon une logique d'économie minimale en visant le plus haut degré d'abstraction possible (être ramenés à des primitives). Ces primitives ont valeur universelle et reflètent l'organisation de l'esprit humain.

L'existence chez l'homme de ce système abstrait structuré reflétant la structure de la pensée, consacre la rupture ontologique entre l'homme d'une part, l'animal et la machine d'autre part. Le système de connaissances symboliques est en outre opaque à l'usage que le locuteur-auditeur réel en fait, en contexte.

<sup>25</sup> L'opposition radicale entre esprit / connaissance symbolique et matière / substance suppose plusieurs niveaux d'interprétation des faits. Un niveau formel, ou structure profonde (qui est l'objet de la phonologie théorique) et un niveau substantiel, physique (qui est l'objet de la phonétique). Ces deux niveaux sont logiquement non isomorphes et les facultés linguistiques sont donc organisées en une architecture de niveaux structurée en modules successifs entretenant des relations hiérarchisées. La psycholinguistique s'articule à un troisième niveau (psychologique). Elle rend compte des procédures psychologiques qui permettent au locuteur d'utiliser la connaissance formelle stockée en structure profonde pour produire et comprendre des objets en structure de surface. Ces procédures et représentations ont un ancrage biologique (quatrième niveau neurologique), génétiquement déterminé.

<sup>26</sup> Les structures universelles visibles dans les langues naturelles reflètent les structures de l'esprit humain et préexistent donc à l'acquisition d'une langue naturelle. La théorie des « idées innées » est donc redéfinie en termes de configuration spécifique innée (GU, grammaire universelle) associée à un dispositif spécifique (LAD= langage acquisition device) permettant l'évolution intrinsèque et autonome de cette GU en Gf par exemple (grammaire du français). Les universaux conditionnent exclusivement le développement, ils ne sont pas appris par expérience, entraînement, mimétisme ou nécessité d'adaptation au milieu mais sont disponibles de manière innée et fournissent des principes d'organisation structurels qui permettent d'apprendre des langues par contact passif avec l'environnement qui sert de déclencheur (« trigger »). Ces facultés linguistiques innées sont spécifiques, modulaires et indépendantes des facultés intellectuelles générales. La maturation des facultés linguistiques innées n'est pas conditionnée par la maturation des autres facultés cognitives.

Enfin les facultés linguistiques innées sont indépendantes de la modalité d'acquisition (orale, gestuelle etc).

<sup>27</sup> Les citations sont tirées de la version française de l'ouvrage traduit par J-C Milner et publié en 1971 aux éditions du Seuil. Les références (pagination, paragraphes ...) sont également faites à partir de cette version française du texte.

La grammaire générative a suscité de nombreux débats internes et externes à la théorie elle-même (Harris, 1993) mais il apparaît que ses fondements se sont élaborés au carrefour de 3 débats, générativisme-structuralisme, Chomsky-Skinner, Chomsky-Piaget qui bornent l'espace épistémologique dans lequel se sont développées la linguistique et la psycholinguistique « cognitive ». J'en vois la pertinence pour mon travail en ce qu'ils s'articulent tous ultimement à la question du réalisme psychologique d'une phonologie interne abstraite et autonome qui est l'interrogation centrale qui a motivé mes travaux de recherche.

Les faits et les discontinuités sont connus et ont donné lieu à la production d'une littérature abondante et parfois très polémique. Je les évoque et les commente ci-dessous, uniquement en ce qu'ils concernent directement mon propos, et en particulier en ce qu'ils conditionnent la manière dont ont été articulées phonologie et psycholinguistique. Les réflexions ci-dessous n'ont pas la prétention d'apporter quelque éclairage épistémologique nouveau à un débat passionnant, polémique et nourri mais elles reflètent plutôt les interrogations qui ont été les miennes concernant la question de la réalité mentale des représentations chez l'adulte et leur émergence chez l'enfant.

## **2. Réalité du modèle/ modèle de la réalité**

### **2.1. L'ambiguïté de la grammaire**

Comme le rappellent très justement Laks (1996, 2005), Carr (2000), Jackendoff (2000), la question cognitive en phonologie n'est pas nouvelle et ne consacre pas en soi la rupture théorique qu'instaure le générativisme. Avant Chomsky, ont été proposées explicitement des formulations cognitives de la grammaire (Hockett, 1948), du signe linguistique (Saussure, 1966) et du phonème en particulier (Saussure, 2001, Sapir, 1933), et l'on voit clairement comment Jakobson (1941) en rapprochant les phénomènes d'acquisition et d'aphasie et les considérant comme deux états incomplets ou dégradés de la grammaire adulte préfigure la

démarche de la psycholinguistique et de la neurolinguistique cognitives qui fleuriront vingt ans plus tard.

Chomsky, en s'inscrivant dans un long héritage « cognitiviste » que l'on trouve déjà au XVIIème chez Nicolas de Beauzée ou la grammaire de Port-Royal, ne pose pas une question fondamentalement neuve ni pour la linguistique, ni pour la psychologie. Ce qui advient de manière complètement nouvelle et radicale avec la grammaire générative est la formulation spécifique du réalisme cognitif telle qu'elle a été proposée par Chomsky (Carr, 2002, Laks, 2005). Avec la grammaire générative et dès *Aspects*, le terme de « grammaire » devient systématiquement et délibérément polysémique et « ambigu » désignant tout à la fois l'objet construit par le linguiste analysant la structure d'une langue et l'objet mental appris par l'enfant et manipulé par le locuteur-auditeur

« La grammaire est un objet *systématiquement ambigu* qui désigne à la fois l'objet mental construit par l'enfant qui apprend sa langue maternelle et l'objet abstrait construit par le linguiste pour rendre compte de cet apprentissage : "Nous employons le terme "grammaire" pour désigner à la fois *le système de règles représenté dans le cerveau du locuteur-auditeur*, système acquis normalement dans la petite enfance et utilisé dans la production et l'interprétation des énoncés, et *la théorie* que le linguiste construit en tant qu'hypothèse concernant la grammaire intériorisée réelle du locuteur-auditeur.

N.Chomsky & M. Halle, *The Sound Pattern of English*, 1968, p.26

La nouveauté ne porte donc pas sur l'éventuelle réalité psychologique voire biologique des objets linguistiques qui a été formulée ailleurs mais sur l'isomorphie stricte entre la grammaire construite par le linguiste et la grammaire interne du locuteur. Ceci donne, par conséquent, un statut directement et *a priori* cognitif et neurologique à tout formalisme linguistique établi sur la base de l'observation des données produites et des comportements d'un locuteur. Ceci suppose donc plus ou moins implicitement, que les objets phonologiques définis par la théorie (par exemple des phonèmes, des syllabes, des éléments) sont simultanément des objets linguistiques et cognitifs.

De là découle un programme de recherche complet où la tâche du linguiste ne consiste plus à décrire la structure des énoncés observables mais à proposer, à partir des données produites un modèle de « la grammaire » interne (I-language : internal, individual, intensional), c'est à dire un modèle des représentations mentales qu'utilise un locuteur-auditeur, voire dans certaines formulations les plus récentes un modèle de « l'organe langage » (Chomsky 2000, Anderson & Lightfoot, 2002). Les propositions concernent un dispositif théorique global rendant compte de l'état initial (GU, grammaire universelle), de l'état final (GF, grammaire de l'adulte), de l'acquisition de « I-language » mettant en relation GU et GF via un LAD (« Language Acquisition Device ») ainsi que de l'utilisation de la « compétence » par le locuteur-auditeur (« performance »).

Ce modèle est donc tout à la fois linguistique, psychologique et ultimement biologique (on suppose neurologique).

« Depuis trente ans l'étude du langage -ou plus précisément, une de ses composantes importantes-est menée dans un cadre qui conçoit la linguistique comme une branche de la psychologie, donc en dernière instance de la biologie humaine »,

N.Chomsky, *RLV*, 19, p. 21

Ceci sous-entend également de manière indirecte, qu'il n'existe pas de linguistique et donc pas de phonologie non cognitive.

On peut considérer que cette conception du réalisme chomskyen s'est imposée plus radicalement (« strong realism ») après *Aspects* et a été précédée d'une version faible (« weak realism ») où la question du réalisme est formulée en termes peu explicites de telle sorte que ne s'imposent clairement ni l'idée d'une « réalité psychologique » de la grammaire, ni celle de l'isomorphie entre grammaire du linguiste et grammaire du locuteur (Seuren, 1998 ; Carr, 2000). Dans cette version faible du réalisme, les concepts linguistiques ne renvoient pas à des objets ou des procédures psychologiques réelles mais seraient plutôt envisagés comme des métaphores de processus psychologiques dont la validité n'entame pas en soi la

justesse de la théorie linguistique<sup>28</sup>.

A partir de *SPE*, est formulée et s'impose la version radicale du réalisme, avec la conséquence énoncée plus haut : les objets linguistiques définis par la théorie sont simultanément des objets linguistiques et cognitifs (et enfin biologiques).

Cette conception radicale du réalisme linguistique (qui est affirmée très régulièrement dans tous les textes chomskyens depuis 1968 mais qui ne correspond pas, en fait, à la pratique réelle de la plupart des phonologues) a déjà fait couler beaucoup d'encre (Harman, 1980 ; Cummins & Harnish. 1980 ; Sober, 1980 ; Mogdil & Mogdil, 1987 ; Botha, 1989 ; Seuren 1998 ; Carr, 2000 ; Devitt, 2003 ; Laks, 1996, 2005 ) et constitue un des points majeurs de remise en question de l'épistémologie chomskyenne. Cette question du réalisme et plus particulièrement de la nature cognitive de la phonologie a recoupé régulièrement les problèmes dont je traite.

## **2.2. Phonétique, phonologie, psychologie, biologie : 4 niveaux de réalité différents.**

Le problème central qui s'est posé à moi est ce qui me semble être l'incohérence logique de la proposition faite par le « strong realism » (Botha, 1989, Seuren 1998, Harman, 1980 ; Cummins & Harnish, 1980 ; Sober, 1980 ; Carr, 2000). Elle porte sur la question du niveau de réalité des modèles génératifs (ou si l'on préfère le degré de métaphorisation), ou plutôt sur la confusion entre les niveaux de réalité qui sont proposés dans l'entreprise de modélisation au sein du paradigme génératif. Ceci entraîne par conséquent un flou dommageable quant à la définition exacte de « GU », « LAD », « compétence », tous concepts non observables empiriquement.

Le recours au mentalisme tel que le formule Chomsky, à partir d'*Aspects*, dans sa version la plus dure, prédit que les représentations phonologiques sont également des objets psychologiques, et des objets biologiques au sein d'une théorie non psychologique et non biologique, donc ne tombant pas *ipso facto* sous le coup des

---

<sup>28</sup> Pour les citations de Chomsky et les commentaires de ces citations, cf Seuren, 1998 p. 253 et sq.

modes de falsification d'une théorie psychologique ou biologique. En ce sens, on peut considérer que les objets sonores sont quadruplement représentés (dans le signal, dans la grammaire, dans l'esprit du locuteur et dans le cerveau du locuteur), ce qui rend leur définition délicate.

Mais répondre à la question en rabattant trois<sup>29</sup> niveaux de réalisme (linguistique, psychologique et biologique) les uns sur les autres ne semble pas pour autant apporter une réponse satisfaisante à ce problème. Cette démarche pose divers problèmes épistémologiques (Cummins & Harnish 1980 ; Harman, 1980 ; Sober, 1980, Devitt, 2003) soulevés dans le numéro spécial de *The Behavioral and Brain Sciences*, 3, 1980 (Ci-après : *BBS*) où la pensée chomskyenne est présentée dans un article intitulé « Rules and Representations » (Ci-après : *RR*) et discutée ensuite par divers philosophes et épistémologues.

Les débats portant sur le réalisme psychologique du modèle linguistique s'articulent autour des deux questions suivantes

- i) Une théorie scientifique « réaliste » se doit de définir clairement le niveau de « réalité » (Devitt, 2003), qu'elle modélise. Comment peut-on, par conséquent, définir explicitement le niveau de réalité des objets formulés par la grammaire générative et la possibilité de leur réalité psychologique ?
  
- iii) Quels sont les modes de falsification qui permettent légitimement de valider ou de falsifier des objets représentés simultanément à plusieurs niveaux de réalité ?

Je traite de la question ii) en 3.2, dans la partie traitant du rapport compétence vs performance.

En ce qui concerne la question i), au sein d'une même théorie, des objets peuvent être, comme le souligne Harman (1980), représentés à des niveaux de réalité

---

<sup>29</sup> Je ne discute pas de l'interface de la phonologie avec la phonétique, qui concerne aussi la question de l'articulation forme / substance et de la réalité du modèle, car tel n'est pas mon champ d'interrogation. Mais il est évident que cette question s'articule aussi au problème de la réalité cognitive des objets perçus et produits.

différents, sans que cela n'empêche pour autant leur véracité. Le fait qu'ils aient ou non une réalité physique n'affecte pas leur degré de réalité au sein du cadre théorique qui les a conçus, ni le degré de véracité du modèle qui les conçoit.

« Chomsky claims that it is pointless to distinguish the question of psychological reality from that of truth. But given any theory we take to be true, we can always ask what aspects of the theory corresponds to the reality and what aspects are mere artefacts of our notation. Geography contains true statements locating mountains and rivers in terms of longitude and latitude without implying that the equator has the sort of physical reality the Mississippi River does. Sometimes we are not sure about the physical reality of some aspect of a theory, even given strong evidence for the truth of the theory. »

G. Harman, « Two quibbles about analyticity and psychological reality », *The Behavioral and Brain Sciences*, 3, 1980, p. 21

En ce sens, il m'apparaît absolument légitime de postuler, comme hypothèse de travail que la GU, le LAD existent, de proposer une dichotomie compétence/performance et de proposer des formalisations complètement abstraites de la compétence phonologique, donc des « consonnes flottantes », des « attaques branchantes », des « traits », ou des « éléments », conceptualisées à partir d'une théorie linguistique et falsifiable par des données linguistiques. Mais leur réalité linguistique n'impose pas qu'ils soient psychologiquement et biologiquement réels, ni qu'ils aient une quelconque réalité physique. Le problème commence à partir du moment où l'on considère que ces trois niveaux de représentations entretiennent le même rapport à la réalité et qu'un objet linguistique est réel, comme un objet psychologique et un objet biologique le sont respectivement (Sober, 1980). De là s'ensuit la possibilité de leur attribuer un degré de réalité relevant d'un autre niveau de théorisation et de falsification que celui où ils ont été définis. Une syllabe peut-être simultanément un objet phonologique et un objet psychologique et même un objet biologique (articulo-motrice, « porteuse de parole », Abry et al. 2002) en tant qu'elle est respectivement conçue, modélisée dans un paradigme phonologique, psychologique et articulo-moteur. Mais elle ne peut être simultanément ces trois objets dans un modèle linguistique unique à moins de considérer



i) que la linguistique est une branche de la psychologie et que tous les objets linguistiques ressortissent ultimement de théories psychologiques. C'est ce qu'affirme régulièrement Chomsky. Or, la théorie de la compétence, n'intègre pas d'hypothèses sérieuses sur la nature véritablement psychologique de la compétence, à partir des connaissances et des théories psychologiques disponibles.

ii) que la linguistique et donc la psychologie sont une branche de la biologie et que tous les objets linguistiques-psychologiques ressortissent ultimement de théories biologiques. Or, on voit mal comment une théorie dualiste séparant par principe matière et esprit peut aboutir à de telles conclusions. L'explicitation de la question faite par Chomsky dans la première section intitulée « mind and body » de *RR* est pour le moins sujette à caution et suppose un corps abstrait, différent de l'esprit mais a-substantiel dont le corps matériel serait la métaphore.

« In fact, we may think of the study of mental faculties as actually being a study of the body-specifically the brain- conducted at a certain level of abstraction »<sup>30</sup>,

N. Chomsky, « Rules and representations », *The Behavioral and Brain Sciences*, 3, 1980, p.2

« Returning to the analogy of the physical body »<sup>31</sup>

N. Chomsky, « Rules and representations », *The Behavioral and Brain Sciences*, 3, 1980, p.3

iii) qu'en tant que discipline autonome, la linguistique, telle qu'elle a été pratiquée depuis toujours par linguistes, grammairiens et philologues n'a pas de légitimité. Ceci signifie que le travail de description et d'analyse des données, l'établissement des régularités typologiques etc. ne sont pas en soi dignes d'intérêt scientifique puisque toute théorie linguistique est une théorie psychologique et finalement biologique. Je pense qu'il ne viendrait à l'idée

---

<sup>30</sup> c'est moi qui souligne

<sup>31</sup> c'est moi qui souligne

d'aucun linguiste de défendre un tel point de vue.

iv) que seules les données psycholinguistiques et ultimement de neuro-imagerie sont aptes à falsifier ou valider la théorie de la compétence. Or ceci est hautement problématique (sur la question de la falsification de la compétence par la performance cf. Ch.I, 3.2).

Il m'apparaît donc que les modèles phonologiques proposés par la Grammaire générative théorisent des objets phonologiques qui sont réels. Ces modélisations formalisent de manière réaliste des hypothèses réelles. La scientificité de l'entreprise n'est absolument pas en cause. La cohérence se construit de manière interne à la proposition et aux présupposés du modèle pour rendre compte de données empiriques et les modes de falsification retenus sont absolument légitimes (Carr, 2000) et pertinents.

Mais tous, sans exception, mettent en forme temporairement, sur la base de choix épistémologiques et de données empiriques une représentation plus ou moins adéquate des structures linguistiques dont dispose le sujet parlant et non pas des structures psychologiques (Devitt, 2003) ou biologiques.

On peut, par ailleurs et simultanément, considérer que la linguistique est une discipline cognitive et que les fonctionnements linguistiques correspondent à des activités cognitives, psychologiques et à des opérations ou états neurologiques chez le locuteur, mais pour autant, la réalité linguistique avérée de ces objets chez le locuteur-auditeur ne leur donne pas automatiquement de réalité psychologique et encore moins neurologique.

L'intérêt de tel ou tel modèle en phonologie ne peut donc actuellement être évalué, comme c'est le cas pour tout modèle scientifique me semble-t-il, qu'en termes de simplification et d'abstraction de l'empirie, d'adéquation explicative (possibilité d'établir des généralisations satisfaisantes en particulier) par rapport aux données considérées dans leur complexité phénoménologique, ainsi qu'en termes de capacité de prédiction des observables possibles à partir des données. Mais la réalité psychologique et /ou neurologique des objets phonologiques est, à l'heure actuelle, vu la manière dont sont construits les formalismes, à partir de la théorie de la

compétence chomskyenne, orthogonale à tout modèle phonologique.

L'établissement de la preuve de la réalité psychologique et neurologique des objets phonologiques proposés, (et ceci est vrai quelle que soit le cadre théorique retenu) ne peut pas être actuellement du ressort d'un modèle phonologique, c'est trop lui demander dès lors qu'il n'intègre pas volontairement et très sérieusement une composante psychologique et une composante neurologique reposant sur une théorie psychologique et une théorie neurologique simultanément, articulées de manière adéquate et pertinente à la théorie linguistique, et nécessitant par conséquent la construction des interfaces entre ces différents niveaux de réalité.

Seul un modèle articulant les différents niveaux de réalité permet de construire de manière rigoureuse et simultanément l'objet au niveau phonétique, phonologique, psychologique et neurologique.

Le problème de l'articulation de la grammaire générative avec la psychologie a été formulée et prise au sérieux dès *Aspects*, par l'établissement de la dichotomie « compétence vs performance ». Par ailleurs, le souci de ne pas abandonner en chemin la problématique cognitive en phonologie s'est exprimée très régulièrement dans SPE et depuis SPE. Il constitue à l'évidence un puissant facteur de renouvellement des propositions théoriques depuis 35ans mais me semble avoir été régulièrement mis en échec.

L'articulation avec la biologie semble au contraire plutôt une position de principe formulée très régulièrement dès *Aspects* et définissant des relations ancillaires entre linguistique et neurologie de manière indirecte et vague.

« On pourrait même supposer au contraire que les études mentalistes seront finalement les plus utiles pour l'exploration des mécanismes microphysiologiques puisqu'elles seules se préoccupent de déterminer abstraitement les propriétés que de tels mécanismes doivent posséder et les fonctions qu'ils doivent assurer. »

N. Chomsky, *Aspects*, 1965, [193-194], p. 13

A de nombreuses reprises, Chomsky parle également d'esprit-cerveau (« mind /

brain ») du locuteur. Et cette expression est depuis régulièrement reprise ici ou là comme allant de soi (Anderson & Lightfoot, 2002) alors qu'il me semble qu'elle est hautement problématique pour un modèle dualiste, qui viserait ultimement l'isomorphie des niveaux de représentations tout en affirmant qu'ils sont distincts par nature (leur distinction est le fondement du dualisme cartésien dont Chomsky se réclame pleinement, 1966).

Doit-on considérer que cette expression est un raccourci qui n'engagerait pas le fond du débat (quelque chose comme : l'esprit dont on postule qu'il a une correspondance anatomique dans le cerveau) ? S'il s'agit de cela, vu les connaissances dont nous disposons actuellement sur les activités mentales de l'être humain et leur localisation cérébrale, c'est une formulation triviale qui n'apporte rien.

Mais il semblerait qu'on ne puisse pas s'en tenir là. Cette expression mettant sur le même plan esprit et cerveau et consistant à postuler que la linguistique est ultimement une branche de la biologie n'est pas, semble-t-il, le fait du hasard. Une autre expression me paraît, dans le même registre, hautement problématique : « l'organe-langage »

« The faculty of language can reasonably regarded as a « language organ » in the sense in which scientists speak of the visual system, or immune system or circulatory system as organs of the body. Understood in this way, an organ is not something that can be removed from the body, leaving the rest intact. It is a subsystem of a more complex structure. »

N. Chomsky, *New horizons in the study of language and Mind*, 2000, p. 4

Le recours au mot « organe » associé ainsi à « langage » m'a toujours choquée par la naïveté qu'il suppose. Il consiste à assimiler un organe à sa fonction et entraîne automatiquement une confusion entre les niveaux anatomiques et fonctionnels (Carr, 2000, Dennett, 1995). Le bébé humain peut être programmé génétiquement pour parler, et le faire de manière irrépensible et non consciente, les seuls organes qu'il utilisera sont le cerveau, le larynx, la langue ... etc. Car, si tout le monde s'accorde sur le fait que le cœur, les poumons, l'estomac sont des organes, il ne viendrait à l'idée de personne de parler pour autant « d'organe-circulation sanguine » ou « d'organe-respiration » ou « d'organe-nutrition » puisque, le coeur, les poumons,

l'estomac sont respectivement les organes de la circulation sanguine, de la respiration, de la digestion, qui sont elles mêmes des fonctionnalités physiologiques du coeur, des poumons, de l'estomac complètement déterminées génétiquement.

Formuler la question des facultés linguistiques en termes d'isomorphie entre les structures anatomiques et les hypothèses fonctionnelles et les désigner comme « organe », laisse supposer que les modèles proposés par la phonologie théorique auraient ultimement une implémentation directe dans la substance cérébrale. Ceci est en soi aberrant, biensûr faux et en outre radicalement incompatible avec une conception dualiste cartésienne qui sépare radicalement ces deux niveaux et leur attribue, par principe, des natures différentes. Ceci est d'ailleurs reflété par la revendication très logique, absolument légitime dans une conception rationaliste d'une phonologie a-substantielle (substance-free) qui, par définition, ne peut et ne doit pas être directement implémentée dans la substance cérébrale, puisque les objets phonologiques sont des objets abstraits et qu'ils doivent le rester.

Rien de ce qui n'a été observé actuellement grâce notamment à la neuro-imagerie ne permet de penser que les « consonnes flottantes », « traits » sont des objets organiques ou qu'il existe des « consonnes flottantes », « traits » etc dans l'organe cerveau. De même, l'activité cérébrale, observable actuellement par la TEP ou les IRMf, se matérialise, pour ce que l'on en sait actuellement, par des échanges électriques et des modifications des flux sanguins et non par des conflits de contraintes ou des rattachements de consonne flottante.

Ce qui a été montré est qu'il existe pour les activités linguistiques, une géographie cérébrale précise (Houdé et al. 2002), qui connaît des variations assez importantes mais qui subit notamment lors de son développement d'importantes contraintes génétiques (Dehaene-Lambertz et al. 2000, Houdé et al. 2002, Rondal et al., 1999).

Ces faits apportent des arguments très consistants qui permettent de considérer qu'il y a une physiologie spécifique du langage, dont nous ne connaissons encore que très peu de choses, et que le langage est le produit d'une des fonctionnalités génétiquement conditionnée du cerveau. Mais ceci ne permet pas d'affirmer que le langage est un organe. Rondal et al. (1999) discutent, d'un point de vue neurologique,

la non-pertinence des notions « d'esprit-cerveau » et « d'organe-langage » et opposent à un « innéisme représentationnel » hautement spéculatif, un « innéisme organiciste » fondant les capacités linguistiques de l'humain sur l'existence d'un substrat neuro-biologique, inné, spécifique et dont le développement et le fonctionnement sont partiellement indépendants des fonctions cognitives générales :

« Ce qui est inné (mais qui vient à maturation), c'est une série de dispositifs cérébraux (encore loin d'être complètement élucidés [...]), sous-tendant une série de capacités et de mécanismes (en partie cognitifs) intervenant dans le développement et le fonctionnement langagier [...]. Ce qu'on doit considérer comme invalidé, c'est l'indication selon laquelle le développement de la grammaire serait entièrement ou de façon majeure sous la dépendance du développement cognitif général. Il n'en est rien. Il semble bien que le développement linguistique dans ses composantes phonologique et morpho-syntaxique, se fasse, au moins en partie de façon intrinsèque »

J.A. Rondal, E. Espéret, J.E. Gombert, J.P. Thibaut, A. Comblain, *Troubles du Langage, bases théoriques, diagnostic et rééducation*, ch 5, p. 153

De la même manière que pour l'articulation à la psychologie, aucun modèle phonologique ne peut actuellement apporter la preuve de la réalité biologique des objets qu'ils propose. Les hypothèses ne peuvent être faites que de manière indirecte par l'étude de l'acquisition phonologique normale et pathologique ou l'étude des aphasies.

Nous sommes, me semble-t-il, arrivés à une étape où l'évolution récente et très prometteuse de la psycholinguistique et neurolinguistique cognitives (Dupoux, 2002), les résultats apportés par les données de perception d'une part, et les travaux en neuro-imagerie cérébrale d'autre part peuvent nous laisser espérer la possibilité de formuler prochainement les 4 niveaux de réalité des objets sonores (phonétiques, phonologiques, psychologiques et biologiques) et de les articuler dans des structures mixtes qui intègrent les différents niveaux de réalité.

Ces réflexions sont programmatiques et ouvrent un espace de travail immense, j'en suis consciente. Par ailleurs, la question est complexe et nous manquons

actuellement de modélisations testables articulant la production et la perception (cf. Peperkamp 2002, pour l'acquisition). Mais je pense qu'il convient surtout et prioritairement de s'attacher particulièrement à l'élaboration des interfaces entre ces disciplines et de ne pas attendre qu'elles émergent d'elles-mêmes. L'idée que chaque discipline définit de manière interne ses objets de recherche et que ce partage des tâches bien compris aboutira à un jour à ce que l'on puisse recoller les morceaux ensemble me paraît relever d'une conception irréaliste qui reconduit et enkyte *ad libitum* les apories du rationalisme chomskyen.

Je pense qu'une phonologie si elle se veut cognitive et rationaliste ne peut plus actuellement se passer d'une réflexion menant à des modèles intégratifs et intégrés sur 4 niveaux de réalité différents et donnant une place centrale à la théorisation des interfaces. C'est à dire en se donnant comme projet de construire ses modélisations en intégrant aussi les interfaces avec les autres composantes du langage, ou du moins de proposer des modélisations qui prennent cette question au sérieux et qui les rendent ultimement possibles.

L'élaboration et la théorisation des interfaces où il s'articule avec la phonétique, la psychologie et la neurologie, est la condition nécessaire du maintien d'un niveau phonologique abstrait, indépendant et structuré qui ne soit assimilé ni à la substance phonétique, ni à des procédures psychologiques. La possibilité que nous avons de définir ce dont la phonologie est la théorie suppose qu'on définisse aussi clairement ce dont elle n'est évidemment pas la théorie. Et je pense que cette tâche doit être faite à partir de la théorisation phonologique, en vue de la construction des interfaces avec les autres disciplines.

C'est dans cet esprit que j'ai entrepris mon travail de recherche.

### **3. Linguistique et psychologie : questions et problèmes**

Le projet chomskyen tel qu'il est énoncé dans *Aspects*, suppose, comme nous l'avons montré, une théorisation nécessaire et inévitable des relations entre

linguistique et psychologie, et ultimement entre linguistique et biologie / neuro-physiologie du langage.

Mais cet impératif tout à la fois épistémologique et méthodologique s'est énoncé dans un contexte historique et scientifique, qui conditionne la manière dont le débat a été théorisé, à partir des repères suivants : l'anti-behaviourisme radical de Chomsky, la formulation de l'articulation compétence *vs* performance et le refus, pour l'acquisition de la psychogenèse piagétienne (dont je discute ch. IV. 1) ainsi que les conséquences sur la dimension cognitive du programme génératif.

### 3.1. La dichotomie compétence *vs* performance

La dichotomie compétence *vs* performance apparaît pour la première fois dans l'appareillage conceptuel posé explicitement dans *Aspects* et va être ensuite reprise très régulièrement jusque dans les textes les plus récents. Les définitions des notions elle-mêmes, comme la définition exacte de ce qu'implique la dichotomie compétence *vs* performance sont selon moi peu claires dès l'origine. Ce flou conceptuel initial explique en partie pourquoi cette dichotomie n'a pas mené à l'émergence d'un programme cognitif intégré et fertile, rendant compte avec succès simultanément des représentations et des procédures phonologiques utilisées en perception et en production.

#### 3.1.1. Compétence

La définition de la compétence qui est donnée régulièrement par Chomsky a très peu varié et est assez claire.

*« L'objet premier de la théorie linguistique est un locuteur-auditeur idéal, appartenant à une communauté linguistique complètement homogène, qui connaît parfaitement sa langue et qui, lorsqu'il applique en une performance effective sa connaissance de la langue, n'est pas affecté par des conditions grammaticalement non pertinentes, telles que la limitation de mémoire, distractions, déplacements d'intérêt ou d'attention, erreurs (fortuites ou caractéristiques).[...] Aspects,*

*[...] la grammaire d'une langue se propose d'être une description de la compétence intrinsèque du locuteur idéal » Aspects,*



« Le terme technique de « compétence » désigne la capacité qu'a le locuteur idéal d'associer son et sens en accord strict avec les règles de sa langue. La grammaire d'une langue comme modèle strict de la compétence idéale établit une certaine relation entre le son et le sens, entre les représentations phonétiques et sémantiques. [...] Le but premier de l'étude d'une langue est de découvrir cette grammaire. N. Chomsky, *La Linguistique cartésienne*, 1966, [398], p.126

De manière explicite, la compétence est désignée comme la « grammaire » abstraite et mentale, mais non consciente, du locuteur-auditeur idéal p.10 [4], ou plus tard par le terme I-Language (Internal, individual, intensional language) mais renvoie toujours au même objet théorique : la somme de connaissances abstraites et symboliques, non substantielles et non contraintes par l'usage et le contexte qu'utilise le locuteur idéal.

### 3.1.2. Performance

La définition claire de la « performance » est beaucoup plus difficile à établir et semble se définir d'abord par défaut comme tout ce qui, dans le langage ne relève pas de la compétence, donc de la grammaire.

« The factors that Chomsky consigns to performance – the things he thinks a theory of linguistic competence should *not* be concerned with- are quite heterogeneous »

R. Jackendoff, *Foundations of language*, 2002, ch. 2, p.30

Elle a en outre évolué et varié au fur et à mesure des formulations (E-language, external language) qui en ont été proposées par Chomsky mais peut être envisagée comme une somme de facteurs variés qui conditionnent l'utilisation en temps réel de la compétence linguistique par un locuteur-auditeur réel en contexte, et qui sont posés a priori comme étant externes aux préoccupations du linguiste

« Nous établissons donc une distinction fondamentale entre la *compétence* (la connaissance que le locuteur auditeur a de sa langue et la *performance* (l'emploi effectif de la langue dans des situations concrètes). [...] L'emploi du langage tel qu'on l'observe ou d'ailleurs les tendances que l'on peut supposer à répondre aux stimuli, les habitudes, etc, tout cela peut fournir une information sur la nature de la réalité mentale mais ne saurait assurément constituer l'objet effectif de la linguistique si celle-ci doit être une discipline sérieuse »

Noam Chomsky, *Aspects*, p. 14

Dans *Aspects*, sont désignés comme relevant de la performance

- la limitation de la mémoire,
- les distractions, perte d'attention
- intonation et prosodie
- le traitement psychologique (ou processing)

A ceci s'ajoute plus tardivement dans d'autres textes, les dispositifs sémantico-pragmatiques et phonético-perceptifs qui en amont et en aval de la grammaire permettent son utilisation en temps réel

« The performance systems appear to fall into two general types : articulatory-perceptual, and conceptual-intentional. If so, it is reasonable to suppose that a generated expression includes two interface levels, one providing information and instructions for the articulatory-perceptual systems, the other for the conceptual intentional system. The properties of these systems, or their existence are matters of empirical fact »

N. Chomsky, *New horizons in the study of language and Mind*, 2000, p. 27-28.

Ou même ce qui pourrait relever de contenus culturels socialement partagés

« La performance, c'est à dire ce que le locuteur-auditeur réalise effectivement, est fondée non seulement sur la connaissance qu'il a de la langue, mais bien sûr d'autres facteurs –comme les limitations de la mémoire, l'inattention, la distraction, les connaissances et croyances non-linguistiques, etc. On peut concevoir l'étude de la compétence comme l'étude de la performance potentielle d'un locuteur idéal qui ne serait pas affecté par des facteurs de ce type qui n'ont pas de pertinence grammaticale».

N. Chomsky & M. Halle, *The Sound Pattern of English*, 1968, p.25

Je ne multiplierai pas ici les citations mais quand on confronte toutes les définitions qui sont données de la performance, il n'est pas très facile en particulier d'établir exactement si la performance

- i) est considérée comme strictement interne ou non ou pas exclusivement

- ii) est considérée comme cognitive et exclusivement cognitive
- iii) est ou non procédurale

Mais ce qui est commun toutes les définitions données dans les textes chomskyens (et c'est avec cette acception qu'il est employé la plupart du temps dans la doxa chomskyenne), le terme de « performance » renvoie aux opérations de traitement psychologiques en perception et production qui sont considérés comme externes à la grammaire.

Cette définition de la performance comme le ou les système(s) qui permettent le traitement de la compétence en temps réel établit de fait une ligne de fracture lisible dès *Aspects* entre d'une part les représentations (knowledge) relevant de la compétence et le traitement (processing) relevant de la performance et ainsi qu'un partage des tâches assez clair entre ce qui revient aux linguistes et ce qui ne relève pas de leur préoccupations.

« Une grammaire générative tente de caractériser de la façon la plus neutre la connaissance de la langue qui fournit sa base à la mise en acte effective du langage par le locuteur-auditeur. Lorsque nous disons qu'une grammaire engendre une phrase pourvue d'une certaine description structurale, nous entendons simplement que la grammaire assigne cette description structurale à la phrase. Lorsque nous disons qu'une phrase a une certaine dérivation du point de vue d'une grammaire générative particulière, nous ne disons rien sur la manière dont un locuteur ou un auditeur pourrait procéder, d'une façon pratique ou efficace pour construire une telle dérivation. Ces questions appartiennent à la théorie de l'acte linguistique – la théorie de la performance »

Il semble qu'on puisse repérer nettement toutes les ambiguïtés mais aussi tous les linéaments des débats qui agitent depuis la phonologie quant à la part respective des représentations et des procédures dans les modèles théoriques (Ch. III, 4.).

### 3.1.3. Les interprétations de la dichotomie compétence vs performance

Quoi qu'il en soit, malgré le flou concernant la définition exacte de la performance et de son rapport à la compétence, on voit quand même se dégager les lignes de force suivantes sur lesquelles on peut considérer qu'il y a un accord

i) Représentations *vs* traitement

Etablissement d'une séparation étanche entre les connaissances linguistiques qu'a le locuteur idéal et la manière dont il les utilise en temps réel et en contexte pour produire et traiter des énoncés réels

ii) Linguistique *vs* psycholinguistique

Etablissement de deux champs disciplinaires : la linguistique s'occupant exclusivement de la compétence, la psycholinguistique de la performance.

Ce partage des tâches suppose un partage sociologique du champ de la recherche, ainsi que l'utilisation de méthodologies et de modes de validation différents à partir d'une phénoménologie différente (productions du locuteurs et intuitions linguistiques *vs* expérimentations)

### 3.1.4. Relations entre compétence et performance, relations entre les deux disciplines

Cette question, pourtant centrale pour le dispositif chomskyen, est également entachée d'un flou dommageable. Il semblerait que deux interprétations concurrentes des relations entre compétence et performance co-existent dans les textes.

Il a été proposé (surtout dans les premiers textes, mais on en trouve trace également dans Chomsky 2000, Bresnan, 1982, Jackendoff, 2000) que la compétence soit définie de manière intrinsèque mais avec comme but ultime une intégration au sein d'une théorie de la performance. L'idéalisation et l'autonomie d'une grammaire radicalement opaque aux usages semblent alors être envisagées comme un état transitoire ne contraignant pas la théorisation du traitement en lui-même

« We are studying a real object, the language faculty of the brain, which has assumed the form of a full I-language and is integrated into performance systems that play a role in articulation, interpretation, expression, of beliefs and desires, referring, telling stories, and so on »

N. Chomsky, *New horizons in the study of language and Mind*, 2000, p. 27

« Sans aucun doute un modèle raisonnable de l'acte linguistique comprendra, comme l'un de ses composants fondamentaux, la grammaire générative où se formule la connaissance que le locuteur- auditeur a de sa langue ; mais cette grammaire générative, en elle-même n'engage ni le caractère ni le fonctionnement d'un modèle perceptuel ou d'un modèle de production de la parole »

Noam Chomsky, *Aspects*, p.20

Mais on trouve quelques pages plus loin l'affirmation contraire supposant une relation d'ancillarité entre ces disciplines, de telle sorte qu'on ne voit plus très bien si le programme scientifique envisagé consiste à construire une théorie de la compétence en tant qu'elle puisse ultimement être une composante d'une théorie psychologique du langage ou une théorie de la compétence en soi où les données de performance n'ont valeur que de validation

« La linguistique mentaliste n'est autre que la linguistique théorique qui prend la performance comme une donnée (jointe à d'autres données : par exemple celles que fournit l'introspection) pour la détermination de la compétence, cette dernière étant l'objet premier de sa recherche. Le mentalisme, en ce sens traditionnel n'a pas à faire d'hypothèses touchant la base physiologique possible de la réalité mentale qu'il étudie. » *Aspects*, [193-194], p. 13

Examiné sous l'angle de la citation qui précède (*Aspects*, [193]), le modèle de la compétence-grammaire proposé est donc quelque peu paradoxal qui s'affirme réaliste et mentaliste, donc avec pour but ultime l'établissement des représentations psychologiques qu'utilise un locuteur-auditeur, tout en s'exemptant, *de facto*, des conditions mêmes de son réalisme qui sont assumées exclusivement par la théorie de la performance.

On voit bien dès lors comment la théorie de la performance devient, dans ce dispositif théorique, le nécessaire garant du réalisme du modèle de compétence retenu

et comment seule la psychologie aura finalement à rendre compte effectivement du réalisme des propositions, alors que la linguistique théorique peut postuler un réalisme *per se* garanti par l'argument de l'ambiguïté de la grammaire. A charge ensuite de la psycholinguistique, si elle le souhaite, de prouver effectivement que les propositions faites par les phonologues par exemple, à partir des données de langues sont effectivement psychologiquement et ultimement biologiquement réalistes. Il se peut qu'elle le soient : en toute logique rationaliste, si les structures linguistiques lisibles dans les données reflètent directement l'organisation des contenus mentaux, elle doivent l'être. Reste que cette prédiction n'est jamais empiriquement vérifiable, ni vérifiée.

Enfin, la psycholinguistique cognitive ayant développé essentiellement des recherches en perception, une conséquence directe de cette dichotomie s'est imposée par les faits eux-mêmes : les modèles de la compétence phonologique sont des modèles de production et les modèles de performance des modèles de traitement.

Il apparaît donc que l'une et l'autre interprétation n'aient pas tenu leurs promesses et qu'aucune réconciliation vraiment satisfaisante entre compétence et performance ne soit actuellement très concrètement réussie. Aucun modèle de performance n'a intégré de manière pertinente un modèle de la compétence, et je pense que depuis l'échec de la théorie de la complexité dérivationnelle (DTC, ci après), l'espoir d'un « paradis de quelque chose comme la DTC » (Bonatti, 2002) a été abandonné tant par les linguistes que par les psycholinguistes. La deuxième interprétation possible des relations entre compétence et performance serait-elle plus satisfaisante ?

### **3.2. Compétence vs performance encore : la question de la falsification de la compétence phonologique par la performance est-elle pertinente ?**

Beaucoup plus que la syntaxe, la phonologie s'est maintenue à distance des modèles de performance et n'a jamais réellement suivi un programme équivalent à la DTC consistant à produire un modèle phonologique qui pourrait être intégré à un

modèle de performance. Par ailleurs, elle s'est assez bien passée des données de performance comme mode de validation de ses propositions théoriques. La question est pourtant très explicitement posée depuis quelques temps par l'émergence de la « phonologie de laboratoire ».

Je discuterai cette question à partir du chapitre de Pierrehumbert et al. (2000), intitulé « Conceptual foundations of Phonology as Laboratory Phonology » qui pose explicitement ce problème. Dans ce chapitre, les auteurs remettent en cause la conception de la compétence défendue par le paradigme phonologique génératif « classique » (rationaliste, descendant, substance-free, dualiste). Les modèles cités explicitement sont la phonologie lexicale, la phonologie métrique, la phonologie autosegmentale, la phonologie déclarative, la phonologie du gouvernement et la théorie de l'optimalité) auxquelles est opposée sur la base d'arguments épistémologiques variés une « phonologie de laboratoire » (empiriste, non dualiste, comportementaliste, non-formelle *a priori*) et appuyant les modélisations proposées sur une logique montante qui consiste à observer et à mesurer de manière objective les données produites et perçues ainsi que la multiplicité des comportements et des usages linguistiques des locuteurs réels pour en tirer ensuite une modélisation cognitivement pertinente.

L'une des questions soulevées intéresse très directement notre propos sur le rapport compétence vs performance et l'exigence de réalisme qui s'impose à l'une et à l'autre. Ce chapitre développe une réflexion assez longue sur les modes de validation respectifs de ces deux types d'approche (formelle / comportementaliste) et met en cause la pertinence des jugements de grammaticalité et intuitions des locuteurs comme mode de validation scientifique

« Much of the generative literature assumes that well-formedness judgements provide the most direct and revealing data about competence [...]. However studies in the sociolinguistics and psycholinguistics literature cast serious doubt on the reliability and predictiveness of well-formedness judgements. Well-formedness judgements are opinions. They are high-level meta linguistic performances that are highly malleable. They do not represent any kind of direct tap onto competence, but rather prone to many types of artefacts, such as social expectations, experimenter bias, response bias, and undersampling »,

*Phonological knowledge, conceptual and empirical issues*, p.289

La démarche consiste ici à interroger, en mettant en cause les modes de validation qu'ils retiennent, la légitimité scientifique des principaux paradigmes phonologiques génératifs cités. Les auteurs revendiquent une épistémologie de type poppérienne, tout en contestant, *a contrario*, et à tort, que la phonologie générative en fasse usage. Par ailleurs, ils lui récusent tout caractère cumulatif qui prendrait en considération les résultats obtenus précédemment par d'autres paradigmes.

Mettant en cause les assises épistémologiques et la pertinence des modes de validation (le recours quasi-exclusif aux jugements de grammaticalité) utilisés par les modèles phonologiques génératifs « classiques », Pierrehumbert et al. suggèrent en conséquence que la conception de la compétence qu'ils proposent n'a que peu de pertinence scientifique puisque ces modèles ne sont pas falsifiables par des méthodes expérimentales objectives et répliquables appuyées sur des données « de performance »

« all data about language come from performance », p.290

et ne peuvent pas être « testés » empiriquement.

« No one has run experiments designed to test any of these frameworks ; it would not be possible to do so » p. 280

Ces remarques appellent divers commentaires.

En premier lieu, on voit comment cette affirmation assimile test à expérimentation et suppose que seules les techniques expérimentales quantifiables peuvent constituer des données pertinentes susceptibles de falsifier les modèles phonologiques.

En l'occurrence, la remarque de Pierrehumbert et al. semble signifier que le seul moyen de falsifier sérieusement un modèle phonologique suppose qu'on ait



recours à des données quantifiables obtenues par mesure des informations substantielles (acoustiques, articulatoires en production) ou des comportements des sujets (taux d'erreurs, temps de réaction). Or, on objectera, que pour de nombreuses sciences, y compris des science dites « dures », tout test n'est pas obligatoirement une expérimentation reposant sur un dispositif technique fournissant des données quantitatives et donc qu'on ne voit pas pourquoi la phonologie tomberait sous le coup d'une mathématisation imposant prioritairement le recours à des données quantifiables et reléguerait *a priori* les jugements de grammaticalité au rang d'opinions non pertinentes scientifiquement. On pourra en effet arguer que, de Panini aux grecs du IV<sup>ème</sup> siècle av. JC, en passant par les grammairiens arabes (Bohas et al., 1989) et les propositions millénaires des linguistes chinois (Casacchia, 1989), la manifestation humaine, universelle et fort ancienne de l'activité et de la conscience grammaticales et, en particulier, dans toutes ces traditions d'une conscience phonologique autonome et structurée, ayant donné lieu entre autres à l'invention de l'écriture, rend le jugement de grammaticalité irréductible à une opinion (Milner, 1989), qui serait exclusivement sociologiquement conditionnée par la sensibilité à une norme.

Carr (2000), dans le même volume, répond d'ailleurs très clairement à cette objection de Pierrehumbert et al. en soulignant que la dimension problématique de la théorie de la compétence proposée par les modèles de phonologie générative ne concerne pas les modes de validation retenus mais le réalisme des modèles

« If one adopt a Popper's philosophy of science, the scientific status of hypotheses in generative phonology depends on their falsifiability *in principle*, not on the mode of falsification. Attempts at falsification (testing) of hypotheses may, but need not involve laboratory techniques [...]. A central problem with generative phonology, with respect to its scientific status is not the testability of its theory but the application of scientific realism to linguistic investigation by Chomsky and thus by generative phonology.»

*Phonological knowledge, conceptual and empirical issues*, p.75

Mais, contrairement à ce que propose Carr, je pense que la mise en cause de la pertinence du mode de validation / falsification constitue un argument indirect mais

très efficace contre la conception de la compétence phonologique retenue par les modèles formels considérés et qui met directement en cause leur réalisme et ce faisant leur légitimité scientifique. L'implicite qui sous-tend ce chapitre de Pierrehumbert et al. est que, dans le cadre théorique génératif, les données de performance devraient obligatoirement, comme nous l'avons explicité plus haut, valider les modèles de la compétence (même si d'autres données comme des jugements de grammaticalité peuvent venir à l'appui du modèle). Or, si le jugement de grammaticalité est le mode de falsification unique retenu et qu'il ne peut en aucun cas équivaloir à une « donnée de performance » et constituer une mesure authentiquement scientifique (objective, constante, répliquable), susceptible de falsifier ou de valider la pertinence psychologique des modélisations proposées par la phonologie générative, ces modélisations perdent du coup toute légitimité scientifique et toute raison d'être, puisque le cadre épistémologique qui les légitime nécessite qu'elles soient psychologiquement réalistes.

Contrairement à ce que dit Carr (2000) les deux questions sont donc logiquement liées par la théorie générative elle-même. Le réalisme des modèles et par conséquent leur légitimité scientifique sont directement mises en cause par l'impossibilité d'une validation par les données de performance, la pertinence intrinsèque des jugements de grammaticalité est secondaire dans ce débat.

Etant non testable par des « expériences » et des « données de performance », la théorie de la compétence proposée est donc considérée comme automatiquement non réaliste, exprimant dès lors des fantaisies théoriques, n'illustrant rien d'autre que l'imagination et le bon plaisir des phonologues et relevant à ce titre d'une épistémologie relativiste kuhnienne, réinterprétée de manière disons quelque peu péjorative par Pierrehumbert et al.

« Much work by relativists, such as Kuhn and Feyerabend, leaves the impression that shifts in scientific thinking are arbitrary outcomes of individual taste and power struggles within the scientific community. Espousal of kuhnian thought has done much to glamorize conceptual upheavals within linguistics.»

p. 275

Reste néanmoins que, comme le reconnaissent d'ailleurs eux-mêmes les auteurs, la confiance aveugle en un positivisme scientifique, qui progresserait par cumul linéaire et trouverait dans les données et les expériences produisant des mesures quantitatives et quantifiables, une manière, plus objective et plus honnête<sup>32</sup> de dépasser exhaustivement l'état donné d'un modèle par l'état suivant, relève d'une naïveté assez confondante, qui fait peu de cas de la complexité de la mise en relation de la phénoménologie observée et du modèle qu'on en fait, et accorde au final peu de crédit à la créativité et à l'audace intellectuelles dans les sciences, revendiquées par l'épistémologie kuhnienne.

En ce qui concerne la phonologie en particulier, nombre d'hypothèses dont la pertinence psychologique a été mise par la suite au jour par le recours à des modes de validation quantifiables avaient été formulées bien antérieurement sur la base des jugements de grammaticalité d'un nombre restreint de locuteurs et de l'intuition du linguiste, indépendamment de toute validation quantifiable et objective.

On en voudra pour exemple les fameuses remarques de Sapir (1933), concernant la réalité psychologique des phonèmes et établies sur la base du jugement de locuteur. Il relate en effet comment des informateurs de nootka, païute, sarcee sont capables, de manière pertinente, d'attribuer des interprétations phonologiques différentes à des formes phonétiques identiques, et, de manifester en particulier une sensibilité aux segments latents, n'ayant donc pas de réalisation substantielle, sur la base de leur seule intuition de locuteur, et sans connaissance orthographique des formes concernées. Ceci l'amène à soutenir 1- que les unités phonémiques ont une réalité psychologique 2- que les locuteurs ont une intuition très sûre des représentations phonologiques de leur langue 3- qu'ils s'en servent lors de tâches psycholinguistiques.

Ces affirmations furent validées, plus de 50 ans après par des dizaines d'expériences réalisées sur des centaines de sujets, et mettant en jeu des paradigmes divers avant de devenir l'objet d'un méta-discours psycholinguistique abondant sur les biais expérimentaux observés et les stratégies déployées par les sujets testés sur la tâche

---

<sup>32</sup> car c'est ultimement en ces termes moralisateurs que la question est posée.

spécifique de détection de phonèmes et les problèmes méthodologiques, théoriques et psychologiques qu'elle pose (Dupoux, 1989).

Il n'est donc pas sûr que la deuxième méthode de validation soit la plus satisfaisante pour clarifier la question de la réalité cognitive des représentations, où la tâche expérimentale utilisée devient l'objet d'interrogation principal, au détriment de l'hypothèse cognitive initialement testée. Il est en tout cas assuré qu'en l'occurrence, elle est nettement plus coûteuse que l'intuition de Sapir ou le recours au jugement du locuteur et qu'il paraît plus judicieux de convoquer, pour juger de la pertinence des propositions faites ici et là, les critères de « fécondité » vs « stérilité » théoriques proposés par Lakatos (1970).

Par ailleurs, on a pu observer la validation psycholinguistique quantifiée, objectivable et répliquable d'effets apportant crédit à l'existence de constituants phonologiques et qui se sont avérés des artefacts expérimentaux.

Les fameuses expériences sur le rôle de la syllabe dans l'accès au lexique (Mehler et al, 1981) ont amené ces chercheurs à proposer la syllabe comme unité de segmentation et d'accès au lexique universelle, réinterprétée ensuite (Cutler et al. 1992) en unité de segmentation et d'accès au lexique en français. L'hypothèse de segmentation proposée par cette hypothèse syllabique n'est pas tenable au regard des phénomènes de liaison enchaînée et d'enchaînement. Si les francophones développaient une stratégie de segmentation alignée sur les frontières syllabiques, ils segmenteraient systématiquement de manière erronée les suites contenant une liaison ou un enchaînement (Wauquier-Gravelines, 1996, 1999). Or, je pense qu'il ne serait, pour autant, venu à l'idée de personne de remettre en cause l'existence de la syllabe phonologique à partir de ces données « de performance » pourtant objectivables et quantitatives, montrant que les locuteurs du français ne s'en servent pas pour segmenter la chaîne parlée.

Enfin, cette attitude semble reposer ultimement sur l'assurance que les données (donc la phénoménologie observable) font « sens » en soi, indépendamment d'une théorisation qui les rendrait pertinentes. Or, il n'y a pas de données qui fassent sens, éclairent les régularités d'un système linguistique ou les fonctionnements

mentaux d'un sujet parlant sans modèle et du système linguistique et du sujet parlant. De même, contrairement à ce que prétendent Pierrehumbert et al., il n'existe pas d'expérimentations brutes ou de mesures acoustiques, ou statistiques qui éclairent un phénomène phonologique, sans hypothèse phonologique théorique concernant ce phénomène. Ce faisant, ces auteurs se réclamant de Popper le ramènent selon moi à une version singulièrement simplifiée : Popper (1973) propose un mode de falsification simultanément montant (data driven) et descendant (theory driven). Le degré exact d'abstraction et de formalisme n'est pas en cause ici, mais pour que le phonologue de laboratoire ait l'idée de mesurer telle ou telle différence acoustique ou de tester des stratégies de traitement dont il prédit qu'elles seront différentes, encore faut-il qu'il sache pourquoi il le fait et de quelle régularité phonologique ces différences acoustiques ou comportementales seront le reflet.

Par ailleurs les propositions phonologiques qui motiveront ces mesures relèvent de choix théoriques plus ou moins explicites, qui ne procèdent pas seulement des données, ni de la validité intrinsèque à la théorie mais également d'une préférence épistémologique et de « themata<sup>33</sup>».

Les divers arguments de Pierrehumbert et al. interrogeant la conception de la compétence retenue par les modèles phonologiques génératifs par la mise en cause des modes de validation qu'ils utilisent, me semblent donc peu éclairants quant aux difficultés que pose la formulation du réalisme cognitif de la compétence dans son rapport à la performance.

Il ne m'apparaît pas malheureusement que la phonologie de laboratoire, même si elle apporte un type de données tout à fait indispensables et permettant d'éclairer différemment la question du réalisme cognitif des représentations phonologiques

---

<sup>33</sup>« L'exemple de la physique, étudié par G. Holton, est éclairant : à chaque étape marquante de son progrès, interviennent des préférences non seulement indépendantes d'une expérience possible, mais indépendantes même de tout projet scientifique. Ainsi, qu'une théorie physique donnée préfère des lois fondées sur l'atomisme à des lois fondées sur la continuité, qu'elle préfère des lois fondées sur la simplicité à des lois fondées sur la diversité, cela, bien évidemment, ne saurait dépendre d'une démonstration empirique ; bien au contraire, ce sont de tels choix qui détermineront par avance ce que la théorie recevra comme démonstration empirique valide, comme indice concluant, comme hypothèse plausible - ou ce qu'elle rejettera comme sophisme, comme exception marginale, comme fantaisie débridée. De plus la préférence qui dicte ces choix relève de tout autre chose que de l'idée de la science en général ; elle peut se retrouver, semblable à elle-même, dans une littérature, dans une esthétique, et, bien plus souvent qu'on ne le dit dans des délires »

J.C. Milner, *Introduction à une science du langage*, 1989, p.30-31.

puisse être une réponse au défi mentaliste chomskyen, plus légitimement scientifique *a priori* que les modèles phonologiques génératifs.

Les arguments développés là pèchent, et c'est un défaut qu'on pourrait pointer souvent dans les formulations empiristes de la question me semble-t-il, par une simplification des enjeux épistémologiques qu'impose la mise en relation de la phénoménologie phonologique et du modèle qu'on en fait. Ils supposent une utilisation simplificatrice de l'épistémologie poppérienne. Ceci trahit une foi, certes appuyée sur la certitude tout à fait légitime qu'une démarche authentiquement scientifique doit se donner les moyens de la mathématisation de ses intuitions. Mais cette foi quand même positiviste et actuellement très en vogue, postule que leur caractère mesurable et quantitatif (statistiques, fréquence etc) assure à certains modes de validation, indépendamment de leur pertinence et de leur adéquation théoriques, les moyens absolus et définitifs de toute mathématisation, une sorte d'objectivité supérieure, qui les légitime *ipso facto*. Il paraît au final un peu naïf de considérer qu'un mode de validation plus « réaliste » devient automatiquement garant du réalisme du modèle testé.

A l'inverse, que les modélisations phonologiques génératives ne puissent pas être directement validées par les données de performance ne constitue pas une preuve en soi qu'elles ne sont pas réalistes, qu'elles sont fausses ou illégitimes ou non scientifiques (Carr, 2000).

Ces faits sont pour moi la preuve du biais épistémologique que constitue le postulat de l'ambiguïté de la grammaire. Comme je l'ai développé plus haut, ces théories doivent être formulées à des niveaux de réalité différents et ce mode de validation n'est pas adéquat aux théories concernées.

Ceci n'invalide par conséquent, ni l'immense intérêt des apports récents et tout à fait fondamentaux de la psycholinguistique (données et méthodes) à la question cognitive, ni par ailleurs la possibilité d'une conception abstraite de la phonologie. Elle signifie, plutôt, selon moi que la dichotomie compétence *vs* performance telle qu'elle est formulée par Chomsky et reprise depuis par la doxa chomskyenne, est inadéquate, et que, par conséquent, elle fausse intrinsèquement la formulation de la question du

réalisme du modèle et la question de la réalité psychologique des modélisations proposées par la phonologie théorique.

### **3.3. Le modèle de la « compétence / performance » comme un rempart antibehaviouriste ?**

A partir des années 1950, le behaviourisme quoique encore socialement dominant est un paradigme déjà scientifiquement dépassé et, dans sa formulation skinnérienne idéologiquement très douteuse, attaqué de toute part par les psychologues, en particulier à Cambridge (Harris, 1993). Chomsky va, par la brutalité intellectuelle qu'il va mettre en œuvre dans le débat, jouer un rôle central dans le discrédit définitif qui va porter sur le behaviourisme skinnérien. Comme le signale Harris (1993), le coup d'envoi fut donné le 11 septembre ( !) 1956, lors d'une conférence organisée à Harvard mais le coup de grâce est assené sans merci dans le très fameux article intitulé « Review of B.F. Skinner, Verbal behavior » paru dans *Language*, en 1959. Les limites de cet article polémique ont été pointées ici ou là (MacCorquodale, 1973 ; Seuren, 1998 ; Harris, 1993), tant il est vrai qu'il se pose comme une véritable machine de guerre dirigée contre la version skinnerienne du behaviourisme et Skinner lui-même, qui s'en trouvent à partir de là définitivement déconsidérés. La dimension idéologique et la violence intellectuelle qui caractérisent ce débat ont été largement commentées (Harris 1993).

Il n'est pas question ici de revenir sur cette question, la pensée skinnérienne comporte effectivement des relents idéologiques peu appétissants et formule les questions cognitives en linguistique (en particulier la question de l'apprentissage / acquisition) de manière particulièrement pauvre, mais je voudrais pourtant rapidement soulever le problème suivant. Il semblerait que la mise à mort de Skinner, non vraiment problématique en soi, tienne quand même un peu du meurtre inaugural. Bruner, commentant « Review of B.F. Skinner, Verbal behavior » dit

« Noam at his best, mercilessly out for the kill, daring, brilliant, on the side of the angels, in the same category as St Georges slaying the dragon ».

J. Bruner, *In search of Mind*, 1983, p.159

Or la violence de l'attaque portée, en caricaturant et donc en discréditant définitivement le behaviourisme met en place un tabou puissant en phonologie générative : toute théorisation linguistique qui intégrerait une réflexion sur la dimension comportementale du langage et qui plus généralement, soulève la question de l'usage, du rôle de l'environnement linguistique et pragmatique, de la variation, en particulier pour ce qui concerne l'acquisition est immédiatement susceptible d'être entachée d'infamie. Chomsky place donc, me semble-t-il, avec cette posture inaugurale, le projet scientifique générativiste dans une perspective non seulement anti-behaviouriste mais également et tout aussi radicalement anti-comportementaliste.

Jackendoff (2000) suggère d'ailleurs que la définition de la compétence trouve son ancrage dans l'anti-behaviourisme de Chomsky. Celui-ci opère assez pragmatiquement selon Jackendoff, en définissant comme il le fait la « compétence », une récupération du travail opéré depuis longtemps par la tradition linguistique et ayant pour but exclusif la description des données. Mais Jackendoff souligne également qu'il cherche d'abord, en lui assignant par le recours au mentalisme une visée explicative, à séparer radicalement la linguistique du behaviorisme et de tout comportementalisme, et en particulier des lectures behavioristes du langage et du mentalisme linguistique proposées par Hockett (1958) et Bloomfield.

« Chomsky makes the competence-performance distinction in part to ward-off alternative proposals for how linguistics must be studied. In particular, he is justifiably resisting the behaviorists [...]. His resistance was particularly urgent at that time because of the overwhelming influence of the behaviorist stance on American structuralist linguistics of the period »

R. Jackendoff, 2002, *Foundations of Language*, p. 30



Chomsky commentant le « mécanisme » bloomfieldien dans *Aspects* [193, 194], rappelle la critique méthodologique que Twaddell fait de Sapir<sup>34</sup> et qui met en cause son recours systématique aux jugements des locuteurs pour établir la réalité psychologique de certains phénomènes phonologiques.

« La position behavioriste n'est pas une question dont on puisse discuter. Elle exprime simplement un manque d'intérêt pour la théorie et l'explication. Cela ressort clairement, par exemple, de la critique faite par Twadell (1935)<sup>35</sup> de la phonologie mentaliste de Sapir, qui employait les réponses et commentaires d'informateur comme des témoignages concernant la réalité psychologique d'un système abstrait d'éléments phonologiques. Pour Twaddell, le projet est sans valeur parce que ce qui l'intéresse est le comportement lui-même « qui se trouve déjà à la disposition de celui qui étudie le langage, bien que ce soit sous une forme moins concentrée ». De manière caractéristique, ce manque d'intérêt pour la théorie linguistique se traduit par la proposition de limiter le terme de « théorie » à un « résumé de données ».

*Aspects*, [193-194], p. 14

La théorie de la compétence *objet mental abstrait* y est présentée comme l'alternative théorique explicative unique, dépassant définitivement la conception behavioriste, définie ici par défaut (c'est à dire ce qu'elle n'est pas, mais supposant *de facto* que l'étude basée sur le comportement n'a aucune visée explicative, alors que de postuler que la théorie faite par le linguiste est isomorphe de ce que contient le cerveau du locuteur l'est). Au-delà de cette formulation spécifique, on voit comment est rejetée, comme étant *de facto* non explicative, toute étude linguistique n'intégrant pas le postulat réaliste et s'appuyant sur la description des données et l'observation des comportements linguistiques et des usages réels du locuteur.

Ce faisant, il apparaît effectivement que le recours au postulat de l'ambiguïté de la grammaire opère un glissement produisant l'argument ultime qui fonde la légitimité d'une conception non-comportementaliste du langage et permet de l'évacuer définitivement hors du domaine d'intérêt de la linguistique. La manière dont

---

<sup>34</sup> Où l'on voit qu'une fois encore, c'est par le débat concernant les modes de validation / falsification que la question est abordée.

<sup>35</sup> Twaddell est référencé (1935) dans les texte cité mais 1934 dans la bibliographie finale de l'édition française ed *Aspects*.

Chomsky reformule le dualisme cartésien par l'élaboration de la théorie de la « compétence » opère, de manière plutôt inattendue dans un modèle dualiste, une transformation de l'objet envisagé : le langage observé en quelque sorte dans ses usages linguistiques réels devient un objet non substantiel complètement idéalisé. Ce faisant, la compétence échappant à l'usage, se pose comme rempart radical à toute proposition qui glisserait vers des élucidations procédurale ou comportementaliste du langage, reléguées au rayon de la performance et proposé à la perspicacité exclusive des psychologues.

Outre la fragilité intrinsèque de l'argument mentaliste tel qu'il est formulé là (cf. section 2), ceci me semble paradoxal pour une conception réaliste où le langage est désigné explicitement comme relevant de la psychologie et où est visé l'établissement d'un modèle psychologique et biologique interne réaliste du locuteur. Car on ne peut, de fait, pas exclure absolument le fait que le langage, dans sa dimension psychologique, tant du point de vue du locuteur réel que du point de vue de l'observation empirique que l'on peut en faire, est aussi un comportement variable, et n'est observable qu'en tant que tel, dans l'usage et les usages variés qu'en font les locuteurs-auditeurs réels aussi bien en production qu'en perception. L'établissement de la dichotomie compétence *vs* performance réalise plus ou moins adroitement une récupération de la dimension comportementale du langage et de la question de l'usage, des usages, mais en la reléguant arbitrairement et définitivement hors de la théorie grammaticale et hors du champ d'action des linguistes.

La phonologie variationniste (Docherty et al. 1997, Encrevé, 1988) a intégré dans la formulation de ses hypothèses théoriques une réflexion sur l'usage et la variation. Dans ces études ainsi que dans les travaux de Paradis et al. sur l'adaptation des emprunts, une attention particulière est portée aux données, à la fréquence des structures produites par le sujets à partir de corpus quantitatifs et statistiques, de même qu'aux variations sociologiquement conditionnées qui sont observables. Les propositions faites par les phonologues variationnistes assouplissent certes la rigueur des modèles formels établissant leurs généralisations sur des formes attestées en tout petit nombre ou normalisées ou standardisées. En ce sens on peut penser que

la couverture empirique réalisée par la phonologie variationniste est meilleure que celle qui est réalisée par les modèles génératifs classiques.

D'autre part, contrairement à ce que dit Carr (2000), même si les propositions faites par les variationnistes restent « theory-driven », la prise en compte des facteurs de variation influence la théorisation du phénomène considéré et réintègre par conséquent la question de l'usage au sein de la grammaire, dans une démarche « data-driven ». L'analyse du non-enchaînement et de la liaison non-enchaînée par Encrevé (1988), l'amène certes à couvrir une phénoménologie plus large, sociologiquement conditionnée à l'évidence, donc à opérer une meilleure description du phénomène mais lui permet aussi de poser l'existence d'une coda flottante et surtout de réenvisager frontalement la question du flottement des positions squelettales qui constitue un enjeu théorique central dans la phonologie autosegmentale et les interrogations qu'elle suscite quant à l'organisation des tires segmentales et syllabiques.

Par ailleurs, j'écarterai la critique de Carr (2000) qui minore l'intérêt de l'approche variationniste en mettant en doute la scientificité des critères sociologiques qui conditionnent le tri des données (« middle-class » par exemple, ne lui paraît n'avoir aucun caractère scientifique intrinsèque et être en plus biaisée par l'appartenance sociale du linguiste adoptant de tels critères de sélection des données). Certaines versions de la phonologie variationniste (Encrevé, 1988) s'inscrivant dans la démarche labovienne ont articulé délibérément leurs hypothèses linguistiques sur une sociologie bourdieusienne très théorisée pourvoyeuse d'un discours théorique et méta-théorique supposant une réflexivité scientifique indéniable et appuyée sur des méthodes absolument scientifiques (enquêtes, traitement quantitatif et qualitatif des données d'enquête). Il me semble qu'elle assure ce faisant, pleinement la fonction de « vigilance épistémologique » qui est sollicitée par le linguiste (Encrevé 1988).

Le problème ne me semble pas par conséquent se poser là, mais consiste en ce que la phonologie variationniste ne se saisit pas vraiment frontalement de la question cognitive, à l'intérieur du cadre génératif où ce problème est pourtant central. La variation est envisagée d'abord et surtout en termes sociologiques, mais sans articulation directe et explicite à une théorie cognitive et aux problèmes

psycholinguistiques que soulèvent l'étude des usages (i) dimension comportementale des usages et ii) réalisme du modèle intégrant les usages).

La question de l'utilité de modéliser la dimension comportementale et pragmatique des usages en termes cognitifs dans la grammaire elle-même (via le recours à la notion de marque par exemple) n'est me semble-t-il jamais posée.

La question du réalisme du modèle a été soulevée, comme je l'ai développé dans ma thèse par Encrevé (1988 p. 153 et sq.). Il justifie en effet psychologiquement le choix du formalisme autosegmental purement représentationnel retenu (Kaye & Lowenstamm, 1984) et le choix du squelette de positions pures. Il avance effectivement, s'appuyant sur Sapir (1933) que la représentation qu'un locuteur a d'une unité lexicale pourrait constituer en fait une connaissance intuitive du nombre d'unités possibles que le locuteur attribue à un mot. Le mot serait donc encodé mentalement sous la forme d'un certain nombre d'unités temporelles linéarisées

« [...], nous croyons utile de poser l'hypothèse que l'encodage mental du lexique comprend un paramètre (dont la matérialisation n'est pas notre objet) qui correspond à la mémorisation du nombre de « places » que peut comporter un élément lexical »

P. Encrevé, *La liaison avec et sans enchaînement, Phonologie tridimensionnelle et usages du français*, 1988, p.154.

Il suggère ainsi que le squelette pourrait être une matrice mentale, un gabarit temporel de stockage et de traitement lexical matérialisant l'une des connaissances basiques qu'un locuteur utilise en perception et en production : la parole est un objet linéaire, temporel, et rythmé : l'ordre de succession des unités dans la chaîne parlée n'est pas indifférent, il est conditionné par la temporalité et les relations de hiérarchie horizontale qui s'établissent entre les constituants et à l'intérieur des constituants.

On voit s'esquisser là sur deux pages, une possibilité d'articulation entre formalisme représentationnel, variation et psycholinguistique qui malheureusement a été peu exploité. Carvalho (1999, 2002) reprend l'hypothèse d'un niveau phonologique subsymbolique constitué de positions temporelles pures non interprétées

segmentalement et qui constituerait la base de la structure phonologique des énoncés bien formés à partir desquelles émergerait l'interprétation segmentale et syllabique par les contraintes de mise en chaîne, mais il ne formule aucune hypothèse cognitive explicite à ce sujet. J'ai néanmoins, dans cet esprit, repris cette idée de l'indépendance du squelette et de l'interprétation segmentale qui en est faite pour rendre compte de l'acquisition de la liaison (Ch.I et II). J'ai aussi retenu l'idée d'un gabarit de positions autosegmentales pures où les mélodies consonantiques et vocaliques émergent séparément et procèdent l'une de l'autre pour rendre compte de phénomènes observés lors de l'acquisition phonologique en français et sur les données de dysphasie des « enfants sans consonnes » (Le Normand et al. 1991). Les données des enfants reflètent en effet une apparente « connaissance » du nombre et de la place d'unités (en particulier des consonnes non encore produites) émergeant à partir des noyaux vocaliques.

Mais, de manière générale, il me paraît dommageable qu'Encrevé (1988) n'ait pas été lu sous cet angle psycholinguistique par les phonologies représentationnelles.

Pour conclure sur ce point, il me semble que la GG s'étant établie de manière polémique, et à juste titre, comme radicalement anti-skinnérienne en 1956, se retrouve de fait, encombrée depuis, de la dimension comportementale et variable du langage et réponde mal à la question de l'usage ou des usages linguistiques.

Je suggérerai qu'on puisse trouver là, l'origine, à la fois de la non-pertinence de l'opposition entre compétence et performance, et de l'échec du programme cognitif unifié qui les articulait originellement l'une à l'autre. J'y vois également les raisons de la résurgence actuelle sous couvert de comportementalisme d'un discours simplifiant à l'extrême la question de l'usage, ramené à ses pures dimensions quantitatives (cf. discussion de cette question Ch.III, 4.). La négligence apportée à la question psycholinguistique du comportement des locuteurs réels, ainsi qu'à une formulation phonologique rigoureuse de l'usage, des usages réels par des locuteurs réels exposait sans doute au retour de balancier.

### 3.4. Faillite du modèle compétence vs performance

En conclusion, il m'apparaît que la dichotomie compétence vs performance, si profondément intériorisée et acceptée par la doxa générativiste, phonologues et psycholinguistes confondus, est obsolète et peu opératoire pour rendre compte du réalisme cognitif en phonologie, pour les raisons que j'ai développées ci-dessus.

Les notions et surtout la définition des rapports qu'elles entretiennent sont dès leur origine formulées de manière floue et peu susceptible de produire un modèle phonologique cognitif unifié et réaliste. Le partage des tâches tel qu'il est organisé par la mise en place de cette dichotomie a d'ailleurs pour l'instant échoué

- 1) à faire converger la théorie de la grammaire interne (I-Language) et la théorie des usages qu'on en fait (E-language). On pourrait même objecter qu'il les a, au contraire, condamnés à être fondamentalement incompatibles
- 2) à contribuer à l'unification de modèles de perception et de production

Est mis à mal et largement compromis me semble-t-il, l'horizon espéré au début des années 70, l'idéal d'une articulation harmonieuse et réciproque de la psycholinguistique et de la phonologie théorique. Malgré l'échec de la DTC en syntaxe, nombre de phonologues et nombre de psycholinguistes cognitivistes dont Mehler (Bonnatti, 2002), Segui (1989) ont défendu ce point de vue et incarné cette attente. Tous bénéficiaient alors d'une situation épistémologique dominante où aucune proposition alternative n'offrait de résultats empiriques, ni d'arguments théoriques assez consistants susceptibles de la remettre en cause. On peut lire dans Piatelli-Palmarini (2002) rendant hommage à Jacques Mehler (Dupoux, 2002), la nostalgie d'un temps où les travaux en cognition se déployaient au sein d' « modèle net, cohérent, passionnant dans lequel [nous étions] installés depuis longtemps et que j'appellerai sciences cognitives « classiques »

Force nous est de constater, à quel point la situation s'est complexifiée et pour nous rapprocher davantage de notre objet, pour ce qui concerne le traitement de la

parole continue, de telle manière, qu'à l'heure actuelle, les deux domaines ont développé des modèles qui rendent compte de données empiriques de manière autonome et en se passant *de facto* des apports de l'autre discipline.

#### **4. De l'opposition compétence vs performance à la question du statut des représentations<sup>36</sup>.**

##### **4.1. Une phonologie exclusivement procédurale est-elle cognitivement plus pertinente?**

Je voudrais faire ici quelques remarques à ce sujet.

###### 4.1.1. Phonologie et fréquence

Je voudrais, dans un premier temps soulever la question du degré d'intégration des informations contextuelles dans la construction des représentations phonologiques que suppose un modèle exemplariste comme celui de Bybee (2001). Si la grammaire et donc la phonologie sont un système de traitement non spécifique, qui catégorise des objets physiques non spécifiques, complètement conditionnés par le contexte, à quoi doit-on limiter l'information pertinente en contexte qui donne lieu à l'émergence des représentations ?

L'usage est envisagé par Bybee exclusivement de manière quantitative et jamais qualitative, ni pragmatique, ni structurelle. Il est ramené à la fréquence d'occurrence pure des objets physiques dans le signal (« token-frequency ») et des structures possibles (par exemple le morphème dérivationnel-ed pour passé en anglais : « type-frequency ») que le sujet va percevoir et utiliser en production. Ce qui suppose qu'il y a une corrélation directe entre la forme de l'objet (le son « ed ») ou de la structure dans laquelle il apparaît (la dérivation du passé) et le nombre de fois où il a donné lieu à une expérience sensorielle de perception ou de production, indépendamment de ses conditions de réalisation, de ses caractéristiques structurelles intrinsèques et des relations qu'il entretient avec d'autres objets de même nature.

---

<sup>36</sup> Cette partie présente un état de réflexion en cours

Ceci suppose donc, car Bybee (2001) n'en dit rien, que toutes les expériences sont équivalentes et que toutes les occurrences d'un « token » ou d'un « type » et même que n'importe quelle expérience sensorielle, toute modalité confondues vont avoir un poids égal dans l'émergence et la réorganisation de la représentation.

Car la logique du modèle, qui est à la fois non spécifique et fonctionnaliste<sup>37</sup> suppose justement une dimension pragmatique, voire psychique implicite dans l'usage des mots et des structures et une influence du contexte dans toutes ses dimensions sensorielles et physiques sur le remodelage de la représentation. La logique du modèle proposé par Bybee nécessite que ces aspects soient modélisés aussi et n'exclut aucunement qu'ils ne puissent jouer aucun rôle dans la constitution des représentations.

La logique de ce modèle nous montre bien comment cette conception de la phonologie –et de la grammaire- repose sur un empirisme excessivement réducteur qui simplifie à l'extrême le rapport entre la phénoménologie et la compréhension qu'on peut en avoir, puisque l'usage est ramené à sa pure dimension quantitative . Que la fréquence, qui est un phénomène effectivement très robuste en psycholinguistique et en particulier en accès lexical, et ce quelle que soit la tâche, ait un effet sur le traitement ou conditionne la structuration des représentations phonologiques abstraites, on peut légitimement le penser. De même qu'on peut également légitimement penser que les modèles phonologiques rationalistes ont négligé, à tort, ces questions liées à l'usage, à la fréquence et à la dimension comportementale du langage dans la construction de leurs hypothèses. Mais la formulation du comportementalisme et le rôle nécessaire et suffisant que la phonologie des exemplaires assigne à la fréquence quantitative me semblent réducteurs puisque le langage est ramené à un comportement objectivable *hic et nunc*, qui remodèle la grammaire en fonction des nécessités de l'adaptation optimale du locuteur aux situations sociales d'échanges verbaux. La boîte noire peut alors être envisagée comme une ressource mémoire illimitée nourrissant des systèmes de traitement non spécifiques qui sont exclusivement conditionnés par une capacité de catégorisation en contexte, assez grossièrement définie et non spécifique. Le langage n'est jamais

---



envisagé comme un système, au sens saussurien du terme, à l'intérieur duquel les objets entretiennent des relations d'oppositions et de constituance. Or, il me semble qu'un locuteur du français connaît effectivement les quatre mots suivants « pull » [pyl], « bulle » [byl], « Paul » [pɔl], « bol » [bɔl], et qu'il peut les avoir stockés avec toutes leurs variantes phonétiques, contextuelles, pragmatiques, etc mais il sait également, sans l'avoir jamais appris, que structurellement « pull » est à « bulle » ce que « Paul » est à « bol » et que sur la base des catégories suivantes [p] [y] [l] [ɔ] [b], il peut construire la suite [byplɔb] et non la suite \*[lpybl] alors que ni l'une ni l'autre n'appartiennent au lexique. Les données d'acquisition, comme les données d'aphasie en production témoignent très clairement de la manifestation d'une logique systémique lors de la mise en place (cf. ch II) ou de la dégradation des représentations phonologiques (Valdois & Nespoulous, 1994), qui est évidemment à rapprocher de la question de la marque phonologique.

Cette logique systémique est également à l'œuvre en diachronie. La logique proposée par les modèles exemplaristes prédit en effet que la productivité lexicale et morpho-phonologique reflète directement les fréquences d'usage

« Since tokens of use map onto existing representations, high-frequency items grow strong and therefore are easier to access. The other side of the coin is that little-used items will tend to fade in representational strength and grow more difficult to access. The conserving effect of high token frequency, which protects high frequency items from regularization on the basis of productive schemas, is represented as lexical strength »

Joan Bybee, *Phonology and Language use*, 2001, p. 28.

Ceci prédit donc que, de manière générale, lors de l'évolution d'un système linguistique, tous les phénomènes diachroniques vont tendre à privilégier les mots, morphèmes, structures etc les plus fréquents, indépendamment de toute logique systémique. Or ce n'est pas le cas. J'illustrerai ceci rapidement avec deux exemples, l'un en français, l'autre en portugais (Carvalho, 2004).

Désirat et Hordé (1988) dans un chapitre concernant l'étude du mouvement lexical en diachronie du français comparent la productivité et la disponibilité des

suffixes verbaux. Deux phénomènes régulièrement attestés par les études détaillées qu'ils présentent (-age / -ement, -on, -oir, -erie, -tion, -isme) infirment les prédictions faites par un modèle exemplariste

Ils montrent d'abord que

« les mouvements ne concernent pas des termes isolés mais des séries suffixales : la progression de suffixes disponibles (-ique, -ien) ou la régression de certains d'entre eux (-eux, -oir) entraînent des rééquilibrations internes du système suffixal » p. 165

ce qui témoigne d'une logique d'évolution systémique car il n'y a aucune raison que tous les noms dérivés dans une langue avec le même suffixe soient de fréquence équivalente (ex : pansage / lavage).

Ensuite, étudiant l'évolution diachronique des suffixes du français, ils montrent que certains d'entre eux peu productifs ont régressé puis ont définitivement disparu comme « ons » permettant de construire une locution adverbiale sur un verbe qui n'a subsisté que dans l'expression « à tâtons ». Mais ils montrent aussi en comparant l'évolution parallèle des deux suffixes nominaux –age et –ement attestés depuis le XII<sup>ème</sup> siècle, que des suffixes peuvent être productifs à certaines périodes, stériles à d'autres, qu'ils peuvent régresser puis s'expanser, se remettre à dériver de nouvelles séries après avoir perdu toute disponibilité. Qui plus est, quand deux suffixes, tels que –age et –ement sont en concurrence, la productivité de l'un conditionne celle de l'autre et entraîne des redistributions morpho-syntaxiques et sémantiques : leur étude montre bien que la disponibilité des suffixes et la productivité lexicale qui en résulte en diachronie, ne dépendent pas exclusivement, ni même prioritairement de la fréquence d'usage, mais qu'une logique systémique complexe s'exprimant au niveau phonologique (forme de la base), au niveau syntaxique (accompli « pavement » « pansement »/ non accompli « pavage » « pansage », verbe transitif « gavage » « abattage »/ intransitif « errements », « abattement »), au niveau sémantique (« claquage » / « claquement ») est lisible.

Carvalho (2004) donne une autre illustration, phonologique du même argument. Etudiant la dérivation du masculin en « o » des adjectifs et noms paroxytons du portugais standard de Lisbonne, il montre qu'ils se divisent en une classe minoritaire où la dérivation est d'origine « apophonique » et une classe majoritaire où elle est d'origine « analogique ». Un modèle exemplariste prédit que la dérivation analogique doit absorber la dérivation apophonique, pourtant la dérivation statistiquement minoritaire dans la langue est toujours (et vraisemblablement en progrès depuis le XIX<sup>ème</sup> siècle) productive. Ceci illustre également la possibilité d'une inversion de la productivité morpho-phonologique complètement contraire à la logique de fréquence quantitative d'usage et exprimant une logique structurelle (en l'occurrence une logique apophonique, purement phonologique et indépendante même de la structure du lexique).

#### 4.1.2. Une phonologie de la perception ?

Bybee propose que la fréquence qui conditionne la disponibilité des exemplaires reflète la fréquence de perception et de production, mais, l'argument qui consiste à construire des modèles cognitifs prétendument plus réalistes à partir des seules données de performance (Pierrehumbert, 2000), amène dans les faits à établir les modèles procéduraux (modèles à exemplaires, modèles stochastiques, OT) à partir des données de perception ou sur la base d'arguments tirés essentiellement des données de perception (Hume & Johnson, 2001, Steriade, 2001, Boersma, 2000).

La psycholinguistique cognitive montre depuis une quinzaine d'années une fertilité croissante et tout à fait indispensable au champ, tant en termes de résultats expérimentaux que de richesse et variété des modélisations, mais elle s'est affirmée essentiellement sur le versant perceptif. Une phonologie appuyée uniquement sur les données de performance tend donc à privilégier une conception bottom-up de la cognition. Les modèles qui en résultent, à l'exception de OT, rendent compte très aisément des traitements perceptifs, beaucoup moins aisément des productions des locuteurs.

Par ailleurs, on peut arguer que la psycholinguistique expérimentale met en place des procédures de validation qui pourraient favoriser une lecture comportementaliste réductrice des phénomènes linguistiques et donc phonologiques observés.

L'étude du langage en tant qu'il est un comportement réel d'un locuteur réel observable en temps réel est le plus souvent réalisée à partir de tâches « on-line » supposées rendre compte du traitement cognitif, à partir d'échantillons de parole enregistrées de manière contrôlée par l'expérimentateur, sur des populations généralement assez homogènes et supposant un dispositif expérimental parfois lourd qui fait perdre beaucoup de réalité et au locuteur et au contexte (peu de travaux par exemple sont réalisés à partir des erreurs de production spontanées, ou de jugements de grammaticalité, tâches off-line et métalinguistique associées à des tâches on-line). Et les données de perception sont rarement mises en regard de la production spontanée des locuteurs testés. Il ne s'agit pas ici de nier l'intérêt ni la validité de ces dispositifs expérimentaux auxquels j'ai moi-même recours. Les psycholinguistes inventant et utilisant des tâches expérimentales ont biensûr absolument conscience des limites et des possibilités d'observation que leur permettent les dispositifs expérimentaux, ils les choisissent et établissent leurs interprétations en conséquence<sup>38</sup>. Il n'empêche que les locuteurs sont le plus souvent, dans des conditions artificielles, exposés à un seul type de stimulus, soumis à une seule tâche spécifique à la fois, donc typiquement dans une situation de type « stimulus-réponse » complètement recontextualisée artificiellement et peu écologique.

On risque donc de voir ces modèles dresser l'écueil inverse de celui dans lequel échouent les modèles phonologiques génératifs. Les modèles phonologiques génératifs ont quasi exclusivement, proposé une conception « top-down » de la cognition. Ils ont par conséquent construit des modèles exclusivement descendants<sup>39</sup>,

---

<sup>38</sup> On peut citer à titre d'exemple le remarquable travail réalisé par Grosjean & Frauenfelder (1996) dans le volume 11 de *Language and Cognitive Processes* qui examine un par un et de manière critique et très informée les dispositifs expérimentaux disponibles en psycholinguistique de l'oral pour l'étude de l'accès lexical. Ce numéro témoigne d'une vraie réflexivité, et partant d'une tentative d'objectivisation des conditions de production des connaissances qui est vraiment impressionnante (et trop rare, selon moi par exemple en linguistique théorique).

<sup>39</sup> C'est d'ailleurs également le cas de la théorie de l'Optimalité qui est pourtant une théorie largement non représentationnelle.

donc des modèles de production, en délaissant la théorisation de la perception-compréhension. Mais le retour de balancier auquel on assiste aujourd'hui, expose après des années d'hégémonie quasi-exclusive des modèles de production à l'émergence d'une phonologie procédurale quasi-exclusivement conditionnée par la perception et non compatible avec la précédente. Que gagne-t-on là ? Et sommes nous en présence d'hypothèses plus réalistes et plus aptes à permettre la mise en place d'un modèle cognitivement réaliste unifié du langage ?

Pour l'heure, la cohabitation entre l'auditeur-acousticien empiriste et le locuteur-autiste idéal risque d'être problématique à moins d'envisager que les systèmes de production et de perception sont radicalement différents (Boersma, 1999).

Il conviendrait donc, me semble-t-il, avant de congédier les représentations abstraites en phonologie au nom de l'usage, de se demander et d'évaluer plus précisément, si les problématiques soulevées par les modèles exemplaristes ou stochastiques- et tout particulièrement l'importance donnée à la fréquence lexicale- ainsi que la remise en cause radicale des représentations phonologiques abstraites ne sont pas en partie générées par le mode d'observation des données et les méthodes retenues, c'est à dire par le fait que ces théories s'élaborent quasi-exclusivement sur des données psycholinguistiques de perception. Ceci permettrait, je crois, d'optimiser l'apport énorme de la psycholinguistique expérimentale depuis une dizaine d'année concernant la perception-compréhension du langage et les possibilités de reformulation théorique de la dimension cognitive de la grammaire et de la phonologie qu'elle permet (Dupoux, 2002 ; Hume & Johnson, 2001).

## **5. Conclusion**

On pourrait suggérer que Chomsky a ceci de paradoxal qu'il a consacré et redonné pleinement ses lettres de noblesse au rationalisme en linguistique, et donc en phonologie, et qu'il en a simultanément proposé une formulation qui fragilise l'héritage qu'on peut en attendre.

Mais, comme je l'ai écrit explicitement (Wauquier-Gravelines, 2004), l'opinion selon laquelle le paradigme chomskyen en phonologie serait dépassé et que la conception représentationnaliste de la phonologie constituerait définitivement une impasse épistémologique me paraît fausse et inintéressante en soi, d'autant qu'elle s'accompagne du sentiment très déplaisant d'être parfois revenu 50 ans en arrière. Si l'on considère la phonologie comme une science « cumulative » (Encrevé, 1999), on se doit de prendre en considération les apories et contradictions du programme de recherche chomskyen, au fur et à mesure qu'elles se font jour, et les conséquences que cela entraîne pour la phonologie théorique. On se doit aussi, puisque question cognitive et biologique il y a, de confronter les modèles phonologiques avec les modèles et données psycholinguistiques et neurologiques. Mais il convient également, en vertu de la logique de cumul, de rendre justice à la fécondité et à la créativité de ce programme et d'en sédimenter les apports pour faire avancer la question cognitive en phonologie. Les hypothèses génératives constituent une reformulation complète, cohérente, fertile jusque dans ses défauts et dans les débats et les oppositions qu'elle suscite, d'une conception rationaliste du langage et de l'esprit. Je pense qu'il faut, en cela, rendre justice à la pensée chomskyenne, en ce qu'elle a consacré définitivement la place de la question cognitive en linguistique et en a assuré la promotion, le développement institutionnel etc... On peut légitimement parler à ce titre de « tournant chomskyen », qu'on adhère ou non à l'épistémologie présentée.

Montrerait-on enfin définitivement que la formulation chomskyenne d'une phonologie rationaliste est inadéquate, ce qui est loin d'être le cas pour l'instant, qu'on ne mettrait en péril ni le paradigme rationaliste en soi, ni la possibilité d'une conception rationaliste de la phonologie, qui, selon moi, n'est pas actuellement susceptible d'être définitivement remise en cause par un retour de positivisme, discrédité par avance par la naïveté de certaines de ses formulations.

Par ailleurs, si les modèles classiques « génératifs » tels qu'ils sont mis en oeuvre, apparaissent actuellement insuffisants, trop idéalisés, épistémologiquement problématiques pour rendre compte de manière vraiment satisfaisante, de la

dimension cognitive de la phonologie, il ne me semble pas non plus, que, pour l'instant, les contre-propositions qui ont été faites apportent des arguments dirimants permettant de congédier le rationalisme et l'existence d'un niveau de représentation abstrait des objets phonologiques, au profit d'une approche exclusivement substantielle ou « usage-based ». Il ne m'apparaît pas que les réponses proposées dans ces paradigmes font avancer la question de manière décisive ou sont radicalement meilleures et définitivement plus convaincantes quant à la question cognitive (cf. ci-dessus). Il convient, à cet égard, d'observer avec toute la distance critique nécessaire, ce qui faut attribuer à des effets de mode, relevant de la sociologie des sciences (Bourdieu, 1997) et non de la linguistique.

La phonologie, conçue comme « science cognitive » est une science jeune. Les modèles phonologiques classiques sont donc amenés à évoluer et à se réinventer. C'est même, me semble-t-il, la condition de leur survie et complètement conforme à la logique scientifique poppérienne dont ils se réclament. Les tentations en miroir de solipsisme ou au contraire de table rase qui les ramène à l'expression arbitraire de choix individuels « glamourisés » (je cite Pierrehumbert et al., 2000), sans objet et sans conséquence, me paraissent symétriquement stériles pour la discipline.

Le même genre de débats, s'appuyant sur les arguments soulevés par exemple par Carr (2000), Burton-Roberts ou Pierrehumbert (2000), Laks (1996, 2005) entre autres, portant sur la réalité du modèle, la légitimité de la GU, les modes de falsification, l'ambiguïté des formulations chomskyennes ont agité la sémantique générative depuis quelque temps déjà (Harris, 1993), sans pourtant menacer aucunement son ancrage rationaliste (Jackendoff, 2000).

Pourquoi en irait-il autrement de la phonologie?

## CHAPITRE IV : L'ACQUISITION DE LA PHONOLOGIE

### LE BEBE ET L'EAU DU BAIN

« **Innées** (idées). – Les blaguer »

G. Flaubert, Dictionnaire des idées reçues,

« Les enfants viennent s'inscrire dans le monde des hommes pour y prendre la parole qui s'y donne. A cette inscription du vivant humain, le seul génie biologique qui transforme et ouvre le corps de la femme pour livrer passage au nouveau-né ne suffit pas. Il n'y a d'humanité dans ce processus qu'à la lumière de l'ouverture et de la transformation du cœur. Cœur ne veut pas dire, ici, émoi sentimental : il signifie le lieu où *la parole qui fait l'homme* est à l'œuvre, dans le champ du langage où l'homme se réalise en tant que corps social dans son rapport aux autres. »

D. Vasse, La chair envisagée : la génération symbolique

### **1. Le dispositif chomskyen : les limites du bébé cartésien**

En toute cohérence intellectuelle, Chomsky, envisageant la question de l'acquisition du langage, élabore des propositions appuyées sur les assises rationalistes et anti-behaviouristes que nous avons décrites plus haut. Le problème posé consiste donc à rendre compte de l'existence de la connaissance linguistique abstraite, spécifique, modulaire et indépendante des facultés intellectuelles générales, qui ne peut pas avoir été apprise par expérience, entraînement, mimétisme ou nécessité d'adaptation au milieu. En d'autres termes, il s'agit de fonder un dispositif théorique permettant à partir d'un nouveau-né normal d'obtenir le « locuteur-auditeur idéal » insensible à son environnement défini dans *Aspects*.

Le modèle proposé, par Chomsky et depuis ici et là dans la doxa générative est celui d'un bébé cartésien, pur sujet de raison, apprenant désincarné insensible à son environnement, travaillant sur sa grammaire avec les mêmes logiques de



généralisation que le linguiste adulte -à ceci près que pour l'enfant elles sont inconscientes- et qui voit se développer en lui de manière irrépessible des facultés linguistiques présentes de manière interne, donc logiquement innées et déclenchées (« triggering») mais non conditionnées par l'environnement linguistique, intellectuel, social, psycho-affectif.

Le modèle d'acquisition repose sur la théorie dite des « Principes et paramètres» (PP). L'enfant vient au monde avec une grammaire universelle (GU), ensemble de propriétés invariantes définissant de manière exhaustive, nécessaire et suffisante, la forme possible de toute langue humaine présente, passée ou à venir. Les principes matérialisent les propriétés et les opérations universelles représentées dans toute langue naturelle. Les paramètres instancient la valeur particulière que prend chaque principe pour une langue donnée. Acquérir une langue consiste donc à fixer la valeur des paramètres à partir de l'input. C'est une expérience passive et inconsciente, un processus de sélection d'une grammaire par contact avec l'environnement, à partir d'un certain nombre de possibles définis par la GU dont l'enfant est génétiquement doté (Chomsky, 1990).

Comme le souligne Jakubowicz (1995)

« Il est évident que la métaphore de Chomsky ne constitue pas un modèle articulé d'acquisition du langage. Faisant abstraction de la dimension temporelle, notamment du fait que l'acquisition se déroule sur un certain nombre d'années, la théorie PP reste neutre quant à certaines questions centrales pour toute théorie de l'acquisition, comme par exemple le problème de l'existence d'étapes successives de développement linguistique, et, le cas échéant, celui du mécanisme responsable du passage de l'un à l'autre »

C. Jakubowicz, « Grammaire Universelle et acquisition du langage », *RLV* 24, 1995, p.13

Cette remarque n'est pas anodine du tout, car effectivement dans la grammaire générative, le problème de l'acquisition n'est envisagé ni d'un point de vue développemental (quel est l'état initial, par quels stades un enfant doit-il passer pour aller de cet état initial à l'état adulte de maîtrise d'une grammaire et quelle est la logique linguistique et cognitive qu'exprime cette évolution, comment évoluent les représentations), ni d'un point de vue maturationnel (un nouveau-né n'est pas un

adulte en réduction mais un être qui change intrinsèquement et au contact de son milieu tout à la fois d'un point de vue physiologique, cognitif, moteur et psychique). Il est envisagé d'un point de vue essentiellement logique (Christophe, 2002) comme élément de validation de tout le paradigme génératif et en particulier comme mode de validation des modélisations de la compétence adulte. Mais on peut considérer qu'il n'y a pas chez Chomsky ce qu'on pourrait appeler une véritable théorie développementale de l'acquisition

La question de l'innéité des représentations linguistiques (GU) est, depuis *Aspects*, le nœud gordien de la grammaire générative. La définition de I-language comme étant, par définition, opaque à tout conditionnement externe à lui même impose logiquement la nécessité de représentations innées ainsi qu'une théorie de l'acquisition où la maturation est strictement linguistique et conditionnée quasi-exclusivement par sa propre logique développementale (cf. fig. 3). Ceci suppose que les facteurs externes (en rouge dans la figure 11) n'aient aucun rôle dans ce processus ou en les réduisant à une information « déclenchant » (« triggering », en vert dans la figure 11), la logique de maturation interne (en bleu dans la figure 11). L'état initial est donc un état qui doit préconditionner les représentations adultes en restant le plus opaque possible aux données externes. Les stades d'acquisition S1, S2, S3 ... entrent en logique de continuité directe avec la GU sans recevoir aucune influence externe qui conditionne leur logique de maturation.

Aux principales réserves que l'on peut avoir à l'égard de ce dispositif théorique (circularité des arguments justifiant respectivement et réciproquement GU et I-Language (Carr, 2000, Rondal, 1999), caractère stipulatif d'une hypothèse non falsifiable), s'opposent des faits et des arguments qu'il convient de mettre en regard mais dont la pertinence pour la phonologie est toute relative puisqu'ils ont été principalement formulés pour l'acquisition syntaxique (pour une revue, voir Pinker, 2000).

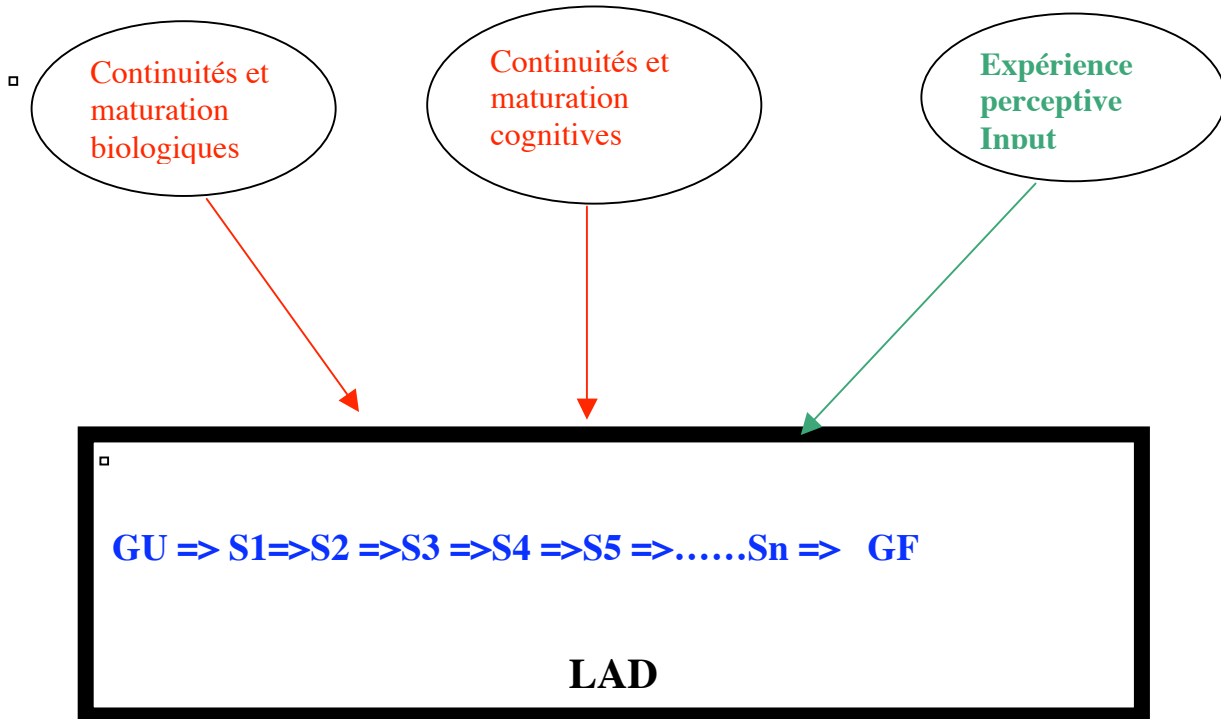


Figure 11 : Modèle du « bébé cartésien »

Mon propos n'est ni de reconvoquer ce débat dans son intégralité ni d'en évaluer toute la dimension épistémologique, philosophique, etc ...

La nature exacte des contenus et facultés linguistiques innés étant indéterminée, le postulat de leur rôle dans la mise en place des représentations phonologiques, l'existence de la « GU » ne doivent, à mon avis, constituer comme c'est trop souvent le cas, ni un dogme, ni un tabou mais une hypothèse de travail à évaluer à partir des faits et des données observables en s'abstenant de positions *a priori* qui en rejetant les apories très réelles du modèle chomskyen pour l'acquisition phonologique rejette *ipso facto*, l'hypothèse d'un dispositif inné spécialement destiné à la mise en place du langage chez l'être humain.

C'est ce que j'ai appelé la question du bébé et de l'eau du bain, qui consiste, en évacuant les bulles chomskyennes, à vider radicalement le bain de l'acquisition phonologique générativiste, en emportant aussi et *a priori* le bébé rationaliste (sur

cette question cf. Carr, 2000). Car aucun empiriste même le plus convaincu ne soutiendrait actuellement que le langage émerge à partir d'une *tabula rasa* cognitive absolue. Ni l'empirisme classique, ni Piaget (1964) Piatelli-Palmarini, (1977), ni la reformulation néo-piagétienne des théories de l'acquisition (Karmiloff & Karmiloff-Smith, 2003), ni les défenseurs d'une continuité biologique phylogénétique entre les primates et les humains reflétée par la transition babillage langage chez le bébé humain (Mac-Neilage, 1998) n'ont jamais refusé de considérer que l'émergence du langage était en partie conditionnée par des facultés cognitives hautement spécialisées et spécifiques à l'espèce humaine, et ce, à partir d'un substrat neuro-physiologique dont on commence à mettre au jour les fonctionnements de manière un peu plus sophistiquée (Rondal, 1999 ; Houdé et al. 2002).

La question concerne donc le conditionnement spécifiquement linguistique de l'acquisition du langage et pour ce qui concerne mon propos, de l'acquisition de la phonologie. Que serait donc la « GU » pour la phonologie, s'il s'avère que « GU » il y a et comment interagit-elle avec les autres facteurs qui conditionnent l'acquisition de la phonologie ?

Car si l'on sait que les contraintes motrices, perceptives, l'émergence des facultés cognitives générales, l'input jouent un rôle dans l'acquisition phonologique, on constate également, dans les données phonologiques, des effets de marque structurels qui ne sont réductibles ni à la fréquence lexicale, ni aux capacités articulatoires ou capacités de traitement ou capacités cognitives générales (Fikkert et al. 2004).

Dans ces pages, je traite de ces questions qui sont au centre du projet de recherche que je mène en acquisition depuis ma soutenance de doctorat. Je soulèverai les points théoriques suivants, qui intéressent directement mon objet de recherche, et concernent donc la question de l'émergence et du statut des représentations phonologiques en acquisition.

- i) Existe-t-il des facultés innées spécifiquement linguistiques, régies par leur logique propre qui conditionnent l'acquisition du langage et plus précisément l'acquisition de la phonologie d'une langue

- ii) Ces facultés interagissent-elles avec d'autres facultés cognitives, motrices, développementales générales et si oui, lesquelles ?
- iii) Ces facultés sont-elles conditionnées par l'input et si c'est le cas, comment ?
- iv) Ces facultés, qui se manifestent via une dynamique développementale, à travers des usages et des comportements variables et évolutifs sont-elles représentationnelles ou procédurales ?

## **2. Pauvreté du stimulus en phonologie : quel est le rôle de l'input ?**

### **2.1. Un problème syntaxique**

L'argument dit de la « pauvreté du stimulus » qui est le pendant logique du postulat de l'innéité de la GU a été formulé par Chomsky exclusivement à partir de faits syntaxiques. Et c'est toujours de ce point de vue qu'il est envisagé. Si l'on regarde en effet le numéro de *Linguistic Review* n° 19 (2002), « A review of the « Poverty of stimulus argument » », on constate qu'il est entièrement consacré à la syntaxe. L'argument consiste à envisager que le stimulus entendu par l'enfant est insuffisant pour lui permettre de mettre en place aussi vite et aussi efficacement une langue sans le recours à des représentations internes innées qui conditionnent prioritairement et quasi-exclusivement l'acquisition. L'insuffisance est envisagée tant d'un point de vue quantitatif (le nombre des phrases possibles dans une langue étant infinie, l'enfant ne peut pas avoir entendu toutes les structures qu'il va produire avant de les produire, il doit donc les construire à partir d'une connaissance interne), que du point de vue qualitatif (l'enfant entend des énoncés variables, plus ou moins acceptables et grammaticaux, conditionnés par le contexte et ne bénéficie jamais d'une correction de l'environnement lui explicitant non seulement les structures permises. L'enfant ne reçoit donc que des évidences positives plus ou moins bien formées (ce que font les locuteurs de son environnement) mais jamais d'évidences

négatives (des informations explicites sur les structures non permises dans sa langue) alors qu'il ne produit jamais certaines structures non permises.

En syntaxe, cette question a donné lieu à des très nombreux débats très intéressants portant sur la validité intrinsèque de l'argument (Cowie, 1999, 2001, Fodor, 2001, Pullum & Scholtz, 2002, Collins, 2003), la circularité argumentale qui lie le postulat de l'existence de la GU et l'argument de la pauvreté du stimulus, la non-falsifiabilité de l'argument (Rondal, 1999 ; Carr, 2000).

Ce qui m'intéresse ici est de discuter la validité de cet argument pour la phonologie. Il a toujours été considéré que la pauvreté du stimulus était un argument qui allait naturellement de soi pour la phonologie alors que pour au moins deux raisons, il est problématique (Scheer, 2004b) : i) le nombre d'unités phonologiques dans une langue est en nombre fini et restreint ; ii) la notion de générativité est non pertinente pour la phonologie : la phonologie est conditionnée par l'existence d'un lexique fini où la générativité est possible (je peux fabriquer de nouveaux logatomes et constater qu'ils sont bien formés (« proti ») ou mal formés \*(« rpoit »)) mais non utilisée dans les faits où le lexique est appris (personne ne va spontanément créer de nouveaux logatomes pour s'exprimer, comme on produit des phrases).

Ces questions, longtemps négligées, sont maintenant centrales dans le cadre du développement des modèles d'acquisition défendant une émergence exclusivement « bottom-up » (Dupoux, à paraître, Pierrehumbert, 2002), directement guidée par la fréquence des occurrences (Bybee, 2001, Chevrot, 2004, Gerken, 2004) et procédant par induction à partir de mécanismes cognitifs généraux (Carr, 2000), Vihman (2000, 2003), Vihman & Mueller-Gathercole (en préparation).

## **2.2. Un stimulus quantitativement riche**

S'il est évident que l'infinité des structures syntaxiques bien formées possibles dans une langue ne sera jamais exhaustivement disponible, un enfant peut au contraire être exposé assez rapidement à toutes les co-occurrences phonémiques et syllabiques possibles dans sa langue. Pierrehumbert (2004) montre par exemple que si un bébé reçoit une heure par jour de parole pendant la première année de vie, il entend 6.570.000 syllabes et 20 millions d'exemplaires de phonèmes, ce qui lui donne un

échantillonnage représentant exhaustivement tous les sons de sa langue et incluant détails phonétiques et variation contextuelle proposés en un nombre d'exemplaires conséquent qui pourrait lui permettre d'utiliser la fréquence de distribution de l'information acoustique pour faire émerger les catégories phonétiques pertinentes dans la langue cible.

Mayé et al. (2002) ont ainsi montré que des enfants exposés à une distribution bimodale de paires de sons où est réalisé un contraste de VOT<sup>40</sup> (vs enfants exposés à un contraste unimodal) établissent plus facilement le traitement catégoriel de paires d'occlusives voisées / non voisées. Ceci montre que la grande disponibilité fréquentielle, dans le signal, de l'information acoustique sur laquelle repose une opposition catégorielle entre deux sons aide les enfants à privilégier cette catégorisation comme étant pertinente dans leur langue. Mayé et al. proposent que les phénomènes de surdité phonologique (perte des contrastes non pertinents dans la langue cible) que l'on peut observer entre 8 et 12 mois (Kuhl et al. 1992, Werker & Tees, 1984) sont phonétiquement conditionnés et reposent sur la sensibilité de l'enfant à la distribution statistique dans l'input des paramètres acoustiques sur lesquels s'appuient les catégorisations phonétiques. Elles excluent que le caractère distinctif de ces catégories dans la langue cible conditionne leur sélection exclusive par l'enfant.

Il a été également montré que cette sensibilité à la fréquence de distribution de l'information est reflétée en production : les enfants répètent plus aisément des diphtonges avec un haut niveau de probabilité que des suites avec un faible niveau de probabilité (Munson, 2001). Sur un tel échantillon, par inférence statistique, l'enfant a donc une probabilité effectivement très élevée de dégager et de produire les co-occurrences phonétiques les plus fréquentes dans sa langue.

---

<sup>40</sup> On remarquera ici que l'expérience a été réalisée à partir du contraste de VOT qui est une réalisation acoustique stable, répliquable et peu variable permettant d'opposer très clairement et simplement deux catégories phonétiques qui sont reflétées massivement dans les systèmes phonologiques de toutes les langues. La plupart des contrastes phonétiques d'une part correspondant à des oppositions distinctives ne sont pas réalisés par la variation linéaire d'un seul critère acoustique mais plutôt par une configuration de critères (acoustiques, mais aussi articulatoires etc) interagissant entre eux. Je ne suis pas sûre que ce design expérimental donnerait le même résultat pour un contraste opposant fricative et occlusive sourde (par exemple p / f) et qu'on pourrait dès lors argumenter que c'est la fréquence de distribution de l'opposition p/f dans une langue qui permet de faire émerger ces deux catégories.

Mais, cette richesse acoustique du stimulus, associée à l'existence de capacités de traitement catégoriel orientées, et à la capacité chez le bébé humain de développer des algorithmes d'apprentissage à partir de cette richesse acoustique, ne répond que partiellement à la question de l'influence de l'input sur l'émergence des représentations phonologiques chez le petit humain.

Pour conclure, s'il est clair que les caractéristiques spécifiques de l'input phonologique (fini et lexicalement conditionné) récuse l'argument de la pauvreté quantitative du stimulus, reste la question de la richesse qualitative de l'information présidant à l'émergence des représentations. L'information pertinente n'est jamais « phonologiquement » réalisée, donc jamais explicitement disponible mais au contraire réalisée sous une forme phonétiquement variable dont l'enfant doit extraire l'invariance. On songe ici à tous les phénomènes phonologiques « classiques » impliquant des alternances lexicales ou s'expansant sur des domaines excédant le « mot » (mutations dans les langues celtiques, harmonies vocaliques), réalisation de structures syllabiques à constituants vides ou flottants (schwa, liaison en français) ainsi que toutes les formes d'allophonies (Peperkamp et al. 2003, Peperkamp et al. à paraître).

Si l'on peut donc plaider en faveur d'une « richesse » quantitative du stimulus, ceci ne suppose pas ipso facto que cette richesse phonétique du signal permette directement l'acquisition phonologique, puisque les réalités substantielles et formelles relèvent, dans notre conception (cf. Ch.I) de deux niveaux de réalité différents.

Se posent alors les deux questions suivantes : si l'enfant procède bien par « calcul inductif » (Carr, 2000) à partir de l'input phonétiquement riche,

- i) qu'est-ce qui conditionne ce calcul. Et comment définir précisément la fréquence pertinente pour un nouveau-né, un enfant de 6 mois, d'un an ou de deux ans ?
- ii) quelle est la genèse des catégories à partir desquelles les enfants procèdent ?



- iii) qu'est-ce que ce calcul inductif ? Sur quels mécanismes psychologiques précis repose-t-il ? En particulier résulte-t-il de l'expérience de perception précoce (1<sup>ère</sup> année) ou est-il également tributaire de l'acquisition lexicale ?

### **3. Stratégies perceptives et rôle de la fréquence : fréquence de quoi ?**

#### **3.1. Un petit bébé très doué**

Parmi les nombreuses inconnues concernant l'acquisition du langage et particulièrement l'acquisition phonologique, subsiste le problème du décalage chronologique entre perception et production. Tous les travaux expérimentaux montrent en effet une antériorité évidente entre capacités de traitement du signal d'entrée déjà très élaborées à la naissance et premières productions linguistiques (vers 12 mois au mieux). Ce décalage évidemment consécutif à des rythmes de croissance différents des organes phonatoires et du système d'audition déjà abouti à 37 semaines de gestation, ne laisse pas d'interroger quant à la maturation cérébrale et cognitive, qui, elle, n'est pas différenciée pour la perception et la production. Se pose alors, en conséquence, la question de la pertinence des informations empiriques qui permettent d'observer l'émergence de représentations phonologiques. Quelles sont les données significatives lorsqu'on observe un enfant et que fait-il pendant la première année de vie ?

On peut en particulier se demander en quoi les données de perception précoces témoignent d'une activité « phonologique » au même sens que les données de production. Mais, on ne peut pas exclure a priori non plus, comme cela a été trop longtemps le cas, que l'expérience perceptive de la première année ne soit déterminante dans l'émergence des « patterns » phonologiques. L'orientation précoce des productions vers la langue cible en témoigne : Boysson-Bardies et al. (1984) Boysson-Bardies & Vihman (1991) ont par exemple montré à partir de données de babillage enregistrées chez des bébés élevés dans des environnements francophones,

arabophones (arabe marocain) et sinophones (chinois cantonais) que dès 10 mois, avant l'apparition des premiers mots, les productions des enfants comportent des patterns acoustiques réguliers orientés vers la langue cible et reconnus comme tels par les adultes francophones, arabophones et sinophones.

Par ailleurs, Gerken (2004) a montré l'existence d'une capacité précoce de généralisation de patterns abstraits à partir d'expériences d'apprentissage de langages artificiels. Elle a familiarisé des bébés de 9 mois à des logatomes de 3 à 5 syllabes lourdes et légères (« do » « ré » « mi » « ton » « fa » « so ») auxquelles étaient appliquées les règles d'accentuation de l'anglais, formulées sous forme de contraintes dans le cadre OT. 4 contraintes A, B, C et D étaient proposées aux enfants dans des suites accentuées illustrant un « ranking » optimal. Elle a ensuite testé les bébés avec une tâche de fixation visuelle. Les résultats montrent que les bébés sont capables de reconnaître dans les patterns accentuels appliqués à des logatomes différents de ceux auxquels ils ont été familiarisés, les patterns qui respectent le « ranking » auxquels ils ont été familiarisés. Ceci suppose donc qu'ils sont capables d'extraire à partir de quelques occurrences particulières qui leur sont explicitement proposées, la régularité structurelle qu'elles partagent. Gerken ne donne pas d'explication au phénomène et envisage que ceci puisse refléter l'utilisation d'une connaissance linguistique universelle innée (GU) (les syllabes lourdes doivent toujours être accentuées, deux syllabes consécutives ne peuvent pas porter l'accent) ou une faculté de traitement opérant une généralisation dans un cadre expérimental spécifique (« ton » doit toujours être accentuée sauf quand il est précédé de « ton » ou suivi de « ton » déjà accentué). Mais dans tous les cas ceci reflète une capacité précoce de généralisation structurelle sur la seule base de l'input perçu.

L'hypothèse d'une mise en place de la phonologie sur la base de l'expérience perceptive suppose qu'acquisition phonologique et lexicale sont indépendantes et qu'il est possible pour l'enfant de construire des représentations avant la phase des 50 mots produits, sur la base exclusive de l'input qui l'entoure. Ce point de vue est retenu par un certain nombre de chercheurs, s'appuyant sur des données de perception précoce ou des expériences d'apprentissage de langues fictives chez l'adulte et chez le

nouveau-né (Morgan & Saffran, 1995, Saffran et al. 1996, Echols et al. 1997, Gomez & Gerken., 1999, Gomez & Gerken, 2000).

Ces modèles ou ces conceptions sont le plus souvent bottom-up supposent que l'enfant extrait les catégories pertinentes à partir du traitement statistique, en s'appuyant sur régularités fréquentielles dans l'input sans recours à une connaissance abstraite du système linguistique.

Peperkamp (2002) propose au contraire un modèle supposant le passage par la connaissance lexicale en intégrant les différents niveaux de traitement possible qui amèneraient le bébé à construire un mini lexique à partir de l'input. Ce modèle postule l'existence d'une représentation *prélexicale* qui est intermédiaire entre le signal acoustique et le lexique. Elle résulte d'une segmentation en groupes intonationnels et groupes phonologiques et est construite à partir d'un codage catégoriel des sons de la langue. Trois types d'information seraient exploités à partir de la représentation prélexicale dans ces tâches, à savoir les régularités distributionnelles, la phonotactique (c'est-à-dire, l'ensemble de contraintes sur la bonne formation phonologique des mots), et la forme typique des mots. Ces informations permettraient de dégager séparément les mots grammaticaux et les mots lexicaux et de construire un premier lexique.

Ce modèle suppose donc, d'une part que la distribution des sons dans le signal acoustique permettrait l'acquisition d'une représentation prélexicale, qui serait ensuite exploitée pour affiner l'acquisition phonologique et pour construire un lexique, par ailleurs que ce traitement de l'input n'est pas guidé par la connaissance de principes universels d'organisation des langues. Les enfants construiraient donc un lexique « muet », disponible ou partiellement disponible, et phonologiquement déterminé et analysable avant qu'ils ne se mettent eux-mêmes à le produire.

Les expériences réalisées avec des petits bébés montrent effectivement, que vers 10 mois, à une période correspondant en général en production à la transition babillage canonique / babillage précoce, les enfants ont acquis la surdité phonologique (Kuhl et al. 1992, Werker & Tees, 1984), ils savent reconnaître les frontières prosodiques majeures (Hirsch-Pasek et al. 1987, Kemler-Nelson et al.,

1989, Jusczyk et al., 1992, Gerken 1994), les patrons accentuels de leur langue (Jusczyk et al. 1993). Mais si nous savons qu'ils sont sensibles à des informations segmentales et suprasegmentales, nous ne savons pas en fait si ces capacités reflètent des possibilités de traitement acoustique de l'information, ni comment les enfants peuvent s'en servir pour construire des représentations phonologiques abstraites.

### **3.2. Traitement phonétique ou phonologique ?**

Tout d'abord, il n'est pas facile - et je pense que pour l'heure personne n'y est parvenu- d'établir quelles sont les informations exactes qu'exploitent les enfants. Dans la variété infinie d'informations disponibles et corrélées qui constituent un signal de parole, il n'est pas très facile de savoir exactement quelles régularités sont exploitées et comment les bébés trient et organisent la variété acoustique à laquelle ils sont exposés. La sensibilité à de très nombreux facteurs acoustiques présentés isolément (comme le VOT) ou corrélés à d'autres a été mise en évidence dans des situations expérimentales assez contraintes, nous l'avons vu, mais dans tous les cas, la question qui reste ouverte est la suivante : peut-on de toute façon, à partir de ces seules informations, sans grammaire et sans lexique, découvrir le système phonologique d'une langue. Il me semble que la réponse est négative et que ces informations permettent de dégager et de mémoriser les catégories phonétiques possibles, les co-occurrences possibles dans cette langue mais ne permettent aucunement de comprendre la logique systémique de l'ensemble et en particulier le rôle distinctif des unités à dégager et leur nature spécifique de 2<sup>ème</sup> articulation, non plus que les phénomènes phonologiques classiques comme les allophonies ou l'existence de constituants vides etc.

Il semblerait d'ailleurs que, plutôt qu'un traitement catégoriel fin précoce annonçant des phonèmes ou des syllabes, les enfants s'intéressent d'abord et prioritairement dans la première année et avant l'apparition du premier lexique au problème de la segmentation de la chaîne parlée et focalisent leur attention sur l'établissement des unités de généralisation possible. Comme le propose Peperkamp (2002), il semblerait que le travail que réalisent les enfants lors de la première année

consiste à construire un proto-lexique, à partir d'étiquettes segmentées et mémorisées mais non analysées. Des travaux réalisés par Saffran et al. (1996) à partir de l'apprentissage de langages fictifs ont montré que des bébés de 8 mois étaient capables de dégager des unités lexicales sur la base de la fréquence des transitions syllabiques. Postulant que les co-occurrences syllabiques systématiques ( $x=>y$ ) sont plus fréquentes à l'intérieur d'un mot que sur une frontière (par exemple « man » derrière « ma » *maman* va avoir une fréquence d'occurrences plus importante que « ma » derrière « li » *jolie maman, joli nounours, joli camion etc.* ils ont présenté des suites de logatomes (tolipubidakudotigotolipu...) avec des fréquences d'occurrences syllabiques différentes, et ont constaté que les enfants opéraient des segmentations sur cette seule base (par exemple ici tolipu). Si on ajoute à cela des informations acoustiques et prosodiques (allongements, accents...), on peut supposer que les enfants réussissent à dégager des formes lexicales régulières.

Pourtant, il semblerait qu'ils opèrent plutôt sur la base d'un traitement holistique et analogique global où des empans de signal largement segmentés et mémorisés comme des étiquettes sonores, sont globalement reconnus avec une logique où l'on voit émerger progressivement des unités de traitement prosodiquement conditionnées de plus en plus petites (Hirsch-Pasek et al. 1987, Jusczyk et al., 1992). Ceci n'implique pas d'analyse linguistique, ni phonologique, ni morphologique fine. Les travaux réalisés en production sur l'acquisition de la liaison et des clitiques (ch.I et II), montrent par exemple qu'en production, pour le français, et vraisemblablement pour d'autres langues romanes (Demuth, 2001), les clitiques ne semblent pas segmentés avant l'explosion lexicale. De nombreux travaux en perception (pour une synthèse cf. Echols & Marti, 2004) ont également montré que les enfants repèrent les mots grammaticaux de manière précoce mais s'en servent, associés à d'autres informations récurrentes et saillantes pour calculer les frontières de mots et réaliser la segmentation du signal en « blocs ».

Par ailleurs, Hallé et Boysson-Bardies (1994) ont réalisé des expériences de reconnaissance de mots familiers et non familiers chez des enfants de 10 et 12 mois à qui ils ont proposé une liste de 12 mots fréquents et 12 mots rares avec une tâche d'écoute préférentielle. Les résultats montrent une très nette préférence pour les mots

fréquents dès 11 mois. Ils ont ensuite présenté la liste de mots familiers en modifiant pour une première liste le voisement, pour une deuxième liste le mode d'articulation de la première consonne remplacée biberon => piperon, viberon, gateau => kateau, jateau etc. Malgré l'importance des déformations pour un locuteur adulte, les enfants ont continué à manifester une préférence pour les mots fréquents déformés et n'ont exprimé aucune préférence quand on leur présente la liste des mots fréquents non déformés et la liste des mots fréquents déformés. Ces résultats laissent penser qu'il n'y a pas à ce stade d'analyse et de décomposition phonologique des unités lexicales, stockées de manière globale ou sous-spécifiée.

Enfin, de nombreux travaux ont montré la difficulté des enfants à réaliser des tâches de reconnaissance et de traitement lexicaux (pour une revue détaillée sur cette question cf. Fisher & al. 2004), et en particulier à distinguer des paires minimales. Stager & Werker (1997) ont mené par exemple une expérience avec des enfants de 14 mois où la discrimination de deux « syllabes » [bih] [dih] peut avoir des valeurs fonctionnelles différentes. Quand les enfants doivent distinguer deux syllabes différentes, leurs performances sont très bonnes, mais quand ces syllabes sont associées à une image, et qu'ils doivent par conséquent reconnaître deux mots, nouveaux pour eux, leurs performances sont moindres. Ceci laisse donc penser que, dans le premier cas leur analyse est strictement phonétique et n'implique aucune analyse linguistique, alors que dans le deuxième cas, l'association des signifiants [bih] et [dih] à des signifiés nécessite que les enfants reconnaissent des signes linguistiques et donc qu'ils réalisent une analyse phonologique de l'input. Ces travaux confirment les résultats obtenus par Barton (1978) qui montrent que la discrimination de paires minimales [bwaR] / [pwaR] est une tâche trop difficile pour des enfants de moins de 20 mois.

Ces résultats amènent à tirer diverses conclusions. A fréquence de distribution égale dans l'input, la fonction de l'information acoustique dans le système phonologique la rend plus ou moins accessible à l'enfant, les mêmes objets acoustiques ne donneront pas lieu au même traitement selon que l'enfant cherche un mot, un morphème vs une syllabe dans l'input. Ceci signifie par conséquent que la

fréquence de réalisation des informations acoustiques dans le signal n'est pas pertinente *en soi*.

Dès lors, on peut effectivement arguer que la fréquence de l'information acoustique joue un rôle important dans l'acquisition (ce qui paraît intuitivement juste voire trivial, plus j'entends un mot plus il m'est familier et plus je le reconnais rapidement), mais il semblerait que si cette régularité fréquentielle permet un traitement phonétique, elle laisse irrésolue la question suivante : fréquence de quoi pour l'acquisition phonologique ?

Quel(s) est (sont) les unités de traitement phonologiquement pertinentes pour un bébé entendant du signal et comment peut-il les détecter sans GU ?

### **3.3. La question de la genèse des catégories abstraites et des représentations<sup>41</sup>.**

Ceci nous amène donc logiquement à la question de la genèse mystérieuse et *sui generis* des catégories abstraites, dès lors qu'on exclut qu'une information ou une faculté innée spécifiquement destinée à l'acquisition du langage conditionne au moins partiellement le traitement précoce de l'input. La formulation strictement « bottom-up » de la genèse des catégories phonologiques, excluant l'apport de facultés spécifiquement linguistiques de traitement à partir de primitives phonologiques nous renvoie toujours, comme le souligne très justement Dupoux (2004) à une logique de type « la poule et l'œuf » : qu'est-ce qu'un phonème pour un sujet n'ayant à sa disposition aucun système phonologique et aucun lexique et comment peut-il le retrouver dans le signal si ce n'est à partir d'un certain nombre d'unités lexicales ? Mais comment ce même sujet peut-il reconnaître des unités lexicales (et des paires minimales) s'il ne sait pas ce qu'est un phonème et qu'il n'a pas de lexique. En d'autres termes, le problème qui se pose là consiste à se demander si l'on peut acquérir la phonologie indépendamment de l'acquisition lexicale.

Comme je l'ai dit au chapitre II, je travaille dans une perspective qui lie acquisition morphologique, phonologique et lexicale. Je présente ci-dessous les

---

<sup>41</sup> Je remercie beaucoup J. Brandão de Carvalho. La plupart des réflexions et des arguments que je développe à ce sujet tout au long de ce mémoire sont issus de discussions informelles avec lui.

arguments et réflexions qui m'ont amenée à concevoir de cette manière l'acquisition phonologique.

Dupoux (2004) soulève directement ce problème du décalage entre traitement phonétique et établissement de catégories abstraites et pose le problème de l'existence d'un objet ou d'une information phonétique qui pourrait favoriser l'émergence des catégories abstraites.

Si l'on ne retient pas le gabarit lexical et que l'on travaille sur des unités de deuxième articulation, la syllabe peut effectivement être, comme il le suggère<sup>42</sup>, un bon candidat favorisant la généralisation des représentations : les exemplaires de [bi] et [ba] stockés expriment suffisamment de régularité acoustique (tous les [b] sont suffisamment identiques entre eux et différents de par exemple [p] ou [d]) pour que la comparaison soit efficace. Par ailleurs ces exemplaires de [b] sont toujours co-occurents d'une unité partageant des particularités spectrales majeures (sonorité car la distribution impose en français à cette place soit une voyelle, soit une liquide) et jamais d'une autre catégorie partageant beaucoup de particularités acoustiques avec eux ([d], [t], [g] ...).

Mais, s'il est vrai que la syllabe est une unité ayant à l'évidence une réalité acoustique, rythmique, articulatoire, perceptive (pour une synthèse sur la question cf. Wauquier-Gravelines, 1996) plus consistante que celle du phonème, elle est également un objet structurel. Il me semble que l'on retombe par conséquent et malheureusement dans le paradoxe énoncé précédemment : pour pouvoir postuler que les enfants recherchent les syllabes dans le signal, il convient de définir au préalable ce qu'on désigne précisément par « syllabe ». Comment traiteront-ils, par exemple, une suite comme « arbre », qui correspond à une seule oscillation mandibulaire<sup>43</sup>, qui est une seule syllabe acoustique et rythmique dans beaucoup de réalisation, mais incontestablement constituée de deux syllabes phonologiques, avec une attaque vide ([lezarbr̥] ) et un noyau vide ([lezarbr̥dypark]). Quelles syllabes (articulatoire, acoustique, métrique, phonologique) chercheront-ils et par conséquent trouveront-ils dans un input non segmenté au préalable ?

<sup>42</sup> Comme le fait également Ramus mais pour en faisant référence à la syllabe comme unité rythmique

<sup>43</sup> Vallée, cp.



Il me semble que quelle que soit l'unité abstraite concernée (syllabe ou phonème), si l'on postule que les bébés peuvent extraire sans difficulté des formes régulières comme [ba] et [bi], ou [i], [p] etc, qu'ils vont ensuite comparer systématiquement, ceci suppose implicitement qu'on admet qu'ils travaillent déjà sur une représentation phonétiquement interprétée correspondant grosso modo à une notation API très large. Or ce n'est pas le cas. S'ils ne disposent pas de primitives leur permettant d'interpréter que tel type de configuration acoustique (VOT, explosion, pente formantique, vitesse de transition) etc. est une syllabe, ou un phonème, par quel mécanisme psychologique peuvent-ils inférer par exemple qu'un diphone particulier est une syllabe : [bi] dans « biberon » et qu'un autre n'en est pas une [ib] dans « ibis » ?

Il apparaît que comme le rappelle Scheer (2004b) opposant « corpus structuraliste » et « corpus générativiste », que les travaux d'acquisition, comme toute la linguistique générative, aient retenu majoritairement une conception du phonème et de la syllabe qui les envisage de manière isolée. Or comme il le souligne,

[...] aucun objet phonologique ne peut être considéré seul : ni en tant qu'entité avec par exemple ses propriétés phonétiques individuelles, ni en tant qu'acteur d'un processus phonologique : deux objets phonétiquement identiques peuvent réagir au même processus de manière totalement différente en fonction de la place qu'ils occupent dans leurs systèmes respectifs. »

T. Scheer, "En quoi la phonologie est vraiment différente", *Corpus*, 3, 2004. p.47

La plupart, pour ne pas dire tous les modèles actuels d'acquisition, qu'ils s'inscrivent dans ou contre la doxa chomskyenne, me semblent considérer que l'acquisition phonologique consiste à mettre en place les catégories phonologiques d'une langue (syllabes, phonèmes) et non un système organisé d'unités conditionnées les unes par les autres.

Ce faisant, on constate qu'un grand nombre d'études en acquisition phonologique ne rendent pas sa place à l'apport structuraliste dont est pourtant issue la réflexion chomskyenne sur le mentalisme linguistique. Une conception de l'émergence des

catégories phonologiques comme des objets isolés que l'enfant découvrirait indépendamment de leur place respective dans un système, ignore l'apport fondateur du structuralisme formulé diversement par Saussure (1879, 1966), : la langue est un système. Elles négligent également l'apport praguois qui ajoute : la langue est un système d'unités distinctives, ainsi que l'apport de Martinet (1949,1957) qui renchérit la langue est un système d'unités distinctives doublement articulées.

Cette conception exclut par conséquent que l'ordre, les procédures d'acquisition phonologique subissent une quelconque pression systémique et que l'enfant utilise la logique systémique à l'œuvre dans sa langue pour acquérir le système. On peut pourtant supposer qu'entendre [ø] pour un petit français en pleine acquisition d'un système ayant 4 degré d'aperture et une série de voyelles d'avant arrondie, n'aura pas les mêmes répercussions que pour un petit espagnol. Indépendamment de la réalisation acoustique de l'objet considéré, la logique des systèmes linguistiques concernés rend la probabilité d'un traitement phonétique (en espagnol) et d'un traitement phonologique (en français) respectivement très élevée. En français le son [ø] a de très fortes chances d'être un phonème puisque dans ce système linguistique existe également [y].

Dès lors, si l'on postule que les enfants procèdent de manière strictement bottom-up, et établissent des généralisations aboutissant à des catégories phonologiques abstraites telles que la syllabe ou le phonème, à partir de la fréquence de distribution de l'information acoustique et donc sans primitives phonémiques et syllabiques de référence avec lesquelles ils peuvent comparer les occurrences possibles, il m'apparaît que la seule forme « motivée » et possible que l'enfant peut reconnaître dans le signal, et ce malgré ces variantes de réalisation, est le mot (au sens « unité lexicale »). C'est uniquement à partir d'une unité ayant un signifié stable et constant lui servant de référence qu'un enfant peut reconnaître et analyser la structure de la suite sonore qui est systématiquement associée à ce signifié et ce malgré ses variantes de réalisation. C'est parce qu'il veut toujours dire la même chose, qu'un mot et par conséquent les unités sonores qui le constituent (syllabes, phonèmes) peuvent

être construits à un niveau de représentation abstrait malgré leurs variétés de réalisation acoustique.

Pour conclure, il est, à l'heure actuelle très difficile de trancher et d'établir clairement le rôle de l'expérience perceptive précoce dans l'émergence de la phonologie, en l'absence d'un lexique. Cette période est à l'évidence essentielle tant du point de vue maturationnel (maturation cognitive, motrice et physiologique) que du point de vue de l'acuité exponentielle et étonnante que les bébés manifestent pendant cette première année pour le traitement linguistique ou pré-linguistique de l'input. Nous connaissons encore mal les mécanismes psychologiques précis qui se développent à cette période qui semble pourtant fondamentale pour l'acquisition phonologique dans la mesure où c'est au cours de cette période que les enfants construisent un premier « lexique ».

Mais le problème central est manifestement constitué par la question de la genèse des catégories abstraites que les enfants pourraient construire uniquement à partir du traitement acoustique du signal et qui est pour l'instant complètement ouverte et irrésolue. Nous n'avons qu'une connaissance très partielle des phénomènes cognitifs qui pourraient être en jeu (Gomez & Gerken, 2000). Que serait en particulier le « calcul inductif » proposé par Carr (2000) qui leur permet la généralisation sur la base de facultés cognitives générales ? L'hypothèse, dès lors qu'elle n'est pas appuyée sur un modèle précis des fonctionnements psychologiques impliqués, qui permettent à partir de mécanismes de catégorisation non spécifiques, de construire des représentations phonologiques qui elles le sont, n'est pas moins stipulative que l'hypothèse GU.

Pour toutes ces raisons et ne retenant pas comme je l'ai dit en 1. la conception chomskyenne d'une GU représentationnelle immédiatement disponible et constituée de principes et paramètres, j'ai fait l'hypothèse dans mon travail (suivant en cela directement M Vihman) que la première année manifestait des aptitudes phonétiques et prélinguistiques permettant l'établissement d'un proto-lexique à partir duquel sont effectivement construites ensuite les représentations phonologiques.

## 4. Lexique et émergence de la phonologie

### 4.1. Quelque chose que « fait » l'enfant ou qui « arrive » à l'enfant ?

La phonologie générative, comme la plupart des travaux en perception que nous avons mentionnés ci-dessus, s'est développée en considérant l'enfant apprenant comme un organisme passif, lieu de développement d'algorithmes d'acquisition ou d'apprentissage conditionnés soit par la maturation cognitive, soit par la logique du modèle (déclenchement de la GU). Or cette vision unifiée de bébés calibrés sur un seul pattern de développement ne correspond ni à la réalité du développement moteur et psycho-affectif des enfants, ni à la variété des patterns développementaux observés. Si l'acquisition de la phonologie est un phénomène universel irrépensible, biologiquement et génétiquement conditionné, chronologiquement assez peu variable (tous les enfants arrivent quasiment en même temps au même résultat et ce quelle que soit la langue), il est aussi pour chaque enfant une aventure individuelle de construction de sa langue maternelle dans un environnement donné, avec des interlocuteurs particuliers avec qui il entretient des relations d'échanges variés et complexes, au sein d'une structure sociale et familiale<sup>44</sup>

En termes cognitifs, on constate en effet une variété non négligeable de comportements phonologiques possibles chez les enfants qui, soit a été masquée dans les études perceptives par le fait de moyenniser les données des sujets, soit dans le cadre génératif par le fait que les généralisations sont faites sur des corpus restreints où les critères de sélection, de traitement des données longitudinales ne sont jamais explicités (pourquoi certaines formes sont retenues et d'autres non, comment sont segmentées les productions précoces etc etc) et où l'objectif visé est d'établir une

---

<sup>44</sup> Je ne parlerai pas ici des questions liées au rôle de l'environnement social, affectif pas plus que des aléas de la construction psychique de l'enfant dans l'émergence du langage, ni de la théorisation très élaborée qu'en a fait la psychanalyse dont c'est un des objets centraux. D'une part parce que je n'ai pas de compétence en la matière, d'autre part parce que s'il est avéré que les maltraitances psychologiques diverses, l'abandon, la misère sociale et affective peuvent entraîner des troubles du langage importants, on ne peut considérer me semble-t-il que ces phénomènes agissent de manière directe et simple sur les phénomènes d'acquisition linguistique. Si c'était le cas, il me semble qu'une bonne partie des humains actuellement vivants seraient muets. Or ce n'est pas le cas. Je pense néanmoins que nous ne pouvons ignorer l'importance de ce champ des sciences humaines dans la compréhension que nous avons de certains aspects de l'émergence du langage chez l'enfant ainsi que de certaines pathologies du langage.

généralisation de portée universelle<sup>45</sup> sur des synchronies successives sans théorisation spécifique de la logique développementale. Ces phénomènes dits de *learning paths* sont selon les auteurs, plus ou moins idiosyncrasiques et diversement contraints. L'une des questions actuellement débattues concerne le rôle de la marque dans la limitation des learning-paths inter-enfants et inter-langues observables (Vihman, 1993, Fikkert et al. 2004, Fikkert et al. 2006).

Les études longitudinales systématiques dans des langues variées ainsi que l'essor de travaux typologiques, le développement d'outils permettant la constitution de bases de données plus larges et phonologiquement analysées (Rose et al. 2005), mettent au jour des différences individuelles qui avaient été pointées depuis longtemps, mais qu'on avait considérées comme négligeables. Parmi les plus connus on retiendra la distinction devenue classique entre enfants prosodiques et enfants segmentaux (Waterson, 1987), l'existence de phénomènes de régression dits de « courbe en U ». Une conception envisageant, tel que le fait la grammaire générative, la mise en place de la grammaire, dans un organisme passif et comme une série linéaire de synchronies est beaucoup plus embarrassée de tels faits empiriques qui, par conséquent sont souvent écartés des données observées.

Ces phénomènes illustrent également un autre aspect de l'acquisition linguistique : l'existence d'une sensibilité différentielle des enfants à l'input, reflétant des logiques développementales variables conditionnées aussi par des motivations idiosyncrasiques, des phénomènes attentionnels ainsi qu'une disponibilité cognitive à l'input qui n'est pas prioritairement contrainte par la fréquence des occurrences entendues. Veneziano (2000) précise relatant les études sur le LAE (Langage Adressé à l'Enfant) et l'évolution de la manière d'envisager le rapport entre LAE et acquisition du langage:

« La position épistémologique sous-jacente aux études est passée d'un modèle simple ou l'input, par le fait même de présenter des caractéristiques facilitatrices, influence directement

---

<sup>45</sup> Ce qui a amené à considérer très souvent que les phénomènes observables chez les bébés américains étaient universels. L'un des meilleurs exemples est celui de l'hypothèse du « biais trochaïque » universel considéré d'abord comme principe de la GU, puis paramètre par défaut (le cas marqué étant « biais iambique ») y compris pour des langues où la notion de pied métrique est non pertinente. Pour une analyse détaillée cf. Fikkert (1994), Braud (2003), ChII du présent mémoire.

l'enfant, à un modèle dans lequel, entre les caractéristiques du LAE et les progrès langagiers, il y a le sujet apprenant. [...]

Même si certaines caractéristiques du LAE sont facilitatrices, il se peut qu'il n'y ait pas de relation linéaire entre leur quantité et les progrès langagiers. En outre, au-delà d'une certaine quantité (qui pourrait être relativement petite), il pourrait ne pas y avoir d'effets supplémentaires. Dans la théorie de « l'événement rare » de K.E Nelson (1981, 1987), peu de matériel, fourni au bon moment, peut suffire pour faire progresser l'enfant »

Edy Veneziano, *Interaction, conversation et acquisition du langage dans les trois premières années*, in M. Kail & M. Fayol (eds), *L'acquisition du langage*, vol1, p. 249-250.

La conception de l'acquisition phonologique comme résultant d'une organisation passive du système cognitif sous l'influence essentielle de la GU et /ou de l'input d'entrée, quel que soit le rôle qu'on leur attribue respectivement sont, dans la tradition cognitiviste, toutes consécutives de l'autre rejet théorique chomskyen, moins violent mais tout aussi radical que l'exclusion du comportementalisme, qui fonde la conception cognitiviste chomskyenne du langage : le rejet de la psychogenèse piagétienne et de l'ontogenèse linguistique qui en découle.

Pour Piaget (1964), le développement linguistique de l'enfant est un processus intégré dans un schéma développemental général (intégrant la perception, la motricité, les différents aspects de la cognition). En cela, le langage est une expression parmi d'autres des facultés cognitives humaines. Il est envisagé comme une manifestation massive de la fonction symbolique.

Les apprentissages linguistiques ne sont donc pas spécifiques et se font par le recours à des mécanismes d'assimilation (mécanismes très généraux qui permettent l'intégration de tout élément nouveau dans les structures mentales de l'enfant sur le plan moteur, perceptif, linguistique, intellectuel) et d'accommodation (adaptation de l'organisme et modification des schèmes de l'enfant sous l'influence de l'information nouvelle qui vient d'être assimilée). Ces deux mécanismes complémentaires et indissociables permettent l'équilibre du sujet dans son milieu, guidé par ses facultés d'adaptation. Le développement de l'intelligence est conçu comme une recherche d'équilibres successifs de plus en plus stabilisés.

A une conception modulaire et opaque de la grammaire entraînant une définition passive de l'enfant subissant l'émergence en lui d'une connaissance innée déclenchée par son environnement, Piaget oppose un organisme auto-régulé en équilibre avec son milieu et construisant ses représentations mentales sous l'influence de facteurs multiples tout à la fois externes (input, interactions etc) et internes (développement cognitif et linguistique).

Cette conception de l'acquisition ne remet donc pas fondamentalement en cause l'innéité de facultés humaines permettant l'acquisition du langage. Lors de la fameuse rencontre de Royaumont mettant en scène Chomsky et Piaget, lors de son discours inaugural (Piattelli-Palmarini, 1977), Piaget concède assez vite qu'il puisse y avoir des contraintes internes innées fortes soutenant l'émergence du langage. Mais elle remet en cause tout à la fois, la modularité du langage et avec elle la spécificité des mécanismes d'acquisition de cette fonction cognitive et leur opacité à l'égard de l'environnement (l'organe-langage se développant comme les bras, *RR*, 1980).

Ecartant Piaget, Chomsky écarte avec lui toute la problématique développementale, les notions d'émergence et d'équilibre du système cognitif au sein d'un environnement multimodal, ainsi qu'une conception constructiviste et dynamique de l'acquisition. On peut remarquer que ce rejet s'appuie sur deux arguments qui sont des arguments conçus à partir de la réflexion concernant l'acquisition de la syntaxe : la pauvreté du stimulus et la récursivité.

Or ces deux notions n'ont que peu de pertinence quand il s'agit de définir en quoi la phonologie est ce qu'elle est, puisque nous l'avons vu, le stimulus phonologique n'est pas quantitativement pauvre et que la récursivité n'existe pas en phonologie (Scheer, 2004b).

L'opposition à Piaget – et tout particulièrement à l'idée centrale de Piaget selon laquelle l'émergence du langage manifeste d'abord l'émergence de la fonction symbolique- résulte donc selon moi d'un biais de pensée qui, dans la grammaire générative, place la syntaxe au centre de la grammaire, pose les représentations et les procédures syntaxiques comme paradigmes des représentations et procédures de toute

la grammaire et par conséquent propose que l'acquisition de la syntaxe reflète le processus d'acquisition linguistique dans son ensemble.

Or ceci me semble faux. Pour un enfant entrant dans le langage, la langue qu'il va devoir s'approprier ne me semble pas devoir être conçue d'abord comme une syntaxe ou sur le modèle formel de la syntaxe ou selon les principes organisationnels de la syntaxe. L'acquisition d'une langue et la possibilité de l'utiliser pour un apprenant ne reposent pas me semble-t-il prioritairement sur la capacité du locuteur à ordonner les mots ou les morphèmes dans une phrase, mais sur sa capacité à catégoriser et à réinterpréter le monde à partir d'unités discrètes associant de manière arbitraire une structure phonique et une structure significative au sein d'un système.

Cette dimension fondamentale de la spécificité systémique des langues humaines constitue me paraît centrale pour l'acquisition. Comme je l'ai souligné elle est sans aucun doute l'une des idées fondatrices majeures de la linguistique scientifique contemporaine, développée et reformulée régulièrement par toute la linguistique structuraliste. J'en retiendrai la formulation de Boltanski, qui, justifiant la conception du langage comme « objet abstrait » souligne

« Que signifie ici « objet abstrait ». Simplement qu'une langue, toute langue, si on prend ne prend pas en compte les substances (le son, le sens) qu'elle met en relation, peut être vue comme une pure structure. Si on admet que ne sont linguistiques que les différences phoniques qui ont un effet dans le domaine conceptuel, il va de soi que ne sont linguistiques que les différences conceptuelles qui ont un effet dans le domaine phonique, ou encore si une langue, comme le montre l'observation la plus élémentaire, ce sont des différences de son qui renvoient à des différences de sens, ou de façon équivalente, des différences de sens qui renvoient à des différences de son, une mise à distance, mettant provisoirement hors jeu les substances concernées, offre la possibilité de saisir la langue (le langage) comme un objet abstrait au sens de « défini » en termes structuraux ». Si cette logique a mis longtemps à être reconnue (comme chacun le sait c'est dans le CLG de Saussure 1915 qu'elle est clairement mise en évidence pour la première fois), c'est qu'elle est paradoxale : chaque domaine est structuré de l'extérieur pour l'autre ; c'est par son entremise qu'il acquiert les unités qui le constituent et cela est vrai de chacun d'eux, puisque dans chaque domaine, les unités sont créées par le contact avec l'autre »

J.E Boltanski, *La révolution chomskyenne et le langage*, 2002, p.125-126.



Il me semble qu'apprendre à parler pour un petit enfant consiste à conquérir, à comprendre, à maîtriser cet « objet abstrait paradoxal » et unique dans le monde sensible, qui s'appuie essentiellement sur l'existence dans le langage de la fonction symbolique manifestée par une structure doublement articulée. C'est d'ailleurs ainsi que Benveniste (1966), reprenant Martinet (1949, 1959) définit la spécificité intrinsèque des langues humaines.

Cet argument plaide très clairement en faveur d'une acquisition de la phonologie et plus généralement du langage par le lexique. Car la phonologie peut bien être envisagée comme « quelque chose d'abstrait » qui arrive à l'enfant sans qu'il en ait conscience, à partir d'un système de traitement spécifique de l'input et sur la base d'une connaissance structurelle innée, mais ce phénomène cognitif émerge à partir d'un autre objet linguistique motivé et acquis volontairement, qui matérialise la relation signifiant / signifié / référent et le paradoxe qu'elle constitue.

C'est également ce qui m'a amenée à retenir le modèle d'acquisition de la phonologie par un gabarit lexical proposé par Vihman (2001) sur la base du même argument. Vihman a été l'une des premières à systématiser la description structurale des données de langues variées au sein d'une unité reflétant l'existence d'une relation signifiant / signifié / référent, sur laquelle les enfants s'appuient pour construire leur système. Nos opinions divergent quant aux hypothèses concernant la manière dont les enfants procèdent pour y arriver.

« The hypothesis that we propose, following many others, is that the starting point for the analysis of linguistic structure should be the sound-meaning link that defines linguistic signs or symbols. This hypothesis does not rule out the possibility that generalizations about linguistic structure, including phonological structure, may be separated from generalizations about their function. Indeed, there is much arbitrariness in language, most notably the arbitrariness of the association of a phonological form with a particular meaning in a particular language. Also, as is well known, the phonological organization of a word into syllables often fails to match the morphological composition of a word. But we will argue below that the basic phonological unit is a word template, specifically defined on a phonological unit that is also a fundamental symbolic unit. We will argue that starting from words can solve certain

theoretical and empirical problems that arise for reasons not directly connected to language function and, furthermore, that this reflects the developmental learning sequence »

M. Vihman & W. Croft, *Phonological development: Toward a 'radical' templatic phonology*, (à paraître), p. 3

#### **4.2. Acquisition phonologique et lexicale : articuler son et sens**

Comme nous l'avons souligné les recherches les plus récentes en perception corroborent ce qui avait été décrit auparavant sur la base d'observations attentives (Grégoire, 1937) : les enfants manifestent des aptitudes lexicales bien plus précoces en perception qu'en production (Vihman, 1996, Boysson-Bardies, 1996, Peperkamp, 2002, Hall & Waxman, 2004).

Si, l'on reprend les échelles chronologiques mettant en regard perception et production (cf. tableau 11, tiré de Wauquier-Gravelines, 2004), à 1 an environ, alors qu'il reconnaît déjà un mini lexique d'une centaine de mots, l'enfant commence à peine à stabiliser le babillage varié et ne produit encore véritablement que quelques rares unités lexicales interprétables pour entourage en contexte (papa, maman, non, son prénom).

Les opérations de segmentation pendant la première année lui ont permis, nous l'avons vu (cf. 3.2) de dégager un lexique d'étiquettes phonétiques à partir d'informations acoustiques variées et saillantes (accents, clauses prosodiques, mots grammaticaux. Pour une synthèse cf. Echols & Marti, 2004).

Divers travaux ont essayé par ailleurs d'établir également quand et comment les enfants sont capables d'associer à l'unité dégagée un référent extra-linguistique.

Comme pour la détermination des formes phonétiques des mots, l'opération de référence consiste en une segmentation progressive de l'environnement qui permet d'extraire peu à peu des invariants correspondant aux étiquettes sonores qu'on leur associe régulièrement. Bruner propose ainsi que l'acquisition lexicale se fasse d'abord en contexte à partir de « formats » (Bruner, 1975), schémas d'action structurés où les situations et les mots qui y sont associés ont un haut degré de prédictibilité (par exemple « tiens, prends le canard » dans le contexte du bain quotidien de l'enfant).

Ceci permet à l'enfant de développer un comportement adapté à la situation et aux informations verbales qui lui sont adressées. Qu'il soit placé dans un « format » libère donc cognitivement l'enfant et lui permet d'expérimenter des hypothèses linguistiques concernant la mise en relation référentielle entre objets, actions, comportements et étiquettes phonétiques. Certains travaux plus récents ont considéré qu'effectivement les premières opérations de référence s'appuyaient sur des heuristiques et approximations en contexte (Landau, 1984) en utilisant des facultés cognitives générales de traitement de l'information sensorielle (masse, poids, couleur, forme) qui reflète une sensibilité précoce des bébés à la permanence des objets, à leurs contraintes de solidité, de continuité, (Baillargeon, 1991, Baillargeon et al. 1985, Spelke et al. 1994, pour une synthèse cf. Landau 2004). Ce point de vue a été contesté par d'autres travaux considérant que l'acquisition lexicale et en particulier l'opposition nom / verbe ne pouvait pas être acquise sans bootstrapping syntaxique (pour une synthèse, cf. Echols & Marti, 2004).

En production, on considère de manière générale (même si cette conception a pu être contestée, Bloom, 2004) que l'acquisition lexicale se fait en deux temps<sup>46</sup>. Après la phase de babillage mixte où les enfants mélangent des suites de syllabes, la production des premiers mots reconnus par les adultes commence entre 12 et 14 mois et ouvre la période dite « des 50 mots ». L'accroissement de ce vocabulaire est lent, les mots sont réalisés avec beaucoup de variabilité et non stabilisés. En d'autres termes peu distinctifs et d'emplois très limités (Boysson-Bardies, 1996). Vihman (1996, 2000, soumis) propose que la sélection de ces 50 premiers mots soit faite par l'enfant à partir d'un nombre de paramètres variés : sélection idiosyncrasique parmi les mots qu'il reconnaît et maîtrise en perception, choix de patterns articulatoires préférentiels exprimés dans le babillage, fréquence lexicale et fréquence de distribution des unités lexicales dans l'input, attention partagée, focalisation, pointages déictiques sur des objets partagés avec les adultes entrant en interaction avec l'enfant. Ces mots sont effectivement le plus souvent déictiques (maman, papa,

---

<sup>46</sup> Je commente ici un cas « théorique » reflétant un pattern d'acquisition « standard » si cette expression a du sens. Il est évident que, comme je l'ai dit plus haut, la chronologie ainsi que le rythme d'acquisition et la forme des premiers mots produits varient considérablement d'un enfant à l'autre.

prénom de l'enfant, de l'animal familier, du doudou etc.) ou utilisés en contexte dans un cadre familier et accompagnés d'un geste déictique de pointage (Vénéziano, 2000) et spécifiques pour chaque enfant. A la suite de cette période qui dure entre 4 et 5 mois environ, apparaît ce qu'on appelle l'explosion lexicale. L'augmentation à la fois quantitative et qualitative du lexique est assez brutale avec une acquisition d'environ 120 à 150 mots en moyenne dans les 4 mois qui suivent (22 mois), puis de 200 mots environ dans les 4 mois suivants (26) et 300 environ au début de la troisième année. Selon les individus, on peut considérer qu'à trois ans les enfants ont un vocabulaire allant de 400 à 700 mots<sup>47</sup>.

Or, à trois ans, la plupart des enfants sont fluents, parfaitement compréhensibles et ont acquis en grande partie la phonologie de leur langue et pour certains également une bonne partie de la syntaxe. C'est donc entre 12 et 25-30 mois que l'on peut voir émerger les phonèmes, les structures prosodiques et syllabiques et que l'on peut observer les phénomènes d'acquisition phonologiques couramment répertoriés dans la littérature (harmonies consonantiques, métathèses, gliding, reduplications, troncations etc (cf. tableau 11 extrait de Wauquier-Gravelines, 2005).

---

<sup>47</sup> Ne dérogeons pas à la tradition de l'anecdote que rapporte toujours le linguiste de terrain et particulièrement en acquisition où les productions des enfants sont régulièrement source d'hilarité ou d'attendrissement collectif quand on les présente aux étudiants en cours : C. (29mois) avait encore un sygmatisme ou seveu sur la langue et quand on l'a enregistré il a donc produit [esato] pour [ɛfato], et quand je lui ai demandé, « mais je ne connais pas ce mot, je ne sais pas ce que c'est [esato] », il m'a répondu excédé « mais pas [esato], [e.sa.to] », ce qui laisse penser que cet enfant disposait à cet âge d'une phonologie structurée, autonome et substance-free.

PERCEPTION	PRODUCTION
0-1 mois perception catégorielle préférence voix de la mère/ langue maternelle sensibilité voix (timbre, chant) et prosodie	0-1mois Sons végétatifs, pleurs, Expression malaise / confort
1-5 mois catégorisation contrastes phonétiques discrimination schèmes intonatifs reconnaissance de syllabes	1-5 mois rires, cris, expressions vocales émotions jeux vocaux, arrehu début contrôle de la phonation
5-7 mois catégorisation voyelles détection des clauses prosodiques grands constituants début perception intermodale	5-7 mois vocalisation variation, imitation intonative trilles
8-10 mois détection frontières syntagmes sensibilité à l'accent début reconnaissance mots connus en contexte	8-10 mois voyelles langue cible babillage canonique (bababa) puis babillage varié orienté langue cible contours intonatifs langue cible
10-12 mois détection frontières mots surdité phonologique reconnaissance mots connus hors contexte compréhension mots (+/- 30 mots) avec référents	10-12 mois syllabes langue cible simplifiées CV/V babillage varié et intonné production formes stabilisées en contexte premiers mots (maman, dodo...)
12-16 mois compréhension +/- 100 mots	12-16 mois stabilisation des formes de babillage varié complexification des structures intonatives 16 mois : stade dit des « 50 mots », noms début acquisition phonologique (réductions, troncations, harmonies, substitutions, )
16-20 mois compréhension +/- 200 mots début sensibilité classes morpho-syntaxiques	16-20 mois 50 / 150 mots premiers verbes acquisition phonologique : pleine période
20-26 mois compréhension mots début parsing syntaxique	20-26 mois explosion lexicale accélération de l'acquisition phonologique et début stabilisation des formes lexicales
26-48 mois compréhension tous mots concrets même hors contextes + mots nouveaux en contexte traitement syntaxique	26-48 mois stabilisation des formes lexicales fin acquisition de la phonologie acquisition morpho-phonologie

Tableau 11 : Principales étapes de l'acquisition phonologique

Pour ce qui nous concerne directement, on peut donc considérer qu'autour de 14-15 mois, les enfants sont dotés d'un mini lexique leur permettant de comprendre et de reconnaître une centaine de mots et d'en produire une cinquantaine et qu'à cette période les mots sont peu distinctifs. Les travaux en perception (pour une synthèse, cf. Werker & Pegg, 2004) ont par ailleurs montré qu'à 14-15 mois les paires minimales ne sont pas distinguées. Ceci laisse donc supposer que ni le traitement perceptif de la première année, ni l'utilisation de ce mini lexique ne reflètent qu'à ce stade, les enfants ont construit des représentations lexicales phonologiquement conditionnées. Tout laisse supposer que ces premières représentations au stade des 50 mots sont encore des étiquettes phonétiques mémorisées et pertinentes exclusivement en contexte.

Les travaux récents en neuro-imagerie (pour une synthèse cf. Houdé et al. 2002) montrent à partir des modèles distribués de Damasio (1996) et du modèle en épicentre de Mezulam (1998) que les représentations linguistiques et lexicales sont distribuées sur plusieurs régions corticales, et que les aires de Broca (fortement impliquée lors de la mise en œuvre de la mémoire de travail) et de Wernicke ne constituent pas tant des zones de production / perception que des « épicentres » permettant l'intégration dynamique en temps réel de représentations fragmentées. La mise en évidence de l'activation de la zone temporo-basale associée à une absence d'activité dans la zone Broca lors de tâches de dénomination d'images est interprétée par Houdé et al. (2002) comme la possibilité de l'existence d'une mémorisation spécifique, dans cette zone, des mots fréquents sans aucune ambiguïté sémantique qui auraient été stockés précocement lors de l'acquisition sans décomposition phonologique.

« Il pourrait donc subsister, dans l'organisation cérébrale de l'adulte, un réseau neural spécifiquement dédié à la dénomination d'objets surappris durant l'enfance, organisé autour de relations univoques entre l'objet et le nom »

O. Houdé, B. Mazoyer, N. Tzourio-Mazoyer, *Cerveau et psychologie*, 2002, p. 467.

S'il est vrai qu'à 14 mois les enfants commencent à produire sur la base d'unités mémorisés sans décomposition phonologique, la question qui m'intéresse est donc celle de la « phonologisation » des représentations lexicales, de l'apparition et des catégories abstraites (syllabes et phonèmes) dans la langue cible et de la logique systémique qui préside à leur organisation.

En admettant que jusque là les données reflètent l'absence de nécessité de l'utilisation de GU, et que les enfants aient procédé sur base inductive à partir de mécanismes cognitifs généraux, comment la phonologisation peut-elle se faire sans connaissance phonologique spécifique ?

Les mécanismes inductifs (et quels sont-ils ?) appuyés sur des capacités cognitives générales et non spécifiquement linguistiques sont-ils suffisants?

Vihman et Velleman (2000), Vihman & Croft (à paraître) suggèrent une progression de l'enfant à partir d'ajustements progressifs entre les patterns dont il dispose (les quelques mots mémorisés et produits) et ce qu'il perçoit des productions adultes. Par approximations progressives, il finit par atteindre les réalisations complètes et bien formées de sa langue. Vihman et Velleman (2000) suggèrent que les enfants procèdent par inférence et comparaison tant sur l'axe paradigmatique que sur l'axe syntagmatique entre ce qu'ils savent et ce qu'ils entendent et que ces tentatives d'ajustement progressive finissent par dégager les unités pertinentes. Le problème qui se pose alors reste celui de la genèse des catégories, les enfants peuvent-ils sur un lexique de 50 mots produits et 100 mots perçus et par comparaison avec l'input inférer la structure phonologique de leur langue ?

Un premier phénomène permet d'éclairer la question, celui de l'explosion lexicale : pourquoi la généralisation systémique apparaît-elle au moment de l'explosion lexicale, c'est à dire à partir de l'expansion exponentielle du lexique qui est accompagnée de la mise en place elle-même exponentielle des représentations phonologiques. Si l'on considère que l'enfant opère les généralisations à partir des quelques 100 mots disponibles en perception ou plus sûrement sur les 50 mots du lexique produit, comment expliquer la simultanéité brutale des deux phénomènes et comment expliquer ce saut quantitatif et qualitatif opéré à cette période. Pourquoi les enfants deviennent-ils simultanément et brutalement aussi productifs et ce, au moment même où ils distinguent les paires minimales en perception et où les erreurs de production témoignant d'une activité de généralisation phonologique apparaissent et se résolvent en quelques mois ?

L'hypothèse de la disponibilité chez l'enfant d'un système spécifique de traitement linguistique de l'input appuyé sur des primitives phonologiques, dégagées et définies au moment des premières associations structurelles systématiques entre son et sens apporte la réponse à cette question.

Si l'on reprend la définition que Boltanski fait de la langue comme « objet abstrait », on peut considérer que c'est par l'association de l'empreinte phonétique avec un signifié, que se matérialise le niveau « abstrait » auquel se construit la

représentation phonologique du mot, puisque c'est dans ce rapport de complémentarité mutuelle que le signifiant et le signifié se définissent mutuellement. Au moment où l'enfant commence à réaliser une interprétation référentielle symbolique du monde extra-linguistique, l'empreinte sonore du mot stocké en mémoire sous la forme d'une étiquette phonétique devient, par son association avec un signifié, un signe, disponible en tant qu'elle est constituée de primitives phonologiques. On peut considérer que ces informations phonologiques sont en réserve de leur révélation ou émergence par le contact avec le signifié. A partir du moment où l'enfant réalise cette opération sur le proto-lexique d'étiquettes stockées telles quelles, il réinterprète chaque unité du proto lexique et peut découvrir la logique systémique à l'oeuvre dans sa langue et commencer des généralisations, à partir des mots dont il dispose. Ceci entraîne logiquement une augmentation rapide du lexique permettant en retour l'accélération des processus de généralisation.

C'est en associant signifiant et signifié et en réalisant l'opération paradoxale de découpage du réel décrite par Boltanski, que seul permet le langage humain, que l'enfant peut dès lors découvrir que le système sonore de sa langue est abstrait, structurel et distinctif et construire un niveau de représentation spécifiquement phonologique, abstrait et « substance free »<sup>48</sup>.

Le passage du niveau analogique au niveau discret ne me semble pas pouvoir être réalisé uniquement à partir d'un calcul statistique qui n'enregistre que des régularités phonétiques, sans le recours à des primitives phonologiques pré-existantes, donc innées rendues disponibles par l'association du signifiant avec le signifié lors de l'opération de référence.

Si l'on écarte le recours à une procédure linguistique de ce genre (la référence et la construction du signe ainsi que la disponibilité des objets phonologiques qui en résulte) permettant l'accès à un niveau de représentation abstrait où l'enfant va construire ses représentations à partir de ses propres expériences motrices et sensorielles, les phénomènes observables ne peuvent être analysés que de la manière suivante : l'émergence de la phonologie est consécutive soit à une augmentation

---

<sup>48</sup> Au sens où j'en ai défendu l'existence chez l'adulte au chapitre III.



significative de la mémoire qui donne brutalement une surcapacité de stockage permettant de faire des inductions sur un corpus beaucoup plus large (l'enfant procède alors comme le postule Chomsky de la même manière que le linguiste), soit à un changement de type de stockage ou à un changement de traitement de l'information lexicale.

C'est l'hypothèse retenue par Vihman & Mueller Gathercole (en préparation). Elles proposent en effet une hypothèse inductive permettant la phonologisation du lexique sans recours à des primitives phonologiques innées mais reposant sur des capacités d'apprentissage innées très sophistiquées

« In discussing the polarization of researchers concerned with language development into empiricists and nativists Braine (1994) came to the logical conclusion that any scientist wishing to account for language learning *without* positing specifically linguistic innate knowledge (i.e., without recourse to the concept of Universal Grammar) would have to posit, instead, powerful 'innate learning mechanisms' »

Appuyant leur hypothèse sur l'existence chez l'être humain de deux systèmes d'apprentissage, l'un implicite (statistique et procédural, non attentionnel)<sup>49</sup> et l'autre explicite (symbolique, attentionnel)<sup>50</sup>, correspondant à deux sites neuronaux de traitement de l'information et de stockage de la mémoire dans l'hippocampus, elles proposent une émergence en trois temps des représentations phonologiques

« We suggest that *three types of learning* should be distinguished for first language learning. Two of these three are the product of 'implicit' processing by the 'non-declarative' system (Squire & Kandel, 1999); the third requires processing by the declarative system. The three types of learning are the following:

- (1) procedural tallying of regularities perceived in sensory data of any kind (e.g., of frequencies of occurrence, sequencing, or rhythmic patterning),
- (2) declarative (categorical, symbolic) registering of arbitrary form-meaning cooccurrences or associations, leading to a mental lexicon of *linguistic items*, and

---

<sup>49</sup> Et j'ajouterais, tel qu'elle le décrit, « algorithmique »

<sup>50</sup> et j'ajouterais, pour le coup, sans doute « heuristique »

- (3) 'secondary' procedural induction (a kind of 'tallying') of the regularities inherent in the linguistic items registered in the mental lexicon, ultimately resulting in abstract knowledge of linguistic system. »

Tout bien considéré, il m'apparaît que cette proposition est assez proche dans une formulation psychologique et neurologique, des hypothèses que j'ai faites ci-dessus dans une formulation linguistique et structuraliste, à partir d'ailleurs de ma lecture conjointe de Boltanski (2002), Carvalho (1999, 2002)<sup>51</sup> ainsi que les propositions de Vihman & Velleman (2000), présentant un modèle de gabarit lexical permettant l'émergence de la phonologie par induction et approximation successive de l'enfant. J'avais cherché à comprendre et à définir, à partir Vihman & Velleman (2000) comment l'enfant procédait pour réaliser ce que Vihman modélise beaucoup plus explicitement ici : le passage à un niveau de représentation discret à partir d'une information analogique. Et il m'avait semblé effectivement que l'interface entre ces deux niveaux de représentation et la possibilité de faire émerger l'un de l'autre passait par le lexique et plus précisément par l'opération de « référence » que suppose la connaissance lexicale<sup>52</sup>.

La proposition de Vihman & Mueller Gathercole qui nécessite, à ce stade, une modélisation psychologique plus détaillée<sup>53</sup> me semble très prometteuse et très intéressante du point de vue du réalisme cognitif. Elles esquissent très clairement la construction d'interfaces articulées et modélisées à un degré théorique sophistiqué entre phonologie théorique abstraite et psychologie / neurologie, et décrivent des mécanismes cognitifs innés et spécifiquement humains permettant d'opérer des généralisations sur l'input linguistique.

---

<sup>51</sup> Je discute la question de l'articulation possible de cette hypothèse de genèse des catégories avec le squelette de positions pures d'Encrevé (1988) et la reformulation émergente qu'en fait Carvalho (1999, 2002) au chapitre III

<sup>52</sup> Je n'ignore pas que la référence peut se manifester par des constructions syntaxiques. On peut même penser que dès la phrase de deux mots qui manifeste la relation « thème », « prédicat » (par exemple : « le chien, dodo »), chien est référencé aussi par « dodo » et la relation syntaxico-sémantique entre ces deux constituants. Mais les premières productions « interprétables » par l'environnement des enfants sont toujours des mots isolés et non des phrases.

<sup>53</sup> Je rappelle que ces citations sont issues d'un « draft » de chapitre que M. Vihman est en train de rédiger et qu'elle a eu l'extrême gentillesse de m'envoyer et qu'elle m'a permis de citer en l'état. Le texte sera peut-être modifié et davantage élaboré avant sa sortie.

Mais ma position diverge de la leur quant à la conception qu'elles ont de la troisième étape (secondary procedural induction) qu'elle définissent de la manière suivante :

Thirdly, once a child has begun to establish a lexicon of words or phrases with both phonological form and semantic content, 'secondary' procedural or distributional learning will automatically occur, as the neocortex again goes to work on recurrent regularities (3). The input to that implicit or procedural processing is now no longer at the level of direct perceptual input but is rather a representational derivative of the cognitive processing that created the lexical entries – hence the term 'secondary'. As Karmiloff-Smith (1992) has put it, 'a specifically human way to gain knowledge is for the mind to exploit internally the information that it has already stored..., by redescribing its representations or, more precisely, by iteratively re-representing in different representational formats what its internal representations represent' (p. 15). Beyond that difference in 'raw material' or input to the processor, however, the learning process itself may be assumed to be the same as in (1). This secondary procedural learning can now be understood as gradually building up the abstract knowledge of system or structure to which we generally apply the term 'grammar' (Pierrehumbert, 2003; Beckman & Pierrehumbert, 2003). This kind of pattern induction may be sufficient to account for the learning of the abstract categories that are the basis for phonological and syntactic systems without the need for innate 'foreknowledge' of linguistic structure in the form of Universal Grammar »

Plusieurs questions (se ramenant peut-être une fois de plus au problème de la genèse des catégories abstraites et de la logique systémique qui préside à leur organisation) se posent ici.

Si cette troisième étape procédurale re-traite explicitement une représentation lexicale résultant d'un traitement probabiliste implicite de l'information, sur quelles unités se fait exactement ce calcul procédural de deuxième type ?

Le passage du niveau analogique au niveau abstrait est formulé d'une manière assez peu précise et comme en creux par un changement de format représentationnel (« re-representationing in different representational formats »). Les unités « phonétiques » stockées par apprentissage probabiliste deviendraient « phonologiques » par réinterprétation entraînant un changement de format. Ceci est effectivement une manière de rendre compte de la « discrétisation » ou de la catégorisation symbolique qu'on peut observer pour n'importe quelle autre information physique (comme une

longueur d'onde lumineuse qui selon la manière dont elle est formatée peut être perçue analogiquement (dans un arc-en-ciel ou lors de la décomposition spectrale dans une goutte d'eau) ou discrètement (une couleur = le rouge).

Mais pour ce qui concerne les unités phonologiques le problème ne peut se formuler de cette manière : changeant de format, elles changent de nature car elles deviennent phonèmes non seulement parce qu'elles deviennent abstraites mais surtout parce qu'elles deviennent des unités distinctives doublement articulées dans le système. Le fait d'être discrétisées ne modifie pas la nature des longueurs d'ondes lumineuses alors que le fait d'être discrétisés change la nature des sons linguistiques. Ce faisant, c'est la fonction qu'elles acquièrent dans la langue qui devient définitoire des unités elles-mêmes. Or ce n'est pas le cas pour les couleurs. Un bébé, peut avoir une idée du rouge indépendamment de la connaissance des autres couleurs fondamentales existant dans le spectre (bleu /jaune). Il opérera ensuite des catégorisations plus fines (rouge vermillon, orange, grenat, rose ...) mais il peut connaître et manipuler sans difficulté un système de couleur où il manque une fondamentale et ne pas savoir que la catégorie « orange » résulte de la fusion de deux primitives. Pour les sons la réalité est différente, la connaissance d'un système phonologique suppose qu'on connaisse les sons qui le constituent mais surtout les relations d'opposition et de constituance qui président à leur organisation dans la langue. Le système de traitement procédural intervenant lors de cette troisième étape nécessiterait donc non seulement une catégorisation des objets individuels mais également une conceptualisation simultanée du calcul procédural permettant de catégoriser les relations qu'entretiennent ces objets.

Par ailleurs, le changement de format qui permet de passer d'un objet physique à une unité abstraite (le rouge) par exemple crée une catégorie qui en linguistique serait comparable (en terme de définition et de fonction) au signe linguistique au sens saussurien du terme associant l'idée d'une expérience sensorielle (ici visuelle) correspondant au signifiant, et l'idée de la couleur qui lui est associée, correspondant au signifié. Mais, comme l'a proposé Martinet (1949, 1957), le langage humain est unique en ce qu'il est doublement articulé et que la phonologie d'une

langue repose sur l'existence d'unités, se définissant exclusivement par leur fonction (à l'empreinte sonore /p/ n'est pas associée l'idée de /p/ qui n'a pas en soi de pertinence puisque le référent extra linguistique de /p/ est [p], un objet sonore). Si ce traitement procédural de deuxième type, n'utilise aucune primitive phonologique, il me semble que nous retombons là sur la question de la poule et de l'œuf car on ne peut inférer la nature distinctive des sons, et donc dégager les phonèmes, que par le recours à la commutation dans une paire minimale, ou au moins par comparaison d'items lexicaux, mais on ne peut savoir ce qu'est une paire minimale si l'on ne sait pas ce qu'est structurellement un phonème.

Enfin, la proposition de Vihman prédit que les catégories peuvent apparaître dans n'importe quel ordre puisque leur acquisition n'est conditionnée par aucune logique structurelle. En ce sens, si l'on reprend l'analogie avec les couleurs, un enfant peut très bien commencer, sans préjudice aucun à apprendre les couleurs en commençant par le rose, le vert et le violet, et chaque enfant les apprendra dans un ordre idiosyncrasique, conditionné par son environnement et l'apprentissage explicite qu'on lui propose (un enfant apprend les couleurs explicitement, il ne les découvre pas seul). Il n'en va pas de même avec les sons dont l'ordre d'apparition n'est pas aléatoire et est conditionné par une logique systémique. Les erreurs répertoriées dans la littérature (métathèses, harmonies) ne sont pas non plus aléatoires. Les enfants par exemple n'inversent jamais occlusive liquide dans une attaque branchante et n'acquièrent, dans aucune langue les fricatives ou les sonantes avant les occlusives.

## **5. Conclusion**

Dans toutes les modalités, les opérations cognitives qui s'appliquent à l'environnement sensoriel permettent une réinterprétation structurée du monde sensible par un filtrage de l'information reposant sur une organisation mentale spécifique qui permet de lire structurellement l'environnement. C'est une opération psychologique primordiale, reposant sur un système unique de discrétisation que seuls les humains possèdent et sur l'existence de la fonction symbolique (s'exprimant également dans d'autres productions humaines comme la peinture par exemple).

Cependant le langage matérialise la fonction symbolique de manière spécifique et singulière puisqu'il repose sur un nombre d'éléments primordiaux purement distinctifs qui ne prennent sens que les uns par les autres. A cet égard, on peut penser que l'observation des procédures de traitement phonologique en temps réel comme l'observation de l'acquisition de la phonologie consistent à définir respectivement comment un sujet parlant utilise la distinctivité et la double articulation en temps réel et comment un bébé la met en place.

J'ai retenu dans mes travaux l'hypothèse d'une émergence de la phonologie liée à l'émergence du lexique en production, effectivement observable au moment de l'explosion lexicale, après la phase des 50 mots. En accord avec Vihman & Velleman, Vihman & Mueller-Gathercole, on peut effectivement envisager que ces phénomènes reposent sur des mécanismes d'apprentissage sophistiqués et procéduraux il me semble que les catégories phonologiques ne peuvent émerger sur la seule base de ces procédures, sans recours à des primitives phonologiques innées et à des opérations linguistiques spécifiques (référence, construction du signe) conditionnées par la structure même du langage.

La phonologie ne me paraît donc être ni donnée de l'extérieur par un input particulièrement riche, ni reçue de l'intérieur par la magie de représentations abstraites universelles génétiquement transmises, mais quelque chose que construit progressivement l'enfant, placé dans un certain milieu à partir d'un substrat neurobiologique inné, spécifiquement humain et orienté vers le langage, qui sur la base d'un nombre de primitives minimales lui permet de traiter efficacement, pour en extraire représentations abstraites, un objet unique dans le monde sensible : une structure sonore doublement articulée

**Sophie Wauquier-Gravelines**

**addendum au mémoire d'HDR, « Statut des représentations phonologiques en acquisition, traitement de la parole continue et dysphasie développementale »**

**Erratum**

p. 17 : Stériade, 2004 => lire Steriade, 2001

p. 39 : Clements & Keyser, 1981 => lire Clements & Keyser, 1983

p 80 : Wauquier-Gravelines 2003 => lire Wauquier-Gravelines 2003 a

p 95 : Wauquier-Gravelines 2003 => lire Wauquier-Gravelines 2003 a

p 134 : Kaye & Lowenstamm 1979 => lire Kaye & Lowenstamm 1982

p.158 : Dupoux 2005 => lire Dupoux 2004

p 160 : Croft & Vihman à paraître => lire Vihman & Croft à paraître

p. 174 : Oller 2002 => lire Oller 2000

p.177 : Encrevé 1999 => lire Encrevé 1997

p178 : Jackendoff 2000 => lire Jackendoff 2002

P. 179 : Carr 2002 => lire Carr 2000

p.206 MacCoquordale 1973 => lire 206 MacCoquordale 1970

p 207 : Jackendoff 2000 => lire Jackendoff 2002

p.218 : Pierrehumbert 2000 => lire Pierrehumbert et al. 2000

p.222 : Jackendoff 2000 => lire Jackendoff 2002

p. 229 : Dupoux à paraître => lire Dupoux 2004

p.227 Mac Neilage 1998 => lire Mac Neilage 1993

p. 250 Baillargeon et al. 1985 => lire Baillargeon et al. 1991

p.252 Werker & Pegg 2004 => lire Werker & Fennell 2004

## Références

- AARSLEFF, H. (1982), *From Locke to Saussure. Essays on the Study of Language and Intellectual History*, University of Minnesota Press, Minneapolis
- ABERCROMBIE, D. (1967), *Elements of General Phonetics*, Edimburgh : Edimburgh University
- ABRY, C., STEPHANUTO, M., VILAIN, A., & LABOISSIERE R. (2002), What can the utterance « tan tan » of Broca's patient Leborgne tell us about the hypothesis of an emergent « babble-syllable » downloaded by SMA? In J. Durand, & B. Laks, (eds) *Phonetics, Phonology and Cognition*, 226-243, Oxford University
- ALLEN G. & HAWKINS S. (1978), The development of Phonological Rhythm, in A. Bell & J. Hooper-Bybee (eds), *Syllables and segments*, 173-185, Amsterdam, North-Holland
- ALLEN G. & HAWKINS S. (1980), Phonological Rhythm : definition and development, in G. Yeni-Komshian, C. Kavannagh & C. Ferguson (eds), *Child Phonology 1 : Production*, 227-256, New-York, Academic Press
- ANDERSON, S. & LIGHTFOOT, D.W. (2002), *The language Organ, Linguistics as Cognitive Physiology*, Cambridge University Press
- ANGOUJARD, J.P. (1997), *Théorie de la syllabe*, CNRS éditions, Paris
- BAILLARGEON, R. (1991), Object permanence in 3.5 and 4.5 month-old infants, *Developmental Psychology*, **23**, 589-594
- BAILLARGEON, R., SPELKE, E.S, & WASSERMAN, S. (1991), Object permanence in five month-old infants, *Cognition*, **20**, 191-208
- BARTON, D. (1978), *The role of perception in the acquisition of phonology*, Bloomington, Indiana University Linguistics Club
- BASSET E. (2000), *La liaison à 3, 7 et 11 ans : description et acquisition*, Mémoire de maîtrise, Université de Grenoble 3, ms.
- BENVENISTE E. (1966), *Problèmes de linguistique générale, tome 1*, Le Seuil, Paris
- BERNHARDT, B., & STEMBERGER, J. (1998), *Handbook of phonological development from the perspective of constraint-based non-linear phonology*, Academic Press, San Diego
- BLOOM, P. (2004), *Myths of word learning*, in G. Hall & S. R. Waxman (eds), *Weaving a lexicon*, 205-225, Cambridge, mass., The MIT Press
- BOERSMA, P. (1999), On the need for a separate perception grammar, ROA # 358-1099, Rutgers Optimality Archive



- BOERSMA, P. (2000), The OCP on perception grammar, ROA # 435-0601, Rutgers Optimality Archive
- BOHAS, G., GUILLAUME, J.P. & KOULOUGHLI D. (1989), L'analyse grammaticale dans la tradition arabe classique, in S. Auroux (ed), *Histoire des idées linguistiques*, 260-283, Mardaga, Bruxelles
- BOLTANSKI, J.E. (2002), *La révolution chomskyenne et le langage*, L'Harmattan, Paris
- BONATTI, L.L. (2002), Les représentations, la réalité psychologique et au-delà, in E. Dupoux (ed), *Les langages du cerveau*, O. Jacob, Paris
- BOTHA, R.P. (1989), *Challenging Chomsky, The generative Garden Game*, Blackwell, Oxford
- BORTOLINI, U. & LEONARD, L. (1996), Phonology and grammatical morphology in specific language impairment : Accounting for individual variation in English and Italian. *Applied Psycholinguistics*, **17**, 85-104
- BOULA DE MAREUIL, P., ADDA-DECKER, M. & GENDNER, V. (2003), Liaisons in French : a corpus based study using morpho-syntactic information, *Proceedings of ICPHS*, Barcelone, 1329-1332
- BOURDIEU, P. (1997), *Les usages sociaux de la science. Pour une sociologie clinique du champ scientifique*, INRA
- BOYSSON-BARDIES DE B., SAGART, L., & DURAND, C., (1984), Discernible differences in the babbling of infants according to target language, *Journal of Child Language*, **11**, 1-15
- BOYSSON-BARDIES DE B., & VIHMAN M. M. (1991), Adaptation to language : Evidence from babbling and first words in four languages, *Language*, **67**, 297-319
- BOYSSON-BARDIES DE B. (1996), *Comment la parole vient aux enfants*, O. Jacob, Paris
- BRADLEY D. & FORSTER, K.I., (1987), A reader's view of listening, in U. H. Frauenfelder, & L.K. Tyler, (eds) *Spoken word recognition*, *Cognition*, **25**, special issue, 71-103
- BRADLEY C & DEJEAN DE LA BATIE, B. (1990), Resolving boundaries in spoken French, in M. Seidl (ed), *Proceedings of the Third Australian International Conference of Speech Science Technology* Canberra : Australian Speech Science & Technology Association
- BRAUD, V. (1998), *Acquisition de l'information phonologique : exemple de la liaison*, Mémoire de maîtrise, Université de Nantes, ms.

- BRAUD, V. (2003), *Acquisition de la prosodie chez les enfants francophones. Les phénomènes de tronctions*, Thèse de doctorat, Université de Nantes, ms.
- BRAUD, V. & WAUQUIER-GRAVELINES, S. (1999), Acquisition de l'information phonologique en français, communication au *VIIIth International Congress for the study of Child Language*, San Sebastian,
- BRAUD, V. & WAUQUIER-GRAVELINES, S. (2004) Approche gabaritique des phénomènes de troncation du français, *Actes des Journées d'Etudes sur la Parole*, Fez
- BRESNAN, J. (1983), *In search of mind*, New York : Harper & Row
- BRESNAN, J. (1982), *The mental representation of grammatical relations*, Cambridge, Mass. : MIT Press
- BRULARD, I & CARR, P. (2001), Consonant substitution in bilingual child : Universal Grammar vs. production templates and strategies, Paper presented at the *3rd International Symposium on Bilingualism*, Bristol, April
- BRUNER J.S. (1975), The ontogenesis of speech acts, *Journal of Child language*, **2**, 1-19
- BYBEE, J. (2001), *Phonology and Language Use*, Cambridge University Press, **94**
- CAO XUAN HAO (1985), Phonologie et Linéarité, Réflexions critiques sur les postulats de la phonologie contemporaine, *numéro spécial du bulletin de la Société d'Etudes Linguistiques et Anthropogiques de France (SELAF)*, **18**, Paris
- CARR, P. (2000), Scientific Realism, Sociophonetic Variation and Innate endowments, in N. Burton-Roberts, P. Carr & G. Docherty (eds), *Phonological knowledge, Conceptual and Empirical Issues*, 67-105, Oxford University Press
- CARVALHO, BRANDÃO DE, J. (1999), Primitives et naturalité, in *Nouvelles Phonologies, Langages*, **125**, 14-34, Larousse, Paris
- CARVALHO, BRANDÃO DE, J. (2002), *De la syllabation en termes de contours CV*, Habilitation à diriger des recherches, EHESS, Paris
- CARVALHO, BRANDÃO DE, J. (2004), L'analogie est-elle un fait fonctionnel ou grammatical ? Le cas de l'apophonie portugaise, in T. Scheer (ed), *Usage des corpus en phonologie, Corpus*, **3**, 101-125
- CASACCHIA, G. (1989), Les débuts de la tradition linguistique chinoise et l'âge d'or de la linguistique impériale, in S. Auroux (ed), *Histoire des idées linguistiques*, 431-449, Mardaga, Bruxelles
- CHEVROT, J.-P. & FAYOL, M. (2000). L'acquisition de la liaison : enjeux théoriques, premiers résultats, perspectives. *Lidil*, **22**

- CHEVROT, J-P. & FAYOL, M. (2001), Acquisition of french liaison and related child errors, in M. Almgren, A. Barreña, M.J. Ezeizabarrena, I. Idiazabal, and B. MacWhinney (eds), *Research on Child Language Acquisition*, vol. 2, 760-774, Cascadilla Press
- CHEVROT, J-P. (2003) Segmentation des mots et formation des représentations lexicales : le cas de la liaison chez des enfants tout venant et dysphasiques, *Séminaire du Laboratoire Cognition et développement*, UMR 8605, 11 Mars 2003, ms
- CHEVROT, J-P. DUGUA, C. & FAYOL, M. (2005), Liaison et formation des mots, un scénario développemental, *Langages, Nouvelles approches de la liaison*, **158**, numéro coordonné par J-P Chevrot, M. Fayol & B. Laks, Larousse, Paris
- CHOMSKY, N. (1957), *Syntactic structures*. The Hague, Mouton
- CHOMSKY, N. (1965), *Aspects of the theory of Syntax*, Cambridge, MA : MIT press
- CHOMSKY, N. (1966), *Cartesian Linguistics*, New York, Harper, Row. Trad franç: (1969) *La Linguistique cartésienne*, Le Seuil, Paris
- CHOMSKY, N. (1968), *Language and Mind*, New York Harcourt, Brace, World. Trad. franç : (1969) *Le langage et la pensée*, Payot, Paris.
- CHOMSKY, N. (1975), *Reflections on Language*, New York, Pantheon Books, Trad franç: (1977) *Reflexions sur le langage*, Maspero, Paris
- CHOMSKY, N. (1980), Rules and representations, *The Behavioral and Brain Sciences*, **3**, 1-15
- CHOMSKY, N. (1990), Sur la nature, l'utilisation et l'acquisition du langage, *Recherches linguistiques de Vincennes*, **19**, 21-45.
- CHOMSKY, N. (2000), *New horizons in the study of language and Mind*, Cambridge University Press
- CHOMSKY, N., & HALLE, M. (1968), *The Sound Pattern of English*, New York : Harper & Row. Trad franç des chapitres 1,2,7,8,9 par Encrevé, P. (1973), *Principes de phonologie générative*, Paris: Seuil.
- CHRISTOPHE, A., (1993), *Rôle de la prosodie dans la segmentation en mots*, Thèse soutenue en vue de l'obtention du doctorat, EHESS, Paris
- CHRISTOPHE, A., (2002), Comment étudier le développement, in E. Dupoux (ed.), *Les langages du cerveau*, Odile Jacob, Paris
- CLEMENTS, N. & KEYSER, J., (1983), *CV Phonology, a generative theory of syllable*, Cambridge Mass., MIT Press

- COLE, R.A & JAKIMIK, J. (1980), A Model of Speech perception, in R.A. Cole (ed), Perception and production of fluent speech, Hillsdale, N.J. : Erlbaum
- COLEMAN, J. (2002), Phonetic Representations in the mental lexicon. In J. Durand & B. Laks (Eds), *Phonetics, Phonology, and Cognition*, Oxford: Oxford University Press
- COLLINS, J. (2003), Cowie on the poverty of stimulus, *Synthèse*, **136**, 159-190
- CONTENT, A., KEARNS, R., & FRAUENFELDER, U. H. (2001), Boundaries vs onsets in syllabic segmentation, *Journal of Memory and Language*, **45**, 177-199
- COTE, M.H. (2002), Between phonology and the lexicon : French liaison revisited, conférence à l'Université de Toronto, ms.
- COTE, M.H. (2005), Le statut lexical des consonnes de liaison, in *La liaison de la phonologie à la cognition, Langages*, **158**, numéro coordonné par J-P Chevrot, M. Fayol & B. Laks, Larousse, Paris
- COWIE, F., (1999), *What's Within ? Nativism Reconsidered*, Oxford University Press, Oxford
- COWIE, F., (2001), On Cussing in Church: In Defence of What's Within, *Mind and Language* **16**, 231-245
- CUTLER, A., MEHLER, J., NORRIS, D., & SEGUI, J. (1986), The syllable's differing role in the segmentation of french and english, *Journal of memory and language*, **31**, 218-236
- CUTLER, A., MEHLER, J., NORRIS, D., & SEGUI, J. (1992), The monolingual nature of speech segmentation by bilinguals, *Cognitive Psychology*, **24**, 381-410.
- CUTLER, A. & CARTER, D.M. (1987), The predominance of strong initial syllables in the English Vocabulary, *Computer Speech and Language*, **2**, 133-142
- CUMMINS, R. & HARNISH, R.H. (1980), The language faculty and the interpretation of linguistics, *The Behavioral and Brain Sciences*, **3**, 18-19
- DAMAZIO, R. M. (1996), A neural basis for lexical retrieval, *Nature*, **380**, 499-505
- DAUER, R. M. (1983) Stress-timing and syllable-timing reanalysed, *Journal of Phonetics*, **11**, 51-69
- DAUER, R. M. (1987), Phonetic and phonological components of language rhythm, *Proceedings of the XI th ICPhS*, Tallinn, Estonia, **vol. 5**, 447-450
- DEHAENE-LAMBERTZ, G., CHRISTOPHE, A. & VAN OOIJEN, B. (2000), in M. Kail & M. Fayol (eds), *L'acquisition du langage*, PUF, 61-93

- DEJEAN DE LA BATIE, B. (1993), *Word boundary ambiguity in spoken French*. Unpublished Doctoral Dissertation. Monash University, Victoria, Australia
- DELAIS, E. (1994), Structure rythmique et prosodique du français : pour une approche psychocognitive, *Linx*, **29**, Université Paris X Nanterre.
- DELAIS, E. (1995), *Pour une approche parallèle de la structure prosodique. Etude de l'organisation prosodique et rythmique de la phrase française*, Thèse de Doctorat. Université de Toulouse-Le Mirail
- DELATTRE, P. (1947), La liaison en français, tendances et classification, *The French Review*, **XXI**, 148-157
- DELATTRE, P. (1966), *Studies in French and comparative Phonetics*, La Haye, Mouton
- DELL, F. (1970), *Les règles phonologiques tardives et la phonologie dérivationnelle du français*, PHD, MIT
- DEMUTH, K. & FEE, E.J (1995), *Minimal prosodic word in early phonological words*, ms. Brown University and Dalhousie University
- DEMUTH, K. (2001), Prosodic constraints on morphological development, J. Weisenborn, & B. Höhle (eds), *Approaches to bootstrapping*, Amsterdam / Philadelphia : Benjamins, 3-21
- DENNETT, D. C. (1995), *Darwin's dangerous Idea*, Harmondsworth : Penguin
- DESCARTES, R. (2002), *Règles pour la direction de l'Esprit*, traduction de J. Sirven, Vrin, Paris
- DESIRAT, C. & HORDE, T. (1988), *La langue française au 20<sup>ème</sup> siècle*, Bordas, Paris
- DEVITT, M., (2003), Linguistic is not psychology, in A. Barber (ed), *Epistemology of Language*, 107-139, Oxford University Press
- DI CRISTO, A., (1999), Le cadre accentuel du français contemporain: essai de modélisation partie 1, *Langues*, **vol. 2, 3**, 184-205
- DOCHERTY, M., MILROY, J., MILROY, L. & WALSHAW, D. (1997), Descriptive adequacy in phonology : a variationnist perspective, *Journal of Linguistics*, **33-2**, 275-310,
- DUGUA, C. (2002), *Liaison et segmentation du lexique en français : vers un scénario développemental*, Mémoire de DEA, Université Stendhal, Grenoble, ms
- DUPOUX, E. (1989), Détection des mots parlés : détection de phonème, unité prélexicale. Thèse présentée en vue de l'obtention du doctorat de l'EHESS, Paris, ms.

- DUPOUX, E. (2002), *Les langages du cerveau*, O. Jacob, Paris
- DUPOUX, E. (2004), The acquisition of discrete segmental categories Data and Model  
In *Proceedings of the 18th International Congress of Acoustics*, Kyoto, April 4-9
- DURAND, M. (1936), *Le genre grammatical en français parlé, à Paris et dans la région parisienne*, Bibliothèque du français moderne, Paris
- ECHOLS, C.H. & MARTI, C.N. (2004), The identification of words and their meanings : from perceptual biases to specific language-cues, in G. Hall & S. R. Waxman (eds), *Weaving a lexicon*, 41-79, Cambridge, mass., The MIT Press
- ENCREVE, P. (1988), *La liaison avec et sans enchaînement, phonologie tridimensionnelle et usages du français*, Le Seuil, Paris
- ENCREVE, P. (1997), L'ancien et le nouveau, quelques remarques sur la phonologie et son histoire, in *Nouvelles Phonologies, Langages* **125**, Larousse, Paris
- ENCREVE, P. & SCHEER, T. (2005), L'association n'est pas automatique, communication au 7<sup>ème</sup> colloque, du Réseau français de Phonologie, Aix-en-Provence, 2-4 juin
- FERGUSON C. A & FARWELL C. B. (1975), Words and sounds in early language acquisition, *Language*, **51**, 419-439
- FERRE, S. (1998), *Les effets du biliguisme sur la paramétrisation des segments flottants : exemple de la liaison en français*, Mémoire de DEA, Université de Paris VIII
- FERREIRA, F. (1993), Creation of Prosody during Sentence Production, in *Psychological Bulletin*, **100**, 233-253
- FIKKERT, P. (1994), *On the Acquisition of Prosodic Structure*, Dordrecht : ICG.
- FIKKERT, P., FREITAS, M. J., GRIJZENHOUT, J., LEVELT, C. & WAUQUIER S. (2004), Syllabic Markedness, Segmental Markedness, Rhythm and Acquisition, *GLOW Phonology Workshop, April 18 th*
- FIKKERT, P., LEVELT, C. & WAUQUIER-GRAVELINES S., (2006), Syllabic Markedness, Segmental Markedness, Rhythm and Acquisition, *Recherches Linguistiques de Vincennes*, **34**, Presses Universitaires de Vincennes, Université Paris VIII
- FISHER, C., CHURCH, B., & CHAMBERS, K. (2004), Learning to identify spoken word, in D.G. Hall D.G & S.R. Waxman, (eds), *Weaving a lexicon*, 3-41, Cambridge mass, The MIT press
- FITZPATRICK, J., & WHEELDON, L.R. (2000). Phonology and phonetics in Psycholinguistic Models of Speech, in N. Burton-Roberts, P. Carr, P. & G.

- Docherty, G. (Eds.), *Phonological knowledge, Conceptual and Empirical Issues*. Oxford: Oxford University Press, 131-161
- FLAUBERT, G., (1978), *Dictionnaire des idées reçues*, Aubier, Paris
- FODOR, J., (2001), Doing Without What's Within: Fiona Cowie's Critique of Nativism, *Mind*, **110**, 99–148
- FORSTER, K.I., (1976), Accessing the mental lexicon, in R.J Wales & E Walker (eds) , *New approaches to language mechanisms*. Amsterdam : North Holland
- FOUGERON, C., GOLDMAN, J.P. & FRAUENFELDER U. H. (2001a), Liaison and schwa deletion in French : an effect of lexical frequency and competition, *Proceedings of Eurospeech Conference*
- FOUGERON, C., GOLDMAN, J.P., DART, A., GUELAT L. & JEAGER C., (2001b), Influence de facteurs stylistiques, syntaxiques et lexicaux sur la réalisation de la liaison en français, in 8<sup>ème</sup> TALN, Tours, 2001
- FOUGERON C. & DELAIS-ROUSSARIE E., (2004), Liaisons et enchaînements « Fais\_en à Fez\_en parlant », *Actes des Journées d'Etudes sur la Parole*, Fez, 2004
- FRANCESCATO, G. (1968), On the role of the word in first language acquisition, *Lingua*, **21**, 144-153
- FRAUENFELDER, U.H., & CONTENT, A. (1999), The role of the syllable in spoken, word recognition. In *Actes des IIèmes Journées d'Etudes Linguistiques*, Nantes: France, 1-8
- FREITAS, M.J., (1996), Onset in the early productions, in B. Beachley, A. Brown & F. Conlin, (eds), *Proceedings of the Boston University Conference on Language Development*, Somerville, MA : Cascadilla Press
- FREITAS, M.J., (2003), The (un)solved mystery of empty onsets in European Portuguese, Second acquisition Workshop, Nijmegen, ms.
- GASKELL, G., SPINELLI, E., & MEUNIER, F. (2002), Perception of resyllabification in French, *Memory & Cognition*, **30 (5)**, 798-810
- GATHERCOLE, S.E. & BADDELEY, A.D., (1990), Phonological memory deficits in language disordered children : Is there a causal connection, *Journal of memory and language*, **29**, 336-360
- GATHERCOLE, S.E. & BADDELEY, A.D., (1993), *Working Memory and Language*, L.E.A Hillsdale USA
- GEE, J.P. & GROSJEAN, F. (1983), Performance Structures : A Psycholinguistic and Linguistic Appraisal, in *Cognitive Psychology*, **15**, 411-488

- GERKEN, L.A., (1994), Young children's representation of prosodic phonology : evidence from English speakers's weak syllables productions, *Journal of Memory and Language*, **33**,19-38
- GERKEN, L.A., (2004), Nine-month-olds extract structural principles required for natural language, *Cognition*, **93**, B89-B96
- GOLDINGER, S.D. (1996), Words and voices: episodic traces in spoken word identification and recognition memory, *Journal of Experimental Psychology: Learning, Memory and Cognition*, **22**, 1166-1183
- GOLDINGER, S.D. (1998), Echoes of echoes? Words and voices: an episodic theory of lexical access, *Psychological Review*, **105**, 251-279
- GOLDINGER, S.D., & AZUMA, T. (2003), Puzzle solving Science : the Quixotic quest for units in Speech perception. In S. Hawkins & N.Nguyen (eds.), *Special Issue on Temporal Integration in the Perception of Speech*, *Journal of phonetics*, **vol. 31 3-4**, 305-321
- GOLDSMITH, J.A. (1976), An overview of autosegmental phonology, *Linguistic Analysis*, **2 (1)**, 23-68
- GOLDSMITH, J.A. (1995), *The Handbook of phonological theory*, Cambridge, Mass and Oxford: Blackwell
- GOMEZ, R. L. & GERKEN, LA. (1999), Artificial language learning by 1-year-olds leads to specific and abstract knowledge, *Cognition*, **70**, 109-135
- GOMEZ, R. L. & GERKEN, LA. (2000), Infant artificial language learning and language acquisition, *Trends in Cognitive Sciences*, **vol. 4, n°5**
- GOPNIK, M. (1994), The Family, in J. Matthews (ed), *Linguistic aspects of familial Language Impairment*, *Special Issue of the Mc Gill Papers in Linguistics*, **10**, 1-4
- GRAMMONT, M. (1914), *Traité pratique de prononciation française*, Paris : Delagrave
- GRAMMONT, M. (1933), *Traité de phonétique*, Paris : Delagrave
- GREGOIRE, M. (1937), *L'apprentissage du langage. Les deux premières années*, Paris, Felix Alcan
- GROSJEAN, F. & DOMMERGUES J.Y. (1983), Les structures de performance en psycholinguistique, *L'année Psychologique*, **83**, 513-536
- GROSJEAN, F., & GEE, J.P. (1987), Prosodic Structure and Spoken Word Recognition, in U. H. Frauenfelder, L.K. Tyler (eds), *Spoken Word Recognition*, *Cognition*, Special Issue, **25**, 135-157



- GROSJEAN, F., & FRAUENFELDER, U. H. (1996), Spoken word recognition paradigms, *Language and Cognitive processes, Special Issue*, **11-6**
- GUERSSEL M. & LOWENSTAMM J. (1996), Ablaut in Classical Arabic measure, Active verbal forms, *Studies in Afroasiatic Grammar*, J. Lecarme, J. Lowenstamm & U. Schlonsky, The Hague, 123-134
- HALL G. H & WAXMAN, S. R. (2004), *Weaving a lexicon*, Cambridge, mass. MIT press
- HALLE P. & BOYSSON-BARDIES DE B. (1994), Emergence of an early receptive lexicon : infant's recognition of words, *Infant behavior and development*, **17**, 119-129
- HARMAN, G. (1980), Two quibbles about analyticity and psychological reality, *The Behavioral and Brain Sciences*, **3**, 21-22
- HARRIS, R.A. (1993), *The linguistic wars*, Oxford University Press,
- HARRIS, J. (1994), *English Sound Structure*, Oxford : Blackwell
- HAYES, B. (1995), *Metrical stress Theory : Principles and Case Studies*, Chicago, The University of Chicago Press
- HAYES B., KIRCHNER R., & STERIADE, D. (2004), *Phonetically-based Phonology*, Cambridge University Press, Cambridge
- HIRSCH-PASEK, K., KEMMLER-NELSON, D.G., JUSCZYK, P.W., WRIGHT-CASSIDY, K., DRUSS B. & KENNEDY, L. (1987), Clauses are Perceptual Units for Young Infants, *Cognition*, **26**, 269-286
- HOCKETT, C. (1948), A note on « Structure », *International Journal of American Linguistics*, **14**, 132-134
- HOCKETT, C. (1958), *A course in Modern Linguistics*, Macmillan, New York
- HOUE O., MAZOYER B. & TZOURIO-MAZOYER N. (2002), *Cerveau et psychologie*, PUF, Paris
- HUME, E. & JOHNSON, K. (2001), A model of the Interplay of Speech perception and phonology, in E. Hume & K. Johnson (eds), *The role of Speech perception in Phonology*, Academic Press, 3-27
- JACKENDOFF, R. (2002), *Foundations of language, Brain, meaning, grammar & evolution*, Oxford University Press
- JAKOBSON, R. (1941), *Kindersprache, Aphasie und allgemeine Lautgesetze*, Uppsala Universitets Arsskrift 9, Uppsala, trad angl. (1968) Child Language Aphasia and phonological Universals, Mouton, trad franç (1969), *Langage enfantin et aphasie*, Paris: Seuil.

- JUSCZYK P.W., HIRSH-PASEK K., KEMMLER-NELSON, D.G., KENNEDY L., WOODWARD A. & PIWOZ J. (1992), Perceptions of Acoustic Correlates of Major Phrasal Units by Young Infants, *Cognitive Psychology*, **24**, 252-293
- JUSCZYK, P.W., A. CUTLER & N.J. REDANZ (1993) Infants' preference for the predominant stress patterns of English words. *Child Development*, **64**, 675-687
- KAIL, R., HALE, C., LEONARD, L., & NIPPOLD, M. (1984). Lexical storage and retrieval in language -impaired children, *Applied Psycholinguistics*, **5**, 37-49
- KARMILOFF, K. & KARMILOFF-SMITH, A. (2003), *Comment les enfants entrent dans le langage*, Retz, Paris
- KAYE, J., & LOWENSTAMM J. (1984), De la syllabicit , in F. Dell, D. Hirst et J.R Vergnaud (eds), *Forme sonore du langage*, Paris : Hermann, 123-159
- KAYE, J., & LOWENSTAMM J. (1982), Syllable structure and markedness theory, in Belletti et al. (eds), *Theory of markedness in Generative grammar*, Pisa, Scuole Normale Superiore
- KAYE, J., LOWENSTAMM, J., & VERGNAUD, J.R. (1985), The internal structure of Phonological Elements : Theory of Charm and Government, in C. Ewen et J. Anderson (eds), *Phonology Yearbook*, **2**, Cambridge University Press, 305-328
- KIRSCHNER, D. & KLATSKY, R. (1985). Verbal rehearsal and memory in language disordered children, *Journal of Speech and Hearing research*, **28**, 556-565
- KLAUSENBURGER, J. (1984), *French Liaison and Linguistic theory*, Stuttgart, Franz Steiner, Verlag
- KLEIN, M. (1991), *Vers une approche substantielle et dynamique de la constituance syllabique : le cas des semi-voyelles et des voyelles hautes dans les usages parisiens*, Th se de doctorat, Universit  Paris VIII, ms.
- KUHL, P., WILLIAMS, K., LACERDA, F., STEVENS, K. & LINDBLOM, B. (1992), Linguistic experience alters phonetic perception in infants by 6 months of age, *Science* **255** (5044), 606-608.
- LAHIRI, A. & MARSLLEN-WILSON W. (1991), The mental representation of lexical form : a phonological approach to the recognition lexicon, *Cognition*, **38**, 245-294
- LAHIRI, A. & REETZ, H. (2002). Underspecified recognition. In C. Gussenhoven & N. Warner (Eds.), *Papers of Laboratory Phonology VII*. Berlin: Mouton de Gruyter
- LAKATOS, I. (1970), Falsification and the methodology of scientific research programmes, in I. Lakatos & A. Musgrave, *Criticism and the Growth of knowledge*, Londres, Cambridge University Press
- LAKS, B. (1996), *Langage et Cognition: l'approche connexionniste*, Paris: Herm s

- LAKS, B. (2005), *Approches cognitives de la phonologie*, in, N. Nguyen, S.Wauquier-Gravelines & J. Durand (Eds.), *Phonologie et Phonétique, forme et substance*, Paris: Hermès
- LANDAU, B. (2004), Perceptual units and their mapping with language : how children can or can't use perception to learn words, in G. Hall & S. R. Waxman (eds), *Weaving a lexicon*, Cambridge, mass., The MIT Press
- LAPOINTE, B. (1994), *Intégrale, chansons, poèmes, inédits*, préface de Jacques Durand, Pézenas, Domens Editions
- LE NORMAND, M.T. & CHEVRIE-MULLER, C. (1991), A follow-up case study of transitory developmental apraxia of speech : 'l'enfant à voyelles', *Clinical Linguistics & Phonetics*, vol. 5, 2, 98-118
- LE NORMAND, M.T., VINTER, S. & KONOPCZYNSKI G. (1993), Early language development in children with consonant apraxia of speech : two case studies, *Scand J Log Phon*, 18, 181-186
- LEONARD, A. (1987), Is specific language impairment a useful construct ? In S. Rosenberg(ed) : *Advances in applied Psycholinguistics*, 1, *Disorders of first-language development*, 1-39, New York Cambridge University Press
- LEONARD, L. (1997), *Children with specific Language Impairment*, MIT Press
- LEVELT, W.J.M. (1989), *Speaking, From Intention to Articulation*, Cambridge : Bradford Book
- LEVELT, W.J.M. (1992), Accessing words in speech production : stages, processes, and representations, *Cognition*, 42, 1-22
- LLEO, C. (1997), Filler syllables, proto-articles and early prosodic constraints in Spanish and German, in Sorace, Heycock, Shillcock (eds) *Language Acquisition : Knowledge, Representation and Processing, Proceedings of GALA 1997*, Edinburgh : Edinburgh University Press
- LOWENSTAMM, J. (1999), The beginning of the Word, In J.R. Renison and K. Kühnammer (eds), *Phonologica 1996 : Syllables !?*, 153-166, The Hague: Thesus
- LOWENSTAMM, J. (2003), A propos des gabarits, in J. Lowenstamm (ed) *Recherches Linguistiques Vincennes*, 32, Presses Universitaires de Vincennes, St Denis
- LUCE, P.A. (1986), *Neighborhoods of words in the mental lexicon*, Doctoral dissertation, Indiana University, Bloomington, Indiana
- LYCHE, C. (2005), Des règles aux contraintes, quelques aspects de la théorie de l'Optimalité, in N. Nguyen, N., S.Wauquier-Gravelines & J. Durand (eds) *Phonologie et phonétique, forme et substance*, Hermès

- MAC CARTHY J. (1982), Prosodic templates, morphemic templates, and morphemic tiers, in H. van der Hulst and N. Smith (eds) *The structure of Phonological Representations, Part I*. Dordrecht Foris
- MAC CLELLAND, J.L. & ELMAN, J.L. (1986), Interactive processes in speech perception : the Trace Model, in PDP research group (eds), *Parallel distributed processing*, vol. 2, 58-121, Cambridge, Mass., MIT press
- MAC CORQUODALE, K. (1970), On Chomsky's review of Skinner's verbal behavior, *Journal of the Experimental Analysis of Behavior*, **13**, 83-99
- MAC GREGOR, K. & LEONARD, L. (1994), Subject pronoun and article omissions in the speech of children with specific language impairment : a phonological interpretation, *Journal of Speech and Hearing Disorders*, **54**, 141-147
- MAC NEILAGE, PF. (1993), The frame/content theory of evolution of speech perception, *Behavioral and Brain Sciences*, **21**, 499-546
- MAC NEILAGE, PF. & DAVIS, B.L. (1993), Motor explanations of babbling and early speech, in B. de Boysson-Bardies, S de Schonen, P. Jusczyk, P. MacNeilage & J. Morton (eds), *Developmental Neurocognition : Speech and face processing in the first year of life*. Dordrecht, Kluwer Academic
- MACKEN, M. (1992), Where's phonology ?, in C.A Ferguson, L. Menn & C. Stoel-Gammon (eds), *Phonological Development*, 249-273, Timonium, MD York Press
- MACKEN, M. (1995), Phonological acquisition, in J. Goldsmith (ed) *The Handbook of Phonological theory*. Cambridge, 671-697, MA, Blackwell
- MALLARME, S. (1945), *Oeuvres Complètes*, Bibliothèque de la Pléiade, Gallimard, Paris
- MARSLEN-WILSON, W.D. (1987), Functional parallelism in spoken word recognition, in U.H. Frauenfelder & L.K. Tyler (eds), *Spoken word recognition*, *Cognition*, **25**, special issue, 71-103
- MARSLEN-WILSON, W.D. & WELSH, A., (1978), Processing interactions during word recognition in continuous speech, *Cognitive Psychology*, **10**, 29-63
- MARSLEN-WILSON, W.D. & TYLER, L. (1980), The temporal structure of spoken word understanding, *Cognition*, **8**, 1-71
- MARTINET, A. (1949), La double articulation linguistique, *TCLC* **5**, 30-37
- MARTINET, A. (1957), Arbitraire linguistique et double articulation, *Cahiers Ferdinand de Saussure*, **15**, 105-107

- MATTER, J.F. (1986), *A la recherche des frontières perdues: Etude sur la perception de la parole en français*. Doctoral dissertation, University of Utrecht, The Netherlands. Amsterdam : De Werelt
- MAYE, J., WERKER, J. & GERKEN L.A (2002), Infant sensitivity to distributional information can affect phonetic discrimination, *Cognition* **82**: B101-B111
- MEHLER, J., DOMMERGUES, J.Y., FRAUENFELDER, U.H., & SEGUI, J. (1981), The syllable's role in speech segmentation, *Journal of Verbal Learning and Verbal Behavior*, **20**, 298-305
- MENN L. (1978), Phonological units in beginning speech, in A. Bell & J. Hooper-Bybee (eds), *Syllables and segments*, Amsterdam, North-Holland.
- MENN L. (1983), Development of articulatory, phonetic and phonological capabilities, in Butterworth (ed) *Language production*, **2**, 1-49, London Academic Press.
- MENN L. & MATTHEI E. (1992), The two lexicon account of child phonology : looking back, looking ahead, in C.A Ferguson, L. Menn & C. Stoel-Gammon (eds), *Phonological development*, 211-247, Timonium, MD : York Press
- MEZULAM, M. (1998), From sensation to cognition, *Brain*, **121**, 1013-1052
- MILNER, J.C. (1967), French Truncation Rule, *Quarterly Progress Report of Research Laboratory of Electronics*, MIT, n°86, 273-283
- MILNER, J.C. (1989), *Introduction à une science du langage*, Paris, le Seuil
- MOGDIL, Y.C. & MOGDIL. C (1987), *Noam Chomsky : consensus and controversy*, Falmer International Master-Minds Challenged, vol. **3**, New York : Falmer Press
- MORGAN, J.L., SAFFRAN, J.R. (1995), Emerging integration of sequential and suprasegmental information in preverbal speech segmentation, *Child Development* **66**, 911-936
- MORIN, S. & KAYE, J.D. (1982), The syntactic bases for french liaison, *Journal of Linguistics*, **18**, 291-330
- MORIN, Y.C. (2003) [1998]. Remarks on prenominal liaison consonants in French. *Living on the Edge, 28 Papers in Honour of Jonathan Kaye*, Stefan Ploch (ed.), Mouton de Gruyter.
- MORTON, J. (1969) Intereaction of information in word recognition, *Psychological review*, **76**, 165-178
- MOSKOWITZ, A. I. (1970), The two-year-old stage in the acquisition of English phonology, *Language*, **46**, 426-441. Reprinted in C.A Ferguson, and D.I Slobin (eds) (1973), *Studies of Child Language development*, New York : Holt, Rinehart & Winston

- MUNSON, B. (2001), Phonological pattern frequency and speech production in children and adults, *Journal of Speech and Hearing research*, vol. 44, 778-792
- NESPOR, M. & VOGEL, I. (1986), *Prosodic phonology*, Dordrecht, Foris
- NGUYEN, N. (2005), La perception de la parole, in N. Nguyen, & S. Wauquier-Gravelines & J. Durand (Eds.), *Phonologie et Phonétique, forme et substance*, Paris: Hermès.
- NYROP, K.R. (1925). *Manuel de phonétique du français parlé (4e ed)*. Copenhagen: Gyldendalske Boghandel Nordisk Forlag
- OLLER, D.K. (1973), The effect of position in utterance of speech segment duration in English, *Journal of the Acoustical Society of America*, 54, 1235-1247
- OLLER, D.K. (2000), *The emergence of the speech capacity*, Mahwah NJ, Lawrence Erlbaum
- PALLIER, C., COLOME, A. & SEBASTIAN-GALLES, N. (2001), The influence of native-language phonology on lexical access exemplar-based vs abstract lexical entries, *Psychological Sciences*, 12 (6), 445-449
- PARADIS, C. & EL FENNE, F. (1995), French Verbal inflection revisited : Constraints, repairs and floating consonants, *Lingua* 95, 169-204
- PARIENTE, J.-C. (1975), Grammaire générale et grammaire générative, *Actes de la recherche en sciences sociales*, 5/6, 36-49
- PASDELOUP, V. (1989), Essai de modélisation acoustique du rythme du français, *Journal d'acoustique*, 2-1, 47-51
- PASDELOUP, V. (1990), *Modèle de règles rythmiques du français appliqué à la synthèse de la parole*. Thèse de 3ème cycle, Université de Provence, Aix-Marseille I
- PASSY, P. (1917), *Les sons du français (8ème ed)*, Paris : Didier.
- PEPERKAMP, S. (2002), *L'acquisition de la phonologie et ses conséquences pour la théorie linguistique*, Habilitation à Diriger des Recherches, EHESS
- PEPERKAMP, S., M. PETTINATO & DUPOUX, E. (2003), Allophonic variation and the acquisition of phoneme categories, in B. Beachley, A. Brown, & F. Conlin (eds.) *Proceedings of the 27th Annual Boston University Conference on Language Development*. Volume 2. Sommerville, MA : Cascadilla Press, 650-661
- PEPERKAMP, S. & DUPOUX, E. (à paraître), The acquisition of abstract phoneme categories, in J. Cole & J. Hualde (éds.) *Laboratory Phonology IX*, Mouton de Gruyter, Berlin.

- PETERS, A. (1985), Language Segmentation : Operating principles for the perception and analysis of language, in D. Slobin (ed) *The crosslinguistic Study of Language Acquisition*, **2**, Erlbaum
- PIAGET, J. (1964), *Six études de psychologie*, Points Seuil, 71, Paris
- PIATELLI-PALMARINI, M. (1977), Avant-propos, in M. Piatelli-Palmarini (ed), *Théorie du langage et théorie de l'apprentissage, le débat entre Jean Piaget et Noam Chomsky*, Points Seuil 138, Paris
- PIATELLI-PALMARINI, M. (2002), Portrait d'un cognitiviste "classique" Ce que j'ai appris auprès de Jacques Mehler, in E. Dupoux (ed.), *Les langages du cerveau*, Paris: Odile Jacob
- PIERREHUMBERT, J. (2002), Word-specific phonetics, in C. Gussenhoven and N. Warner (eds), *Laboratory Phonology VII*, Mouton de Gruyter, 101-140
- PIERREHUMBERT, J. (2004), *Fréquence et structure*, Communication, Colloque du Réseau français de Phonologie, Orléans, 3 juin 2004
- PIERREHUMBERT, J., BECKMANN, M.E, & LADD, D.R. (2000), Conceptual foundations of phonology as a laboratory science, in N. Burton-Roberts, P. Carr & G. Docherty (eds), *Phonological knowledge, Conceptual and Empirical Issues*, 273-305, Oxford University Press
- PINKER, S. (2000), *L'instinct du langage*, Odile Jacob, Paris
- PLENAT, M. (1984), Toto, Fanfa, Totor et même guiguite sont des anars, in F. Dell, D. Hirst & J.R. Vergnaud (eds.), *Forme sonore du langage*, Paris: Hermann.
- PLENAT, M. (1999), Prolégomènes à une étude variationniste des hypocoristiques à redoublement en français, *Cahiers de grammaire*, **24**, 183-219
- POPPER, K.R. (1973), *La logique de la découverte scientifique*, Paris, Payot
- POURIN, D. (2003), *Etude phonologique déclarative des semis-voyelles du français. Aspects synchroniques et diachroniques*, Thèse de doctorat, Université de Nantes, ms.
- PULLUM, G. K. & SCHOLZ, B. C. (2002), Empirical assessment of stimulus poverty arguments, *The Linguistic Review*, **19**, 9-50
- RAMUS, F., NESPOR, M., & MEHLER, J. (1999), Correlates of linguistic rhythm in the speech signal. *Cognition*, **73(3)**, 265-292
- RAMUS, F., HAUSER, M. D., MILLER, C., MORRIS, D., & MEHLER, J. (2000), Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, **288**, 349-351

- RIALLAND, A. (à paraître), Problems in the adaptation of perceptual category boundaries to context and in the acquisition of liaison consonants by French-speaking children with SLI : Comments on the articles by Hallé, Wauquier-Gravelines, and Sauzet, in, C. Jakubowicz, L. Nash. & K. Wexler (Eds.), *Essays on Syntax, Morphology and Phonology in SLI*, MIT Press
- ROACH, P. (1982), On the distinction between ‘stress-timed’ and ‘syllable-timed’ languages, in D. Crystal (ed.) *Linguistic controversies*, London: Edward Arnold, 73-79
- RONDAL, J.A., ESPERET, E. GOMBERT, J.E, THIBAUT, J.P. & COMBLAIN, A. (1999), Développement du langage oral, in J.A Rondal & X. Séron, *Troubles du langage, bases théoriques, diagnostic et rééducation*, 107-179, Mardaga, Bruxelles
- ROSE, Y. (2000), *Headedness and prosodic Licensing in the L1 acquisition of phonology*, PhD, unpublished, Mc Gill, University, Montréal
- ROSE, Y. (2003), ChildPhon: A database solution for the study of child phonology, *Proceedings of the 27th Annual Boston University Conference on Language Development*, Somerville, MA, Cascadilla Press, 674-685
- ROSE, Y., BYRNE R., WAREHAM H., HEDLUND G. ET O'BRIEN P. (2005), Phon: A tool for transcription and analysis of phonological corpora, Unpublished manuscript, Memorial University of Newfoundland
- ROUSSELOT, ABBE (1924), *Principes de phonétique expérimentale*, Paris
- SAPIR, E. (1933), La réalité psychologique des phonèmes, in *Journal de psychologie normale et pathologique*, **30**, 247-265, (texte publié originellement en français), trad. de J.E. Boltanski (1968) *Linguistique* Paris : Minuit, à partir du texte original en anglais
- SAFFRAN, J. R., ASLIN, R. N., & NEWPORT, E. L. (1996), Statistical learning by 8-month-old infants, *Science*, **274** (5294), 1926–1928
- SAUSSURE DE F. (1879), *Mémoire sur le système primitif des voyelles*, Paris : Payot
- SAUSSURE DE F. (1966), *Cours de Linguistique générale*, Paris : Payot
- SAUZET, P. (1999), Linéarité et consonnes latentes, *Recherches Linguistiques de Vincennes*, **28**, 59-87, Presses Universitaires de Vincennes, Université Paris VIII
- SCHANE, S.A. (1968), *French phonology and morphology*. Cambridge, MA: The MIT Press
- SCHANE, S.A. (1974), There is no French truncation rule, in R.J. Campbell (ed), *Linguistic Studies in Romance Languages*, **89-100**, Washington, DC: Georgetown University Press



- SCHEER T. (2004a), *A lateral Theory of Phonology vol.1: What is CVCV and what it should be ?*, Mouton de Gruyter, Berlin.
- SCHEER, T. (2004b), En quoi la phonologie est vraiment différente, in T. Scheer (ed), *Usage des corpus en phonologie, Corpus*, **3**, 5-84
- SEGERAL, P. & SCHEER T., (1998), A generalized theory of Ablaut: the case of modern german strong verbs, in A. Ortmann, R. Fabri & T. Parodi (eds), *Models of Inflection*, 28-59, Tübingen: Niemeyer.
- SEGUI, J. (1989), Psycholinguistique sans linguistique: un pari perdu. *Bulletin de Psychologie*, **39**, 504-512
- SEUREN, P.A.M. (1998), *Western Linguistics, an historical introduction*, Blackwell, Oxford
- SOBER, E. (1980), Representation and psychological reality, *The Behavioral and Brain Sciences*, **3**, 38-39
- SPELKE, E.S., KATZ, G., PURCELL, S. EHRlich, S., BREINLIGER, K. (1994), Early knowledge of object motion : continuity and inertia, *Cognition*, **51**, 131-176
- SPINELLI, E., MCQUEEN J., G., & CUTLER, A. (2003), Processing resyllabified words in french, *Journal of Memory and Language*, **48**, 233-254
- STAGER, C.L. & WERKER, J.F. (1997), Infants listen for more phonetic detail in speech perception than in word-learning tasks, *Nature*, **388**, 381-382
- STERIADE, D. (2001), Directional assymetries in place assimilation : a perceptual account, in E. Hume & K. Johnson, (eds), *The role of Speech perception in Phonology*, 219-251, Academic Press
- SUET-BOURET D. (2002), *Inventaire phonologique et analyse du premier lexique dans le corpus d'une enfant francophone entre 23 et 26 mois : le yod*. Mémoire de maîtrise, Université de Nantes, ms.
- TOMASELLO, M. (2000), The item-based nature of early syntactic development, *Trends in Cognitive Science*, **4(4)**, 156-163
- TRANDEL, B. (1981a), *Concreteness in generative phonology*, Evidence from french, Berkeley and Los Angeles, University of California Press
- TRANDEL, B. (1981b), The treatment of french liaison : descriptive, methodological and theoretical implications, in H. Contreras & J. Klausenburger (eds), *Papers in Romance*, **vol.3**, Seattle, 261-281
- TRANDEL, B. (1990), On suppletion and French Liaison , *Probus* **2 (2)**, 169-208

- TRANEL, B. (1992), The representations of french final consonants and related issues, in *Proceedings of the 22nd Linguistics Symposium on Romance Languages*, University of Texas at El Paso
- TRANEL, B. (1995), Current Issues in french Phonology : liaison and position theories, in J. A. Goldsmith, *The handbook of Phonological Theory*, Blackwell
- TULLER, B. (2003), Computational models in speech perception, *Journal of Phonetics* **31**, 503-507
- TYLER L. K. & FRAUENFELDER, U. H. (1987), The process of spoken word recognition : an introduction, in U.H. Frauenfelder & L.K. Tyler (eds) *Spoken word recognition, Cognition*, **25**, special issue, 1-21
- VAISSIERE, J. (1983), Language independent prosodic features, in A. Cutler & R. Ladd (eds), *Prosody : models and measurements*, Springer series in Language and Communication, **14**, Berlin, Heidelberg, Springer-Verlag
- VAISSIERE, J. (1995), Phonetic explanations for cross-linguistic prosodic similarities, *Phonetica*, **52**, 123-130
- VALDOIS, S. & NESPOULOUS, J.L. (1994), Perturbations du traitement phonétique et phonologique du langage, in X. Séron, X. & M. Jeannerod (eds), *Neuropsychologie humaine*, Mardaga, Bruxelles.
- VALLEE, N & BOE, L.J. (2000), *L'acquisition des structures sonores du langage par l'enfant et les tendances phonologiques des langues du monde*, ms.
- VALLEE, N. (2004), *Le français dans la typologie des langues, structure et organisation phonologiques et syllabiques*, Communication au colloque PFC, décembre 2004, Paris. Document Power point (cp.)
- VASSE, D. (1988), *La chair envisagée, la génération symbolique*, Le Seuil, Paris
- VENEZIANO, E. (2000), Interaction, conversation et acquisition du langage dans les trois premières années, in M. Kail & M. Fayol (eds), *L'acquisition du Langage*, PUF
- VIGARIO, M, FROTA, S. & FREITAS, M.J. (2003), in B. Beachley, A. Brown & F. Colin (eds), *Proceedings of the 27th Annual Boston University Conference on Language Development*, Sommerville: Cascadilla Press, 809-821
- VIHMAN, M.M., (1993), Variable Paths to early words production, *Journal of Phonetics*, **vol. 21**, 61-82
- VIHMAN, M. M., (1996), *Phonological Development :The origins of language in the child*, Blackwell, Oxford

- VIHMAN, M. M. (2001), *Word Templates and the origins of Phonology*, Paper presented as part of the symposium on radical construction Grammar, 7th International Linguistics Conference, July, 26th
- VIHMAN, M. & VELLEMAN S. (1989) Phonological reorganization A case study. *Language and Speech*, **32**
- VIHMAN, M., VELLEMAN, S. & MC CUNE L. (1994), How abstract is child phonology ? Towards an integration of linguistic and psychological approaches in M.Yavas (ed), 9-31, *First and second language Phonology*, San Diego : Singular publishing Group
- VIHMAN, M. & VELLEMAN S. (2000) Phonetics and the origins of Phonology, in N. Burton-Roberts, P. Carr, P. & G. Docherty, G. (Eds.), *Phonological knowledge, Conceptual and Empirical Issues*. Oxford: Oxford University Press
- VIHMAN, M. M. & CROFT, W. (à paraître), Radical Templatic Phonology and phonological development. *Handbook of Cognitive Linguistics*. Oxford : Oxford University Press, ms
- VIHMAN, M. M. & MUELLER-GATHERCOLE (en préparation), ms, cp.
- WATERSON, N. (1987), *Prosodic phonology : the theory and its application to language acquisition and speech processing*, Newcastle, Grevatt and Grevatt
- WAUQUIER-GRAVELINES, S. (1996), *Organisation phonologique et traitement de la parole continue*. Thèse de doctorat en linguistique théorique et formelle, UFRL, Paris 7
- WAUQUIER-GRAVELINES, S. (1999), Segmentation lexicale et parole continue: la linéarité en question. *Recherches Linguistiques de Vincennes*, **28**, 133-157, Presses Universitaires de Vincennes, Université Paris VIII
- WAUQUIER-GRAVELINES, S. (2002), Statistical learning or phonological generalisation The case of acquisition of liaison in French, Special Session « Acquisition » of Tenth Phonology Meeting of Manchester, 23 mai 2002
- WAUQUIER-GRAVELINES S., (2003a), Troncation et reduplication, peut-on parler de gabarits dans le lexique précoce ?, Troisième forum international de Morphologie, Université de Lille, 19-21 septembre
- WAUQUIER-GRAVELINES S., (2003b), La question du réalisme des formalisations phonologiques contemporaines : que nous apprennent les données d'acquisition, in J.-P. Angoujard & S. Wauquier-Gravelines S. (eds), *Phonologie : Champs et perspectives*, ENS-éditions, Lyon
- WAUQUIER-GRAVELINES S. (2004), Les liaisons dangereuses. Psycholinguistique et phonologie : une interface complexe, in L. Ferrand & J. Grainger (eds) *Psycholinguistique cognitive, Essais en l'honneur de Juan Segui*, De Boeck Université, Collection Neurosciences et Cognition, Bruxelles,

- WAUQUIER-GRAVELINES, S. & BRAUD, V. (2005), Proto-déterminant et acquisition de la liaison obligatoire en français, *Langages, Nouvelles approches de la liaison*, **158**, numéro coordonné par J.P. Chevrot, M. Fayol & B. Laks, Larousse, Paris
- WARREN, P. & MARSLEN-WILSON, W. (1987), Continuous uptake of acoustic cues in spoken word recognition, *Perception & Psychophysics*, **41**, 262-275
- WARREN, P. & MARSLEN-WILSON, W. (1988), Cues to lexical choices : discriminating place and voice, *Perception & Psychophysics*, **43**, 21-30
- WENK, B. & WIOLAND F. (1982), Is French really syllable-timed?, *Journal of Phonetics*, **10**, 193-216
- WERKER, J., & TEES R. (1984) Cross-language speech perception : Evidence for perceptual reorganization during the first year of life, *Infant Behavior and Development*, **7**, 49-63

## TABLE DES MATIERES

<b>Ouverture .....</b>	<b>3</b>
<b><u>CHAPITRE I : SYNTHÈSE DES TRAVAUX .....</u></b>	<b>9</b>
<b>1. Introduction.....</b>	<b>9</b>
1.1. Linéarité, temporalité, segmentations et troncations .....	9
1.2. Données et prises de données.....	13
1.3. Position théorique.....	15
<b>2. Organisation phonologique et traitement de la parole continue : liaison et accès au lexique .....</b>	<b>19</b>
2.1. Le problème : segmentation de la parole continue et accès lexical.....	19
2.1.1. Contraintes de la parole continue.....	19
2.1.2. Modèles lexicaux et séquentiels : aligner lexique et signal.....	20
2.1.3. Modèles prélexicaux : trouver l'unité d'accès au mot dans le signal ....	23
2.2. La liaison et l'accès au lexique .....	25
2.2.1. La liaison, phénoménologie.....	25
2.2.2. Trouver les mots en contexte de liaison .....	26
2.3. Expériences et résultats.....	28
2.3.1. Coût cognitif à la reconnaissance de mot.....	29
2.3.2. La consonne de liaison est-elle une consonne comme les autres ?.....	30
2.4. Analyses.....	32
2.4.1. Rôle de l'information montante.....	32
2.4.2. Rôle de l'information lexicale : quelles représentations ? .....	35

2.4.2.1. <i>L'effacement</i> .....	36
2.4.2.2. <i>L'épenthèse</i> .....	38
2.4.2.3. <i>Le flottement</i> .....	39
2.4.2.4. <i>La supplétion</i> .....	43
2.5. Hypothèse de traitement de la parole continue en temps réel.....	44
<b>Travaux traitant de ces questions</b> .....	<b>48</b>
Publications.....	48
Communications.....	48
<b>3. Acquisition de la liaison</b> .....	<b>50</b>
3.1. Introduction.....	50
3.2. Premières données, premières analyses.....	52
3.2.1. Données et résultats : enfants monolingues.....	53
3.2.1.1. <i>Expériences</i> .....	53
3.2.1.2. <i>Résultats globaux</i> .....	54
3.2.1.3. <i>Résultats détaillés des enfants de 3 ans</i> .....	55
3.2.2. Données et résultats : enfants bilingues.....	58
3.2.3. Premières analyses.....	59
3.2.3.1. <i>Généralisation du phénomène : âge et type d'erreurs</i> .....	59
3.2.3.2. <i>Première interprétation : l'absence de sous-spécification</i> .....	60
3.2.3.3. <i>Deuxième interprétation : l'épenthèse</i> .....	64
3.3. Acquisition de la liaison : acquisition phonologique ou lexicale ?.....	66
3.3.1. Données et analyse nouvelles.....	67
3.3.2. Critique de l'approche lexicaliste de la liaison.....	70
3.4. Nouvelle analyse : la liaison comme un cas de variation.....	75
3.4.1. Données longitudinales.....	75
3.4.2. Acquisition de la liaison : détacher deux mots.....	78
3.4.2.1. <i>Premières formes produites</i> .....	78
3.4.2.2. <i>Acquisition de la liaison : nouvelle hypothèse</i> .....	80
3.4.3. Conclusion.....	84
<b>Travaux traitant de ces questions</b> .....	<b>85</b>

Publications.....	85
Communications .....	85
<b>4. Troncations et reduplications. ....</b>	<b>87</b>
4.1. Troncation des mots longs en dénomination d'images (données obtenues sur la liaison) .....	87
4.2. Troncations et reduplications chez les enfants francophones .....	89
4.3. Synthèse et conclusion.....	97
<b>Travaux traitant de ces questions.....</b>	<b>99</b>
Publications.....	99
Communications .....	99
<b>5. Travaux en dysphasie développementale .....</b>	<b>100</b>
5.1. Données et analyses.....	101
5.1.1. Expérience .....	101
5.1.2. Résultats .....	102
5.1.2.1. Résultats des enfants D 1.....	103
5.1.2.2. Résultats des enfants D 2.....	103
5.1.2.3. Résultats des enfants D 3.....	104
5.2. Synthèse et discussion des résultats .....	107
5.2.1 Synthèse des résultats.....	107
5.2.2. Limite des constituants prosodiques pour les enfants de 3ans, D 2 et D 3 .....	107
5.2.1.1 Déficit de l'encodage phonologique dans un modèle de production .....	109
5.2.1.2. Déficit de la boucle phonologique .....	110
5.3. Conclusion .....	112
<b>Travaux traitant de ces questions.....</b>	<b>113</b>
Publications.....	113
Communications .....	113
<b><u>CHAPITRE II : TRAVAUX EN COURS.....</u></b>	<b>114</b>

<b>1. Introduction.....</b>	<b>114</b>
<b>2. Acquisition de la liaison, nouvelle proposition .....</b>	<b>116</b>
2.1. Acquisition de la liaison et variation .....	116
2.2. Vers une explication unifiée de la variation.....	118
2.2.1 Non-association des segments flottants .....	118
2.2.2. Propositions pour l'acquisition de la liaison .....	119
2.2.2.1. <i>Liaisons obligatoires vs facultatives : deux conditionnements</i> <i>différents</i> .....	119
2.2.2.2. <i>Acquisition de la liaison, nouvelle proposition</i> .....	121
<b>Travaux en cours.....</b>	<b>126</b>
Communications .....	126
<b>3. Traitement de la liaison chez l'adulte, nouvelles perspectives.....</b>	<b>127</b>
3.1. Introduction .....	127
3.2. Modèles abstractionnistes vs exemplaristes.....	128
<b>Travaux en cours.....</b>	<b>133</b>
Communications .....	133
<b>4. Acquisition syllabique, <i>learning paths</i> et marque.....</b>	<b>134</b>
4.1. Introduction .....	134
4.2. Attaques vides en portugais et en français.....	135
4.3. Facteurs possibles contrariant ou complétant l'effet de marque.....	139
4.3.1. Fréquence de distribution syllabique dans l'input adulte.....	139
4.3.2. Structure du lexique .....	141
4.3.3. Influence du rythme .....	143
4.4. Existence d'un <i>learning path</i> en yod.....	148



4.4.1. Etude distributionnelle .....	148
4.4.2. Analyse.....	150
4.4.3. Pourquoi yod ?.....	151
4.4.3.1. <i>Comparaison avec les données d’Alice</i> .....	151
4.4.3.2. <i>‘Palatal pattern’ et chemin apophonique</i> .....	153
4.5. Conclusion .....	155
<b>Travaux en cours.....</b>	<b>156</b>
Publications.....	156
Communications .....	156
<b>5. Vers un modèle gabaritique de l’acquisition en français.....</b>	<b>157</b>
5.1. Introduction.....	157
5.2. L’ hypothèse gabaritique.....	159
5.2.1. Gabarit et acquisition, le modèle de Vihman .....	160
5.3. Hypothèse d’un gabarit prosodique pour l’unité primitive en français.....	163
5.4. Avantages de l’hypothèse gabaritique .....	166
5.4.1. Proto-déterminant et resegmentation .....	166
5.4.2. Stratégies d’allongement et de diversification par syllabes internes ...	167
5. 5. Conclusion .....	168
<b>6. Propositions de travail.....</b>	<b>169</b>
6.1. Etude de la liaison encore .....	169
6.2. Acquisition des déterminants en français : articulation phonologie / syntaxe et sémantique .....	172
6.3. Projet typologique sur l’acquisition de la phonologie .....	172
6.4. Travaux sur la dysphasie.....	173
6.4.1. Troubles prosodiques dans la dysphasie de développement .....	173

6.4.2. Les enfants sans consonnes .....	173
<b>Travaux en cours.....</b>	<b>174</b>
Communications .....	174
<b><u>CHAPITRE III : PHONOLOGIE ET COGNITION, PROBLÈMES ET QUESTIONS.....</u></b>	<b>175</b>
<b>1. Introduction.....</b>	<b>175</b>
<b>2. Réalité du modèle/ modèle de la réalité .....</b>	<b>178</b>
2.1. L'ambiguïté de la grammaire .....	178
2.2. Phonétique, phonologie, psychologie, biologie : 4 niveaux de réalité différents .....	181
<b>3. Linguistique et psychologie : questions et problèmes .....</b>	<b>190</b>
3.1. La dichotomie compétence vs performance.....	191
3.1.1. Compétence .....	191
3.1.2. Performance.....	192
3.1.3. Les interprétations de la dichotomie compétence vs performance .....	194
3.1.4. Relations entre compétence et performance, relations entre les deux disciplines.....	195
3.2. Compétence vs performance encore : la question de la falsification de la compétence phonologique par la performance est-elle pertinente ? .....	197
3.3. Le modèle de la « compétence / performance » comme un rempart antibehaviouriste ?.....	206
3.4. Faillite du modèle compétence vs performance .....	213

<b>4. De l'opposition compétence vs performance à la question du statut des représentations.</b> .....	<b>214</b>
4.1. Une phonologie exclusivement procédurale est-elle cognitivement plus pertinente?.....	214
4.1.1. Phonologie et fréquence .....	214
4.1.2. Une phonologie de la perception ?.....	218
<b>5. Conclusion</b> .....	<b>220</b>
<b><u>CHAPITRE IV : L'ACQUISITION DE LA PHONOLOGIE, LE BÉBÉ ET L'EAU DU BAIN.</u></b>	<b>223</b>
<b>1. Le dispositif chomskyen : les limites du bébé cartésien</b> .....	<b>223</b>
<b>2. Pauvreté du stimulus en phonologie : quel est le rôle de l'input ?</b> .....	<b>228</b>
2.1. Un problème syntaxique .....	228
2.2. Un stimulus quantitativement riche.....	229
<b>3. Stratégies perceptives et rôle de la fréquence : fréquence de quoi ?</b> .....	<b>232</b>
3.1. Un petit bébé très doué .....	232
3.2. Traitement phonétique ou phonologique ? .....	235
<b>4. Lexique et émergence de la phonologie</b> .....	<b>243</b>
4.1. Quelque chose que « fait » l'enfant ou qui « arrive » à l'enfant ? .....	243
4.2. Acquisition phonologique et lexique : articuler son et sens .....	249
<b>5. Conclusion</b> .....	<b>260</b>
<b>Références</b> .....	<b>262</b>

