

Joaquim Brandão de Carvalho

Université de Paris 8, UMR 7023

Sophie Wauquier

Université de Paris 8, UMR 7023

APPROCHES INDUCTIVES ET THEORIE PHONOLOGIQUE : VRAIS ET FAUX PROBLEMES

Résumé

Cet article traite de l'intérêt théorique des approches *usage-based* de la phonologie qui substituent aux représentations autonomes, symboliques et invariantes un listage d'"exemplaires" encodant la variation à partir du calcul probabiliste opéré sur les données jusqu'à un niveau de "détail phonétique fin". Il montre que l'apport de ces modèles inductifs se situe essentiellement dans un registre psycholinguistique, où ils permettent une meilleure modélisation de certains phénomènes de traitement. Au regard de la phonologie, ces approches, de par leur point de vue fondamentalement anti-*SPE*, conduisent, au mieux, à une reformulation inavouée de concepts structuralistes ; au pire, elles reproduisent les apories théoriques d'une conception linéaire des représentations qu'elles critiquent par ailleurs.

Mots-clés: phonologie, modèles à exemplaires, fréquence, variation

Abstract

What do usage-based models provide to theoretical phonology? Is there any advantage to replace the autonomous, symbolic and invariant representations assumed by phonological theory with stochastic storage of “exemplars” encoding any phonetic variation and “fine phonetic detail”? Though exemplar models contribute to a better comprehension of some psycholinguistic issues, their theoretical contribution is much less relevant for phonology proper. At the best, their radical opposition to *SPE*-based views leads them to unavowed reformulations of previous structuralist concepts; at the worst, they perpetuate some aporias from the linear views of phonology they reject nevertheless.

Keywords: phonology, exemplar theory, frequency, variation

[Funes] era casi incapaz de ideas generales, platónicas. No sólo le costaba comprender que el símbolo genérico *perro* abarcara tantos individuos dispares de diversos tamaños y diversa forma ; le molestaba que el perro de las tres y catorce (visto de perfil) tuviera el mismo nombre que el perro de las tres y cuarto (visto de frente).

Jorge Luis Borges, *Funes el memorioso*.

1. Introduction

Depuis une dizaine d'années, le développement de la psycholinguistique expérimentale et la mise à disposition de données d'acquisition et d'adultes reposant sur des travaux en perception ainsi que la critique des postulats de la grammaire générative ont réactivé, dans une espèce de retour de balancier après les disparitions successives du behaviorisme et des "phonologies naturelles" des années 70, une conception empiriste en phonologie revendiquée par les approches *usage-based*, probabilistes et exemplaristes (Bybee, 2001 ; Pierrehumbert, 2001, 2006 ; Coleman, 2002). Cette réaction est loin de constituer un courant homogène. Elle ne repose, ni sur l'étude d'une phénoménologie particulière, ni sur un corps de doctrine clairement identifiable. Mais, malgré leurs divergences, toutes ces conceptions partagent la caractéristique suivante : elles considèrent que la nature, l'émergence et le format des catégories phonologiques sont prioritairement conditionnés par l'usage des structures matérialisées (*embodied*) dans le signal acoustique ainsi que par leur traitement probabiliste en perception par le locuteur-auditeur. En découle, une réorientation aristotélicienne de la phonologie, par laquelle l'"idée du son" remontant à Baudouin puis Sapir, et partiellement reprise par Troubetzkoy, irait jusqu'à laisser la place à un nuage d'exemplaires centré sur un prototype. Ceci entraîne une conception univoque des relations entre forme phonologique et substance sonore formulée en termes d'inférence directe et présupposant que la substance engendre et conditionne exclusivement les représentations catégorielles et invariantes traditionnellement considérées en phonologie, pour leur substituer des "exemplaires" et "prototypes" encodant la variation allophonique et toutes les variations phonétiques, caractéristiques du locuteur, informations socio-pragmatiques, etc., etc., sous la forme de "détails phonétiques fins".

On peut arguer que ce revirement empiriste trouve légitimement ses assises dans les limites d'une phonologie héritée de Chomsky & Halle (1968 ; dorénavant *SPE*) et formulée sur la base de la conception chomskyenne du rationalisme linguistique. A ce titre, les modèles "à exemplaires" présentent de nombreux avantages et avancent des propositions leur permettant de rendre compte de manière efficace des phénomènes psycholinguistiques observés dans le traitement et l'acquisition de la parole

naturelle (Pierrehumbert, 2006 ; Nguyen *et al.*, 2007), chose que les modèles phonologiques générativistes sont en général peu capables voire incapables de faire. Par ailleurs, les approches empiristes appuient leur critique sur la mise au jour de contradictions épistémologiques dans la formulation chomskyenne et du rationalisme linguistique et de la question de “la réalité du modèle et du modèle de la réalité” (Carr, 2000 ; Devitt, 2003 ; Wauquier-Gravelines, 2005).

Ces approches sont-elles pour autant réellement neuves ? Apportent-elles, par ce tournant empiriste revendiqué, des solutions satisfaisantes aux apories des modèles formels et à l’arbitraire qu’elles attribuent aux théories symboliques ? Constituent-elles une solution alternative à ces formalismes qu’elles dénoncent et ambitionnent de dépasser (Pierrehumbert *et al.*, 2000 ; Bybee, 2001) ? Nous ne le pensons pas. Ces conceptions mettent en cause, selon nous, une conception de la langue bien plus ancienne que *SPE* reposant sur le principe d’invariance systémique qui sous-tend aussi bien la notion classique de phonème que le concept bloomfieldien de lexique repris par la grammaire générative. La conséquence en est — même si elle n’est pas toujours délibérément assumée — que cette critique, comme l’a fait avant elle le connexionnisme (Laks, 1996), pose la question de la légitimité théorique et empirique d’une phonologie formelle qui postule l’existence d’un niveau de représentations symboliques autonome. Nous verrons que si les problèmes soulevés sont réels et les questions pertinentes (§ 2), les solutions avancées ne le sont pas (§ 3). Nous nous attacherons à montrer que la mise en cause d’une phonologie autonome, symbolique et invariante par les approches empiristes est largement biaisée en ce que les critiques qui lui sont adressées concernent essentiellement les présupposés du modèle génératif classique introduit par *SPE*, et qu’elles ne tiennent aucun compte ni de ce qui a précédé ce modèle — notamment le structuralisme européen — ni de développements ultérieurs — par exemple, l’hypothèse autosegmentale.

2. Pourquoi une réorientation empiriste de la phonologie ?

Les arguments théoriques qui sont régulièrement avancés en faveur d’une démarche empirique se déclinent selon nous en deux axes majeurs que nous examinons successivement.

2.1. L’absence de réalisme cognitif

La critique empiriste la plus récurrente de la phonologie se nourrit d’une divergence majeure quant à la conceptualisation des relations entre compétence et performance et donc quant à la place exacte que la phonologie doit attribuer aux données conditionnées par ce qu’on appelle communément les “facteurs de performance” □ données de perception, d’acquisition, de

variations de tous types □ dans la construction d'un modèle mentaliste "réaliste". Le reproche concerne le peu d'intérêt (voire le désintérêt total) des phonologies *SPE* et post-*SPE* contemporaines à l'égard de cette phénoménologie ainsi que l'absence d'une théorisation explicite de la performance associée à la théorie de la compétence, en particulier l'absence de réponse sérieuse à la question suivante : que sont et que font les données de performance pour une théorie de la compétence (Pierrehumbert *et al.*, 2000). Pour dire les choses autrement, les phonologies de laboratoire et *usage-based* critiquent le fait que la phonologie générative a, dans ses divers avatars, toujours privilégié □ alors même que ce paradigme scientifique revendique ouvertement une position mentaliste et réaliste □ l'élaboration d'une théorie des représentations symboliques du locuteur idéal, en remettant à d'autres le soin d'en garantir et l'ancrage substantiel et le réalisme psychologique.¹

On ne peut sur ce point leur donner tort. A partir des années 80, le développement de la psycholinguistique, de la phonétique expérimentale et des recherches sur les facultés cognitives humaines dans leur ensemble ont mis au jour un nombre important de faits nouveaux concernant les facultés mnésiques et catégorielles générales de l'être humain, la mémorisation et la catégorisation de l'information linguistique chez le nouveau-né, l'enfant et l'adulte, la perception et le traitement de la parole par l'adulte, les capacités perceptives et les facultés d'apprentissage du petit bébé. Or ces découvertes tracent le portrait d'un locuteur-auditeur au système cognitif de moins en moins compatible avec celui du locuteur-idéal chomskyen. De nombreux arguments empiriques concernant la sensibilité des auditeurs à la structure détaillée du signal acoustique, au "détail phonétique fin" plaident en faveur des modèles à exemplaires (pour une synthèse, cf. Nguyen *et al.*, 2007). Des travaux récents montrent en effet que les "détails phonétiques fins" marquant des variations allophoniques (Marslen-Wilson & Warren, 1994 ; Allen & Miller, 2004 ; Gaskell, 2003), des variations sociophonétiques, des caractéristiques propres au locuteur (âge, sexe, catégorie sociale : Goldinger, 1998), influencent la perception et le traitement de la parole continue. Or les modèles à exemplaires ont de ce point de vue l'avantage d'opposer à une représentation symbolique unique, arbitrairement posée, une conception des représentations résultant d'un stockage des variantes et des formes alternantes d'une même unité donnée encodant de nombreuses spécificités de l'occurrence (phonétiques, sensori-moteurs, sémantiques, pragmatiques) et modélisant, à ce titre, la variété des traitements et des productions inter et intra-locuteurs.

Par ailleurs, l'activation des exemplaires et le stockage en temps réel par le système de toute nouvelle variante comme exemplaire peut donner lieu à la formation d'une trace générique abstraite (un prototype), ce qui repose sur l'hypothèse d'un re-calcul dynamique du prototype sous l'influence des nouveaux percepts. De tels modèles sont donc a priori plus à même de rendre

compte efficacement des variations inter et intra-individuelles liées à l'âge, le sexe, l'origine socio-culturelle du locuteur (Pierrehumbert, 2006 ; Foulkes & Docherty, 2006), ainsi que de leur évolution chez un même locuteur (Harrington, 2006) ou en diachronie (Bybee, 2001).

Pourtant, ces modèles posent, selon nous, divers problèmes théoriques et laissent plusieurs questions centrales irrésolues.

Le premier problème concerne la nature exacte des exemplaires et des prototypes. Cette question sera discutée en § 3.

Le deuxième problème concerne l'existence même d'un niveau encodant des représentations de nature phonologique, lequel n'est pas nécessairement retenu dans tous les cas. Si, dans certains cas (Pierrehumbert, 2001, 2006), l'existence d'un niveau phonologique spécifique est clairement posée, dans d'autres (Bybee, 2001) des unités de petite taille et de grande taille peuvent être simultanément activées avec un biais naturel en faveur des unités les plus larges. Certains modèles considèrent en effet que les exemplaires n'ont pas de structure phonologique interne et sont formés par des représentations auditives non décomposées (Hawkins, 2003 ; Johnson, 1997). Or il a été montré par de très nombreux travaux que, même lors de tâches d'accès au lexique, les auditeurs peuvent être sensibles à différentes unités infra-lexicales (segments, constituants syllabiques).

De manière générale, leur lien générique avec les modèles d'accès au lexique, ainsi que l'importance qu'ils donnent à la fréquence de distribution de l'information dans le signal (et en particulier à la fréquence lexicale : cf. § 2.2), amène la plupart des modèles à privilégier de manière assez uniforme le "mot" au détriment d'un niveau proprement phonologique. Ce faisant, ils s'inscrivent plus ou moins explicitement dans une conception de la grammaire n'établissant pas de différence entre grammaire et lexique (Bybee, 2001). Or cette conception supposant l'absence d'un niveau phonologique autonome entre en contradiction avec un principe systémique qui gouverne l'organisation phonologique de toute langue humaine (Wauquier-Gravelines, 2005) : les systèmes phonologiques humains sont universellement décrits comme des systèmes reposant sur des unités dites de deuxième articulation dont les unités linguistiques signifiantes ("les signes" au sens saussurien du terme) résultent par combinaison. Ces unités phonologiques sont définies non seulement par leurs caractéristiques propres mais aussi par les relations d'opposition qu'elles entretiennent à l'intérieur du système (x est x mais il est également non- y). On doit par conséquent considérer que les représentations phonologiques reposent sur la modélisation des catégories phonologiques *per se*, mais également sur la modélisation et des relations entre les unités dans le système et de l'organisation systémique entre ces catégories qui en garantit la distinctivité. Les modèles à exemplaires n'apportent à l'heure actuelle aucune réponse à de telles questions.

2.2. L'usage et la fréquence

La deuxième question régulièrement convoquée en défense d'une approche empiriste est la négligence de la phonologie vis-à-vis de la question de l'usage. Aux phonologies formelles sont opposées des conceptions "usage-based" (Bybee, 2001). Là aussi la critique est juste même s'il nous apparaît, que, comme pour le reste, elle est essentiellement conditionnée par la position chomskyenne² et ne motive pas, à ce titre, la remise en cause d'une conception représentationnelle de la phonologie, bien antérieure à la GG. La position antibehaviouriste radicale adoptée par la GG pose un tabou puissant en phonologie générative : toute théorisation linguistique qui intégrerait une réflexion sur la dimension comportementale du langage et qui, plus généralement, soulèverait la question de l'usage, du rôle de l'environnement linguistique et pragmatique, de la variation tant d'un point de vue psycholinguistique que sociolinguistique, pour les données d'adultes comme pour les données d'acquisition, est immédiatement susceptible d'être entachée d'infamie. Chomsky place donc, avec cette posture inaugurale, comme le souligne avec justesse Carr (2000), le projet scientifique générativiste dans une perspective non seulement anti-behaviouriste mais également et tout aussi radicalement anti-comportementaliste.

Or on ne peut, de fait, pas exclure absolument le fait que le langage, dans sa dimension psychologique et sociale, tant du point de vue du locuteur réel que du point de vue de l'observation empirique que l'on peut en faire, *est un comportement* et qu'il 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. Faire l'économie radicale d'une réflexion sur ce sujet³ ne pouvait être envisagé que de manière transitoire et uniquement à condition que la théorie de la compétence abstraite puisse ultimement être articulée à une théorisation du comportement psycho-social du locuteur. L'anti-comportementalisme de principe posé par la GG exposait nécessairement à ce retour de balancier.

Ce qui nous semble pourtant problématique et dont nous discutons ci-dessous est la manière dont les modèles à exemplaires se sont saisis de cette question. Les propositions formulent la question de l'usage d'un point de vue probabilistique en considérant que la *fréquence* d'usage conditionne de manière déterminante et exclusive la genèse et la nature des représentations phonologiques (Pierrehumbert, 2001 ; Bybee, 2006).

"In contrast, work on sound structure in intellectual circles outside of generative linguistics proper has used probabilistic models for many decades. This line of research has established *that the cognitive representation of sound structure is probabilistic, with frequencies playing a crucial role in the the acquisition of phonological and phonetic competence, in speech production and perception, and in long-term mental representations*" (Pierrehumbert, 2001 : 195).⁴

2.2.1. Robustesse de l'effet de fréquence

Tous les travaux en psycholinguistique de l'adulte ont mis en évidence, très rapidement, l'existence d'effets de fréquence d'abord sur des tâches lexicales et infra-lexicales. En perception auditive de la parole, ces phénomènes ont été montrés (Howes, 1957) et répliqués à de très nombreuses reprises avec des tâches on-line et off-line : les mots plus fréquents sont perçus et reconnus beaucoup plus vite que les mots rares. A cela s'ajoute que ces effets persistent pour les tâches infra-lexicales n'impliquant pas nécessairement l'accès au sens du mot (reconnaissance et détection de phonèmes, de syllabes, de clusters, etc.), et qu'ils sont extrêmement robustes : l'effet de fréquence résiste en effet à l'augmentation du débit, à la dégradation acoustique du signal, aux diverses caractéristiques du locuteur et à tout ce qu'il est convenu d'appeler les "effets de contexte" (contexte syntaxique, sémantique qui vont favoriser ou au contraire gêner la reconnaissance de mots : pour une synthèse, cf. Tyler & Frauenfelder, 1987).

En production, les effets de fréquence lexicale ont également été mis en évidence et attribués non à une facilité articulatoire, ni à la disponibilité sémantique de l'item lexical à récupérer, mais à la morphophonologie du mot (Jeschaniak & Levelt, 1994 ; Segui & Ferrand, 2000). De manière moins convaincante, Jurafsky (2002) et Bybee (2006) soutiennent que l'allomorphie est conditionnée par la fréquence lexicale (pour un contreexemple, cf. Carvalho, 2004). L'approche probabilistique s'appuyant sur l'existence des effets de fréquence s'est ensuite développée de manière importante à partir de la fin des années 90 (pour une synthèse, cf. Jurafsky, 2002).

2.2.2. Traitement probabilistique chez le bébé

Ces résultats ont trouvé un corrélat en psycholinguistique de l'enfant où la question du rôle de la fréquence d'exposition à certaines structures dans la genèse des catégories est un des domaines actuellement les plus étudiés. (Morgan & Saffran, 1995 ; Saffran *et al.*, 1996 ; Echols *et al.*, 1997 ; Gomez & Gerken, 1999 ; Peperkamp & Dupoux, sous presse; Werker & Fennell, 2004). Ces divers travaux interrogent notamment la nécessité d'une composante phonologique de la grammaire universelle (GU) visant à compenser la pauvreté du stimulus, considéré en syntaxe comme inapte à fournir à l'enfant des stimuli de bonne qualité en quantité suffisante pour qu'il puisse développer sa compétence linguistique par le seul contact avec l'environnement. En phonologie, la question de la pauvreté du stimulus ne peut pas être envisagée sous un angle quantitatif, puisque l'ensemble des objets qu'il faut acquérir est (i) fini, (ii) présenté de manière très redondante à l'enfant dès ses premières années de vie⁵. Il a donc été envisagé que les

enfants effectuaient un traitement probabilistique très précoce leur permettant d'opérer progressivement les généralisations nécessaires à l'émergence de la composante phonologique de la grammaire.

Parmi les travaux les plus cités prennent place ceux de Saffran *et al.* (1996) qui, à partir de l'apprentissage de langages fictifs ont, parmi les premiers, 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 ($x \Rightarrow y$) 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*). Ces résultats ont été répliqués.

Des travaux plus récents ont tenté d'observer l'existence de cette capacité de calcul probabilistique chez le bébé et son rôle éventuel dans la genèse des catégories phonologiques. Ainsi, Mayé *et al.* (2002) ont montré que des enfants exposés à une distribution bi-modale de paires de sons où est réalisé un contraste de VOT (*vs* des 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. 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).

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. 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.

Ces études apportent des arguments invitant effectivement à reconsidérer le dispositif d'acquisition chomskyen au profit d'un apprentissage reposant sur le calcul probabilistique et la mémorisation dynamique de toutes les informations de l'environnement sous la forme d'exemplaires encodant les détails phonétiques et servant à la généralisation.

2.2.3. *Fréquence de quoi ?*

Le constat étant fait de l'existence de robustes effets de fréquence, d'une sensibilité des locuteurs aux détails phonétiques fins et de traitements probabilistiques de l'information tant chez l'adulte que chez l'enfant, restent pourtant irrésolues un certain nombre de questions. Parmi elles, celle qui consiste à se demander pourquoi il semble nécessaire de faire de ces effets de traitement probabilistique le principe causal qui conditionne l'émergence et le format des représentations phonologiques.

En ce qui concerne le bébé, il a d'abord été contesté que les enfants aient une sensibilité très précoce aux détails phonétiques fins (Hallé & Boysson-Bardies, 1996 ; Werker *et al.*, 2002) et qu'ils puissent traiter de manière exhaustive la totalité de l'information présente dans le signal. Il a été montré que les tâches de reconnaissance précoce s'appuient plutôt sur une reconnaissance holistique n'incluant pas les variations détaillées des formes lexicales ou syllabiques discriminées.

Ensuite, en admettant que les bébés réussissent à établir une catégorisation systématique de l'input sonore, dégagent-ils pour autant des catégories phonologiques distinctives ? En d'autres termes, peut-on, à partir du seul calcul statistique, sans grammaire et sans lexique, découvrir le système phonologique d'une langue ? Tous les travaux actuels montrent que ce ne semble pas être le cas (pour une synthèse sur cette question cf. Wauquier-Gravelines, 2005 ; Wauquier, 2006 ; Vihman & Sunnari, 2006). Les enfants manifestent d'abord ce qui semble être un traitement du signal sur la base de catégorisations phonétiques alors que le traitement phonologique proprement dit n'intervient qu'à partir de 14 mois, quand la dimension distinctive des catégories dégagées commence également à émerger. Les différents résultats obtenus à ce sujet (Stager & Werker, 1997) montrent en effet qu'à fréquence de distribution égale dans l'input, la fonction de l'information acoustique (ou du détail phonétique fin) le rend plus ou moins accessible à l'apprenant (adulte ou enfant). On peut supposer que les mêmes détails phonétiques fins ne donneront pas lieu au même

traitement selon que l'apprenant cherche un mot, un morphème ou 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*, et que là non plus on ne peut définir *en soi* le détail phonétique fin qui sera inmanquablement fonctionnel pour construire les catégories phonologiques pertinentes dans une langue donnée. Dès lors, on peut effectivement arguer que la fréquence de l'information acoustique et des détails phonétiques fins joue un rôle important dans l'émergence des représentations phonologiques ; mais passé l'énoncé de cette généralité, il semblerait qu'on ne soit pas à l'heure actuelle en mesure de modéliser précisément comment. On tombe *ipso facto*, si l'on réduit le modèle d'apprentissage à cette capacité de traitement précoce, dans un paradoxe de type "la poule et l'œuf" (Dupoux, 2004). Les travaux récents montrent que les modèles d'apprentissage excluant le recours à des primitives phonologiques innées requièrent la modélisation de processus d'apprentissage puissants qui ne pouvant évidemment pas se réduire au seul calcul statistique (Vihman & Sunnari, 2006).

Pour ce qui concerne l'adulte, nous avons soulevé la question du rapport entre détails phonétiques fins et format des représentations en § 2.1. A cela nous voudrions ajouter la remarque suivante qui nourrit les réflexions que nous développons en § 3. Le rôle de la fréquence de distribution des informations du signal dans le traitement en temps réel opéré par les locuteurs existe et a été montré par un nombre substantiel de résultats psycholinguistiques. Mais cette question est formulée de diverses manières selon les auteurs. Alors que certaines phonologies probabilistes en font un principe d'accès (Pierrehumbert, 2001) ou de calcul des représentations ou d'émergence (Coleman, 2002) ou d'apprentissage du système (Peperkamp & Dupoux, sous presse ; Gerken, 2004), d'autres conceptions en font la caractéristique centrale de la composante phonologique (Bybee, 2001, 2006), rejoignant en cela certaines modélisations psycholinguistiques où les systèmes de traitement appuyés sur une mémoire dynamique ne nécessitent plus l'existence d'une composante phonologique (Goldinger & Azuma, 2003). Ceci dessine ultimement le tableau d'une "grammaire" qui serait une mémoire dynamique constituée de "morceaux de signal" de taille et de nature variable (morphèmes, mots, NP, VP, locutions, agrégats de mots), enregistrés de manière détaillée, en des milliers d'exemplaires, à différents niveaux de représentations (acoustique, sémantique, lexical), et dont le stockage est essentiellement conditionné par la fréquence de co-occurrence des éléments qui les constituent (co-occurrence de segments, de syllabes, de morphèmes, de mots). Qu'un psycholinguiste tel que Goldinger retienne ce point de vue est absolument cohérent et conforme à son assise épistémologique, au modèle qu'il propose et aux modes de validation qu'il met en œuvre, mais la question qui nous semble devoir être posée est la suivante : en va-t-il de même pour la phonologie ?

La sensibilité des auditeurs au détail phonétique n'est effectivement pas en soi incompatible avec l'existence de représentations phonologiques discrètes et symboliques (Nguyen *et al.*, 2007). Il semblerait donc que la question posée par les modèles à exemplaires concerne surtout le stockage vs l'absence de stockage des détails phonétiques fins et la modélisation des effets de fréquence dans la mémoire à long terme. Des données empiriques plaident là encore en faveur d'un stockage du détail phonétique dans la mémoire à long terme (Nguyen *et al.*, 2007). Mais ces résultats ne distinguent pas, comme nous l'avons souligné, les variations allophoniques prédictibles par l'auditeur, en vertu de sa connaissance de la grammaire de la langue concernée, et les variantes liées aux caractéristiques du locuteur ou de l'environnement, qui sont aléatoires et donc imprédictibles. Si dans les deux cas on suppose le stockage à long terme des détails phonétiques fins correspondant à ces différentes variations sur la base de leur fréquence distributionnelle, on devrait s'attendre à ce qu'elles donnent lieu à des traitements identiques. Or les travaux de Luce & McLennan (2005) ont montré que variantes allophoniques et variantes interindividuelles n'étaient pas traitées de manière identique par les auditeurs.

Ces résultats invitent donc à considérer que tout détail phonétique fin, fût-il ultra fréquent, et en admettant qu'on puisse le définir précisément, ne donne pas davantage lieu chez l'adulte que chez le bébé à un traitement et à un stockage phonologique uniforme et univoque reposant exclusivement sur l'enregistrement de sa fréquence de distribution dans le signal.

Pour conclure, il nous apparaît qu'il faudrait avant toute chose répondre aux deux questions suivantes.

Primo, la fréquence et le traitement probabilistique constituent-ils, en phonologie, la seule formulation pertinente de l'usage ?

Secundo, jusqu'à quel point la phonologie doit-elle intégrer le rôle de la fréquence (qui a un effet psycholinguistique observable lors du *traitement phonologique*) comme reflet de l'"usage" dans sa modélisation des *représentations*, et pourquoi est-il plus légitime et pertinent qu'elle le fasse — et sous cette forme — que l'inverse ? Nous ne nions pas que cela puisse être le cas, mais il nous paraît important que les présupposés d'un tel choix soient explicités et explicitables, que les solutions alternatives (Encrevé, 1988 ; Labov, 1994, 2001) qui ont été écartées soient mentionnées (cf. Pierrehumbert, 2006), que les prédictions qui en découlent soient empiriquement testables et que le gain en termes de "réalisme d'une phonologie cognitive", puisque c'est sous cet angle que les arguments sont avancés, soit réel et puisse être évalué. Il ne nous semble pas que ce soit clairement le cas.

3. Phonèmes et prototypes : quatre faux problèmes

3.1. Catégories phoniques et universaux

Nous venons de voir en quoi le prétendu “réalisme” des approches empiristes est loin d’être évident sur la base même des arguments psycholinguistiques qui en constituent le socle fondateur. Penchons-nous maintenant sur la question suivante : jusqu’à quel point ces propositions théoriques obligent-elles à remettre en cause les concepts clés de la phonologie ? Cette question est largement indépendante de celle du réalisme cognitif des thèses empiristes : elle se pose quelle que soit l’appréciation que l’on a de celui-ci, ne serait-ce que parce que l’histoire de la phonologie, ne l’oublions pas, a commencé bien avant *SPE*, principal objet des critiques formulées par les théories inductives, et que le courant de pensée inauguré par *SPE* a connu d’autres remises en cause que celles dont il a été question ici. Il convient, pour mesurer la portée théorique des thèses empiristes, de partir à la fois de la façon dont celles-ci sont justifiées et envisagées par leurs défenseurs, et des critiques adressées par ceux-ci à la phonologie.

S’agissant des propriétés mises en cause dans les modèles phonologiques contemporains, nous nous attacherons au point suivant :

- (1) “the relationship between outputs of the phonological grammar to physical properties of the speech signal lies outside of implicit phonological knowledge proper, being instead describable as a universal — language-independent — interface.” (Pierrehumbert, 1999).

A l’encontre de cette hypothèse, les tenants des approches empiristes rappellent qu’une masse de données expérimentales démontre l’existence de “language-specific phonetic patterns down to extremely fine levels of detail, most naturally described using continuous mathematics rather than an inventory of phonetic categories such as the IPA” (Pierrehumbert, 1999).

Il y a ici, nous semble-t-il, un faux problème en partie lié à un pur artefact introduit par *SPE* (§ 1) : l’idée d’une “phonétique universelle”, de “catégories phonétiques universelles”. La phonologie naissante, de Sapir (1925) à Troubetzkoy (1939), n’a jamais nié l’existence d’un continuum substantiel, ni celle de patrons phonétiques particuliers à chaque langue. Bien au contraire, l’affirmation de la nature non-discrète et variable du donné phonétique est, par définition, indissociable de celle, fondatrice, de l’existence de catégories discrètes et invariantes — phonèmes ou traits —, dont le caractère phonétique et la portée universelle sont par ailleurs subordonnés à leur nature foncièrement abstraite et relativiste pour le premier, et à la notion de marque pour la seconde. Autrement dit, si catégories il y a, elles ne sont précisément pas “phonétiques”, au sens où chacune d’entre elles serait pourvue de corrélats physiques précis et invariants de langue à langue (voire au sein de la même langue). D’autre part,

elles ne sont pas davantage “universelles”, au sens où toutes existeraient dans toutes les langues ; seules les catégories dites non-marquées (par exemple, les obstruantes “non-voisées”) sont universelles à ce titre, puisque le caractère plus ou moins répandu des primitives phonologiques dépend de leur degré de marque.

Ce patrimoine épistémologique commun à la plupart des structuralistes survit jusqu’au système de traits binaires proposé par Jakobson, Fant & Halle (1952) où, s’il est toujours exclu de s’intéresser pour elles-mêmes aux propriétés non distinctives des “dix ou douze variétés de *a*, par exemple, que comptent les manuels de phonétique” (Jakobson & Waugh, 1980 : 77), une inflexion s’opère pourtant dans la définition de l’invariant : le trait distinctif correspond désormais aux “particularités communes” aux variantes d’un phonème, et se doit d’avoir un fondement phonétique constant ; il se veut, à ce titre, une catégorie phonétique.⁶ Ainsi, là où Martinet (1974 : 133-136 [1957]) retient la voix comme trait pertinent de /b d g.../ en français, en dépit de sa dégradation voire totale disparition dans certaines positions, Jakobson définira ces consonnes comme des “douces”, propriété physique censée être commune à toutes leurs réalisations, y compris dans des langues où la voix joue un rôle très réduit voire nul dans tous les contextes. Remarquons l’ambivalence de la position jakobsonienne : il est certain que parler de “douces” évite d’avoir à postuler un voisement virtuel dans *gibecière*, *médecin*, etc. ; d’un autre côté, plus l’invariant est général, plus son fondement phonétique est imprécis.

La filiation est évidente entre cette conception substantialiste du trait et le premier chapitre de *SPE*, où il est postulé un inventaire fini de catégories phonétiques universelles. Elle se double cependant d’une rupture, encore plus fondamentale que celle opérée par Jakobson, qui ôte définitivement aux traits de la phonologie générative le caractère de purs invariants formels et *language-specific* qu’avaient les traits pragois : l’abandon du niveau “phonétique systématique”, c’est-à-dire du concept de phonème.

La “forme sous-jacente” de la phonologie générative se compose d’unités morphophonologiques dont la nécessité ne provient plus directement de la réponse à la question “quelles sont les unités distinctives de la langue ?”, mais répond plutôt à cette autre question : “quel est le signifiant de tel morphème ?” Obtenu par le rapprochement des allomorphes, mêlant souvent des alternances dont la nature synchronique est des plus diverses, le signifiant lexical ne se ramène plus, dans son principe, à la concaténation des catégories phoniques distinctives propres à telle langue.

Par ailleurs, entre la forme sous-jacente et la forme de surface, il y a autant de niveaux que de règles ordonnées, les faits allophoniques ne constituent plus qu’un sous-ensemble des régularités considérées, et les traits effacés ont le même statut formel que ceux qui sont ajoutés : dans ce continuum dérivationnel, il n’y a pas de stade équivalent à un niveau phonémique désormais aboli.

En l'absence d'un principe théorique explicite qui justifie de poser tel ou tel trait dans telle ou telle langue, il n'est pas étonnant qu'on ait eu recours à l'hypothèse d'un inventaire fini de catégories phoniques universelles : elle donnait l'illusion d'un système clos qui ne pouvait plus s'appuyer sur la pertinence phonémique de la phonologie structurale. Nul besoin donc de s'appuyer sur la critique empiriste de *SPE* pour refuser l'idée d'un inventaire universel de catégories phonétiques : une telle notion n'a tout simplement aucun sens ; le concept même de catégorie n'a de validité qu'au sein d'un système donné d'objets dont la valeur, on le rappellera en § 3.2, découle tout autant de l'ensemble de leurs réalisations que du réseau de distinctions dans lequel ils s'insèrent.

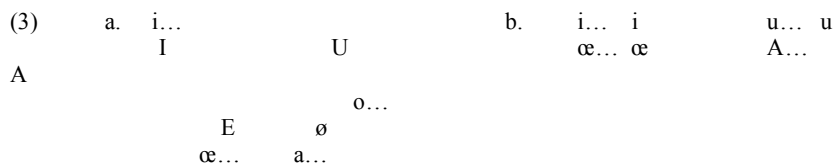
3.2. Variabilité et invariants

Faut-il pour autant déduire de l'inexistence d'une "phonétique universelle" que l'analyse des "language-specific phonetic patterns" dont parle Pierrehumbert doit descendre jusqu'à "un niveau de détail extrêmement fin" ? Rien de moins sûr. Les cas de variabilité de langue à langue qui, selon les thèses empiristes, impliqueraient de renoncer aux catégories discrètes de type phonémique au profit d'une analyse phonétique fine se laissent ramener à deux types que l'on peut schématiser de la façon suivante :



Dans (2a), une même catégorie — ou "label" (Pierrehumbert 2001 : 199) — phonique correspond, selon les langues, à diverses réalisations — ou, pour employer un terme très utilisé dans certaines de ces approches, à divers "prototypes" (cf., par exemple, Pierrehumbert, 1999 ; Bybee, 2001 ; Bybee & Hooper, 2001). Dans (2b), à l'inverse, un même signal de la parole constitue, selon les langues, la réalisation de diverses catégories phoniques.

3.2.1. Le cas de figure en (2a) renvoie à un vieux débat dans la phonologie prégénérative : "what phonemes does a language 'have' ?" (Lass, 1984 : 134). Hockett (1955 : 76 suiv.) donne les "allophones caractéristiques" des voyelles du fox en (3a), et leur interprétation "phonémique" en (3b) :



(3b) est le fruit de la “normalisation” des prototypes en (3a) : n’ont été retenues que les propriétés distinctives.⁷ Il en résulte (i) que le système comporte le plus petit nombre possible de “cases vides” ; (ii) que rien ne distingue le système ainsi normalisé de celui d’une langue qui offrirait, par exemple, les réalisations prototypiques en (4) :

(4) i... i u... u
 æ... æ a... a

Les caractéristiques spécifiques des voyelles du fox ont été gommées en (3b) au nom d’un principe d’invariance. Jusqu’à quel point ce principe est-il justifié ? La voyelle [o...] peut-elle être “phonémiquement haute” (/u.../) ? Il faut bien avouer que la plupart des linguistes n’ont apporté qu’une réponse pragmatique à ces questions : le degré de réalisme dans la représentation des segments/phonèmes fut et reste conditionné par les possibilités notationnelles fournies par l’API, bien plus que par une définition précise de ce qui, en phonologie, constitue un invariant. Le rejet plus ou moins explicite du concept de phonème a, de plus, empêché, à partir de la fin des années 60, de tirer de ces questions des conclusions théoriques sur ce qui, en phonologie, constitue un invariant. Essayons pourtant d’y voir plus clair.

Le principe d’invariance phonémique résulte de la prise en compte des deux axes syntagmatique et paradigmatique. Le phonème (dans sa version pragoise : cf. § 3.3) est conçu comme un ensemble de traits, dits distinctifs :

- (5) a. communs à un certain nombre de phones en distribution complémentaire,⁸
 b. nécessaires et suffisants pour distinguer cet ensemble de tous les autres pouvant figurer dans les mêmes contextes.

Or l’élagage des traits non pertinents n’a ni la même ampleur ni la même justification selon qu’il est dicté par (5a) ou par (5b). La première condition aboutit à un inventaire du type de celui en (3a) : sont ici écartés des traits imputables au contexte. C’est la condition en (5b) qui produit le système hautement abstrait en (3b). Et c’est cette condition qui peut prêter le flanc à une double critique.

D’une part, ce n’est pas le caractère abstrait en soi que l’on peut reprocher aux phonèmes de (3b) ; c’est le caractère *arbitraire* de cette abstraction qui est en cause. Ainsi, comme le remarque Lass (1984 : 137), qu’est-ce qui empêcherait, pour le fox, l’hypothèse moins radicale en (6), dût-elle être moins “économique” que celle en (3b) ? Il n’y a pas de réponse à cette question.⁹

(6) i... i
 e... e o... o
 a... a

D'autre part, Jakobson (1931) puis Martinet (1970 [1955]) ont voulu expliquer, en partie, la variation et le changement phoniques par une tendance — confirmée sur des bases statistiques par Clements (2003) — au “rendement fonctionnel” optimal des traits. Les systèmes phonologiques seraient soumis à un principe d'économie qui les pousse à former le plus grand nombre de phonèmes à l'aide du plus petit nombre de traits. Les entorses à ce principe, phonétiquement conditionnées, tendraient à être corrigées par l'élimination des “cases vides” correspondant à des combinaisons de traits inexistantes, comme */u... u/ dans (6). Or, si ce principe d'économie se veut explicatif, il est clair que les systèmes ne sauraient présenter par définition, contrairement à (3b), un degré d'intégration optimal. La théorie diachronique ne peut être prouvée (ou infirmée) que si la réalisation phonétique des phonèmes reflète leur processus d'intégration, allant d'une étape où, comme en (3a), celle-ci paraît loin d'être achevée, vers un stade où, comme en (4), la contrainte d'économie est pleinement satisfaite. Autrement dit, si (3b) peut constituer la cible d'une stratégie évolutive possible pour le système en (3a), celui-ci ne saurait représenter la simple réalisation des unités de celui-là, sous peine de vider de toute portée explicative une théorie du changement phonétique qui a le mérite de reposer sur une propriété vérifiable (et quantifiable) des systèmes phonologiques : leur caractère relativement intégré.

Il s'ensuit que le critère paradigmatique, ajouté en (5b) à la définition du phonème et responsable de la normalisation radicale des systèmes comme celle en (3b), n'est pas une condition nécessaire à la définition d'une catégorie phonologique. Or, si tel est le cas, alors on en arrive à une conclusion remarquable : *phonème et prototype sont pratiquement le même objet théorique*.¹⁰ On ne peut donc pas, contrairement à ce que prétendent les thèses empiristes, inférer de la variabilité des réalisations d'un “label” selon les langues qu'il faille renoncer aux catégories discrètes de type phonémique ; bien au contraire, le concept de phonème *comprend* celui de prototype, qui n'en est en général qu'un cas particulier — celui qui découle de (5a) —, et, par là-même, il en fournit, comme on le verra en § 3.3, une définition plus rigoureuse et plus motivée que celle qu'en donnent les approches inductives.

3.2.2. Peut-on pour autant affirmer que la définition du phonème/prototype peut et doit s'affranchir de tout conditionnement paradigmatique ? Nous pensons que non, et qu'une des raisons en est d'ailleurs fournie par les tenants des thèses empiristes eux-mêmes. Intéressons-nous à l'autre type de variation entre les langues qui, selon ces thèses, impliquerait de subordonner l'idée de catégories discrètes à l'analyse fine du continuum phonétique : celui, schématisé en (2b), où un même signal de la parole correspond à des “labels” distincts selon les langues. Et prenons l'exemple de la variabilité du

voice onset time (VOT). Il est exact, comme le rappelle Pierrehumbert (2001 : 196), qu'un "speech signal that constitutes a /p/ in one language may provide an example of /b/ in another". Mais ceci ne pose aucun problème particulier à une pensée phonologique non biaisée par la "phonétique universelle" de *SPE*. C'est ici qu'intervient nécessairement la dimension paradigmatique de la catégorisation phonémique. Dans les deux cas, on a affaire au terme *lenis* d'une opposition de force, mais sa valeur de marque diffère : là où le signal constitue une réalisation de /p/, la consonne *lenis* correspond au terme non-marqué de la distinction, en s'opposant à un terme *fortis* marqué (l'aspirée /p^h); là où le signal est interprété comme la réalisation de /b/, la consonne *lenis* correspond au terme marqué de la distinction, en s'opposant à un terme *fortis* non-marqué (/p/ à voix modale). Autrement dit, la différence de catégorisation implique de faire l'hypothèse de consonnes différentes dans le paradigme : si X = /p/, alors il existe /p^h/ et /b/ n'est pas nécessaire ; si X = /b/, alors il existe /p/ et /p^h/ n'est pas nécessaire.

Les catégories phonémiques revêtent dès lors à la fois un aspect universel — il y a dans les deux systèmes une opposition de force, dont X instancie le terme *lenis* — et un aspect paramétrique — le contenu des "labels" *fortis* et *lenis* découlant de la marque : dans un cas *lenis* = non-aspirée, dans l'autre *lenis* = voisée. On retrouve là les deux niveaux d'analyse, "réaliste" et "abstrait", proposés en (3) pour le système vocalique du fox. Toutefois, contrairement au fox, ici c'est la définition du phonème la moins abstraite qui découle de la prise en considération des relations paradigmatiques, soit de la condition (5b). Celle-ci est donc nécessaire si l'on veut rendre compte des différences de catégorisation entre les langues. La condition (5b) pouvant aussi s'avérer inadéquate comme on l'a vu en § 3.2.1, il s'ensuit que beaucoup reste à faire en vue de déterminer la véritable portée des relations paradigmatiques dans la définition du phonème. La théorie de la spécification contrastive fondée sur le Successive Division Algorithm (Dresher, 2003) pourrait être un premier pas dans cette direction.

Quoi qu'il en soit, s'il est vrai que le prototype équivaut à la définition moins abstraite du phonème, alors on en arrive à deux autres conclusions remarquables qui viennent préciser celle à laquelle on était arrivé en § 3.2.1. *Primo*, le prototype peut se ramener à un simple *trait* : ce qui est prototypique de la consonne *lenis* est, dans le premier cas ci-dessus, l'absence d'aspiration, et, dans le second, son caractère voisé. *Secundo*, le prototype n'est pas, par principe, "incarné" (*embodied*), au sens où il serait systématiquement représenté dans le signal par un allophone caractéristique : si X est commun aux deux langues et que celles-ci diffèrent dans sa catégorisation, alors c'est que chacune sélectionne dans X un ensemble de traits différent qui constituera un prototype littéralement abstrait.

3.3. Prototypique inductif *versus* prototypique formel

3.3.1. C'est dans cette notion d'ensemble de traits que réside un troisième faux problème. Non seulement les approches empiristes se trouvent avoir fait l'impasse, à la suite de la théorie générative qu'elles critiquent, sur le concept de phonème (cf. § 3.2), mais encore, pour autant qu'elles en aient retenu quelque chose, elles ne se réfèrent implicitement qu'à une tradition. Or celui-ci n'a jamais relevé d'une définition unique et unanimement partagée. Deux conceptions sont à distinguer, qui se sont plus souvent ignorées qu'opposées : la première unit la théorie du phonème de Jones au courant néobloomfieldien et à l'école de Leningrad issue de la pensée de Chtcherba ; l'autre rapproche l'école de Prague et celle de Moscou. Selon la première conception, le phonème apparaît avant tout comme un *ensemble de sons* ; selon la seconde — celle que nous avons suivie en § 3.2 —, le phonème est d'abord un *ensemble de traits*.

Ce clivage théorique est à l'origine tant du principe de biunivocité ("*once a phoneme, always a phoneme*")¹¹ et du refus de l'idée de neutralisation et d'archiphonème chez la plupart des distributionalistes que, de l'autre côté, de la "typologie logique des oppositions phonologiques" et de la notion de marque proposées par Troubetzkoy. Dans le premier cas, la catégorie phonémique est une sorte de "substance seconde" aristotélicienne, qui n'existe que par ses diverses manifestations, le phonème ayant une fonction fondamentalement prédicative : "tel son appartient (ou n'appartient pas) à tel phonème". Or de tels énoncés, courants dans la tradition distributionaliste, ne jouent que marginalement dans la tradition européenne. Ici c'est l'invariant pur qui prime sur ses manifestations et qui constitue le sujet dont on décrit les propriétés : "tel phonème se réalise de telle façon dans tel contexte".

Est-ce un hasard ou l'effet d'une filiation principalement anglo-saxonne ? Les actuelles théories des prototypes sont fondées sur le nuage allophonique déduit du signal, dont émergerait la variante prototypique : là encore, la catégorie est d'abord représentée par un de ses avatars (cf. Bybee, 2001 : § 3). L'ironie de l'histoire, selon nous, est que de telles approches peinent précisément à justifier l'existence et l'identité du prototype : qu'est-ce qui fait que, parmi toutes les variantes d'un phonème, il y en a une dotée d'un statut particulier, et comment la définir ? Les théoriciens de la perception fondent, en général, leur réponse à ces questions sur la fréquence statistique des variantes : est censée être prototypique la variante la plus fréquemment attestée dans le signal (cf. Bybee, 2001) ; c'est sa fréquence qui, selon Kuhl (1991), lui confère le pouvoir d'exercer un "effet magnétique" tel qu'il y aura autant de nuages d'exemplaires que de catégories phoniques. Or l'argument de la fréquence recouvre, en vérité, un raisonnement fondamentalement circulaire (cf. § 2.2.3) : ce qu'il faudrait

expliquer c'est pourquoi telle variante est précisément la plus fréquente et est donc à même de jouer le rôle d'un prototype.

3.3.2. C'est là, paradoxalement, que la conception pragoise du phonème, qui n'est qu'indirectement fondée sur ses réalisations, peut pourtant fournir une réponse directe à la question de l'existence et de l'identité des prototypes. Dans les cas les plus simples, s'il est vrai que le phonème implique de considérer l'ensemble des traits communs à ses allophones, alors, s'il y a un allophone qui ne présente *que* ces traits communs, celui-là équivaut à un prototype (cf. § 3.2.1). En établissant un lien entre forme (invariante) et substance (allophonique), une telle définition a un pouvoir de généralisation plus élevé que celle que proposent les thèses inductives substituant le prototype au phonème : la définition du prototype en tant qu'intersection des variantes découle de la définition générale du phonème ; c'en est un cas particulier au sens logique du terme. Elle a ainsi une caractéristique qui la distingue de la plupart des thèses jusqu'ici avancées sur la notion de prototype phonique : elle en implique l'*existence* au moins théorique, puisqu'on peut toujours définir un ensemble de traits obéissant à la condition en (5a).

Nous avons vu en § 3.2.1 l'exemple des voyelles du fox. Donnons-en deux autres. Dans une langue, /d/ se réalise [ɰ] à l'intervocalique, [d] ailleurs. L'occlusion étant, par hypothèse, le mode consonantique par défaut, le prototype est [d] = {coronal, voisé} □ [ɰ] = {coronal, voisé, approximant} = {coronal, voisé} = [d]. Second exemple : dans de nombreuses langues, /p t c k/ sont voisées entre voyelles et sourdes ailleurs. Soient, par exemple, [p] = {labial} et [b] = {labial, voisé}, la distinction de voisement étant, là encore, privative, le prototype de /p/ est [p] □ [b] = {labial} = [p].

Dans ces exemples, le prototype correspond à l'allophone non marqué du phonème. Est-ce toujours le cas ? Certains faits suggèrent qu'il y a une limite à la sous-spécification du prototype. En malais, la dorsale /k/ se réalise [ʔ] à la coda et [k] ailleurs. Il est clair que [k] □ [ʔ] = [ʔ], si l'on considère, avec Clements (1985), que la glottale n'est que le résultat du déliage du nœud dominant les traits de lieu d'articulation (*place features*). Or il est pour le moins contre-intuitif de voir en [ʔ] le prototype de la catégorie en question. Deux raisons possibles à cela, qui jouent de façon complémentaire, situées qu'elles sont sur des axes, syntagmatique et paradigmatique, distincts : (i) est-ce à cause du caractère minoritaire de [ʔ] vis-à-vis de [k] ? ; (ii) est-ce à cause du moindre "potentiel distinctif" de [ʔ] vis-à-vis de celui de [k] ?

Concernant (i), il est certain, en tout cas, que la moindre fréquence de la glottale s'ensuit du fait distributionnel suivant : ici, à l'opposé des deux exemples précédents, *c'est l'allophone sous-spécifié ([ʔ]) qui est circonscrit à un contexte marqué* (la coda), lui-même minoritaire par rapport à la position d'attaque syllabique. S'agissant de (ii), on remarquera, en effet, que les traits résiduels de la glottale équivalent à ce qui est commun à *toute une série de*

phonèmes de la langue (/p t k.../). S'il faut exclure l'allophone sous-spécifié de la définition du prototype, une condition paradigmatique doit donc être introduite encore une fois (cf. § 3.2.2). Nous proposons de compléter la définition du prototype de la manière suivante :

- (7) Est prototypique d'un phonème X l'ensemble des traits :
- a. communs à ses allophones,
 - b. suffisants pour opposer X aux autres phonèmes *de la même classe sur une base équipollente*.

On le voit, (7) est en tous points conforme à la plus classique des définitions du phonème pragois, dont la réalisation prototypique évite la redondance comme la sous-spécification.

3.3.3. Il existe un second avantage à une définition formelle (*phoneme-based*) du prototype : concevoir un tel objet sur la base de conditions du type de celles en (7) revient à faire une prédiction sur son *identité* qui reste, en principe, indépendante du critère, circulaire, de la fréquence. On peut supposer que cette définition des prototypes [p], [d] et [k] sera confirmée par une étude sur la fréquence d'usage respective des allophones [p] vs. [b], [d] vs. [ʃ] et [k] vs. [ʎ]. Il reste qu'elle n'en dépend nullement. Le rapport entre prototype et haute fréquence ne tient ici que s'il est vrai que le terme non-marqué (allophone ou contexte) est plus récurrent que le terme marqué. Or, si les prototypes définis par induction découlent directement de la fréquence, le concept de marque en est théoriquement indépendant.¹² Ce n'est donc pas parce que telle variante est la plus fréquemment attestée dans le signal qu'elle constitue un prototype ; c'est, au contraire, parce qu'elle est, en vertu de (7), un prototype qu'elle est, en principe, la variante la plus fréquente.

Si le pouvoir prédictif de cette conception des prototypes est, par définition, supérieur à celui de toute thèse qui les abstrait directement du signal, est-elle pour autant plus adéquate au plan empirique ? Prenons le cas de la consonne /r/. En portugais brésilien, les réalisations, en variation dite "libre", du /r/ "fort" vont d'un *trill* apical à des productions dorsales et pharyngales au degré de voisement variable : [r x X ©]. Les traits communs à l'ensemble de ces allophones ne suffiraient pas à caractériser un prototype plausible, étant aussi partagés par la totalité des autres consonnes du portugais. En vertu de la condition (7b), une spécification supplémentaire est donc nécessaire. Or le choix du trait "fricatif" (sur la base des usages les plus — et de plus en plus — fréquents, ainsi que du nombre d'allophones qui en sont pourvus) ne suffirait toujours pas à distinguer, sur une base équipollente, cette catégorie des consonnes /f v s z S Z/, qui, elles aussi, existent en portugais ; il violerait de plus la condition (7a), puisque le *trill* apical compte aussi au nombre des variantes. Quelle spécification choisir ? L'ajout au trait fricatif d'un trait "dorso-pharyngal" lèverait, certes, la première difficulté : il

suffirait à distinguer /r/ de /f v s z S Z/. Reste que la condition (7a) n'est toujours pas respectée et on voit mal comment elle pourrait l'être.

Conclusion : de même que le prototype d'un phonème peut se réduire à seulement une partie des propriétés d'un segment (cf. § 3.2.2), de même un phonème peut avoir *plus d'un prototype* : c'est le cas du /r/ en portugais, dont les prototypes sont la liquide [r] et l'uvulaire [X] (à "équidistance" de [x] et [©]). Il est possible qu'une étude statistique de ces variantes aurait conduit à la même conclusion. Il est, en tout cas, *certain* qu'une définition formelle du prototype y mène. Et c'est là toute la différence entre les deux approches.

3.4. Allophonie et linéarité

Un quatrième faux problème posé par les approches empiristes des catégories phoniques tient à un autre oubli dans l'histoire de la pensée phonologique : après celui du phonème pragoïse, celui de la plurilinéarité des représentations posé par la théorie autosegmentale des années 70-80. Cet oubli sous-tend à la fois la critique souvent adressée aux modèles formels par de nombreux tenants des thèses inductives, à savoir que la phonologie opère avec un "inventory of phonetic categories such as the IPA" (Pierrehumbert, 1999 ; cf. § 3.1), et une bonne partie du fondement empirique de l'importance accordée par ces thèses au "détail phonétique" dans la connaissance phonologique des locuteurs (cf. § 3.2). Que cet argument repose au moins partiellement sur une aporie, c'est ce que montre, en diachronie, la morphogénèse des catégories phoniques. Posons-nous la question : *quand et comment un allophone devient-il un prototype*, c'est-à-dire, virtuellement, un phonème distinct ?

Le vieux russe avait une voyelle antérieure haute brève (écrite *ь* et que nous noterons *i*), qui palatalisait la consonne précédente (représentée par *t*), entraînant, en bonne phonologie classique, les règles allophoniques suivantes :

- (8) a. /t/ → [tʲ] / __ i
 b. /t/ → [t] ailleurs

Il s'ensuit qu'à ce stade il n'y avait qu'un unique phonème /t/ : la différence phonétique entre [tʲ] et [t] n'était pas perçue comme telle par les locuteurs ; elle était imputée à la présence de /i/ dans le premier cas et à son absence dans le second. Plus tard, nous dit-on, /i/ disparut ; son effet palatalisant, cependant, demeura ; dès lors, puisque la perte de /i/ rendait la consonne [tʲ] imprédictible, il s'ensuivit l'émergence d'une opposition /t/ : /tʲ/ par scission du */t/ unique initial.

Or cette seconde étape du changement est absurde en termes structuralistes. Si le contexte d'un allophone vient à changer, l'allophone doit changer avec lui. Imaginons un verlan espagnol : /un+lago/ "un lac", réalisé

[unlão], les obstruantes voisées se spirantisant à l'intervocalique, donnera [uÑgola] et non *[uÑola]. De la même façon, s'il est vrai que le caractère palatal du [tʲ] du vieux russe était imputé par les locuteurs au phonème /i/, la perte de celui-ci aurait dû conduire à la dépalatalisation de la consonne. La préservation de [tʲ] est donc incompréhensible.

Pourtant, les faits contredisent ces prédictions et semblent confirmer la thèse insoutenable de la “transphonologisation” de la palatalité. On sait, en effet, que */ti/ a donné [tʲ] en russe : [i] > Ø n'a donc pas entraîné la dépalatalisation de */t/. Comment un tel changement a-t-il été possible ? Il y a trois réponses possibles à cette question : la première montre les limites de la phonologie classique ; la deuxième, celle des approches empiristes ; seule la dernière est satisfaisante, qui vient amender la vision linéaire du phonème tout en en conservant l'aspect discret et invariant.

Première hypothèse : d'après les principes structuralistes et la plupart des auteurs ayant abordé ce paradoxe (cf. Carvalho, 2005 : 368-370), si le changement a été possible c'est que [tʲ] était *déjà* un phonème avant la perte de [i]. Or ceci aboutirait à une contradiction dans les termes : étant données les règles en (8), nous aurions là deux phonèmes (phonétiquement voisins) en distribution complémentaire ! C'est pourtant à une telle position que s'en sont tenus la plupart des phonologues, qui ont parlé à ce sujet de phonèmes “marginaux”, “secondaires” ou de “quasi-phonèmes”. Malheureusement, comme le souligne Janda (2003 : 409), de tels concepts “provide neither any motivation for *why* nor any mechanism for *how* certain [...] allophones which are in complementary distribution could become phonologized”. En effet, contrairement à ce que semble suggérer Bybee (2001 : 149 suiv.), il n'existe aucune unité de mesure intuitive de la similarité phonétique permettant de décider si deux sons appartiennent ou non au même phonème.

Une deuxième solution, diamétralement opposée à la précédente, est fournie par des études récentes, par exemple Booij (2002) et Janda (2003), qui, à l'instar des postulats des approches inductives, déduisent de ce type de paradoxe diachronique que les propriétés allophoniques — le fameux “détail phonétique” — doivent être stockées dans le lexique pour pouvoir survivre après la perte du contexte qui en était la cause. Cette thèse se heurte toutefois au même problème que la précédente : si tout est dans le lexique, comme le prétend notamment Bybee (2001 : § 3), *pourquoi se fait-il que certains allophones survivent après la perte de leur contexte, alors que d'autres non* (voir l'exemple espagnol *supra*) ?

Il est assez ironique de constater que le point de vue empiriste, sous-jacent à la deuxième hypothèse, est, tout compte fait, victime du même “mirage alphabétique” que celui qu'il ne cesse de dénoncer chez les défenseurs d'une conception formelle de la phonologie, lorsque ceux-ci sont accusés de s'en tenir, dans leurs représentations, à un “inventaire de catégories phonétiques tel que l'API”. Dire que le caractère palatal de [tʲ] est un trait allophonique de /t/ et est stocké comme une propriété de la consonne

revient à reprendre à son compte, *nolens volens*, la vieille thèse de la linéarité des représentations, pur artefact induit par l'écriture, et à raisonner selon un schéma des plus classiques : ainsi, comme en (8a), “/t/ se réalise [tʰ] devant /i/”. Or, si vraiment la palatalité de /t/ était attribuée par les locuteurs à la catégorie /i/, pourquoi ne pas envisager les choses à rebours et, en revenant sur ce point à une conception qui était, dès 1881, celle de Baudouin de Courtenay (Cao, 1985 : 165, n. 26), dire plutôt : “/i/ se réalise [iʰ] en combinaison avec /t/” ? Cela implique de renoncer non au phonème en tant que *Gestalt*, définie par le principe d'invariance, mais simplement à son *format* linéaire classique. Telle est l'une des possibilités qu'offre aujourd'hui la théorie autosegmentale. S'il est vrai que /i/ est d'abord un ensemble de traits, alors la raison de la préservation de la palatalité de la consonne après la perte de la voyelle tient, en quelque sorte, à ce que “la catégorie /i/ est restée malgré la chute de la voyelle [i]”, ce qui est, bien sûr, inexprimable, dans les modèles classiques faute de distinction entre *mélodies* (*ti*) et *positions* (*CV*) :

(9)	a.	[tʰi]	b.	[tʰ]
		t i		t i
		g l g	>	g l
		C V		C V

La “chute de la voyelle” a consisté en la perte de l'association de la mélodie /i/ à la position V : cf. (9b). A l'inverse de ce qui s'est produit en russe, la dépalatalisation de /t/ aurait impliqué la disparition de la mélodie elle-même.

Ainsi, contrairement à ce que prétendent les critiques empiristes de la phonologie formelle, les représentations du type /xxxx/, toujours courantes dans l'usage de la plupart des auteurs,¹³ ne sont, au mieux, que des raccourcis notationnels commodes mis pour des structures qui, à l'image de (9), ne ressemblent guère à des “letter-like tokens” et où la dimension temporelle, résultant de l'interaction entre mélodies et positions, a au moins autant d'importance que la dimension symbolique. Or seules ces structures permettent d'éviter les écueils rencontrés par les deux premières tentatives d'explication du paradoxe diachronique dont nous parlions. D'une part, il n'y a pas deux phonèmes /t/ et /tʰ/ en distribution complémentaire puisque /i/ est toujours là, fût-il partiellement flottant comme dans (9b). D'autre part, à l'encontre des thèses inductives, si la palatalité de /t/ est bien dans le lexique et a donc été préservée en russe, c'est qu'elle n'est précisément pas un trait allophonique de la consonne, du “détail phonétique”, mais un trait distinctif de la voyelle!¹⁴

4. Conclusion

On le voit, une bonne partie de l'argumentation développée par le retour de balancier empiriste repose sur la construction d'un adversaire formaliste, largement biaisée en ce qu'elle est essentiellement dirigée contre la tradition phonologique directement issue de *SPE*.¹⁵ Plus généralement, les arguments avancés en défense d'une phonologie à exemplaires non symbolique se laissent ramener à une critique directe de la conception de la grammaire définie par la "compétence" chomskienne. Objet psychologique envisagé sous le seul angle "cognitif", cette grammaire est dénoncée comme arbitraire, réductrice, non réaliste, coupée de la substance, non contrainte par l'usage et le contexte.

Quel que soit le fondement de cette critique — par ailleurs souvent justifiée — adressée à la linguistique chomskienne, et dès lors que l'on considère les notions et les arguments avancés par les thèses inductives — variation, prototypes, fréquence, catégorisation — sous l'angle d'une conception de la phonologie qui ne se réduirait pas à la formulation *SPE*, il nous apparaît que ce qu'il y a d'intéressant dans ces thèses n'est guère original et ce qu'il y a d'original n'intéresse pas prioritairement la phonologie.¹⁶ Au mieux, ce revirement empiriste constitue une "révolution" toute relative, réactivant des formulations structuralistes des questions soulevées. Pour le reste, les modèles encodant toute la variation phonétique dans les représentations établissent un rapport nullement nécessaire entre ce qu'ils critiquent et ce qu'ils proposent. Les arguments avancés sont insuffisants pour que l'on substitue une approche probabiliste à exemplaires "incarnés" à une conception qui, tout en maintenant les fondements formels et symboliques de la phonologie, s'emploie à échapper aux travers générativistes.

Bibliographie

- ALLEN, Jont ; MILLER, Joan (2004), "Listener sensitivity to individual talker differences in voice-onset-time", *Journal of the Acoustical Society of America* 115 : p. 3171-3183.
- ANDERSON, Stephen R. (1985), *Phonology in the twentieth century : theories of rules and theories of representations*. Chicago : University Of Chicago Press.
- BOOIJ, Geert (2002), "The balance between storage and computation in phonology". Dans Nootboom, S., Weerman, F., Wijnen, F. (eds), *Storage and computation in the language faculty* : p. 115-138. Dordrecht : Kluwer.
- BYBEE, Joan L. (2001), *Phonology and language use*. Cambridge : Cambridge University Press.
- BYBEE, Joan (2006), *Frequency of use and the organization of language*. Oxford : Oxford University Press.

- BYBEE, Joan L. ; HOOPER, Paul (2001), "Introduction to frequency and the emergence of linguistic structure". Dans Bybee, J., Hopper, P. (eds), *Frequency and the emergence of linguistic structure* : p. 1-24. Amsterdam : Benjamins.
- CAO Xuan Hao (1985), *Phonologie et linéarité. Réflexions critiques sur les postulats de la phonologie contemporaine*. Paris : SELAF.
- CARR, Philip (2000), "Scientific realism, sociophonetic variation and innate endowments". Dans Burton-Roberts, N., Carr, P., Docherty, G. (eds), *Phonological knowledge, Conceptual and Empirical Issues* : p. 67-105. Oxford : Oxford University Press.
- CARVALHO, Joaquim Brandão de (2004), "L'analogie est-elle un fait fonctionnel ou grammatical ? Le cas de l'apophonie portugaise", *Corpus* 3 : p. 101-125.
- CARVALHO, Joaquim Brandão de (2005), "Building on old foundations : From phonemic theory to C/V-segregation", *Folia linguistica* 39 : p. 363-384.
- CARVALHO, Joaquim Brandão de (2007), "From features to contours : why forms, not acoustic signals, should be modelled", *Mathématiques et sciences sociales*. [A paraître.]
- CHOMSKY, Noam ; HALLE, Morris (1968), *The sound pattern of English*. New York : Harper & Row.
- CLEMENTS, George N. (1985), "The geometry of phonological features", *Phonology yearbook* 2 : p. 225-252.
- CLEMENTS, George N. (2003), "Feature economy in sound systems", *Phonology* 20 : p. 287-333.
- COLEMAN, John (2002), "Phonetic representations in the mental lexicon". In Durand, J., Laks, B. (eds), *Phonetics, Phonology, and Cognition* : p. 96-130. Oxford : Oxford University Press.
- DEVITT, Michael (2003), "Linguistics is not psychology". Dans Barber, A. (ed.), *Epistemology of language* : p. 107-139. Oxford : Oxford University Press.
- DOCHERTY, Gerard ; MILROY, James ; MILROY, Lesley ; WALSHAW, David (1997), "Descriptive adequacy in phonology : a variationist perspective", *Journal of Linguistics* 33 : p. 275-310.
- DRESHER, B. Elan (2003), "Contrast and asymmetries in inventories". Dans di Sciullo, A.-M. (ed.), *Asymmetry in grammar, Volume 2 : Morphology, phonology, acquisition* : p. 239-257. Amsterdam : Benjamins.
- DUPOUX, Emmanuel (2004), "The acquisition of discrete segmental categories : data and model", *Proceedings of the 18th International Congress of Acoustics*, Kyoto, April 2004.
- ECHOLS, Catharine H. ; MARTI, C. Nathan (2004), "The identification of words and their meanings : from perceptual biases to specific language-cues". Dans Hall, G. ; Waxman, S. R. (eds), *Weaving a lexicon* : p. 41-79. Cambridge, Mass. : The MIT Press.
- ENCREVE, Pierre (1988), *La liaison avec et sans enchaînement : Phonologie tridimensionnelle et usages du français*. Paris : Seuil.
- FOULKES, Paul ; DOCHERTY, Gerard (2006), "The social life of phonetics and phonology". *Journal of Phonetics* 34 : p. 409-438.
- GASKELL, Gareth (2003), "Modelling regressive and progressive effects of assimilation in speech perception", *Journal of Phonetics* 31 : p. 447-463.

- GERKEN, Lou Ann (2004), "Nine-month-olds extract structural principles required for natural language", *Cognition* 93 : p. 89-96.
- GOLDINGER, Steven (1998), "Echoes of echoes? Words and voices: an episodic theory of lexical access", *Psychological review* 105 : p. 251-279.
- GOLDINGER, Steven ; AZUMA, Tamiko (2003), "Puzzle solving science : the Quixotic quest for units in speech perception", *Journal of phonetics* 31 : p. 305-321.
- GOMEZ, Rebecca L. ; GERKEN, Lou Ann (1999), "Artificial language learning by 1-year-olds leads to specific and abstract knowledge", *Cognition* 70 : p. 109-135.
- HALLE, Morris (1959), *The sound pattern of Russian*. The Hague : Mouton.
- HALLÉ, Pierre ; BOYSSON-BARDIES, Bénédicte de (1996), "The format of representation of recognized words in infants' early receptive lexicon", *Infant Behavior & development* 19 : p. 465-483.
- HARRINGTON, Jonathan (2006), "An analysis acoustic of 'happy tensing' in the Queen's Christmas broadcast", *Journal of Phonetics* 34 : p. 439-457.
- HAWKINS, Sarah (2003), "Roles and representations of systematic fine phonetic detail in speech understanding", *Journal of Phonetics* 31 : p. 373-405.
- HOCKETT, Charles (1955), *A manual of phonology*. Baltimore : Waverly Press.
- HOWES, David (1957), "On the relation between the intelligibility and frequency of occurrence of English words", *Journal of the acoustic society of America* 29 : p. 296-305.
- JAKOBSON, Roman (1976 [1931]), "Principes de phonologie historique". Dans Troubetzkoy (1976) : p. 315-336.
- JAKOBSON, Roman ; FANT, Gunnar ; HALLE, Morris (1952), *Preliminaries to speech analysis*. Cambridge, Mass. : MIT Acoustics Laboratory.
- JAKOBSON, Roman ; WAUGH, Linda (1980), *La charpente phonique du langage*. Paris : Minuit.
- JANDA, Richard D. (2003), "Phonologization" as the start of dephoneticization". Dans Joseph, B. D., Janda, R. D. (eds), *The handbook of historical linguistics* : p. 401-422. Malden, Mass. : Blackwell.
- JESCHENIAK, Jörg D. ; LEVELT, William, J. (1994), "Word frequency effects in Speech production : Retrieval of syntactic information and of phonological form", *Journal of experimental Psychology, Learning, Memory and Cognition* 20 : p. 824-843.
- JOHNSON, Keith (1997), "The auditory/perceptual basis for speech segmentation", *Ohio State University Working Papers in Linguistics* 50 : p. 101-113.
- JURAFSKY, Dan (2002), "Probabilistic modeling in psycholinguistics : linguistic comprehension and production". Dans Bod, R., Hay, J., Jenedy, S. (eds), *Probabilistic linguistics* : p. 39-45. Cambridge, Mass. : The MIT Press.
- KUHL, Patricia K. (1991), "Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not", *Perception and Psychophysics* 50 : p. 93-107.
- KUHL, Patricia K. ; WILLIAMS, Karen A. ; LACERDA, Francisco ; STEVENS, Ken N. ; LINDBLOM, Bjorn (1992), "Linguistic experience alters phonetic perception in infants by 6 months of age", *Science* 255 : p. 606-608.

- LABOV, William (1994), *Principles of linguistic change : Internal factors*, Vol. 1. Oxford : Blackwell.
- LABOV, William (2001), *Principles of linguistic change : Social factors*, Vol. 2. Oxford : Blackwell.
- LAKS, Bernard (1996), *Langage et cognition : l'approche connexionniste*. Paris : Hermès.
- LASS, Roger (1984), *Phonology. An introduction to basic concepts*. Cambridge : Cambridge University Press.
- LUCE, Paul A. ; MCLENNAN, Conor T. (2005), "Spoken word recognition : The challenge of variation". Dans Pisoni, D. R., Remez, R. (eds), *The Handbook of Speech Perception* : p. 591-609. Malden, Mass. : Blackwell.
- MARSLÉN-WILSON, William ; WARREN, Paul (1994), "Levels of perceptual representation and process in lexical access : words, phonemes, and features", *Psychological Review* 101 : p. 653-675.
- MARTINET, André (1970), *Economie des changements phonétiques. Traité de phonologie diachronique*. 3^{ème} éd. Berne : Francke.
- MARTINET, André (1974), *La linguistique synchronique*. Paris : PUF.
- MAYE, Jessica ; WERKER, Janet ; GERKEN Lou Ann (2002), "Infant sensitivity to distributional information can affect phonetic discrimination", *Cognition* 82 : p. 101-111.
- MORGAN, James L. ; SAFFRAN, Jenny R. (1995), "Emerging integration of sequential and suprasegmental information in preverbal speech segmentation", *Child development* 66 : p. 911-936.
- MUNSON, Benjamin (2001), "Phonological pattern frequency and speech production in children and adults", *Journal of speech and hearing research* 44 : p. 778-792.
- NGUYEN, Noël ; WAUQUIER, Sophie ; TULLER, Betty (2007), "The dynamical approach to speech perception : from fine phonetic detail to abstract phonological categories". Dans Chitoran, I., Coupé, C., Marsico, E., Pellegrino, F. (eds), *Approach to phonological complexity*. Berlin : Mouton de Gruyter. [Sous presse.]
- PEPERKAMP, Sharon ; DUPOUX Emmanuel (sous presse), "Learning the mapping from surface to underlying representations in an artificial language", *Laboratory phonology* 9.
- PIERREHUMBERT, Janet B. (1999), "What people know about sounds of language", *Studies in the linguistic sciences* 29 : p. 111-120.
- PIERREHUMBERT, Janet B. (2001), "Stochastic phonology", *Glott International* 5 : p. 195-207.
- PIERREHUMBERT, Janet B. (2006), "The next toolkit", *Journal of Phonetics* 34 : p. 516-530.
- PIERREHUMBERT, Janet B. ; BECKMANN, Mary E. ; LADD, D. Robert (2000), "Conceptual foundations of phonology as a laboratory science". Dans Burton-Roberts, N., Carr, P., Docherty, G. (eds), *Phonological knowledge : Conceptual and empirical issues* : p. 273-305. Oxford : Oxford University Press.
- SAFFRAN, Jenny R. ; ASLIN, Richard N. ; NEWPORT, Elissa L. (1996), "Statistical learning by 8-month-old infants", *Science* 274 : p. 1926-1928.

- SAPIR, Edward (1925), "Sound patterns in language", *Language* 1 : p. 37-51.
- SEGUI, Juan ; FERRAND, Ludovic (2000), *Leçons de parole*. Paris : Odile Jacob.
- STAGER, Christine L. ; WERKER, Janet F. (1997), "Infants listen for more phonetic detail in speech perception than in word-learning tasks", *Nature* 388 : p. 381-382.
- TROUBETZKOY, Nikolay S. (1976 [1939]), *Principes de phonologie*. Paris : Klincksieck.
- TYLER, Lorraine K. ; FRAUENFELDER, Uli H. (1987), "The process of spoken word recognition : An introduction". Dans Frauenfelder, U. H., Tyler, L. K. (eds), *Spoken word recognition* : p. 1-21. (*Cognition*, special issue.)
- VIHMAN, Marilyn ; SUNNARI, Sari (2006), "The source of phonological knowledge, a cross-linguistic perspective", *Recherches linguistiques de Vincennes* 35 : p. 133-162.
- WAUQUIER-GRAVELINES, Sophie (2005), *Statut des représentations phonologiques en acquisition, traitement de la parole continue et dysphasie développementale*, Mémoire d'HDR, EHESS, Paris.
- WAUQUIER, Sophie (2006), "Du son au sens : acquérir ou apprendre la phonologie", *Recherches linguistiques de Vincennes* 35 : p. 5-30.
- WERKER, Janet F. ; FENNEL, Christopher T. (2004), "Listening to sounds versus listening to words : early steps in word learning". Dans Hall G., Waxman, S. (eds), *Weaving a lexicon* : p. 79-109. Cambridge, Mass. : The MIT Press.
- WERKER, Janet F. ; FENNEL, Christopher T. ; CORCORAN, Kathleen M. ; STAGER, Christine L. (2002), "Infants' ability to learn phonetically similar words : effects of age and vocabulary size", *Infancy* 3 : p. 1-3.
- WERKER, Janet F. ; TEES, Richard (1984), "Cross-language speech perception : Evidence for perceptual reorganization during the first year of life", *Infant behavior and development* 7 : p. 49-63.

¹ "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." (Carr, 2000 : 75).

² On remarquera à cet égard que Bybee reproduit la confusion fréquente entre les dichotomies "compétence"/"performance" et "langue"/"parole". Alors que la "compétence" chomskyenne renvoie à la grammaire interne du locuteur idéal et qu'elle est ultimement un objet psychologique, la "langue" chez Saussure est définie comme "un produit social de la faculté de langage et un ensemble de conventions nécessaires adoptées par le corps social pour permettre l'exercice de cette faculté chez les individus" (*CLG*, p. 25). La langue, dans la définition saussurienne que Bybee convoque sans la connaître, est donc conditionnée, quoi qu'elle en dise (2001 : 5), par l'usage.

³ La linguistique variationniste (Labov, 1994, 2001) en phonologie (Docherty *et al.*, 1997 ; Encrevé, 1988) a développé une réflexion systématique sur l'usage et a intégré dans la formulation de ses hypothèses théoriques une réflexion sur la variation. Une attention particulière y est portée aux données, à la fréquence des structures produites par le sujet à partir de corpus quantitatifs et statistiques, de même qu'aux variations

sociologiquement conditionnées observables. Le problème ne nous semble donc pas se poser là ; il 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 soulève l'étude des usages linguistiques.

⁴ C'est nous qui soulignons ici.

⁵ Pierrehumbert (2004) a montré qu'un bébé recevant une heure par jour de parole pendant la première année de vie 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. Un nombre d'exemplaires conséquent, donc, 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.

⁶ On sait aussi combien le fondement auditif du système de Jakobson a pu jouer en faveur d'une conception substantialiste des traits, l'acoustique étant censée fournir l'invariant physique qui ferait défaut à l'articulatoire. Cette thèse était, il est vrai, séduisante à une époque marquée par l'essor des recherches en phonétique acoustique des Bell Laboratories, et l'illusion iconique induite par le *visible speech*. Un sonagramme apporte à la fois le réalisme qui manque à un schéma articulatoire et la stabilité dont est dépourvue une séquence filmée de l'acte phonatoire.

⁷ Nous avons modifié la transcription phonémique de Hockett en conformité avec la "normalisation" qu'il propose.

⁸ On laissera ici de côté la variation dite "libre", qui, seule à être perçue *en tant que variation*, est d'une autre nature.

⁹ On peut essayer de montrer que le [o...] du fox "se comporte" comme le [i...], ce qui justifierait (3b). Mais que conclure si cela s'avère empiriquement impossible ?

¹⁰ Martinet (1970 [1955] : 178-179) est d'ailleurs peut-être le premier à avoir introduit, dans ce sens-là, le terme de "prototype" en phonologie.

¹¹ Qui veut que, dans une suite de segments donnée, à un phonème correspond un son et un seul, et, réciproquement, à un son correspond un phonème et un seul. C'est là un principe sinon nécessaire, du moins naturel dans l'esprit behavioriste de cette phonologie, et dont l'inadéquation, démontrée par Halle (1959), contribua à l'abandon de la notion de phonème par la phonologie générative. On le sait, l'argument de Halle n'aurait pas tenu face à la théorie pragoise (cf. Anderson, 1985 : 110-112).

¹² On sait le problème que pose /r/ aux thèses inductives. C'est peut-être la consonne la plus fréquente en français. Or c'est l'une des dernières dont il maîtrise la production, ce qui est d'autant plus intéressant que, si /r/ est la plus fréquente des consonnes, c'est aussi l'une des plus marquées.

¹³ La primauté de la théorie de l'optimalité depuis une douzaine d'années n'a rien fait pour changer cet usage, l'étude des représentations n'étant pas l'objet central de ce courant de pensée.

¹⁴ Il s'ensuit un défi intéressant pour la phonologie : si le locuteur perçoit tout "détail phonétique" associé aux variantes combinatoires car celui-ci traduit, en fait, la propagation de qualités distinctives, alors *toute* allophonie contextuelle implique une

“assimilation” et un contour temporel du type (9a) ayant des conséquences cruciales sur le nombre et la nature des traits (cf. Carvalho, 2007).

¹⁵ Le procédé n’est pas nouveau : la phonologie générative a eu recours aux mêmes méthodes à ses débuts, en faisant passer la théorie distributionaliste du phonème pour la théorie du phonème tout court (cf. la n. 11).

¹⁶ On remarquera le revirement opéré sur cette question par Pierrhumbert (2006 : 521) □ “Given the myriad attractions of exemplar theory, it is a good idea to recall why linguistic theory has phonology at all” □, suggérant de manière presque explicite que les modèles à exemplaires ne sont pas phonologiquement pertinents.