

**Aux frontières de l'écoute.
Durée des échantillons et choix des auditeurs :
deux variables déterminantes dans la construction des tests de perception¹**

Morange, Séverine

EA 1483, *Recherches sur le français contemporain* (Univ. Paris 3)
LCPE (LAM, UMR 7190, *Langages, Cognition, Pratiques, Ergonomie* – CNRS, Univ. Paris 6, Minist.Culture et Comm.)
morange@lam.jussieu.fr

Candea, Maria

EA 1483, *Recherches sur le français contemporain*, (Univ. Paris 3)
maria.candea@univ-paris3.fr

1. Introduction

Cet article s'inscrit dans la tradition réflexive qui se pratique régulièrement en sciences humaines, et vise à questionner le mode de constitution des données d'expérience, et ce à partir d'une analyse des frontières des variables sélectionnées pour l'élaboration de tests de perception. En effet, dans l'espace pluridisciplinaire qui nous concerne ici, celui qui met en relation linguistique et psychologie cognitive, il nous a semblé intéressant de discuter des variables descriptives des phénomènes langagiers impliqués dans les protocoles expérimentaux, dans la mesure où les frontières de ces variables sont données implicitement, d'évidence, et peuvent s'avérer parfois problématiques. Nous prendrons l'exemple de tests d'écoute – ou tests de perception –, que nous avons été amenées à pratiquer dans des études antérieures, et nous mettrons en avant les interrogations que nous avons rencontrées dans la mise en place de ces expérimentations. Les tests dont nous traitons consistent le plus souvent à faire écouter des échantillons sonores à un groupe de sujets humains. Après chaque écoute, ces derniers sont impliqués dans une tâche à partir de consignes précises ; souvent ils doivent répondre à des questions sur ce qu'ils entendent afin de tester des hypothèses sur les processus généraux de perception, d'identification, et d'évaluation.

Dans la mise en œuvre des tests d'écoute liés au champ de la linguistique nous avons dû procéder à la sélection du matériel sonore, des sujets à interroger, et à la formulation des consignes. Nous proposons de mettre l'accent ici sur deux difficultés auxquelles sommes confrontées à chaque étude:

- le choix d'une variable descriptive du « matériel », à savoir plus précisément la durée des « stimuli » soumis à l'écoute
- la définition des catégories de sujets susceptibles de tester les hypothèses.

La présente proposition s'inscrit dans le cadre plus vaste d'une réflexion sur la construction du savoir scientifique, telle que celle menée par Edgar Morin : « *la théorie, quelle qu'elle soit et de quoi qu'il s'agisse, doit rendre compte de ce qui rend possible la production de cette théorie elle-même, et si elle ne peut en rendre compte, elle doit savoir que le problème demeure posé* » (1982 : 172-173). Convoquer cette question épistémologique et fondatrice du mode de construction des connaissances, concernant les protocoles des tests d'écoute, ne vise pas à faire une critique qui serait stérile de l'utilisation de ces tests et de l'illusion d'une certaine scientificité qu'ils peuvent parfois induire. La question qui nous intéresse est la prise en compte de la complexité des informations relatives à l'objet traité (ici

¹ Nous remercions Y. Grinshpun et les membres du LCPE (LAM – CNRS, Paris 6) - tout particulièrement D. Dubois, J. Poitevineau, G. Delepaut, E. Mahler - pour la lecture et les commentaires d'une première version de cet article.

les langues) lors de son introduction dans le champ de la psychologie cognitive à travers la construction de protocoles expérimentaux. Plus précisément, dans le cadre des tests utilisant des variables mesurables, nous essaierons d'interroger ce qui relève de l'identification des frontières de variables pertinentes et ce qui peut être imputable au « contexte » créé par le test lui-même, contexte trop souvent gommé ou neutralisé, par routine ou par évidence. Il nous a semblé intéressant de confronter ici deux expériences, pratiques et bibliographiques, qui tant à travers la variété des matériaux et procédures, que des objectifs et interrogations, permettent de mener un questionnement relatif aux frontières des catégories présidant à la mise en place de l'expérimentation en psycholinguistique.

De nombreuses variables sont à définir et à choisir dans l'élaboration d'un test. En effet, il faut choisir les extraits (paroles reconnaissables, bruits, sons dé-sémantisés), le profil des sujets (le milieu socio-professionnel, leur degré de connaissance du sujet, leur âge, décider si on les groupe selon des frontières plus ou moins discutables - par exemple 18-25 ou 18-35 ans ?), choisir les tâches à accomplir par les sujets (discrimination, identification, tri par dissemblance / ressemblance, par préférence, description, mémorisation...), le protocole de réponse (catégorisation non verbale, questionnaire, notation sur échelles), choisir le lieu de l'expérience (laboratoire, chambre sourde hautement technologisée, ou bien *in situ*, c'est-à-dire au domicile, sur le lieu de travail du sujet, ou dans un cadre quotidien), etc.

Nous discuterons ici de manière restrictive de deux variables que nous avons été amenées à « contrôler » dans nos recherches respectives, à savoir : la durée des échantillons sonores utilisés et la définition d'un groupe « expert » parmi les sujets interrogés. Dans les deux cas, nous nous attacherons à discuter de la frontière à construire pour constituer les catégories de variables pertinentes (entre les extraits courts et les extraits longs, entre les sujets experts et les sujets non experts) et son importance dans la production des résultats et leur interprétation. Notre étude vise à montrer à travers ces deux exemples le caractère problématique de la définition des « stimuli » ou des « sujets », les choix théoriques sous-jacents et le plus souvent implicites qui s'y manifestent, en particulier dans le transfert d'une discipline à une autre, et qui relèvent dès lors souvent davantage d'une évidence de sens commun que d'une expertise disciplinaire. Nous espérons contribuer ainsi aux réflexions en cours sur les méthodologies dans les disciplines visées que recouvrent la linguistique et la psychologie.

2. Durée des échantillons dans les tests : y a-t-il un seuil pour produire du sens ?

Parmi les différents paramètres qui se trouvent contrôlés dans les tests d'écoute, la durée des échantillons peut sembler non problématique dans la mesure où il suffit de la mesurer physiquement sur une échelle temporelle. Et ce contrairement au degré d'expertise des sujets, dont la variable, que nous aborderons plus loin relève alors du domaine social et psychologique. Cette durée est de fait aisément communiquée dans la description des variables indépendantes des tests de perception comme propriété « objective » des stimuli.

2.1 Quelques repères sur les pratiques courantes

Qu'il s'agisse de travaux en acoustique, en sociolinguistique ou psycholinguistique notamment, les frontières de durées utilisées sont généralement comprises dans des intervalles assez restreints. Ainsi, les extraits sonores utilisés dans les tests relevant de nos centres d'intérêt et ayant fait l'objet d'une publication dans les trente-quarante dernières années durent, à notre connaissance, entre 200 ms et 30 secondes. Citons quelques exemples récents où les extraits sonores durent entre 200ms et 1 seconde (Vasilescu, Candea, Adda-Decker,

2005, Tran *et alii* 2005 ; Jacquier, Meunier, 2005), y compris lorsqu'il s'agit de perception audio-visuelle, comme dans Traunmüller, Öhrström (2007) ; entre 4 et 15 secondes maximum par exemple dans Gendrot (2004) ou Boula de Mareuil, Brahim, Gendrot (2004). Il existe une longue tradition d'articles, comme par exemple celle dans laquelle s'inscrivent Schroder, Aubergé, Cathiard (1998), où la durée des échantillons utilisés dans le test n'est pas précisée et encore moins questionnée ; les lecteurs peuvent déduire qu'il s'agit d'extraits de quelques syllabes. Très rares sont les travaux qui font écouter des séquences dépassant 30 secondes (Morange S. *et ali*, 2006), ou la minute (Morange 2005 ou Vieru-Dimulescu 2007).

Dans le choix d'extraits très courts, un objectif explicite est celui d'éliminer le traitement d'un contenu sémantique, pour ne permettre que le traitement « pre- »sémantique, et ce dans le cadre d'une conception du traitement du langage essentiellement sériel de la perception périphérique jusqu'à l'interprétation. Ainsi, l'étude serait limitée à la perception de voyelles, de consonnes, de frontières syllabiques ou d'autres types de frontières, autrement dit l'identification de phénomènes prosodiques dans des séquences modifiées acoustiquement et sans contenu segmental¹. En revanche, les extraits plus longs, de parole, de musique, ou même de bruit, laissent de fait place à l'hypothèse de la saisie possible d'un sens, à l'interprétation et à la possibilité d'identifier une écoute en contexte, plus proche de l'écoute ordinaire. C'est sans doute la raison pour laquelle on trouve des extraits longs lorsqu'il s'agit d'obtenir le portrait perceptif des locuteurs, la caractérisation de la voix et de sa perception comme processus identitaire (Morange 2005), d'une personne, d'un accent (Vieru-Dimulescu, 2007) ou encore la perception du chant diphonique par des Européens (Barras, Gouiffès, 2007).

Bien entendu, il existe des travaux qui, non seulement discutent le choix de la durée des échantillons sonores, mais testent explicitement les effets du déplacement de cette frontière de durée sur les résultats. Par exemple, dans le test d'écoute de parole modifiée en vue de garder seulement le suprasegmental, Ohala & Gilbert (1979) montrent, dans un article qui a fait date, que les auditeurs identifient mieux la langue des échantillons présentés si on leur fait écouter des segments de durées de 15-25 secondes, plutôt que de 10 secondes. Plus récemment, Shötz (2005) montre également que l'augmentation de la durée des échantillons augmente le taux d'identification correcte de l'âge des locuteurs.

Il nous semble cependant que le choix de la durée se présente le plus souvent dans les travaux comme allant de soi, tout comme le fait qu'une durée courte éliminerait les processus d'interprétation. Certes, ce choix peut être clairement justifié : lorsqu'on veut tester la discrimination entre deux voyelles, il est difficile de penser un protocole qui utiliserait des extraits de parole longs dans lequel, de fait, les processus de discrimination vocalique que l'on cherche à identifier seraient dépendants du contexte. Il est donc plus simple d'utiliser des paires minimales de mots ou des pseudo-mots² pour étudier la perception du son cible. Cependant, la définition même du stimulus comme unité dans le champ de l'expertise linguistique fait l'impasse sur la pertinence de cet extrait pour le sujet humain qui traite « l'information » censée y être contenue. Il n'y a donc aucune exploration d'une possible construction de sens par le sujet allant à l'encontre du postulat formulé. Or, le choix d'une courte durée ne garantit pas en soi, dans le champ psychologique, que la tâche en apparence simple et univoque que les sujets effectuent est celle prédite par le chercheur, ni que les sujets accordent au stimulus le même statut que celui prévu par l'expérimentateur. Rien ne permet de prouver parfois que les sujets n'utilisent pas les stimuli comme indices de quelque chose que l'expérimentateur n'a pas prévu, qui relèverait de l'attribution d'un sens à un « objet » totalement inédit et hors de toute pratique courante pour les sujets.

¹ Autrement dit des séquences qui sont censées n'avoir aucun sens identifiable.

² Des mots qui n'existent pas mais qui pourraient exister compte tenu des contraintes phonotactiques d'une langue (exemples tirés de Content *et alii* 2004 : un « grouque », « cas légal/qualégal », etc.).

Ainsi, une récente étude qui aborde la distinction phonétique entre deux diphtongues très proches montre que le fait d'associer la prononciation de simples paires minimales avec des photos de personnes connotées différemment du point de vue social (en disant aux sujets de regarder la photo de la personne qui a prononcé les paires minimales écoutées) change la perception de ces diphtongues (Hay, Warren, Drager, 2006). Ce constat implique que l'écouteur prend en compte la représentation qu'il se fait de la personne qui parle même lorsque l'échantillon écouté est tout à fait minime. L'objectif de ces auteurs n'est pas simplement de montrer que la perception auditive de la parole est modifiée par le fait de regarder ou non une image ou une vidéo associée en même temps, comme c'est le cas par exemple de Traunmüller, Öhrström (2007) dont l'expérience centre l'attention sur les effets de la perception audio-visuelle de la parole produite. L'approche qui nous intéresse ici, pour son point de vue « sociophonétique », s'attache à mettre en évidence que l'identité sociale du sujet « écoutant » et celle prêtée au locuteur enregistré pour le test influencent les résultats obtenus, et ce même avec des échantillons extrêmement courts qui sont censés être dépourvus de tout sens.

Aussi, à la lumière de ces quelques constatations résultant le plus fréquemment de tests de phonétique acoustique où les échantillons sont très courts et très ciblés, on peut penser que les sujets convoquent dans leurs réponses à la fois ce qu'ils s'imaginent être les attentes de l'enquêteur et ce qu'ils s'imaginent des personnes dont ils entendent la voix dans le cadre du test. Ce dernier peut être envisagé, de fait, comme une interaction complexe entre sujet et expérimentateur. Autrement dit, le traitement cognitif met en jeu des représentations préalables qui interagissent avec les processus perceptifs à des « niveaux » très bas et dans des délais très courts.

Nous avançons alors l'hypothèse que, paradoxalement, les protocoles visant à neutraliser le maximum de représentations individuelles et sociales sont peut-être ceux qui en intègrent le plus, car ces paramètres s'insèrent librement dans les aspects les plus difficiles à contrôler, dans la mesure où ils sont considérés par leurs concepteurs comme neutralisés.

2.2 Etude de cas : test d'écoute révisité

Prenons l'exemple d'un test que nous avons publié récemment portant sur la discrimination de huit langues – allemand, anglais américain, arabe du Moyen Orient, chinois mandarin, espagnol, français, italien, portugais – à partir de l'écoute de la voyelle d'hésitation autonome (équivalent du « euh » français) utilisée par des locuteurs de ces langues (Vasilescu, Candea, Adda, 2005). Si les échantillons sonores avaient inclus des durées plus longues, avec autre chose que la voyelle d'hésitation concernée, on peut supposer que les auditeurs auraient utilisé les indices essentiellement phonotactiques et lexicaux pour identifier la langue et non le timbre vocalique des « euh », « huh », « eh » que nous souhaitions tester précisément. Dans le cadre d'une telle étude, il paraît donc difficile d'estimer un effet de la durée sur les résultats obtenus. Néanmoins, cela ne veut pas dire nécessairement que :

- la durée courte a permis de se concentrer uniquement sur le timbre de la voyelle, tel que la consigne le demandait, sans prise en compte, par exemple, de la qualité de la voix entendue ;

- la durée courte a empêché les sujets d'avoir une représentation sociale de la personne qui avait produit les voyelles d'hésitation entendues.

D'ailleurs, si on se penche sur les interactions occasionnées par le déroulement de ce test, un détail qui n'a pas été exploité dans l'analyse des résultats peut toutefois être évoqué : la plupart des sujets éclataient de rire au début de l'expérience, et montraient souvent de l'étonnement en écoutant les voyelles dont ils devaient deviner la langue. Cela ne pourrait pas signifier qu'ils construisaient du sens en dehors du timbre des voyelles, objet du test ? Et si

cela était exact, pourrait-on légitimement se demander quel serait ce sens, quelle serait la raison de leur hilarité comme première réaction ? Était-ce le fait qu'ils écoutaient des hésitations, qu'ils s'imaginaient que nos locuteurs auraient été « piégés » par nos enregistrements dans un moment délicat, etc ...) D'autre part, les sujets ont pu s'imaginer des locuteurs pourvus des caractéristiques stéréotypées attribuées aux natifs des langues cibles. Cette hypothèse renvoie à une observation que nous avons faite lors de notre expérience : des voyelles de français prononcées avec une voix « craquée » (*creaky voice*) ont été attribuées massivement à l'arabe dans le test de discrimination arabe-français. Nous suggérons donc qu'un questionnaire permettant l'identification des représentations stéréotypées des sujets à propos des langues en question puisse être intégré à ce type de test, afin de préciser leur éventuelle interaction avec les indices purement acoustiques visés par la recherche menée.

Les pratiques les plus courantes en phonétique acoustique n'intègrent pas les résultats issus des études de socio- ou de psycholinguistique : nous pensons, pour notre part, que la mise en place de tests d'écoute avec des échantillons sonores de très courte durée ne dispense pas d'intégrer l'analyse des représentations sociales des sujets testés, à la fois celles renvoyant à la tâche qui leur est demandée, aux attentes des expérimentateurs et aussi à l'origine « incarnée » socialement des extraits de voix données à écouter. Nous suggérons que la frontière de durée de quelques secondes (ou même moins) ne constitue pas un barrage robuste qui permet d'isoler la perception d'un phénomène phonétique ciblé. On peut dès lors penser que de nombreuses autres études en dehors de la nôtre, qui se contentent généralement de mentionner par une simple allusion certaines difficultés liées aux réactions imprévues des humains impliquées dans les expérimentations, se prêteraient à une ré-analyse intéressante de leurs résultats si elles étaient enrichies d'un questionnement des sujets sur un certain nombre de variables « contextuelles ».

3. Frontière entre experts et non experts en SHS et dans les protocoles d'écoute

Les notions d'« expert » et d'« expertise » sont interrogées dans divers domaines des sciences humaines et sociales depuis plusieurs décennies. Contrairement à la durée d'échantillons qui constitue une donnée quantifiable selon une échelle partagée, l'expertise en tant que propriété du sujet ne se mesure pas quant à elle, sur une échelle physique. Par ailleurs les frontières entre plusieurs niveaux d'expertise peuvent dépendre de facteurs divers.

Nous aborderons par conséquent la question de la catégorie experts en parcourant diverses disciplines des sciences humaines, en vue de donner à voir la complexité même de cette catégorie et de ses contours. Nous discuterons ensuite de ces aspects au regard de la constitution des groupes d'experts pour des tests de perception.

3.1 Quelques repères sur la catégorisation expert / non-expert dans divers champs disciplinaires

Force est de constater de prime abord la variabilité de dénominations correspondant à « expert » et « tout venant » dans la diversité des champs disciplinaires concernés par la question, depuis l'anthropologie des sciences à la sociologie de l'expertise, en passant par l'analyse sensorielle et la psychologie cognitive. Dans le même temps, l'imprécision et l'implicite de leurs définitions renforcent le sentiment d'hétérogénéité. Selon les domaines ou les auteurs, il peut être question d'*experts*, de *professionnels*, de *spécialistes*, de *référents*, face à des *non-experts*, des *novices*, des sujets *naïfs*, *ordinaires*, *profanes*. La distinction peut porter sur sujets *entraînés* et *non entraînés*, *experts sensoriels* et *experts produits* (dans la terminologie de l'analyse sensorielle). Pour ce qui concerne les études sur la parole, on

différencie souvent les sujets *natifs* des *non natifs*, les *bilingues* des *monolingues*, les *phonéticiens*, *musiciens*, *acteurs*, *professionnels de la parole* des *tout venant*.

Globalement, deux groupes émergent en parallèle, appréhendés de façon hétérogène :

- des sujets considérés comme spécialistes, connaisseurs d'un domaine particulier et alors doués de compétences singulières, particulièrement aptes à la résolution de problèmes, à la « *distinction de l'essentiel, du conjoncturel, et de l'aléatoire* », déterminants dans la « *qualité de l'information qui sera produite* »³ (en linguistique, on trouvera par exemple le nourrisson de parents bilingues ou trilingues, l'étudiant en langues étrangères, l'orthophoniste) ;

- des sujets considérés comme « naïfs », ou « neutres », confrontés à l'expérience ordinaire, doués de représentations de sens commun, de connaissances collectives (prises en compte notamment dans les travaux de sociolinguistique), et de connaissances individuelles (prises en compte dans certaines approches de la psychologie cognitive).

Notre démarche emprunte bien entendu à l'anthropologie des sciences qui s'attache à comprendre comment les sciences fabriquent du savoir ou « comment les sciences parlent du monde ? » (Latour, 1993 :188). Cette discipline place l'expert au cœur de son observation, tout en discutant le rôle qui lui est dévolu : « *il ne s'agit pas de constater des faits puis de les passer aux politiques qui sont supposés décider, mais de rendre aux scientifiques une fonction de décideurs de fait qu'ils ont d'ailleurs toujours eue et qui n'est pas la seule de leurs compétences* » (Latour, 2003 : 48).

Plus loin de notre domaine, la notion d' « expert » est au cœur de la sociologie de l'expertise qui s'intéresse à la compétence en s'employant à comprendre « *comment les expériences perceptuelles sont raccordées à des représentations collectives* » (Bessy & Chateauraynaud, 1995). Signalons ici brièvement les travaux de ERASE⁴ qui attirent précisément l'attention sur la construction sociale des savoirs.

Nous arrêtant sur des données de l'expertise sensorielle, nous ne développerons pas ici les diverses approches de la psychologie, ni les analyses issues des recherches cognitives (psychologie, intelligence artificielle) souvent concernées par les « systèmes experts » en tant qu'outils d'aide à la décision et à l'action, eux-mêmes constitués à partir de l'identification de la diversité des connaissances expertes (Dubois dir., 1992).

3.2 Statut du sujet dans les études sur la perception sonore en (psycho)linguistique

A travers les travaux de linguistique ou de psycholinguistique, et malgré la richesse des réflexions menées dans les autres disciplines pourtant proches, la catégorie du sujet expert / non-expert, ou ses équivalents comme natifs / non natifs notamment, est rarement problématisée dans les protocoles d'écoute courants.

Une première observation fait ressortir l'hétérogénéité des savoirs et savoir-faire mobilisables selon des facteurs différenciés. Ainsi, la variation du domaine d'expertise dans des tâches de perception pourra-t-elle être mise en relation avec i) la situation, ou ii) le domaine d'étude. A titre d'exemple :

i) Les niveaux d'expertise dans une tâche de discrimination d'une langue étrangère chez des hispanophones varient selon que la situation de communication est naturelle ou artificielle

³ Ghiglione R., Richard J.F (dir), 1995, *Cours de psychologie*, n°2, *Bases, méthodes, épistémologie*, Dunod, p.374.

⁴ Equipe de Recherche d'Anthropologie et de Sociologie de l'Expertise (Université Paul Verlaine – Metz)

(Gaillard, Magnen, Billières, 2006). On peut toutefois imaginer que les frontières entre situations *in situ* ou en laboratoire ne sont pas étanches car des facteurs jouant dans telle situation peuvent jouer également dans l'autre.

ii) Les niveaux d'expertise dans la perception du timbre musical chez des étudiants issus de trois écoles différentes varient selon leur formation (Castellengo, Dubois, 2005). Ici aussi, on peut s'interroger sur les frontières entre les groupes d'étudiants dans leur pratique des sons musicaux.

Par ailleurs, d'autres expériences d'écoute mettant en scène des paradigmes experts / novices, s'attacheront plutôt à l'âge ou au mono- / bilinguisme. Par exemple, Ramus (1999), a procédé à de nombreux tests sur des nourrissons et des adultes afin d'étudier la perception du rythme dans les langues. De même, Barras & Gouiffès (2007), ont choisi des enfants (âgés de 10 et 15 ans) en vue d'identifier différents types de chant diphonique. On peut se demander comment se partagent les individus entre experts et novices dans ces conditions d'expérience, lorsqu'il s'agit de nourrissons, enfants ou adultes. Dans un autre exemple Ohala & Gilbert (1979), cités plus haut, ont sélectionné des auditeurs natifs et non natifs dans des tâches de discrimination entre différentes langues. Dans ce cas, ce sont les frontières entre sujets monolingues / bilingues / trilingues qui semblent loin d'être évidentes.

En interrogeant la catégorie experts / non-experts et les frontières souvent difficiles à délimiter d'un groupe à l'autre, émerge en fait la question de la place même accordée au sujet dans les expériences. De nombreux cas de figures peuvent être représentés, dépendant des cadres théoriques et des instruments de mesure. Le sujet peut en effet être envisagé comme un exécutant de la tâche, au fond un instrument de mesure, désensibilisé et dépersonnalisé, qui réagit à un stimulus, à un système de contraintes, et ce type de procédure permet d'échafauder des hypothèses de recherche. Mais selon d'autres approches, le sujet peut au contraire être considéré comme un être sensible qui perçoit, pense, a eu des expériences, a des connaissances, interagit, est placé face à un expérimentateur connu ou inconnu, est face à un objet du monde, qui lui est aussi complexe et relève d'une pratique sociale, plus ou moins ritualisée (la parole, la musique...). En ce sens, Mondada (1999 : 20) constate que la catégorisation des locuteurs, et par là même la frontière entre groupes de locuteurs constitués est souvent « *d'avantage présumée que problématisée* ». On peut même se demander avec Dubois (1996 : 91) : « *dans quelle mesure les questions posées à travers les consignes, sans prise en compte de l'identité (non seulement individuelle mais sociale) des sujets n'imposent-elles pas aux sujets de s'inscrire dans les mêmes cadres que ceux qui ont présidé à l'élaboration du matériel* ».

3.3 Etude de cas : construction revisitée d'un groupe d'experts

Nous proposons de (re)placer l'individu au centre de l'expérience, et d'octroyer un rôle nodal à l'identification de ses représentations potentielles et préalables dans la mise en œuvre du protocole expérimental, quel que soit le type d'interface considéré (homme-machine, homme-homme, homme-animal, animal-machine). Ainsi nous nous proposons de revisiter ici une recherche pluridisciplinaire à laquelle nous avons participé (Morange, Fontaine, Vogel, Poitevineau, Dubois, 2005 et Morange, Dubois, Fontaine, 2008). Un test d'écoute avait été mis en place en vue de préciser les critères de jugement d'un public d'auditeurs à propos de traitements de restauration d'enregistrements musicaux proposés lors de ré-éditions. Pour cette expérience, le choix de sujets considérés comme experts ou comme non-experts n'a pas été sans susciter de discussions au sein de notre groupe de chercheurs issus de disciplines académiques différentes.

Le test proposait d'écouter des extraits d'opéra, puis de les classer (tâche de catégorisation non verbale), et de commenter les classements effectués (cette dernière tâche était requise en vue d'une analyse linguistique). Les sujets sélectionnés comprenaient, d'abord un groupe de professionnels, principalement des acousticiens et ingénieurs du son, et d'étudiants en métiers du son, considérés alors comme experts. Un autre groupe a été considéré comme non-expert : il comprenait des sujets n'ayant pas de formation sur le son, mais se présentant comme mélomanes. Les résultats des analyses catégorielles et psycholinguistiques ont révélé deux caractéristiques remarquables distinguant experts et non-experts :

- des catégories (regroupement des échantillons sonores) distinctes⁵
- un discours différent dans la relation à l'objet perçu⁶.

Ces résultats ont ainsi fait émerger des tendances, mais à y regarder de près la frontière expert / non-expert demeure assez fragile. Tout d'abord, malgré la tendance générale, les classements catégoriels ne sont pas toujours très différents entre un expert et un non-expert (il n'est pas rare qu'un expert et un non-expert classent dans une même catégorie les mêmes extraits donnés). Ensuite, les experts parlent eux aussi de leur ressenti⁷, et réciproquement les mélomanes parlent eux aussi du son comme un objet extérieur⁸. Ce constat suscite de nouvelles questions : sur quoi les sujets s'appuient-ils pour constituer les catégories et pour en parler ? Comment se donner les moyens de savoir que nous n'avons pas commis une erreur en catégorisant *a priori* un sujet comme non-expert, par exemple, alors qu'il se comporterait comme « expert » pour le domaine testé ?

La discussion reste donc ouverte quant à la définition des critères à partir desquels sélectionner un expert. Ici, un acousticien, un ingénieur du son, un musicien, un mélomane, un collectionneur de vieux disques, un collectionneur de phonographes, un détenteur de matériel haute définition, sont-ils *a priori* plus ou moins experts ? Sans compter qu'un ingénieur du son peut être musicien, collectionneur... Autrement dit, les sujets ont des identités plurielles et une diversité de modes d'écoute. Ces identités sont variables selon l'objet de catégorisation, selon la façon dont les sujets se présentent, sous la façon de considérer leur statut, mais également selon la situation, leur(s) interlocuteur(s), etc.

L'idée d'un degré d'expertise variable selon le type de connaissances (encyclopédiques, savantes, empiriques...) semble être opérante car finalement tout sujet est constamment dans un double ou triple ancrage, dans une pratique d'usager et une pratique d'expert, dans une position d'observateur, de scientifique, d'acteur associatif... La frontière du degré d'expertise s'avère ouverte, polyvalente. Ainsi, lorsque certains « non experts » obtiennent des résultats identiques à ceux des « experts », c'est bien la frontière entre les deux groupes qui se trouve interrogée. Aussi, dans le cadre de cette enquête, il aurait pu être utile de procéder à des pré-tests en vue de faire un premier repérage de différents niveaux d'expertise.

En résumé, nous pensons n'avoir ici fait qu'entrouvrir la discussion sur l'aspect problématique de ces frontières dans la mise en place d'expérimentations : entre les catégories de sujets, entre les degrés d'expertise, entre les types d'expertise, entre discours d'expert et discours de sens commun.

5. Discussion : élaboration des tests et implications théoriques

⁵ Analyses par algorithmes de calculs de similitudes.

⁶ Les experts évoquant souvent des caractéristiques physiques et acoustiques du son (ex : « *contenu spectral assez complet, pas de distorsions* »), les non-experts se situant davantage dans le ressenti ou l'expérience passée, par la référence à des aspects qualitatifs (ex : « *criard, plus chaud, plus sérieux, moins pleurnichard, plus charnel* ») ou culturels du son (ex : « *on dirait l'aiguille d'un vieux pick-up* »).

⁷ Exemple : « *grosse voix dont on perd l'intimité de la tendresse* ».

⁸ Exemple : « *trop de bruits de surface qui masquent le son original* ».

Les tests d'écoute offrent un côté rassurant pour leurs utilisateurs. Ils permettent, pour le moins théoriquement, de fonder une hypothèse non sur leur seule intuition ou la tradition expérimentale qui relève d'une conception positiviste de la psychologie, mais sur les résultats obtenus après consultation d'un groupe de sujets pas (ou peu) informés de l'objet poursuivi par la recherche. En outre, l'approche expérimentale permet une exploitation statistique des données par l'utilisation de tests de signification. Ces deux arguments expliquent le succès des tests de perception, ainsi que leur validation institutionnelle dans les critères appliqués pour accepter une publication. Néanmoins, les critères apparents de scientificité sous couvert de procédure expérimentale, de résultats « objectifs », ou encore de traitements mathématiques, ne doivent pas occulter une réflexion en amont sur le sens et le statut des variables « manipulées » ni, en aval, sur les « mésusages » des tests statistiques (Poitevineau (2004), Lecoutre & Poitevineau (2000)).

Aussi séduisante que puisse paraître l'expérimentation dans le domaine des sciences du langage, son inflation actuelle ne va pas sans soulever des problèmes épistémologiques et théoriques, abordés tant de l'extérieur par les anthropologues des sciences comme Latour et Morin, que de l'intérieur par les concepteurs eux-mêmes de tels protocoles (Grossen, 1988 ; Dubois, 1996). On pourra s'étonner du fait que lors de la constitution du groupe de sujets sollicités pour passer le test en phonétique ou psycholinguistique, leur identité et leurs connaissances ne fassent que rarement l'objet d'une prise en compte. En particulier, lorsque les résultats visent la perception du public tout venant, l'essentiel du contrôle s'effectue sur les variables du « matériel » et la mise en place de la procédure. La plupart des enquêteurs, font appel à leurs étudiants, à des collègues ou amis, ou encore à des étudiants sélectionnés sur des critères de disponibilité ou de voisinage non spécifiés. Il paraît dès lors difficile de savoir dans quelle mesure le groupe de sujets serait un échantillon représentatif de la population, et quelle serait cette population (étudiants, Européens, citadins, *etc.*).

Nous nous sommes également penchées sur le choix des échantillons sonores, bien souvent brefs, fabriqués ou censés être dépourvus de sens (logatomes, pseudo-mots, parole inversée, voyelles isolées, parole dégradée, modifiée, bruitée, filtrée, artificielle, ...).

D'un côté, cette pratique du test rend la tâche d'écoute parfois artificielle et fort éloignée des situations de la vie quotidienne du sujet, de sorte qu'il est difficile de considérer les processus cognitifs en jeu comme « standard » et d'interpréter les résultats obtenus comme généralisables. Il est bien difficile alors de décider si les résultats concernent la perception de la parole en général, ou bien la possibilité par les sujets de mettre en œuvre des capacités cognitives spécifiquement mobilisées dans cette tâche inédite pour eux.

D'un autre côté, nous l'avons vu, les sujets peuvent réagir de manière imprévue et complexe. Parfois les expérimentateurs mentionnent sans approfondir l'éviction de certains sujets ou de « mauvaises réponses » dans le traitement des résultats, sur le motif que les consignes données n'auraient pas été comprises ou que les sujets n'auraient pas réussi à respecter le protocole (découragement, manipulations trop « difficiles » / « complexes », blocages informatiques), ce qui signifie que les sujets recevables doivent appartenir à un certain groupe, qu'on se garde bien de définir. On se garde encore plus souvent de s'interroger plus avant sur les possibilités laissées aux sujets de donner un sens à ce qu'ils écoutent, fût-il différent de celui que l'expérimentateur attendait.

Si l'on se focalise non pas sur la tâche demandée mais sur le groupe de sujets sélectionnés, nous pensons avoir montré, à travers la frontière fragile et discutable de la catégorisation expert / non-expert, que là encore les résultats obtenus sont configurés par un choix de départ. Celui-ci, souvent traité comme une évidence, peut voler en éclats si l'on fait entrer en ligne de compte le sujet dans sa complexité et son rapport intersubjectif avec

l'expérimentateur, dans le cadre socialement codifié que représente le test de perception pratiqué par les chercheurs.

Nous pouvons par conséquent envisager l'hypothèse que dans le cadre des tests de perception tels qu'ils sont habituellement pratiqués, il n'est pas exclu que « *le culturel et le social que l'on croyait avoir évacué par la construction rigoureuse de matériels expérimentaux fondés sur les métriques scientifiques de leur réalité mondaine se trouvent [...] massivement présents et sans contrôle, dans le questionnement expérimental* » (Dubois, 1996 : 97).

6. Conclusion

Les frontières de variables utilisées généralement dans les expérimentations en sciences humaines sont souvent implicites et davantage fondées sur un usage qui se reproduit, plutôt que sur une réflexion épistémologique au sujet du statut des variables « contrôlées ». Ces dernières sont rarement repositionnées en regard des champs disciplinaires dans lesquelles elles sont constituées : en particulier, pour ce qui nous concerne, le champ de la physique qui définit objectivement la durée des stimuli, ou de la linguistique qui instaure les « unités » pertinentes du discours comme entités existantes en soi et indépendantes du contexte et des traitements cognitifs qu'elles suscitent. En vue d'augmenter la fiabilité de l'interprétation des résultats, il serait souhaitable de préciser les contours disciplinaires des concepts en jeu dans les expériences d'écoute, et de restituer la richesse des variables identifiées dans ces divers champs en interaction (social, psychologique, linguistique, physiologique, cognitif) dès qu'il s'agit d'étudier les processus cognitifs qui se trouvent sollicités dans des expérimentations.

En interrogeant la catégorie de durée des échantillons et la catégorie d'experts (en contraste avec non-experts), nous avons essayé de dessiner les limites des tests de perception. Ceux-ci ne représentent pas, selon nous, un « instrument de mesure » du subjectif, mais bien un mode d'investigation, dont le caractère expérimental contrôlé permet d'identifier, entre autres, les frontières de nos interprétations.

Nous avons en particulier développé l'idée que, quelle que soit l'expérience, il est réducteur de considérer que l'on manipule du mono-dimensionnel, du tout construit, du neutralisé, du tout objectivé ou objectivable, et que ces propriétés sont primitives ou élémentaires, y compris pour le sujet questionné. En effet, nous avons rappelé que tout traitement ordinaire est multidimensionnel, complexe, sinon global (holistique), non décontextualisé, non abstrait dans son caractère expérientiel, et donc variable d'un sujet à l'autre, et *a fortiori* d'un groupe de sujets à l'autre. Il est difficile, y compris en situation expérimentale, de dissocier le méta- de l'infra-cognitif. Cette problématique qui dépasse d'ailleurs le cadre des études sur l'humain, est également abordée dans l'étude du vivant en général, et en particulier des animaux. Ainsi, Despret (2007 :167) questionne les liens entre l'observé et l'observateur dans la relation homme / animal lors des expériences d'éthologie : « *la science n'est-elle qu'une machine à produire des lois générales ? Ce serait plutôt le propre de certaines sciences, les sciences du non-vivant. Dès que l'on étudie des êtres, on est confronté à leurs particularités (...) la réponse des sujets est bien plus riche et plus intéressante que leur réaction* ».

Notre étude contribue, nous l'espérons, à préciser les articulations pluridisciplinaires qu'il est indispensable d'envisager et à faire progresser les protocoles des tests et leurs interprétations ; nous espérons contribuer à améliorer la « *validité écologique* » de la situation expérimentale par rapport aux traitements cognitifs ordinaires⁹.

⁹ Cf. Gibson, J. J., 1979, *The ecological approach to visual perception*, Boston, Houghton Mifflin

Références

- Barras M-C, Gouiffès A-M, 2007, The reception of overtone singing by an unformed listener, *Proceedings of the third Conference on Interdisciplinary Musicology (CIM07)*, K. Maimets-Volk, R. Parncutt, M. Marin & J. Ross (Eds.), Tallinn, Estonia, 15-17 august 2007, 5 pages ; en ligne: www-gewi.uni-graz.at/cim07/
- Bessy C., Chateauraynaud, 1995, *Experts et Faussaires. Pour une sociologie de la perception*, Métailié.
- Boula De Mareuil Ph, Brahim B, Gendrot C., 2004, Role of segmental and suprasegmental cues in the perception of Maghrebien-accented French, *ICSLP04*.
- Candea M., Vasilescu I., Adda-Decker M., 2005, Inter- and intra-language acoustic analysis of autonomous fillers, in *Proceedings of DISS 2005*, Aix-en-Provence.
- Castellengo, M., Dubois, D., 2005, Timbres ou timbres ? Propriété du signal, de l'instrument, ou construction(s) cognitive(s)?, (CIMO5), *Cahiers de la Société Québécoise de Recherche en Musique*.
- Content A., Bagou O., Fraunfelder U., Fougeron C., 2004, Indices de frontières lexicales avec des pseudo-mots : une étude préliminaire, in *Actes des Journées d'Etudes sur la Parole JEP 2004*.
- Despret V., 2007, Animal et humain, d'individu à individu, *La recherche*, n°405, février 2007, p. 64-67.
- D. Dubois (dir), 1992, Connaissances et rationalités. Questions ergonomiques et recherches cognitives, n°spécial, *Intellectica*, 15.
- Dubois D., 1996, Matériels et consignes : un type de questionnement social dans la recherche expérimentale en psychologie, in *Le questionnement social*, Numéro Spécial des Cahiers de Linguistique Sociale, dir. Richard-Zappella J., p.89-98.
- Dubois D. 2002, Informations, représentations, connaissances, et significations : des objets en question dans les sciences cognitives, in *Du sujet : théorie et praxis*, MSH, Paris, 97-106
- Gaillard P., Magnen C., Billières M., 2006, Expertise phonétique : une réflexion vers la notion de « surdité fonctionnelle », 4^{ème} Journée du Sensolier, *Les expertises sensorielles*, 12 oct.2006, p. 11-14.
- Gendrot, C., 2004, Indices prosodiques et segmentaux pour l'identification de sa langue natale. *MIDL : Identification des langues et des variétés dialectales par les humains et par les machines* – Paris (France), nov. 2004, p. 97-100
- Grossen M., 1988, *L'intersubjectivité en situation de test*, Delval.
- Hay J., Warren P., Drager K., 2006, Factors influencing speech perception in the context of a merger-in-progress, *Journal of Phonetics*, 34, p. 458-484.
- Jacquier C., Meunier F., 2005, Perception of time-compressed rapid acoustic cues in French CV syllables, *Interspeech*, (CD-Rom) Lisbonne.
- Latour B., 1993, Le topofil de Boa-Vista. La référence scientifique : montage photo-philosophique, *Raisons Pratiques*, p.187-216.
- Latour B., 2003, *Un monde pluriel mais commun (Entretiens avec F. Ewald)*, L'Aube, Poche Essai.
- Lecoutre B., Poitevineau J., 2000, Aller au-delà des tests de signification traditionnels : vers de nouvelles normes de publication, *L'Année psychologique*, p.683-713.
- Mondada L., 1999, L'accomplissement de l'étrangeté » dans et par l'interaction : procédures de catégorisation des locuteurs, in *Interaction et langue étrangère*, par Arditty J. et Vasseur M.T., *Langages*, juin 1999, n°134, Larousse, p.20-34.

- Morange S., 2005, *Approches structurale, prosodique, psycho-cognitive de quelques propriétés linguistiques subjectives de la parole : l'exemple de trois Montmartrois âgés*, doct. Sciences du Langage, Univ. Paris 3, dir. M-A. Morel, non publié.
- Morange S., Fontaine J-M, Vogel C., Poitevineau J., Dubois D., 2005, Appartenances catégorielles d'un objet complexe, la voix : approches linguistique et psycho-cognitive, *3ème Journée du Sensolier*, oct. 2005, Ivry, CNRS, p.14-17, <http://www.risc.cnrs.fr/actu.php>
- Morange S., J-M. Fontaine et D. Dubois, 2008, Perception of Recording Singing Voice Quality and Expertise : cognitive linguistics and acoustics approaches, *Journal of Voice*, Elsevier, The Voice Fondation & Intern. Ass. of Phonosurgery, 9 p. (à paraître).
- Morin E., 1982, *Science avec conscience*, Fayard.
- Morin E., 1981, *Pour sortir du 20^{ième} siècle*, Seuil.
- Morin E., Motta R., Ciurana E-R., 2003, *Eduquer pour l'ère planétaire, la pensée complexe comme méthode d'apprentissage dans l'erreur et l'incertitude humaine*, Balland.
- Ohala J.J., Gilbert J.B., 1979, Listeners' Ability to identify Languages by their Prosody, in Problèmes de prosodie Vol.II, Expérimentations, modèles et fonctions, *Studia Phonetica 18*, Léon P. et Rossi M. dir., Didier.
- Poitevineau J., 2004, L'usage des tests statistiques par les chercheurs en psychologie : aspect normatif, descriptif et prescriptif, *Math. & Sciences humaines*, n°167 (3), p. 5-25.
- Ramus F., 1999, *Rythme des langues et acquisition du langage*, doctorat en Sciences du Langage, dir. Mehler J., EHESS.
- Schroder M., Aubergé V., Cathiard M-A., 1998, Can We Hear Smile?, in *ICSLP98 : Prosody and Emotion 2*.
- Shötz S., 2006, Stimulus Duration ant Type in Perception of Female and Male Speaker Age, *Proceedings of Interspeech*, Lisbonne, p. 529-532.
- Tran DD. *et ali*, 2005 Influence of F0 on Vietnamese syllable perception, *Interspeech*, (CD-Rom) Lisbonne.
- Traunmüller H, Öhrström N, 2007, Audiovisual perception of openness and lip rounding in front vowels, *Journal of Phonetics*, Volume 35, 2, p. 244-258.
- Vasilescu I., Candea M., Adda-Decker M., 2005, Perceptual salience of language-specific acoustic differences in autonomous fillers across eight languages, *Proceedings of Interspeech*, Lisbonne.
- Vieru-Dimulescu B. 2007, [thèse en cours, Limsi-CNRS], *Aspects segmentaux et suprasegmentaux dans la perception d'un accent étranger*.

