



HAL
open science

Catégorisation de distorsions vocaliques produites par un apprenant hispanophone adulte en français L2

Cynthia & Magnen, Pascal Gaillard

► **To cite this version:**

Cynthia & Magnen, Pascal Gaillard. Catégorisation de distorsions vocaliques produites par un apprenant hispanophone adulte en français L2. 4e Congrès Mondial de Linguistique Française, Jul 2014, Berlin, Allemagne. pp.1329 - 1343, 10.1051/shsconf/20140801366 . hal-01586658

HAL Id: hal-01586658

<https://hal.science/hal-01586658>

Submitted on 13 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Catégorisation de distorsions vocaliques produites par un apprenant hispanophone adulte en français L2

Magnen, Cynthia & Gaillard, Pascal

Octogone, Université de Toulouse 2 le Mirail
cynthia.magnen@univ-tlse2 et pascal.gaillard@univ-tlse2

1 Introduction

La perception des distorsions phonétiques en langue maternelle (L1) constitue un thème de recherche peu exploité ou seulement dans le domaine des pathologies (cf. par exemple Blumstein et al. 1980 ; Blumstein 1990). Pourtant, un bon exemple de ce type de parole nous est offert si l'on considère les productions d'un apprenant en cours d'acquisition d'une L2. Sur les plans phonétique et prosodique, une distorsion des catégories phonémiques et des patrons accentuels de la L2 au profit de catégories natives de la L1 ou de catégories intermédiaires créées lors du stade constitutif de l'interlangue (Selinker, 1972) est fréquemment observée. À ce stade de l'apprentissage, les distorsions phonétiques ainsi produites peuvent ne pas être reconnues comme catégories natives par les usagers de la langue cible lors d'interactions verbales. La difficulté à identifier ce type de production peut alors entraver la communication à différents degrés.

Le stade de l'interlangue est généralement caractérisé par le développement de *systèmes intermédiaires successifs et dynamiques* permettant d'approcher au mieux la langue cible (cf. Gaonac'h, 1991 ; Vogel, 1995). La manifestation de propriétés et de règles communes, non seulement à la L1 mais aussi à la langue cible, a également été décrite (Corder, 1975 ; Besse & Porquier, 1984). Aussi, le stade de l'interlangue comme phase d'adaptation à un système linguistique nous est apparu comme un cadre d'analyse pertinent pour mieux comprendre les processus cognitifs impliqués dans le traitement des unités phonologiques d'une langue. Partant des résultats de recherche dans le domaine de l'acquisition des catégories phonologiques en L1 et L2, nous considérons en effet que l'étude de la catégorisation de distorsions acoustico-phonétiques devrait permettre d'affiner les connaissances sur les processus d'identification de catégories phonémiques.

La présente étude concerne la catégorisation de distorsions de la voyelle orale antérieure arrondie /y/ produites en français L2 par un apprenant hispanophone de niveau A2. À travers une tâche de catégorisation libre, nous observons comment des natifs francophones perçoivent plusieurs formes de distorsion de cette voyelle produite dans un même contexte phrastique « *C'est bucolique* ». Notre étude a pour objectif de tester deux hypothèses : 1/ la distorsion est perçue en fonction de catégories natives et dans ce cas, des phénomènes d'assimilation (cf. Best, 1995) des distorsions phonétiques à des catégories phonologiques de L1 devraient se produire en fonction de caractéristiques communes ; 2/ la distorsion est très éloignée de la cible vocalique à atteindre ; dans ce cas, les distorsions sont traitées comme des réalisations phonétiques particulières n'entrant dans aucune catégorie phonologique native.

1.1 Catégorisation de contrastes sonores en L1 et L2

Même lorsque le niveau de compétences lexicales et syntaxiques des locuteurs apprenant une L2 est satisfaisant, la perception des variations dans la prononciation se présente comme une tâche complexe et difficile pour les auditeurs natifs de la langue cible. Étant communément établi que les déviations par rapport aux normes de la prononciation de la langue cible sont largement influencées par la structure phonologique de la L1 (cf. Strange, 1995 pour une revue détaillée sur le sujet), il s'en suit différentes variations allant de déviations phonétiques légères à des substitutions segmentales, des phénomènes

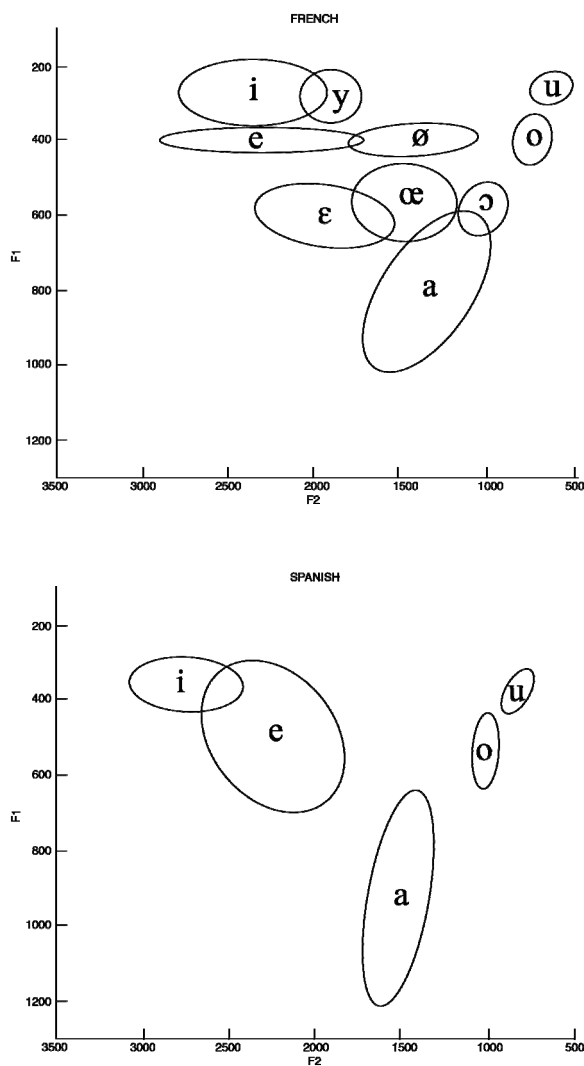
d'élision et d'insertion, et des modulations prosodiques (Weber et al. 2011). Notre problématique de départ est de comprendre comment cette variabilité est traitée, et comment le locuteur parvient à interpréter un percept stable à partir des différents degrés de variation de la cible normée vers laquelle il tend.

Pendant longtemps, les études menées sur la gestion de la variabilité phonétique en L1 se sont orientées vers les *prototypes*, figurant des instances phonémiques centrales et relativement stables, autour desquels se structureraient les catégories phonologiques de la L1 (voir Rosch, 1976 pour une première étude sur les prototypes, puis Kuhl, 1991 pour une adaptation aux catégories de la parole). L'aptitude à discriminer entre différents contrastes sonores serait conditionnée par le *degré de typicalité* d'un exemplaire au sein d'une catégorie. Les réalisations phonétiques s'approchant d'un prototype seraient en effet plus difficiles à discriminer que des exemplaires moins typiques (Grieser & Kuhl, 1989 ; Guenther & Gjaja, 1996 ; Iverson & Kuhl, 2000). De leur côté, les nombreuses études portant sur la perception de contrastes non natifs ont mis en évidence que les habitudes sélectives permettant d'organiser et d'acquérir la phonologie de la langue maternelle affectent l'aptitude à percevoir de nouveaux sons (e.g., Best, 1995 ; Flege, 1995 ; Kuhl & Iverson, 1995). Plusieurs recherches ont cependant permis de nuancer ces résultats en montrant que sous certaines conditions d'écoute, les auditeurs manifestent des performances de discrimination de contrastes non natifs comparables à celles des auditeurs natifs (voir par exemple, Dupoux et al., 1997 ; Polka, 1991 ; Werker & Tees, 1984). Plus récemment, le modèle ASP (Automatic Selective Perception) (Strange, 2002 ; 2011) propose qu'un plus large éventail de facteurs variationnels soient inclus dans l'analyse perceptive. Parmi eux, non seulement la finalité et les intentions de la situation de communication, les connaissances des sujets recrutés lors d'une tâche expérimentale et la complexité du matériel linguistique sont à observer avant de conclure sur les résultats. Le traitement perceptif est dès lors envisagé comme un acte dynamique et adaptatif reflétant différents modes de traitement privilégiés en fonction des différents facteurs variationnels (*cf.* également la notion de « filtres dynamiques » introduite par Werker & Curtin, 2005).

1.2 Distorsions vocaliques à l'étude

La présente étude propose de se centrer sur la catégorisation des distorsions de la voyelle orale antérieure arrondie /y/ produites en français L2 par un apprenant hispanophone de niveau A2. Les stimuli proposés sont issus d'une précédente étude longitudinale destinée à évaluer les effets de la Méthode d'intégration phonétique Verbo-Tonale (MVT) sur la production de contrastes non natifs par un apprenant hispanophone en Français L2. L'énoncé « *C'est bucolique* » a spécifiquement été sélectionné pour la présente étude, car l'entourage phonétique de la voyelle /y/ est extrêmement productif en termes d'erreurs réalisées par l'apprenant (pour plus de détails sur le choix de l'entourage phonétique de la voyelle, *cf.* Billières et al., 2006). Parmi les différentes réalisations choisies de l'énoncé, les distorsions acoustico-phonétiques de la voyelle /y/ sont influencées à la fois par les modèles proposés par l'enseignant utilisant la MVT (pour une présentation détaillée de la méthode, voir par exemple Guberina, 1991 ; Renard, 1979), destinés à accentuer les caractéristiques du français oral, et par les habitudes perceptuelles sélectives de la L1 de l'apprenant. Ce type de parole est en règle générale difficile à proposer en contexte expérimental, car elle subit un grand nombre de variations difficilement contrôlables. Or, la sélection d'un même énoncé produit au cours de diverses séances de correction phonétique offre l'avantage, et non des moindres, d'avoir accès à ses conditions de production et aux règles auxquelles il est soumis.

Afin de mieux rendre compte de la cible vocalique à atteindre et des difficultés que peut rencontrer un apprenant hispanophone en français L2, nous proposons une représentation des zones possibles de réalisation des voyelles de la L1 et de la L2 selon un plan biformantique (Figures 1 et 2) (Meunier et al., 2003). Cette représentation acoustique des deux systèmes permet de témoigner du peu de liberté dont bénéficie le système du français puisque chaque espace est dédié à une réalisation particulière. Cela n'empêche cependant pas certaines catégories vocaliques de se chevaucher. En se concentrant sur la zone formantique du son [y], on peut en effet observer dans chacun des deux systèmes que : (a) l'espace formantique de la réalisation [i] chevauche celui de la réalisation [y] en français, (b) comme celui de la réalisation [i] chevauche la réalisation [e] en espagnol.



Figures 1 et 2. Représentations bifonctives des aires de réalisation des voyelles orales du français (en haut) et de l'espagnol (en bas) (Meunier et al., 2003).

Ces brefs points de comparaison permettent d'imaginer la difficulté que les apprenants hispanophones peuvent rencontrer face au système vocalique francophone et plus spécifiquement face au phonème /y/. Par leur système phonologique qui n'inclut pas la série des voyelles antérieures arrondies du français, les Hispanophones éprouvent souvent des difficultés à percevoir et à situer le son [y] qui partage soit le trait d'arrondissement, soit le trait d'antériorité des sons [i] et [u] présents dans leur système. Aussi, une éventuelle transposition de la zone de réalisation du [y] dans le système français sur le système vocalique espagnol permet d'observer un recouvrement des zones de réalisations des voyelles [i] et [e]. À partir de ce raisonnement, on peut également présumer des difficultés à percevoir le phonème [e] français qui ne se réalise pas de la même manière dans les deux langues (cf. Baqué & Cañada, 2005).

1.2.1 Mesures formantiques des distorsions de la voyelle /y/

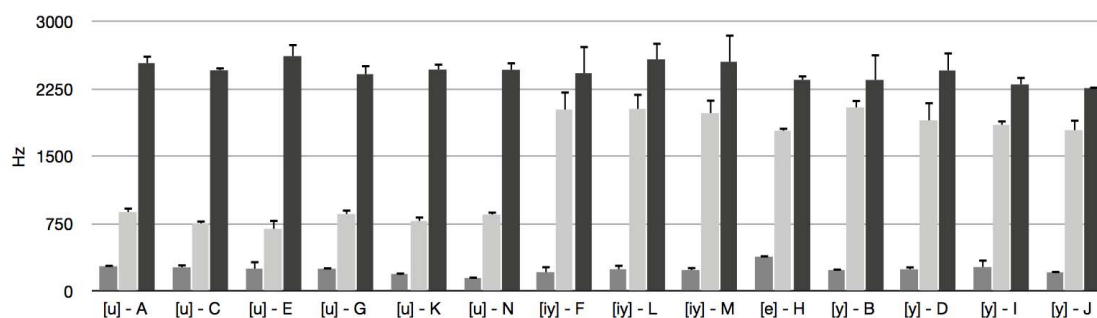
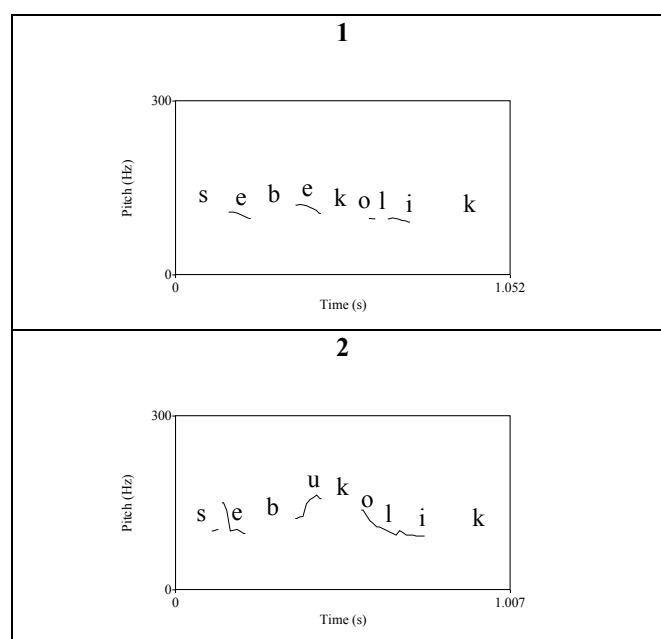


Figure 3. Moyennes et écart-types des valeurs de F1, F2 et F3 (Hz) des distorsions phonétiques de la voyelle /y/ produites par un apprenant hispanophone en français L2 (niveau A2) (analyse LPC sous PRAAT).

Dans le détail, les distorsions produites sont de différents ordres. En se référant à la Figure 3 ci-dessus, on dénombre en effet des déviations phonétiques plus ou moins légères (*cf.* écarts-types des stimuli B, D, I, J), des substitutions segmentales ([e] et [u], *cf.* stimuli H et A, C, E, G, K, N), des phénomènes d'insertion qui produisent des réalisations diphtonguées de la voyelle cible (*cf.* stimuli F, L, M).

1.2.2 Modulations de la F0 et durée vocalique

Le Tableau 1 présente les différents types de modulations de la F0 de l'énoncé « *C'est bucolique* ». Il permet de visualiser trois types de modulations appliquées lors de la remédiation phonétique en MVT. La voyelle cible est proposée soit (1) sans modulation précise (en guise de premier modèle généralement), (2) en modulation ascendante afin de faire percevoir les caractéristiques perceptives dites « claires » de la voyelle, (3) en creux mélodique, ce qui constitue un patron non facilitant incitant l'apprenant à produire plus facilement une distorsion.



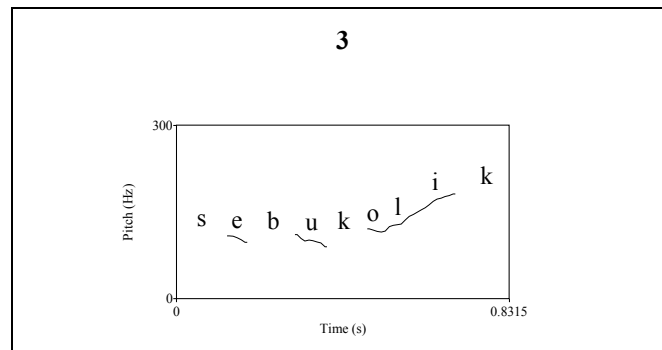


Tableau 1. Types et exemples de modulations prosodiques produites (analyse de la F0 sous PRAAT).

Les procédures de MVT peuvent également consister à ralentir le débit afin que l'apprenant puisse percevoir l'ensemble des caractéristiques sonores et diminuer sa tension articulaire. Ceci a pour effet d'allonger la durée des segments vocaliques. Le Tableau 2 reporte l'ensemble des durées vocaliques et précise le type de modulation de F0 à laquelle un énoncé est associé.

	[u] - A	[u] - C	[u] - E	[u] - G	[u] - K	[u] - N	[y] - F	[y] - L	[y] - M	[e] - H	[y] - B	[y] - D	[y] - I	[y] - J
Début	0,36	0,38	0,35	0,33	0,37	0,31	0,34	0,41	0,34	0,36	0,43	0,36	0,32	0,31
Fin	0,41	0,42	0,41	0,40	0,41	0,35	0,47	0,55	0,48	0,46	0,55	0,48	0,46	0,39
Durée totale	0,05	0,04	0,06	0,07	0,04	0,04	0,13	0,14	0,14	0,10	0,12	0,12	0,14	0,08
Modulation F0	2	2	3	2	2	3	3	2	2	1	1	2	1	2

Tableau 2. Durée en secondes des distorsions vocaliques produites (analyse sous PRAAT).

1.2.3 Hypothèses

En tenant compte de la complexité du matériel expérimental et des connaissances préalables de la population testée, nous envisageons nos hypothèses en nous basant sur le contexte théorique que nous avons décrit en amont. Nous considérons que le traitement des stimuli lors d'une tâche de catégorisation libre (*cf.* section suivante) est dépendant du jugement des adultes francophones natifs (donc experts de la langue cible) sur les distorsions produites. Ainsi deux modes de traitement sont à envisager pour la catégorisation, selon deux grandes hypothèses :

1. (a) Les classes de l'auditeur sont réalisées selon un mode phonologique incluant la validité lexicale des énoncés, c'est-à-dire que les règles phonologico-lexicales se constituent comme critères de catégorisation (*cf.* modèle ASP, Strange 2011). Dans ce mode de traitement, l'auditeur considère les exemplaires B-[y], D-[y], I-[y] et J-[y] comme des productions acceptables de l'énoncé malgré leur grande variabilité (*cf.* Figure 3 et Tableau 1) et les assimile à la catégorie phonologique native /y/. Deux grandes classes seront constituées de cette façon: les énoncés produits avec la voyelle cible et les autres réalisations.

1. (b) La classification selon un mode phonologique décrite en 1. (a) peut toutefois subir quelques variations si l'auditeur considère dans sa catégorisation les contextes prosodiques en tant que mouvements intonatifs ainsi que les catégories phonologiques natives autres que celles de la voyelle cible, c'est-à-dire sans se limiter à la congruence de la voyelle dans l'énoncé. Ainsi, l'exemplaire H-[e] et les distorsions arrondies de la voyelle cible en [u] pourraient par exemple constituer deux classes supplémentaires. Dans ce cas l'incongruité lexicale n'est pas prise en compte, les critères phonologiques sont restreints aux unités syllabiques et ne s'étendent plus au mot lexical.

2. L'auditeur ne juge pas les distorsions les plus proches du modèle comme acceptables. Les critères perçus ne sont pas suffisants pour assimiler les différents exemplaires à une catégorie phonologique connue. Les variations des énoncés (segments phonétiques, durée vocalique, modulations de la F0,...) sont

classées selon un mode phonétique et leurs caractéristiques devraient permettre de réaliser des classes en fonction de critères acoustiques fins.

2 Méthodologie

Un test de catégorisation libre (TCL) est proposé aux sujets afin de répondre aux hypothèses posées en section précédente. Le TCL repose sur les principes de classification définis par Rosch et qui mettent en évidence la tendance naturelle de l'être humain à percevoir les informations du monde de manière structurée et non aléatoire. Cette aptitude à la catégorisation permet une appréhension et un apprentissage du monde efficace et reposerait sur un principe général d'économie cognitive (Rosch, 1978 ; 1983). Si elle n'a été introduite que récemment dans le champ sensoriel, la procédure de catégorisation libre a couramment été utilisée en psychologie sociale et cognitive (pour une revue détaillée, voir Coxon, 1999) pour mettre en évidence - grâce à des analyses statistiques - la structure des espaces perceptuels autour des stimuli et pour pouvoir interpréter les dimensions sous-jacentes et définitoires de ces espaces (Valentin et al., 2012). La méthode est de plus en plus répandue dans les recherches en perception auditive. Elle a notamment été appliquée à l'évaluation perceptive de sons environnementaux (Dubois et al., 2006, Guastavino, 2007), de sons du quotidien (Guyot et al., 1997), de la qualité d'enceintes acoustiques (Lavandier et al., 2005) et de la qualité des sons en général (Parizet, 2012). Cette étude se présente comme la seconde application de la méthode à l'analyse perceptuelle de la parole (pour la première étude, cf. Magnen et al. 2005).

2.1 Participants

Locuteur: 1 étudiant hispanophone (bilingue castillan/catalan avec langue prédominante catalan) âgé de 20 ans et inscrit en Licence 3 Géographie à l'Université de Toulouse le Mirail. Les premiers enregistrements de ses productions ont été réalisés alors qu'il résidait depuis 5 mois en France dans le cadre du programme d'échange européen *Erasmus*.

Auditeurs: 25 étudiants francophones natifs de l'Université de Toulouse le Mirail (8 femmes et 17 hommes) âgés entre 18 et 30 ans et ne présentant pas de troubles de l'audition.

2.2 Matériel

14 réalisations de l'énoncé "*C'est bucolique*" produites par un locuteur hispanophone apprenant le français en L2, dans le cadre de séances d'intégration phonétique selon la MVT. Les énoncés ont été sélectionnés par un phonéticien et un psychoacousticien francophones natifs au sein d'un corpus oral constitué de 7 séances de correction phonétique de 45 minutes effectuées sur une période de 1 mois. Les énoncés retenus mettent en évidence les différents types de distorsions produites de la voyelle /y/ reflétant les difficultés éprouvées par le locuteur hispanophone dans son processus d'apprentissage de la prononciation d'un nouveau son. Nous renvoyons le lecteur à la Figure 3 pour le détail de l'analyse formantique de chacune des réalisations.

2.3 Procédure

Les participants effectuent une tâche de catégorisation libre (TCL). Comme son nom l'indique, la procédure de catégorisation proposée consiste à catégoriser les stimuli auditifs en fonction de critères libres non imposés par l'expérimentateur. Les stimuli sonores sont présentés (1) d'un point de vue visuel, par des icônes numérotées aléatoirement qui apparaissent simultanément sur un écran d'ordinateur (2) et d'un point de vue auditif, via un casque audio. Les participants peuvent interagir de deux façons sur les icônes sonores : ils peuvent écouter le son en cliquant sur l'icône qui lui correspond et former des groupes de sons semblables en les déplaçant sur l'écran (TCL-LabX, <http://petra.univ-tlse2.fr/tcl-labx/>) (Gaillard, 2009). Les sujets doivent écouter les 14 stimuli sonores et les catégoriser. Ils sont informés que les stimuli mis dans un même groupe sont jugés similaires. Ils peuvent faire autant de groupes qu'ils veulent. Ils

peuvent écouter les stimuli autant de fois qu'ils le souhaitent et dans l'ordre désiré. La tâche n'est pas contrainte par le temps.

2.4 Analyse des résultats

Nous utilisons la méthode FAST - pour Factorial Approach for Sorting Task data - développée par l'équipe du département de Mathématiques Appliquées de l'Agrocampus de Rennes en France (pour une présentation détaillée, voir Cadoret et al., 2009). La méthode FAST est basée sur une analyse à correspondance multiple (*Multiple Correspondance Analysis* ou MCA) (Greenacre, 1984) permettant de proposer une représentation des objets et des catégories sur un plan à 2 dimensions. L'usage d'une MCA pour analyser les données de catégorisation a déjà été éprouvé par Van der Kloot & Van Herk (1991) qui obtiennent des résultats similaires en comparant une MCA à un positionnement multidimensionnel (*MultiDimensional Scaling* ou MDS) habituellement plus répandu en analyse catégorielle. L'intérêt de FAST est de compléter les outils d'analyse offerts par la MCA par une représentation des sujets et de fournir des éléments de validation grâce à une représentation en ellipses de confiance des catégories moyennes formées par les sujets. Les sous-sections suivantes ont pour objet de présenter ces outils d'analyse et les représentations engendrées. Pour le détail du calcul et de la méthode permettant d'obtenir ces représentations, nous renvoyons le lecteur à l'article de présentation de la méthode FAST suscitée (Cadoret et al. 2009).

2.4.1 Analyse préliminaire

En premier lieu, les résultats peuvent être analysés avec des histogrammes permettant de décrire le nombre de groupes par participant et le nombre de stimuli par groupe. Ensuite, les associations entre stimuli sont synthétisées dans une matrice de co-occurrences entre les stimuli de dimension $I \times I$. Cette matrice permet de voir les associations les plus fréquemment réalisées entre stimuli par l'attribution de valeurs propres égales à 0 ou 1 (1 si deux stimuli ont été classés dans le même groupe et 0 dans le cas contraire).

2.4.2 Représentation des participants

La représentation des participants est obtenue en utilisant dans un premier temps une analyse factorielle multiple (*multiple factor analysis* ou MFA) (Escofier & Pagès, 1998). Une équivalence entre MCA and MFA est ensuite réalisée sur la base des mêmes variables en considérant chaque variable comme un groupe (Pagès, 2002). La représentation des groupes de variables en MFA devient une représentation des participants par leur matrice individuelle de co-occurrence de dimension $I \times I$. Dans cette représentation, le produit scalaire entre deux participants exprime la ressemblance entre les partitions associées à deux variables quantitatives. La MFA fournit une représentation optimale des participants en termes de dimensions en lien avec la représentation des stimuli. Le plan à deux dimensions (Axes 1 et 2) représente les solutions les plus importantes (critères utilisés par les participants pour classer les sons) pour la classification. Le pourcentage associé à chaque dimension représente la contribution de cette dimension à la variance totale. Sur ce plan, chaque participant est représenté par un point et deux participants sont d'autant plus proches qu'ils ont effectué les mêmes regroupements de stimuli.

2.4.3 Carte perceptuelle des stimuli

Une MCA est réalisée à partir d'un tableau disjonctif représentant chaque participant par un ensemble de variables nominales, où chaque variable correspond à un groupe : une variable nominale associée à un groupe est égale à 1 si le stimulus appartient au groupe et à 0 dans le cas contraire. L'ensemble des stimuli peut alors être représenté dans un espace K-dimensionnel où les distances entre stimuli illustrent deux principes : (a) deux stimuli se superposent dans l'espace s'ils ont été classés ensemble par tous les participants; (b) deux stimuli sont d'autant plus proches qu'ils ont été classés ensemble par le plus grand nombre de participants. Réciproquement, deux stimuli sont d'autant plus distants qu'ils ont souvent été

classés dans deux groupes distincts. Plus précisément, la distance d'un groupe sera inversement proportionnelle à sa taille, c'est-à-dire que le classement d'un stimulus dans un groupe de petite taille éloigne ce stimulus des autres. Les dimensions représentées sur les figures obtenues sont les mêmes que les dimensions utilisées dans la représentation des participants.

2.4.4 Stabilité de la représentation des stimuli

La méthode FAST permet de représenter la variabilité inter-individuelle par des ellipses de confiance. Les ellipses ont pour objectif d'évaluer à quel point le positionnement des stimuli aurait été identique si l'étude avait été réalisée avec un autre groupe d'auditeurs. Cette évaluation se base sur une technique de rééchantillonnage de l'ensemble des participants de l'étude en effectuant un tirage aléatoire avec remplacement. Ce traitement de l'ensemble initial des participants génère un panel virtuel composé d'autant de participants que le premier groupe. Les participants de ce panel sont ensuite introduits comme éléments supplémentaires dans l'analyse et permettent de calculer de nouvelles coordonnées de stimuli. Ce procédé est répliqué plusieurs fois jusqu'à obtenir des ellipses de confiance incluant 95 % des représentations des stimuli virtuels.

3 Résultats

3.1 Analyse préliminaire

En tout, les participants ont constitué entre 3 et 6 groupes composés généralement de 1 à 7 stimuli. La matrice de co-occurrences entre les stimuli comporte les valeurs représentant le nombre de fois que deux stimuli ont été associés. Elle présente un bon nombre de cellules nulles indiquant un consensus entre les sujets sur la dissociation entre des groupes de stimuli. Certains stimuli ont souvent été isolés et ont ainsi été perçus comme des stimuli particuliers : B-[y], F-[iy] et H-[e]. Les paires C-[u] et A-[u] ; C-[u] et K-[u] ont été classés 23 fois dans le même groupe, ce qui représente la fréquence la plus importante. Les paires L-[iy] et M-[iy] ; A-[u] et K-[u] ; G-[u] et K-[u] ; E-[u] et N-[u] sont également très fréquentes, car elles ont été classées 22 fois dans le même groupe. Ceci indique un fort consensus sur les groupes comportant les distorsions [u] et [iy].

3.2 Analyse des classes

L'arbre hiérarchique (Figure 4) montre les classes réalisées par l'ensemble des participants. Les trois grandes classes sont représentées par les stimuli K, A, C, G, N, E et F pour la première classe, D, I, J, B, M et L pour la deuxième et H seul pour la troisième. On constate la présence de deux classes avec un nombre équivalent de stimuli (7 et 6) et une classe distincte avec un seul stimulus. La première classe peut se subdiviser en deux sous-classes : K, A, C et G, et N, E et F. Celles-ci sont également équivalentes en nombre d'éléments. La deuxième grande classe se divise pour sa part en trois sous-classes : D, I et J ; M et L ; puis B seul. L'ensemble des participants a donc procédé à une classification comportant des nombres équivalents de stimuli dans les classes proposées, à l'exception notable des stimuli H, et dans une moindre mesure B.

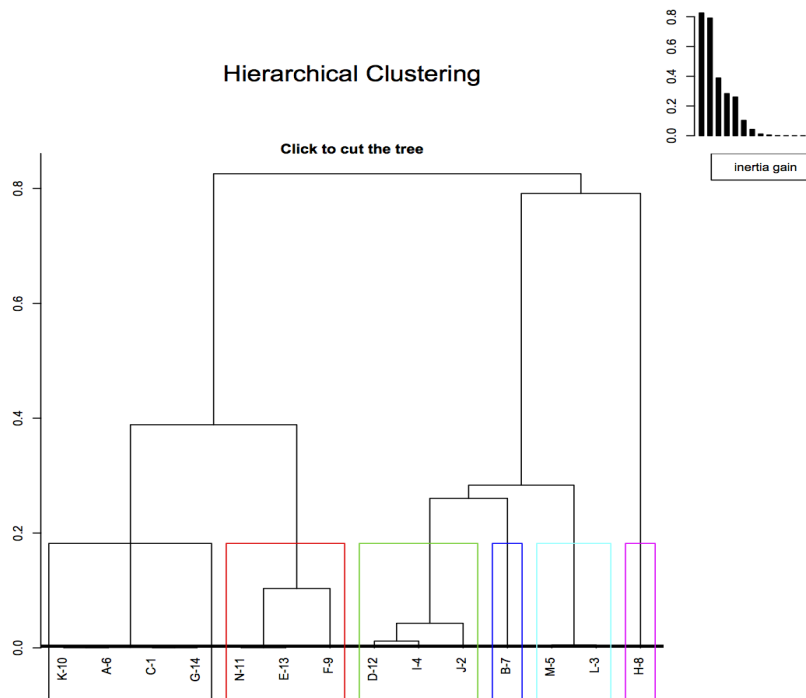


Figure 4. Analyse hiérarchique des classes pour les 25 participants.

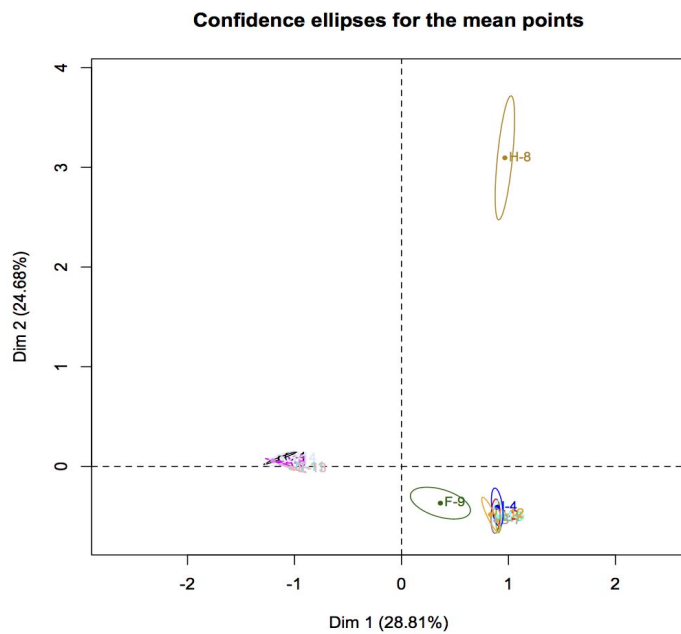


Figure 5. Ellipses de confiance des groupes de stimuli réalisés par les 25 participants.

Dans la Figure 5, on peut observer une distribution de points sur deux axes représentant les deux premiers critères utilisés par les participants pour classifier les stimuli. Sur l'axe de la Dimension 1, deux groupes se distinguent (à gauche et à droite du 0). Le premier à gauche du zéro comporte les stimuli K, A, C et G, le deuxième à droite du zéro comporte F, D, I, J, B, M, L et H. Seuls les stimuli F et H sont

représentés avec une ellipse importante. Nous pouvons en conclure que ces deux stimuli ne font pas consensus entre les participants en fonction des critères choisis pour leur classification. Le stimulus F a une ellipse qui s'étend horizontalement indiquant que cette hésitation a eu lieu sur le classement en fonction de la Dimension 1. L'ellipse du stimulus H est plutôt verticale, illustrant une hésitation davantage présente sur la deuxième dimension du graphe. Si on analyse les caractéristiques phonétiques des stimuli à droite du zéro et à gauche du zéro, on constate que la distinction entre ces deux groupes porterait a priori sur les traits d'antériorité et d'arrondissement (*cf.* les écarts types relatifs aux moyennes des F2 au F3 sur la Figure 3).

La Dimension 2 permet de distinguer trois groupes : l'un en bas du zéro composé de F, D, I, J, B, M et L. Le deuxième groupe se situe au-dessus du zéro et ne comporte qu'un seul élément, le stimulus H. Le troisième groupe est positionné quasiment sur le zéro et compte les stimuli K, A, C, G, N et E. Cette position semble indiquer que la Dimension 2 n'a pas été discriminante pour la classification de ce groupe. L'analyse des groupes de stimuli sur cet axe suggère que le trait d'ouverture a été discriminant pour distinguer le stimulus H des autres. Le groupe de stimuli positionné sur le zéro de la Dimension 2 ne semble pas concerné par ce trait distinctif. Ce qui explique probablement sa position neutre sur cette dimension.

3.3 Représentation des participants

Sur les deux dimensions retenues préalablement, nous pouvons positionner les différents participants. La Figure 6 représente les participants en fonction de l'usage qu'ils ont fait des deux critères de classification susmentionnés.

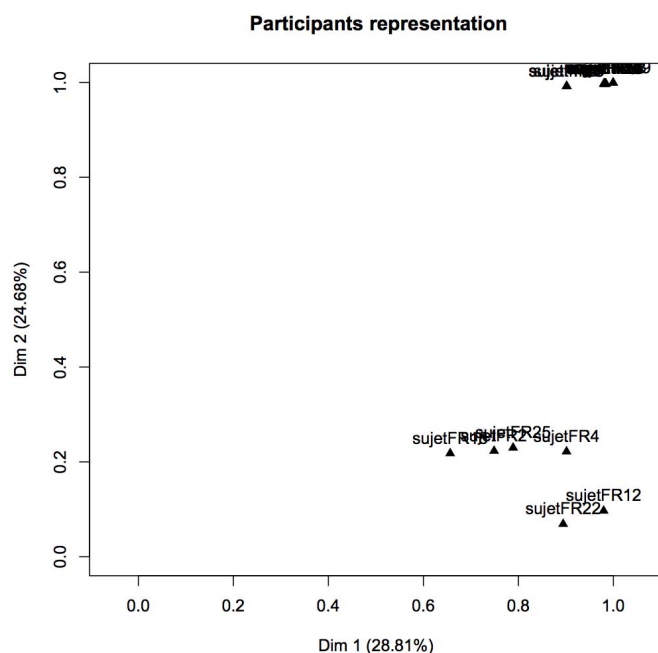


Figure 6. Représentation des participants sur un plan défini par les Dimensions 1 et 2.

Les sujets positionnés entre 0.8 et 1 sur l'un et/ou l'autre des axes indiquent qu'ils ont utilisé le critère représenté par l'axe. Ainsi, on peut observer que sur le Dimension 1, la quasi totalité (92 %) des participants se situent entre 0.8 et 1, indiquant par là-même que les classes réalisées répondent à ce critère. Par contre, sur la Dimension 2, on peut voir que seuls 19 sujets sur 25 (76 %) semblent avoir utilisé ce critère. L'analyse détaillée des classes des participants ne partageant pas le consensus de la

Dimension 2 révèle qu'ils ont tous classé le stimulus H avec d'autres stimuli tandis que les participants restant l'ont classé à part. Nous pouvons en conclure que le premier critère (Dimension 1) fait plus consensus pour l'ensemble des sujets que le second.

4 Discussion générale

Dans cette étude, nous avons proposé une tâche de catégorisation libre à des francophones natifs. Les stimuli à catégoriser étaient différents types de distorsions de la voyelle orale antérieure fermée non-arrondie [y] du français produites dans l'énoncé "*C'est bucolique*" par un apprenant adulte hispanophone en français L2. Rappelons que les instructions se rapportant à la tâche n'incitent aucunement les participants à catégoriser les stimuli en fonction des distorsions à l'étude.

Le consensus constaté entre les participants, en particulier dans la matrice de co-occurrences des stimuli, permet d'affirmer que la tâche proposée engage des processus de traitement des stimuli similaires pour la quasi-totalité des participants. Dans l'ensemble, plusieurs tendances se dégagent des analyses effectuées. Tout d'abord, l'analyse hiérarchique des classes et le graphique des ellipses de confiance (Figures 4 et 5) semblent révéler que les catégories phonologiques natives aient constitué un critère de classification prédominant permettant de distinguer efficacement trois grandes catégories natives parmi les distorsions vocaliques entendues: /u/, /y/ et /e/. Aussi, les distorsions donnant lieu à des réalisations diphtonguées de la voyelle cible [iy] ont été associées aux distorsions les plus proches de la voyelle [y]. C'est-à-dire que les distorsions [y] considérées a priori par les expérimentateurs comme les plus proches du modèle cible (même si très variables) n'ont pas été considérées par les participants comme des exemplaires particulièrement représentatifs de la catégorie native. L'inclusion de ces éléments diphtongués [iy] dans la catégorie laisse également supposer que dans cette tâche particulière la catégorie /y/ créée tolère une variabilité importante.

Les deux autres classes semblent avoir été formées sur les mêmes principes : les stimuli comportant les distorsions [u] et la distorsion [e] paraissent avoir également été perçus comme des exemplaires plus ou moins bons de catégories natives connues. Notons toutefois la classification particulière du stimulus H (76 % des participants l'ont classé isolément) qui rappelle le caractère unique de ce stimulus : il constitue en effet la seule distorsion consistant en une substitution de la voyelle cible par la voyelle [e]. Il apparaît légitime de se demander si ce simple statut ne constituerait pas un critère suffisant pour distinguer ce stimulus dans une classe particulière. Cependant, le fait qu'il ait été dissocié de la catégorie /y/ en particulier (cf. Dimension 2 sur la Figure 5) et que la catégorie /u/ ne soit absolument pas concernée par cette distinction, nous fait dire que le critère du trait d'ouverture particulièrement distinctif en français entre les catégories phonologiques /y/ et /e/ a permis de distinguer entre les réalisations associées aux deux catégories phonologiques natives.

Ces observations nous permettent d'infirmer l'hypothèse 1.a. définie en amont (cf. section 1.2.3.) selon laquelle une classification pourrait s'opérer sur la base d'un jugement fondé sur la congruité lexicale. Ainsi, et pour cette tâche particulière, un mode phonologique de classement semble avoir permis aux participants d'opérer une classification en fonction des connaissances phonético phonologiques de leur langue là où les connaissances lexico-phonologiques n'ont pas semblé pertinentes. Ces conclusions pourraient suffire à valider l'hypothèse 1.b. posant un traitement phonologique des stimuli dépourvu de jugement sur la congruence lexicale. Ce mode de traitement des stimuli induit une focalisation de l'analyse sur la structure syllabique, voire phonémique, en l'isolant de tout contexte lexical.

Pour autant, les sous-classes observées dans la Figure 4 nous permettent de penser que les associations entre stimuli reposent sur des critères de classification plutôt phonétiques car faisant appel à des propriétés acoustiques fines des stimuli. Par exemple, les deux sous-classes de la catégorie /u/ constituées des stimuli : K, A, C, G et N, E, F font état de similarités dans les modulations de la F0 (cf. Tableau 2). Si on peut penser à première vue que les stimuli répondant aux modulations de type 3 (voyelle située en creux mélodique par la modulation ascendante de la dernière syllabe du mot) peuvent être interprétées comme des questions, et donc selon une perception « sémantique », ce n'est pas le cas pour les modulations de type 2 (voyelle située en modulation ascendante afin de faire percevoir les

caractéristiques dites « claires » de la voyelle et qui ne renvoie, elle, à aucune interprétation particulière). Nous pouvons par conséquent imaginer que le critère employé dans ces sous-classes est plus phonétique (détection des variations de la F0) que phonologique. Nous pouvons constater la même chose dans l'autre grande classe (en mettant à part le cas du singleton H), avec également l'introduction du critère de durée vocalique qui permettrait de distinguer encore plus finement les exemplaires à l'intérieur des sous-classes. Notons qu'aucun des 25 sujets n'a choisi de critère phonétique comme critère principal. Nous observons par conséquent deux types de critères de classification, l'un phonologique et l'autre phonétique utilisées avec les mêmes stimuli, mais pour des fonctions différentes : le premier (phonologique) permettant de distinguer efficacement entre plusieurs grandes classes, et le deuxième (phonétique) permettant de continuer la tâche de catégorisation. La sous-catégorisation se fait à l'intérieur de la grande classe préalablement réalisée selon les critères phonologiques.

Finalement, deux processus de traitement sont convoqués pour cette tâche proposant aux participants de catégoriser des stimuli complexes en langue native selon des critères qu'ils définissent librement. Les hypothèses 1.b. et 2 semblent donc confirmées, mais contrairement à ce que nous avons préalablement défini, l'existence d'un mode de traitement n'exclut pas forcément l'autre. Ces résultats sont à rapprocher des modèles récents sur la perception de la parole qui intègrent le caractère dynamique et interactif des processus cognitifs sous-jacents. Par exemple, l'architecture du modèle PRIMIR (Processing Rich Information from Multidimensional Interactive Representations) (Werker & Curtin, 2005) fait coexister trois plans multidimensionnels de traitement des informations (Plan Perceptuel Général, Plan des Formes Lexicales et Plan Phonémique). Comme observé dans nos résultats, l'accès à ces différents plans ne serait pas exclusif et permettrait de recourir à différentes informations en fonction des différentes tâches langagières. De son côté, et comme décrit dans la section 1.1., le modèle ASP (Automatic Selective Perception) (Strange, 2011) prend également en compte les différentes connaissances préalables des locuteurs natifs d'une langue et précise les conditions de recours aux modes de traitement phonologique et phonétique. Aussi, les « routines de perception sélective » (Selective Perceptual Routines - SPRs) (Strange 2006), c'est-à-dire les processus automatiques, rapides et très robustes sur lesquels repose le processus de catégorisation de la parole seraient convoqués lors de conditions expérimentales proposant des stimuli complexes et contraignantes d'un point de vue cognitif. Finalement, lors d'une situation expérimentale, des conditions nouvelles et complexes inciteraient les auditeurs à recourir à des mécanismes de traitement linguistique de haut niveau tandis que la simplification de la tâche rendrait possible l'écoute des caractéristiques acoustiques fines de la parole par le recours aux capacités sensori-auditives originelles sur lesquelles s'appuie le processus auditif en général.

5 Conclusion

La coexistence des traitements phonétique et phonologique mise en évidence au sein de la même tâche a des implications directes sur le plan méthodologique concernant l'élaboration de tâches de perception de contrastes phonétiques natifs et non natifs.

Le protocole utilisé a en effet permis de mettre en lumière une variabilité à la fois inter participants et intra participants. Concernant la variabilité inter participants, les ellipses de confiance permettent de repérer les consensus existants ou non dans la classification des stimuli, tandis que la répartition des sujets sur un plan à dimension permet de comprendre où se situent les différences observées. Par ailleurs, les caractéristiques de la tâche permettent de montrer les différentes stratégies ou modes perceptifs employés en fonction de l'évolution de son traitement. Face à la complexité initiale de la tâche (stimuli complexes, interface nouvelle, liberté dans le choix des critères de catégorisation), les participants semblent avoir privilégié le recours à des processus automatiques de traitement par la simplification de la variabilité aux simples catégories phonologiques connues. Cet allègement de la tâche par le choix d'un premier critère, et la possibilité d'écouter les stimuli autant de fois que souhaité sans contrainte de temps, a ensuite permis un tri inter stimuli en fonction de critères acoustiques fins. C'est une des qualités de la tâche de catégorisation libre qui permet à des sujets au cours de l'expérimentation de pouvoir changer de critères et de modes perceptifs en fonction des besoins que le sujet s'assigne lui-même.

Il n'y a donc peut-être pas d'opposition à faire entre les différents modes perceptifs, autant en termes de capacités que d'exclusivité, dans la mesure où, sur une même tâche, des participants sont susceptibles de changer d'idée et d'objectifs en cours de traitement, en convoquant l'ensemble de leurs savoirs. Il est évident que cette tâche de catégorisation libre ne suffit pas à refléter les conditions réelles de traitement quotidien de distorsions produites par un locuteur, mais elle présente l'avantage de révéler la possibilité d'utiliser les différents modes perceptifs disponibles en fonction des besoins. Nous restons conscients de la nécessité de valider cette méthodologie par d'autres expérimentations permettant de faire varier les conditions expérimentales. Cette variation pourrait notamment inciter les participants à réaliser la tâche exclusivement selon l'un ou l'autre mode de traitement perceptif.

Références bibliographiques

- Baqué, L., & Cañada, M. D. (2005). Structuration de l'opposition [ə]-[e] en FLE : incidence du degré de focalisation sur la forme. *Revue Parole*, 34-35-36, 101-130.
- Besse, H., & Porquier, R. (1984). *Grammaire et Didactique des Langues*. Paris : Hatier (CREDIF). Collection « LAL ».
- Best, C. (1995). A direct realistic view of cross-language speech perception. In *Speech Perception and Linguistic Experience : Issues in cross-language research* (pp. 171-204). Timonium, MD : York Press.
- Billières, M., Gaillard, P., & Magnen, C. (2006). Étude expérimentale de la progression phonétique et des stratégies d'acquisition d'un hispanophone en français. In *Les deuxièmes langues étrangères dans le système éducatif : du primaire à l'université* (Vol. Repères &, p. CD-ROM). Bellaterra, Institut de Ciències de l'Educació, Universitat Autònoma de Barcelona: Campà, A. Baqué, L. Sanchez, M. O.
- Blumstein, S E. (1990). Phonological deficits in aphasia : Theoretical perspectives. In A. Caramazza (Ed.), *Cognitive neuropsychology and neurolinguistics* (pp. 33-53). Hillsdale, N. J. : Lawrence Erlbaum.
- Blumstein, S. E., Cooper, W. E., Goodglass, H., Sattlender, S., & Gottlieb, J. (1980). Production deficits in aphasia : A voice-onset-time analysis. *Brain and Language*, 9, 153-170.
- Cadoret, M., Lê, S., & Pagès, J. (2009). A Factorial Approach for Sorting Task data (FAST). *Food Quality and Preference*, 20(6), 410-417.
- Corder, S. P. (1975). Error Analysis, Interlanguage and Second Language Acquisition. *Language Teaching*, 8(04), 201-218.
- Coxon, A. M. (1999). *Sorting Data: Collection and Analysis* (Sage Unive., p. 07_127). Thousand Oaks, CA : SAGE.
- Dubois, D., Guastavino, C., & Raimbault, M. (2006). A cognitive approach to urban soundscapes : Using verbal data to access everyday life auditory categories. *Acta Acustica united with Acustica*, 92(6), 865-874.
- Dupoux, E., Pallier, C., Sebastián-Gallés, N., & Mehler, J. (1997). A destressing « deafness » in French ? *Journal of Memory and Language*, 36, 406-421.
- Escofier, B., & Pagès, J. (1998). *Analyse factorielle simple et multiple*. Dunod.
- Flege, J. E. (1995). Second Language Speech Learning Theory, Findings, and Problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience* (York Press., pp. 233-246). Timonium, MD.
- Gaillard, P. (2009). Laissez-nous trier ! TCL-LabX et les tâches de catégorisation libre de sons. In D. Dubois (Ed.), *Le Sentir et le Dire : Concepts et méthodes en psychologie et linguistique cognitives* (pp. 189-210). Paris : L'harmattan.
- Gaonac'h, D. (1991). *Théories d'apprentissage et acquisition d'une langue étrangère* (CREDIF - H., p. 239). Paris.
- Greenacre, M. J. (1984). *Theory and applications of correspondance analysis*. London : Academic Press.
- Grieser, D., & Kuhl, P. K. (1989). Categorization of speech by infants: Support for speech-sound prototypes. *Developmental Psychology*, 25, 577-588.
- Guastavino, C. (2007). Categorization of environmental sounds. *Canadian Journal of Experimental Psychology*, 61(1), 54-63.

- Guberina, P. (1991). Rôle de la Perception Auditive dans l'Apprentissage Précoce des Langues. *Le Français dans le Monde, Recherches et Applications, N° spécial*, 65–70.
- Guenther, F., & Gjaja, M. (1996). The perceptual magnet effect as an emergent property of neural map formation. *Journal of Acoustical Society of America*, 100, 1111–1121.
- Guyot, F., Castellengo, M., & Fabre, B. (1997). Étude de la catégorisation d'un corpus de bruits domestiques. In *Catégorisation et cognition: de la perception au discours* (pp. 41–58). Paris : Kimé.
- Iverson, P., & Kuhl, P. K. (2000). Perceptual magnet and phoneme boundary effects in speech perception: do they arise from a common mechanism ? *Perception & Psychophysics*, 62, 874–886.
- Kuhl, P. K. (1991). Human adults and human infants show a perceptual magnet effect for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, 50, 93–107.
- Kuhl, P. K., & Iverson, P. (1995). Linguistic Experience and the « Perceptual Magnet Effect. » In W. Strange (Ed.), *Speech Perception and Linguistic Experience* (pp. 121–154). Baltimore : York Press.
- Lavandier, M., Meunier, S., & Herzog, P. (2005). Perceptual and physical evaluation of differences among a large panel of loudspeakers. In *Forum Acusticum*. Budapest, Hungary.
- Magnen, C., Billières, M., & Gaillard, P. (2005). Surdit  phonologique et catégorisation. Perception des voyelles françaises par les hispanophones. *Parole*, 33, 9–30.
- Meunier, C., Frenck-Mestre, C., Lelekov-Boissard, T., & Le Besnerais, M. (2003). Production and perception of foreign vowels : does the density of the system play a role? In *the 15th International Congress of Phonetic Sciences*. Barcelona, Spain.
- Pagès, J. (2002). Analyse factorielle multiple appliquée aux variables qualitatives et aux données mixtes. *Revue de Statistiques Appliquées*, 50(4), 5–37.
- Parizet, E., & Koehl, V. (2012). Application of free sorting tasks to sound quality experiments. *Applied Acoustics*, 73(1), 61–65.
- Polka, L. (1991). Cross-language speech perception in adults : Phonemic, phonetic and acoustic contributions. *Journal of the Acoustical Society of America*, 89, 2961–2977.
- Renard, R. (1979). *Introduction à la méthode verbo-tonale de correction phonétique*. (Cipa, Ed.) (3ème ed.). Paris : Didier.
- Rosch, E. (1976). Classification d'objets du monde du réel: origines et représentations dans la cognition. *Bulletin de psychologie, La mémoire sémantique, n° spécial*, 242–250.
- Rosch, E. (1978). Principles of Categorization. In Eleanor Rosch & B. Lloyd (Eds.), *Cognition and Categorization* (pp. 27–48). Hillsdale, N. J. : Lawrence Erlbaum.
- Rosch, E. (1983). Prototype classification and logical classification : The two systems. In E. Scholnick (Ed.), *New Trends in Cognitive Representation : Challenges to Piaget's Theory* (pp. 73–86). Hillsdale, NJ : Lawrence Erlbaum Associates.
- Selinker, L., (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, 10(3), 209.
- Strange, W. (1995). Speech perception and linguistic experience : issues in cross-language research. Baltimore : York Press.
- Strange, W. (2002). Speech perception and language learning : Wode's developmental model of speech perception revisited. In P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development : Paper in honor of Henning Wode* (Wissenscha).
- Strange, W. (2006). Second-language speech perception : The modification of automatic selective perceptuel routines. *Journal of Acoustical Society of America*, 120, 3137.
- Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech : A working model. *Journal of Phonetics*, 39(4), 456–466.
- Valentin, D., Chollet, S., Lelièvre, M., & Abdi, H. (2012). Quick and dirty but still pretty good : a review of new descriptive methods in food science. *International Journal of Food Science & Technology*, 47(8), 1563–1578.

- Van der Kloot, W. A., & Van Herk, H. (1991). Multidimensional scaling of sorting data : A comparison of three procedures. *Multivariate Behavioral Research*, 26(4), 563–581.
- Vogel, K. (1995). *L'interlangue : la langue de l'apprenant*, trad. de Jean-Michel Brohée et Jean-Paul Confais (p. 313). Toulouse : Presses Universitaires du Mirail.
- Weber, A., Broersma, M., & Aoyagi, M. (2011). Spoken-word recognition in foreign-accented speech by L2 listeners. *Journal of Phonetics*, 39(4), 479–491.
- Werker, J. F., & Curtin, S. (2005). PRIMIR : A Developmental Framework of Infant Speech Processing. *Language Learning and Development*, 1(2), 197–234.
- Werker, J. F., & Tees, R. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, 75, 1866–1878.