



HAL
open science

**Asymétries en perception et traitement de bas niveau :
traces auditives, mémoire a court terme et
représentations mentales**

Charalampos Karypidis

► **To cite this version:**

Charalampos Karypidis. Asymétries en perception et traitement de bas niveau : traces auditives, mémoire a court terme et représentations mentales. Linguistique. Université de la Sorbonne nouvelle - Paris III, 2010. Français. NNT : 2010PA030015 . tel-00487590v2

HAL Id: tel-00487590

<https://theses.hal.science/tel-00487590v2>

Submitted on 26 Nov 2013

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.

UNIVERSITÉ PARIS 3 – SORBONNE NOUVELLE
ÉCOLE DOCTORALE 268 “LANGAGE ET LANGUES”
LABORATOIRE DE PHONETIQUE ET PHONOLOGIE (LPP)

Asymétries en perception et traitement de bas niveau : traces auditives, mémoire à court terme et représentations mentales

thèse de Doctorat nouveau régime
présentée par Charalampos KARYPIDIS
discipline : Phonétique

Directeur de thèse : M. Nick CLEMENTS

soutenue le 15 janvier 2010 devant le jury composé de :

M. René CARRÉ, Examineur — Directeur de Recherche Émerite

M. Noël NGUYEN, Rapporteur — Professeur

Mme Annie RIALLAND, Examinatrice — Directrice de Recherche

M. Jean-Luc SCHWARTZ, Prérapporteur — Directeur de Recherche

M. Willy SERNICLAES, Prérapporteur — Directeur de Recherche

Mme Jacqueline VAISSIÈRE, Examinatrice — Professeure

Résumé

L'effet d'ordre de la présentation des stimuli a un impact sur leur discriminabilité. Dans cette thèse, nous étudions le phénomène d'asymétrie et comment celui-ci réagit avec divers facteurs cognitifs localisés dans différentes étapes du traitement. Nous nous focalisons sur le rôle de la périphéricité acoustique, de la dégradation auditive et des représentations mentales permanentes. Des données sur la perception visuelle complètent la partie principale de notre travail. Nous envisageons aussi la variabilité inter- et intra-sujet dans la perception et dans la production des voyelles. Nos données révèlent que les effets d'ordre sont un mécanisme complexe qui ne peut pas être réduit à un seul facteur déclencheur.

Mots-clé : effet d'ordre ; asymétrie ; voyelles ; périphéricité ; neutralisation ; traces auditives ; mémoire à court- et à long-terme ; représentations mentales

Abstract

It has been known for 150 years that the order of presentation of stimuli has an impact on their discriminability. In this thesis, we have examined the phenomenon of asymmetries in vowel perception and how it interacts with various cognitive factors located at different steps of stimulus decoding. We have focused on the role of acoustic peripherality, auditory decay and permanent mental representations. Data from visual perception complements the main body of our work. Inter-subject variability in vowel perception and production is also dealt with. Our data shows that order effects are a complex mechanism which cannot be reduced to a single trigger factor.

Keywords : order effect ; asymmetry ; vowels ; peripherality ; neutralisation ; auditory traces ; short- and long-term memory ; mental representations

Remerciements

La rédaction de cette thèse et l'exécution du travail qui y figure fut un véritable plaisir. Travaillant sur les asymétries depuis mon DEA, je suis devenu obsédé par ce phénomène intrigant et opaque à la fois.

Étant donné le nombre de tests que j'ai effectué pour cerner le phénomène des asymétries, j'ai eu l'occasion de travailler et de me mettre en contact avec un grand nombre de personnes. Je m'excuse pour toute omission.

Je considère que ceux qui méritent le plus grand merci sont les maints auditeurs et locuteurs qui ont participé à mes expériences. Sans leur dévouement et leur bonne humeur, ce travail n'aurait jamais vu le jour.

Nick Clements a été un directeur idéal à bien des égards. Il a respecté mon rythme et style de travail et m'a toujours donné des conseils précieux. Il n'a malheureusement pas eu l'occasion de lire ces lignes.

Merci à Antonia-Colazo Simon pour avoir participé à tous les tests que je lui ai proposés, pour avoir fait passer des tests à plus de 100 personnes pour nos différents projets, et surtout pour son amitié.

Et que dire de Soha Safaï? Elle a toujours accepté de passer des tests avec enthousiasme, sans se plaindre. Merci pour toutes les balades rafraîchissantes à la fin desquelles je rentrais à la maison avec du désir pour travailler.

Merci à Angelica V. Costagliola pour sa collaboration et pour avoir participé à l'étude des cas.

Merci à Samia Bouzouita pour son amitié, pour sa collaboration et pour toutes les pauses-déjeuner.

Merci aux auditeurs qui ont pris part à mon étude des cas : Audrey Acher, Johanne Akpoussan, Antonia Colazo-Simon, Angelica V. Costagliola, Stéphanie Borel, Samia Bouzouita, Wieslawa Cholewa, Christelle Exare, Iran Farkhondeh, Paul Isambert, Soha Safaï, Karim Shoul. Un merci spécial à Anne-Gaëlle Toutain, qui, sans me connaître, a répondu à mon appel à participation et a accepté de se déplacer plusieurs fois à l'ILPGA.

Merci à Leïla Boutora pour la collaboration sur la Langue des Signes Française. Ce projet m'a beaucoup apporté et a ravivé mon intérêt pour la linguistique.

Merci à tous les amis à Paris 3 et à Paris 7 pour leur amitié et pour tout ce qu'ils m'ont inspiré. Mes 'bagarres' nocturnes avec Mballo sur les *a priori* de la psycholinguistique m'ont donné plein d'idées. Ismaïl Benali, Paul Isambert, Audrey Acher, Jean-Marc Beltzung, David Ravet, votre aide ne passe surtout pas inaperçue.

Encore un merci au Laboratoire de Phonétique et Phonologie pour m'avoir octroyé une allocation de recherche. Merci aussi au Laboratoire de Phonétique de Paris

7 (ARP) pour m'avoir permis de travailler dans des conditions très agréables. Pouvoir rester jusqu'à 0h30 et parfois toute la nuit au labo, les lumières éteintes et absorbé par le travail, c'est un réel plaisir.

Le travail de correction de Christelle Exare a été précieux et méticuleux. Toute faute demeure la mienne.

Un grand merci aux membres du jury d'avoir consacré un peu de temps à ma recherche.

Abréviations

DC	Détection Correcte
FA	Fausse Alarme
IIS	Intervalle Interstimulus
MCT	Mémoire à court terme
MLT	Mémoire à long terme
n.a.	Note de l'auteur

Conventions

- Utilisation exclusive de guillemets anglais (quotes)
- Guillemets simples pour les termes, guillemets doubles pour les citations et les consignes des différentes tâches
- Notation américaine des nombres (point au lieu de virgule pour les décimaux)
- Citations anglophones non traduites
- Quand les statistiques complètes ne sont pas données dans le texte, le lecteur peut les retrouver dans le DVD
- Les références bibliographiques concernant des présentations (orales ou affichées) apparaissent dans la bibliographie sous la forme suivante : ‘Auteur’ (‘Année’). ‘Titre’. Dans ‘*Nom du colloque*’, ‘*Lieu*’.
- Le lecteur trouvera dans le DVD joint à ce document les scripts pour les expériences, les stimuli utilisés, les données brutes, les statistiques complètes et le corpus audio du chapitre 7. Les logiciels nécessaires à l’exécution des scripts sont également référencés.

Sommaire

Table des figures	11
Liste des tableaux	15
Introduction	18
I Partie théorique	23
1 Discriminabilité et modules mémoriels	25
1.1 Fechner et la psychophysique	25
1.2 Les modules de la mémoire	27
1.3 Catégorisation et modèles de discriminabilité	28
1.4 Théorie de la détection du signal	36
1.5 Conclusion	40
2 Asymétries et hypothèses	41
2.1 Asymétries et stimuli non-linguistiques	41
2.2 Asymétries et stimuli linguistiques	42
2.3 Modèles génériques	54
2.4 Recadrage des théories et hypothèses formulées	57
II Partie expérimentale	61
3 Asymétries et expérience linguistique	63
3.1 Expérience 1 : Catégorisation	65
3.2 Expérience 2 : Discrimination	70
3.3 Expérience 3 : Catégorisation AXB	78
3.4 Conclusions générales	84
4 Asymétries et configurations manuelles	85
4.1 Expérience 1 : Catégorisation	86
4.2 Expérience 2 : Discrimination	90
4.3 Conclusions générales	94

5	L'hypothèse de la neutralisation	95
5.1	Expérience 1 : Discrimination	95
5.2	Conclusions générales	105
6	La variabilité inter-sujet	107
6.1	Sujets	108
6.2	Stimuli	108
6.3	Procédure	109
6.4	Résultats	111
6.5	Conclusions générales	130
7	Asymétries et cibles acoustiques	145
7.1	Expérience 1 : Production de logatomes CVCV	146
7.2	Expérience 2 : Production de triptongues $V_1V_2V_1$	153
7.3	Conclusions générales	154
	Discussion générale	156
	Perspectives	159
	Bibliographie	163
	Table des matières	175

Table des figures

1.1	La psychophysique telle que conçue par Fechner (Ehrenstein & Ehrenstein, 1999). Contrairement à la psychophysique externe qui emploie les outils des sciences dures pour manipuler la nature du stimulus, la psychophysique interne — limitée par la technologie de l'époque — demeure encore au niveau conceptuel et emprunte ses méthodes à la psychophysique externe pour raisonner sur la nature du traitement sensoriel et neurobiologique du stimulus.	26
1.2	La psychophysique moderne (Ehrenstein & Ehrenstein, 1999). Les outils neurophysiologiques permettent la mesure objective de l'activité neuronale et la définition de la relation entre les données psychophysiques et les corrélats neuronaux de la perception.	26
1.3	Un modèle du traitement de l'information contenant les composantes les plus importantes (adapté de Haber & Hershenson, 1980).	27
1.4	Exemple théorique de la perception catégorielle. La ligne continue représente les scores d'identification pour les deux catégories examinées dans la tâche. La ligne pointillée illustre les scores de discrimination pour $n-1$ et $n+1$. La fonction de discrimination est parfaitement prédite par l'identification.	29
1.5	Le modèle de discriminabilité de Fujisaki (1979) pour un paradigme ABX. MCT : mémoire à court terme; MLT : mémoire à long terme.	34
1.6	Le modèle de discriminabilité à deux étapes de Kroll & Hershenson (1980).	36
1.7	Les quatre possibilités dans la théorie de la détection.	37
1.8	Les étapes exécutées pendant une tâche de détection.	37
1.9	Proportion des items ayant concédé le niveau d'excitation indiqué en abscisse.	38
1.10	Proportion des items ayant concédé le niveau d'excitation indiqué en abscisse et le critère de décision. DC : détection correcte; FA : fausse alarme; RC : rejet correct.	39
1.11	Distributions du bruit (à gauche) et du signal avec bruit (à droite) dans trois conditions différant dans l'emplacement du critère.	39
2.1	Le modèle de changement de qualité vocalique, adapté de Cowan & Morse (1986). Au moment temporel 1, la représentation mentale du stimulus est précise et étroite; au moment 2, la dégradation de la trace mémorielle du stimulus déplace le centre de la 'région de confiance', qui est élargie vers un point central dans l'espace.	45

2.2	Les 45 stimuli de Repp & Crowder (1990). Au centre de chaque croix se trouve le prototype de la catégorie respective.	46
2.3	Représentation graphique des asymétries retrouvées dans les résultats de Repp & Crowder (1990). Les flèches indiquent l'ordre dans lequel la discrimination a été plus facile. Flèches simples : effet d'ordre supérieur à 10% ; flèches doubles : effet supérieur à 20%.	48
2.4	Les oppositions phonémiques étudiées dans les travaux ci-indiqués et la direction de changement dans laquelle la discrimination a été plus facile. (a) : études d'enfants ; (b) études animales. Les valeurs formantiques sont indicatives.	50
2.5	Les 37 qualités vocaliques basiques (des segments oraux, non-longs, monophthongaux, à voix normale) retrouvées dans UPSID (Maddieson, 1991). Dans chaque ovale, le symbole API de la voyelle et le nombre de langues (sur 451) qui possèdent la voyelle en question. Adapté de Karypidis & Montagu (2009).	52
3.1	Les systèmes vocaliques des quatre langues étudiées.	64
3.2	Les configurations de l'appareil phonatoire proposées par VTCALCs pour [i] et [e] respectivement. À gauche : la cavité labiale.	66
3.3	Fonctions d'identification pour les quatre groupes. La ligne horizontale désigne le point de hasard.	69
3.4	Fonctions de discrimination pour les quatre groupes. Pro : ordre progressif ; Reg : ordre régressif.	73
3.5	Les effets d'ordre en % pour chaque paire avec une erreur standard. Les astérisques indiquent une différence significative telle qu'indiquée avec un test de rang de Wilcoxon. * : $\alpha = 0.05$, ** : $\alpha = 0.01$. À gauche : un pas ; à droite : deux pas. Du haut au bas : portugais brésilien, italien, français et espagnol.	75
3.6	Comparaison des scores de discrimination pour les deux directions de changement vocalique. * : $\alpha = 0.01$ avec un test Wilcoxon.	76
3.7	Représentation spectrale des dix stimuli.	77
3.8	Représentation graphique des valeurs de F2' estimées en utilisant différentes formules de conversion Hertz-Bark.	78
3.9	Représentation sur le plan F1/F2 des trente-trois stimuli prototypes (Schwartz <i>et al.</i> , 1997).	81
3.10	Pourcentage auquel les 19 stimuli ont été jugés plus similaires à la voyelle antérieure de leur série. Numérateur : ordre AXB, dénominateur : ordre BXA avec A= voyelle antérieure et B= voyelle postérieure. En italiques : les différences significatives à $\alpha = .01$ avec un test Wilcoxon.	82
3.11	Les deux triades dans lesquelles le stimulus [y] peut être présenté. Les flèches démarquent la position de la voyelle antérieure standard à laquelle [y] est comparée et les chiffres correspondent aux scores pour [y], comme il a été jugé similaire à /i/ par les grikoïphes.	83
4.1	Les trois configurations manuelles localisées sur les extrémités des continua U-V et V-X.	87

4.2	Proportions des réponses ‘V’ pour le continuum U–V en fonction du contexte. $n, n+1, n-1$: position relative sur le continuum du stimulus précédant la voyelle à catégoriser. $N = 16$	88
4.3	Proportions des réponses ‘X’ pour le continuum V–X en fonction du contexte. $n, n+1, n-1$: position relative sur le continuum du stimulus précédant la voyelle à catégoriser. $N = 16$	89
4.4	Scores de discrimination pour les deux directions sur le continuum U–V à un pas. * : $\alpha = 0.05$ avec un test Wilcoxon ($N = 16$).	92
4.5	Scores de discrimination pour les deux directions sur le continuum U–V à deux pas. * : $\alpha = 0.05$ avec un test Wilcoxon ($N = 16$).	92
4.6	Scores de discrimination pour les deux directions sur le continuum V–X à un pas. * : $\alpha = 0.05$ avec un test Wilcoxon ($N = 16$).	93
4.7	Scores de discrimination pour les deux directions sur le continuum V–X à deux pas. * : $\alpha = 0.05$ avec un test Wilcoxon ($N = 16$).	93
5.1	Les 50 stimuli synthétisés sur le plan F1/F2 (en Bark).	98
5.2	Les effets d’ordre attestés dans nos données. Les flèches marquent la direction dans laquelle la discrimination a été plus facile. Flèche simple : effet d’ordre supérieur à 10% ; flèche double : effet supérieur à 20%. . . .	102
5.3	Les effets d’ordre attestés dans les données de Repp & Crowder (1990). Les flèches marquent la direction dans laquelle la discrimination a été plus facile. Flèche simple : effet d’ordre supérieur à 10% ; flèche double : effet supérieur à 20%.	103
5.4	Représentation vectorielle des effets d’ordre retrouvés dans nos données.	104
5.5	Représentation vectorielle des effets d’ordre retrouvés dans Repp & Crowder (1990).	105
6.1	Les résultats d’identification et de discrimination pour les tâches iAXi-iAXe et dAXi-dAXe. $n = 10$	112
6.2	Les effets d’ordre pour les tâches dAX1r et dAX2r. $n = 7$	118
6.3	Les effets d’ordre pour les tâches dAX2r et dAX2r_fil. $n = 7$	121
6.4	Les effets d’ordre pour les tâches dAX2r et dAX2f. $n = 7$	125
6.5	Les effets d’ordre pour les tâches dAX2r_250, dAX2r_500 et dAX2r_1000. $n = 5$	128
6.6	Les effets d’ordre pour les tâches dAX1r et fli. $n = 7$	131
6.7	Les effets d’ordre pour les tâches dAX2r et flii. $n = 7$	134
6.8	Les effets d’ordre pour les tâches dAX1r et f2i. $n = 7$	136
6.9	Les effets d’ordre pour les tâches dAX2r et f2ii. $n = 7$	138
6.10	Les effets d’ordre pour les tâches dAX1r et f3i. $n = 7$	140
6.11	Les effets d’ordre pour les tâches dAX2r et f3ii. $n = 7$	142

Liste des tableaux

2.1	Matrice de confusion des voyelles anglaises présentées avec un ratio signal/ bruit à -14 dB (Benkí, 2003).	43
2.2	Matrice de confusion des voyelles anglaises présentées avec un ratio signal/ bruit à -12 dB (Miller & Nicely, 1955).	44
2.3	Études en perception enfantine et animale ayant indiqué des effets d'ordre. Les flèches indiquent la direction dans laquelle la discrimination a été plus facile (Polka & Bohn, 2003). Entre parenthèses, les oppositions pour lesquelles l'effet d'ordre n'a pas été significatif.	51
2.4	Les traits lexicaux et de surface des voyelles /i e y/ de l'allemand et du turc dans le cadre du modèle FUL (Lipski <i>et al.</i> , 2007).	54
2.5	Les hypothèses formulées dans le cadre des différents modèles de discriminabilité.	59
3.1	Nombre (N) d'auditeurs par langue, étendue d'âge, moyenne d'âge (M) et déviation standard (ds).	65
3.2	Les paramètres articulatoires des dix stimuli synthétisés. Les chiffres correspondent à des valeurs standardisées des paramètres articulatoires. HM : Hauteur de la mâchoire; PL : Position de la langue; FL : Forme de la langue; PA : Position de l'apex; HLe : Hauteur des lèvres; PLe : Protrusion des lèvres; HLa : Hauteur du larynx	66
3.3	Valeurs de F1 et F2 (en Hertz) et différence moyenne entre des stimuli consécutifs (dernière ligne) pour les trois continua. Les paramètres modifiés à chaque fois ont été : la hauteur de la mâchoire (colonnes 2 et 3), la position de la langue (colonnes 4 et 5), ou les deux (colonnes 6 et 7). . .	67
3.4	Valeurs formantiques (F1–F4) en Hertz (colonnes 2–5) et en Bark (colonnes 6–9) des dix stimuli synthétisés.	67
3.5	Les étapes du test de catégorisation.	68
3.6	Tableau des moyennes de la frontière sur le continuum pour les quatre groupes.	69
3.7	Résultats de PLSD ($\alpha = 0.01$) pour comparaison de frontière entre les quatre groupes.	70
3.8	Les quatre groupes de paires de stimuli et leurs caractéristiques : ordre, différence en pas sur le continuum, étiquette (la façon dont le groupe est référencé dans le texte) et nombre de paires dans chaque groupe.	71
3.9	Les étapes du test de discrimination.	71

3.10	Résultats d'ANOVA (mesures répétées sur deux facteurs : Ordre et Paire) pour la condition d'un pas ; ddl= degré de liberté.	72
3.11	Résultats d'ANOVA (mesures répétées sur deux facteurs : Ordre et Paire) pour la condition de deux pas ; ddl= degré de liberté.	72
3.12	Résultats des tests χ^2 2x2 pour les quatre langues : nombre de réponses 'différent', valeur χ^2 et valeur p	74
3.13	Les valeurs de F2' estimées pour les dix stimuli en utilisant différentes formules de conversion Hertz-Bark. Les en-têtes des colonnes font référence aux formules 3.1-3.5 présentées plus haut, correspondant à celles proposées par Fourcin <i>et al.</i> (1977), Schroeder <i>et al.</i> (1979), Traunmüller (1990), Tjomov (1971) et Zwicker & Terhardt (1980) respectivement. . .	77
3.14	Groupe linguistique, étiquette sous laquelle le groupe apparaît dans le texte, nombre (N) d'auditeurs par langue, étendue d'âge, moyenne d'âge (M) et déviation standard (ds).	78
3.15	Les valeurs des quatre formants pour les trente-trois stimuli synthétisés (Schwartz <i>et al.</i> , 1997).	80
3.16	Résultats des tests χ^2 pour les quatre langues : valeur χ^2 de Yates corrigée pour continuité, degrés de liberté, taille de l'échantillon et valeur p	83
4.1	Moyenne d'âge (M) des seize participants de chaque groupe, déviation standard (ds) et étendue d'âge.	86
4.2	Les comparaisons de contexte par paire pour U-V, significatives à $\alpha = .05$ avec un test de rang de Wilcoxon ($N = 16$).	89
4.3	Les comparaisons de contexte par paire pour V-X, significatives à $\alpha = .05$ avec un test de rang de Wilcoxon ($N = 16$).	90
5.1	Moyenne d'âge, déviation standard (ds) et étendue d'âge des 23 auditeurs de chaque séance.	95
5.2	Les valeurs formantiques en Hertz des 50 stimuli synthétisés. P : le prototype de chaque catégorie ; S1-S4 : les quatre satellites d'un prototype.	96
5.3	Les étapes du test de discrimination.	99
5.4	Les scores de discrimination pour les paires identiques (premier bloc de valeurs) et non identiques dans l'ordre P-S _i (deuxième bloc) et S _i -P pour les dix catégories. Le dernier bloc de valeurs représente la différence entre les deux ordres.	100
6.1	Identifiant des auditeurs, langue maternelle et âge.	108
6.2	Résultats des tests χ^2 ($ddl = 1, n = 200$) pour les tests iAXi-iAXe (colonnes 2 et 3) et dAXi-dAXe (colonnes 4 et 5). * : le chiffre dans au-moins une cellule de l'analyse était inférieur à 5.	115
6.3	Résultats des tests χ^2 pour les tâches iAX1r ($ddl = 2, n = 196$) et iAX2r ($ddl = 2, n = 182$). * : le contenu d'au-moins une cellule était inférieur à 5.	116
6.4	Résultats des tests χ^2 2x2 sur les tâches dAX1r et dAX2r : nombre de détections correctes, valeur χ^2 et valeur p	117
6.5	Résultats des tests χ^2 2x2 sur les tâches dAX2r et dAX2r_fil : nombre de détections correctes, valeur χ^2 et valeur p	123

6.6	Résultats des tests χ^2 2x2 sur les tâches dAX2r et dAX2f : nombre de détections correctes, valeur χ^2 et valeur p	124
7.1	Identifiant des auditeurs, langue maternelle et âge.	146
7.2	Plafonds optimaux pour chaque voyelle prononcée par nos locuteurs. Colonnes 2-3 : résultats obtenus avec la méthodologie de Escudero <i>et al.</i> (2009). Colonnes 4-5 : résultats obtenus avec notre protocole adapté. . .	148
7.3	Les valeurs des bandes critiques et des largeurs de bande critiques en fonction de la fréquence (Zwicker & Terhardt, 1980).	148
7.4	Résultats des analyses multivariées (MANOVA) deux à deux pour les 19 locuteurs.	151
7.5	Résultats de l'analyse discriminante. Nous indiquons les 4 bandes critiques qui contribuent le plus à la fonction canonique.	152
7.6	Les vingt triptongues de notre corpus.	153
7.7	Valeurs p extraites d'une analyse <i>a posteriori</i> de Tukey.	154

Introduction

La perception de la parole est un système de processus extrêmement complexe. Il est, par conséquent, attendu que son étude abonde en théories et hypothèses. Cette incertitude est largement due à deux problèmes inhérents à la parole : l'absence de frontières abruptes entre les segments consécutifs — contrairement à ce qui se passe dans la parole écrite — et le manque d'invariance dans des indices acoustiques des phonèmes. Les indices dépendent du contexte segmental, du style de locution et du locuteur. Parmi les théories de la perception les plus connues, nous trouvons la théorie motrice (Liberman *et al.*, 1957), la théorie réaliste directe (Best, 1995; Fowler, 1983), le modèle 'fuzzy-logical' (Massaro, 1989) et la théorie exemplariste (Johnson, 1997; Kruschke, 1992).

Ce qui semble peu réaliste dans la plupart des modèles proposés est leur négligence des processus de traitement de l'information. Comment le signal acoustique se transforme en connaissance ou en information est souvent passé sous silence. Quel est le rôle de la mémoire à court terme et des traces mémorielles ? Quelle est l'interaction entre les segments adjacents ? Comment la mémoire gère-t-elle le flux d'informations ?

Nous partons du postulat que, pour qu'un modèle de la perception soit réaliste, il faut un cadre théorique solide et formalisation des différents processus mémoriels. Mais pour établir un tel modèle, il est nécessaire d'implémenter un certain nombre de paramètres, et plus particulièrement, de contraintes. Il faut donc établir une liste des contraintes imposées par les différents modules de la mémoire : degré et direction de la dégradation des traces mémorielles, masquage auditif, nombre maximal d'informations pouvant être traitées à la fois (Miller, 1956), et nature du traitement (sériel *vs.* parallèle).

Dans cette thèse, nous étudions un phénomène qui nous paraît être ancré dans les contraintes mémorielles ci-mentionnées, à savoir l'effet de l'ordre de la présentation. Ceci consiste à dire que l'ordre dans lequel deux stimuli sont présentés a un effet sur leur discriminabilité. Les diverses théories utilisent différents concepts et entités pour expliquer ce biais. Afin de mieux comprendre ces hypothèses et pouvoir préparer des expériences adéquates, nous avons besoin d'un modèle précis des modules de la mémoire.

Dans le chapitre 1, nous explicitons notre domaine d'investigation, qui est l'intersection de la psychoacoustique externe et de la psychologie cognitive. Nous étudions le rapport entre le stimulus d'input et le patron comportemental des sujets sans faire d'inférence quant à l'activation neuronale et comment celle-ci se transforme en information sur laquelle le sujet agit. Nous présumons l'exclusion d'un traitement

top-down, c'est-à-dire l'apport d'informations syntaxiques, morphologiques et sémantiques. Ensuite, nous présentons les différents niveaux de la mémoire, leur nature et leur fonction d'après le modèle de Haber & Hershenson (1980). Ce modèle nous permet de mieux cerner les théories sur les asymétries que nous présentons dans le chapitre 2. Nous donnons ensuite une description des mécanismes implémentés dans les tâches de discrimination. Cinq modèles sont présentés : perception catégorielle, prototypes, modèle du processus duel, trace-contexte et modèle à deux étapes. Nous finissons avec une courte introduction à la théorie de la détection.

Le chapitre 2 fait un bilan de la littérature sur les asymétries en perception. Nous découvrons que les formes géométriques, les couleurs, les tons modulés et les accords musicaux sont quelques-unes des catégories des stimuli perçus de façon asymétrique. Quant aux stimuli linguistiques, les matrices de confusion révèlent souvent que les erreurs de reconnaissance ou de catégorisation dévoilent un biais en faveur de certaines catégories phonémiques. Quatre hypothèses sur l'origine des effets d'ordre en perception vocalique sont présentées : l'hypothèse de la neutralisation, la théorie de la périphéricité, l'effet des aimants perceptifs et la théorie de la sous-spécification. L'illustration de deux modèles génériques qui tentent d'expliquer les asymétries suggère qu'une approche plus centrée sur la nature des voyelles doit être adoptée. L'adoption de ces modèles génériques exige un grand nombre d'*a priori* sur les indices acoustiques grâce auxquels le sujet perçoit les voyelles. Nous réanalysons les modèles de discriminabilité présentés dans le chapitre 1 dans le cadre des asymétries et formulons les hypothèses que ces modèles génèrent en termes d'effets d'ordre.

Dans le chapitre 3, nous examinons le rôle que l'expérience linguistique joue sur le mécanisme de prise de décision, en menant deux études. La première étude utilise un continuum [i]-[e] synthétisé avec un modèle articulatoire et fait appel à des auditeurs de quatre langues. Le but de cette étude est de voir si des asymétries sont attestées chez les sujets des quatre groupes. Si oui, ceci apportera des indications en faveur d'un biais universel. Si des asymétries sont retrouvées chez les sujets d'une partie des langues en question, ceci nous amènera au postulat que les effets d'ordre ne sont pas localisés au niveau de la mémoire précatégorielle et qu'ils sont déclenchés par un traitement auditif de bas niveau (mémoire à court terme) ou par l'interférence de la mémoire à long terme. La deuxième étude fait intervenir des locuteurs de quatre langues dont les systèmes vocaliques partagent la même structure. Des stimuli dispersés dans l'espace acoustique de façon symétrique sont utilisés dans une tâche de catégorisation AXB. Notre but est d'envisager la région de l'espace vocalique où la perception est asymétrique. Nous nous demanderons si 1. le poids relatif des voyelles antérieures ou postérieures est plus important, et si 2. les effets d'ordre sont retrouvés dans tout l'espace vocalique ou s'ils dépendent de la hauteur vocalique.

Dans le chapitre 4, nous envisageons de nouveau le facteur de l'expérience linguistique. Cette fois, nous utilisons des stimuli visuels. Nous avons considéré deux continua constitués d'unités de la dactylogogie de la Langue des Signes Française (LSF), U-V et V-X. Nous incluons deux groupes de sujets : des signeurs sourds, natifs de la LSF, et des non-signeurs entendants n'ayant pas d'expérience en LSF. Nous avons ainsi tenté d'étudier le rôle de l'expérience linguistique en éliminant l'apport des représentations linguistiques pour un des deux groupes. Cette méthode nous per-

met de mieux cerner le concept des points de référence cognitifs. Nous voulons donc comprendre si les deux groupes de sujets utiliseront des points d’ancrage propres aux continua utilisés ou si les signeurs utiliseront des catégories permanentes et les entendants des points de référence sensoriels. Deux types d’effets sont analysés : les effets de contexte en catégorisation et les effets d’ordre en discrimination.

Dans le chapitre 5, nous répliquons l’étude de Repp & Crowder (1990) sur l’hypothèse de la neutralisation selon laquelle les asymétries dans la perception des voyelles sont provoquées par la dégradation du premier stimulus dans une paire vers le centre de l’espace vocalique. Le protocole originel a introduit des inconvénients qui ont amené les auteurs à des résultats contradictoires. Nous adaptons le protocole pour le français de sorte qu’il nous permette de vérifier à la fois trois hypothèses : la dégradation vers [ə], la dégradation vers le centre de la dispersion des stimuli et la périphéricité.

Le chapitre 6 est consacré à la variabilité intra-sujet. Il s’agit d’une mini-étude de cas qui exploite quelques pistes sur les diverses sources d’asymétries. Nous menons une série de tests qui nous permettent d’explorer les effets de contexte et de contraste à travers le continuum [i]-[e]. Nous étudions l’intervalle interstimulus, la distance acoustique, les formants supérieurs, le paradigme expérimental, et comment ces facteurs peuvent influencer les effets d’ordre. Un autre aspect des stimuli que nous envisageons concerne les fréquences formantiques. Partant de l’hypothèse que les effets d’ordre peuvent être déclenchés par des fréquences localisées dans une région spécifique du spectre, nous avons construit trois continua contenant des tons purs dont les fréquences correspondent aux fréquences des trois premiers formants des stimuli du continuum originel.

Malgré les nombreuses études dans le domaine, les effets d’ordre restent toujours dissociés des stratégies de production de la parole. Il est possible que les asymétries en perception aient un effet sur le degré de coarticulation des segments. Notre idée est que quand la discriminabilité de deux segments est plus considérable dans un ordre spécifique de présentation, le déplacement des organes phonatoires sera moindre. Si, par exemple, la discrimination des voyelles /i/ et /e/ est plus facile dans l’ordre /e-/i/, le sujet a besoin d’un déplacement des organes moins important pour atteindre le seuil de discriminabilité. Dans le chapitre 7, nous étudions les asymétries en production pour la paire /i-/e/. De même, nous nous penchons sur le concept de la stabilité de la cible acoustique d’une voyelle en fonction de sa position dans l’espace vocalique afin d’établir un lien entre les mécanismes de la production et de la perception.

Nous sommes persuadé que l’étude du phénomène d’asymétrie est un outil précieux pour le déchiffrement de la structure interne des catégories. C’est un pas considérable vers la définition des processus cognitifs activés pendant le traitement d’information, et vers une représentation plus précise des espaces de décision. Notre travail montre qu’il n’est pas efficace d’aborder le phénomène des asymétries en supposant la robustesse des représentations mentales. Il est possible que le processus de similarité, à savoir la comparaison des stimuli, implémente différents points de référence en fonction de la tâche expérimentale. Il est à noter aussi que le rôle de l’expérience linguistique est un argument en faveur d’un traitement *top-down*. La saillance perceptive n’est donc pas le seul facteur déclenchant des asymétries.

Première partie
Partie théorique

Chapitre 1

Discriminabilité et modules mémoriels

"The task did not at all originally present itself as one of finding a unit of mental measurement; but rather as one of searching for a functional relationship between the physical and the psychical that would accurately express their general interdependence."

Gustav Fechner, *Elemente der Psychophysik*¹

1.1 Fechner et la psychophysique

La psychophysique, la première discipline du domaine plus large de la psychologie expérimentale, est née avec les travaux de Gustav Theodor Fechner et son ouvrage devenu classique, *Elemente der Psychophysik* (Fechner, 1860). Rejetant la distinction dualiste cartésienne entre esprit et corps, et persuadé que ces derniers ne constituent que deux aspects du même existant, Fechner soutient que toute nature peut être envisagée aussi bien du côté physique que du côté psychique (Wozniak, 1999).

Suite donc à une longue série d'expériences traitant de la perception des poids, de la luminance, et des distances visuelles et tactiles, Fechner conçoit deux 'visages' de la psychophysique, chacun décrit dans un volume séparé des *Elemente* en 1860.

Le premier volume aborde la psychophysique externe (angl. 'outer psychophysics'), à savoir la relation entre l'aspect physique d'un stimulus et la sensation que celui-ci engendre chez l'individu. Cette thèse est consacrée à l'étude de cette relation. Le deuxième volume des *Elemente* traite la psychophysique interne. Les figures 1.1 et 1.2 présentent une description de la psychophysique telle que perçue par Fechner et de la psychophysique moderne. Le but de Fechner est de fournir une définition de la relation fonctionnelle entre ces deux niveaux. Néanmoins, alors que l'aspect physique de la nature peut être abordé à travers les sciences 'objectives' telles que les mathématiques, la physique et la chimie, son caractère psychique comprend des procédures internes, subjectives, qui ne sont pas tangibles et mesurables au premier abord.

1. Traduit en anglais par Wozniak (1999).

FIGURE 1.1: La psychophysique telle que conçue par Fechner (Ehrenstein & Ehrenstein, 1999). Contrairement à la psychophysique externe qui emploie les outils des sciences dures pour manipuler la nature du stimulus, la psychophysique interne — limitée par la technologie de l'époque — demeure encore au niveau conceptuel et emprunte ses méthodes à la psychophysique externe pour raisonner sur la nature du traitement sensoriel et neurobiologique du stimulus.

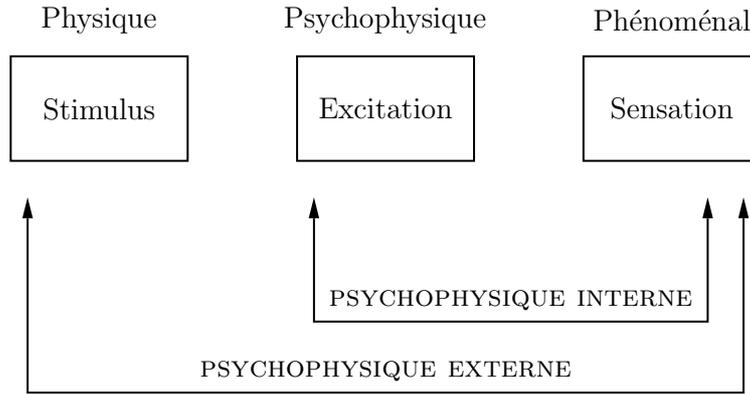
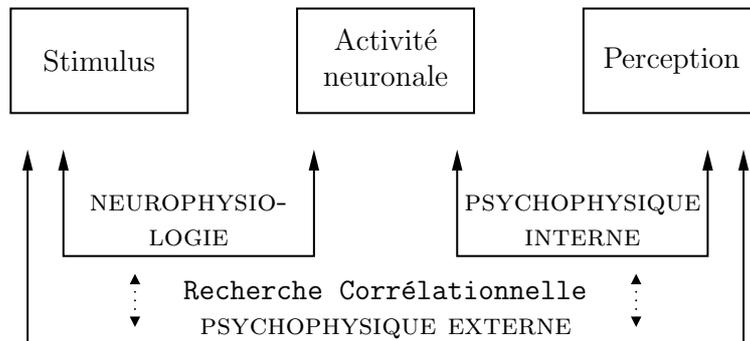
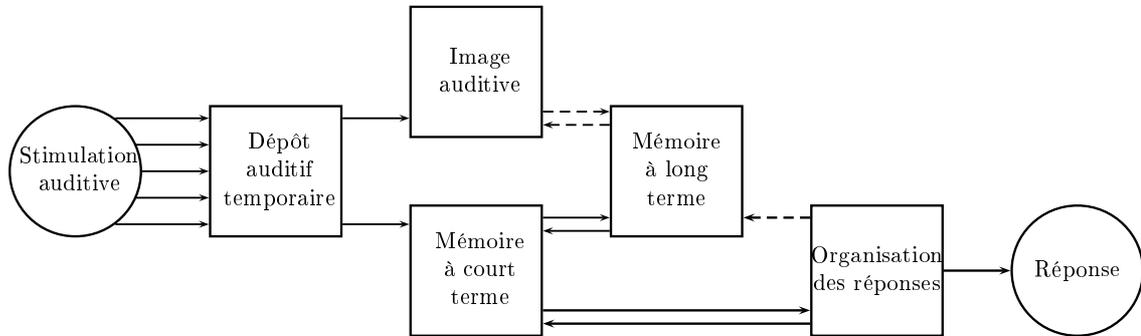


FIGURE 1.2: La psychophysique moderne (Ehrenstein & Ehrenstein, 1999). Les outils neurophysiologiques permettent la mesure objective de l'activité neuronale et la définition de la relation entre les données psychophysiques et les corrélats neuronaux de la perception.



Dans ses *Elemente*, Fechner introduit la notion du ‘time-order erreur’ (nous proposons de traduire ce terme par ‘l’erreur de l’ordre temporel’). Il suggère que quand deux stimuli sont comparés, l’un des deux est souvent sous- ou surestimé. Ce phénomène est largement étudié depuis les années 1930 mais ses mécanismes restent toujours obscurs. Nous avons décidé d’étudier ce phénomène tout en restant dans le cadre de la psychophysique externe. Néanmoins, nous avons opté pour une approche combinant les paradigmes expérimentaux de la psychophysique classique avec les outils de la psychologie expérimentale. Contrairement au grand corps des travaux psychophysiques, nous abordons les processus impliqués dans les tâches de discrimination et de catégorisation en les localisant au niveau des modules de la mémoire et non pas à travers un modèle computationnel tel que la théorie de l’information de Shannon & Weaver (1949). Nous avons, par conséquent, besoin d’un modèle explicite

FIGURE 1.3: Un modèle du traitement de l'information contenant les composantes les plus importantes (adapté de Haber & Hershenson, 1980).



de la mémoire et de ses modules.

1.2 Les modules de la mémoire

Les travaux dédiés à la modélisation de la mémoire et aux processus cognitifs sont abondants et sont basés sur les données de la cognition visuelle (Atkinson & Shiffrin, 1968; Broadbent, 1958; Norman, 1970; Sperling, 1967, 1963). Néanmoins, le traitement des informations auditives et visuelles semble très similaire. Nous présentons dans cette section le modèle proposé par Haber & Hershenson (1980, p.298) et illustré dans la figure 1.3.

L'énergie provenant du signal acoustique arrive aux nerfs auditifs — nous ne décrivons pas ici les premières étapes du système auditif périphérique. Cette information est immédiatement stockée dans le dépôt auditif temporaire (angl. 'pre-categorical acoustic storage'), aussi connu en tant que mémoire 'sensorielle', 'échoïque' ou 'précategorielle'. Il s'agit de la partie du mécanisme mémoriel la plus primitive, celle qui reçoit l'information brute, sous une forme analogique. Si l'information est nécessaire pour l'exécution d'une tâche, elle est transférée dans la mémoire à court terme (MCT). En même temps, le sujet en déduit une représentation mentale (image auditive). Si l'information ne sert pas au sujet, elle est remplacée ou masquée par l'input suivant. La dégradation est imminente et commence environ 250 ms après la présentation du signal.

À partir du début du stimulus, le sujet en construit une image mentale qui est possiblement assujettie aux règles d'organisation perceptives. Ces règles sont localisées dans la mémoire à long terme (MLT), ce qui fait que les deux étapes sont liées. Puisque des stimulations auditives continuent à arriver à l'oreille, l'image mentale concerne la scène auditive entière. Les deux étapes, le dépôt temporaire et l'image mentale, sont difficiles à distinguer et à décrire. Pour cela, une série de travaux a été consacrée à leur nature et à leur fonction (Crowder & Morton, 1969; Massaro, 1972, 1970; Morton, 1969, 1970, entre autre).

Les informations stockées dans la MCT ne sont retenues que pour une durée limitée, celle-ci ne dépassant en général pas trente à soixante secondes selon le modèle

adopté. Ces informations peuvent être : a) des items récemment récupérés de la MLT ; b) le résultat d'un processus mental récent (probablement lié à la mémoire de travail) ; c) le produit d'encodage de la mémoire sensorielle. Cette capacité restreinte peut être comparée à celle de la mémoire à long terme dans laquelle un nombre virtuellement illimité d'informations peut être emmagasiné pendant des décennies. Il est possible de prolonger la viabilité des informations stockées dans la MCT au moyen d'une boucle articulatoire, à savoir une répétition consciente ou inconsciente de ces informations.

La mémoire à long terme et ses composantes très complexes ne sont pas discutées ici. Les éléments de la MLT — qu'il s'agisse de lettres, d'images, de phonèmes ou de sons — sont probablement structurés par une représentation sémantique qui lie ces éléments entre eux.

La dernière étape du traitement concerne l'organisation de la réponse. Il n'est pas possible d'extraire des informations de façon directe à partir des différents modules. Nous ne pouvons pas, par exemple, voir directement le contenu de la mémoire précatégorielle d'un individu. Le sujet doit exécuter certaines commandes motrices qui contrôlent les organes qui donneront la réponse qui peut être écrite ou verbale.

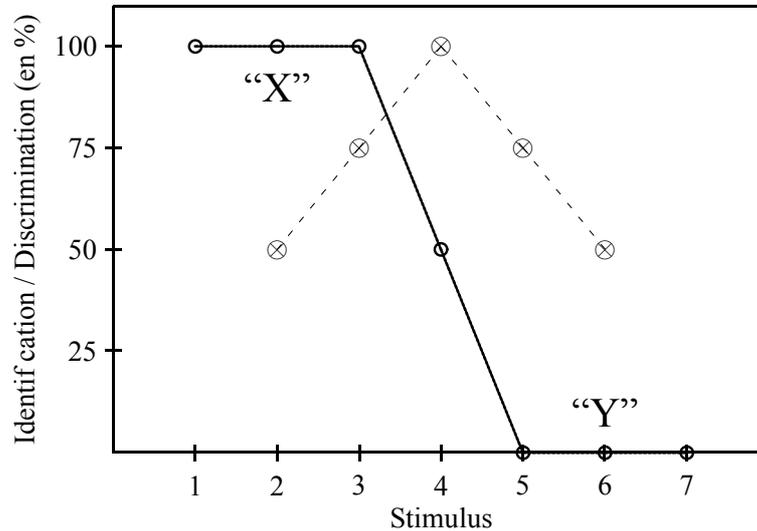
Dans la section suivante, nous présentons quelques modèles de traitement offrant une explication des mécanismes employés dans les tâches 'même-différent'.

1.3 Catégorisation et modèles de discriminabilité

Il est souvent soutenu que la perception de la parole est fondée sur le concept de catégorisation. Il s'agit tout simplement d'étiqueter un stimulus, de le classer dans une catégorie linguistique (phonème, morphème, etc.). Le concept de 'catégorie' évolue ces dernières années. Les premiers travaux (voir Smith & Medin, 1981, pour une revue de ces concepts) considèrent que les frontières entre catégories sont très nettes : soit un objet contient les attributs nécessaires d'une catégorie, soit non. Cette notion d'appartenance discrète a été abandonnée par les psychologues qui se sont dirigés vers des modèles flous (Zadeh, 1965), dans lesquels aucune caractéristique n'est nécessaire et où les frontières sont incertaines. Prenons le champ sémantique de la catégorie 'chien'. Si nous menons une enquête auprès de sujets et que nous leur demandons d'indiquer les attributs de cet animal, les réponses les plus récurrentes seront : 'quatre pattes', 'queue', 'aboie', etc. Cette hiérarchie de traits, pourtant, ne nous empêchera pas de classer un chien sans queue dans cette catégorie. Comme nous le verrons plus bas, que les frontières catégorielles soient nettes ou floues, de nombreux travaux ont montré que les catégories sont structurées (Rosch, 1978; Tversky, 1977).

Il est à noter ici que notre objet n'est pas les mécanismes généraux de la parole mais seulement la catégorisation de stimuli (présentés en isolé ou en contexte) et la (dis)similarité entre les stimuli. Notre décision de travailler sur cet aspect de la perception est fondée sur notre supposition qu'étant donné le caractère multidimensionnel et, par là, difficilement cernable des segments, il est impossible d'aborder les mécanismes de la perception de la parole de façon directe. Nous croyons alors qu'il est d'une importance primordiale de commencer à dresser un plan des contraintes mémorielles afin d'améliorer les modèles préexistants de perception. De même, nous

FIGURE 1.4: Exemple théorique de la perception catégorielle. La ligne continue représente les scores d'identification pour les deux catégories examinées dans la tâche. La ligne pointillée illustre les scores de discrimination pour $n-1$ et $n+1$. La fonction de discrimination est parfaitement prédite par l'identification.



ne prétendons pas que les mécanismes de discrimination correspondent nécessairement aux mécanismes de perception de phonèmes. Nous sommes donc d'accord avec Schouten & van Hossen (1992) qui, en expliquant leur modèle de discriminabilité, explicitent que le modèle de processus duel et celui de trace-contexte sont des théories d'identification et de discrimination plutôt que des théories de la perception des phonèmes (p.1841).

Les différences principales entre les modèles ci-après se déclinent selon l'importance qu'ils attribuent aux représentations mentales, la nature de ces représentations et les modules mémoriels qu'ils implémentent. Nous allons revisiter les modèles directement liés à nos hypothèses.

1.3.1 Perception catégorielle

Lieberman *et al.* (1957) mènent une tâche de discrimination ABX en utilisant un continuum synthétique /ba/-/da/-/ga/. Ils postulent ensuite qu'il est plus facile de discriminer des stimuli appartenant à différents percepts que des stimuli assignés à la même catégorie phonémique. Ils suggèrent que les auditeurs ne prêtent pas attention aux détails fins du signal et qu'ils font abstraction de l'information perçue. Ce phénomène où la discrimination intra-catégorielle est minimale et la performance inter-catégorielle est maximale est appelé la 'perception catégorielle (PC)'.

La figure 1.4 présente un exemple idéal de perception catégorielle d'un continuum. La transition d'une catégorie à l'autre est plutôt abrupte (ligne continue), en passant de la catégorie 'X' (stimuli no. 1–3) à la catégorie 'Y' (stimuli no. 5–7). Le stimulus 4 est situé à la frontière catégorielle. La discrimination est parfaitement prédite par la tâche d'identification.

Fry *et al.* (1962) testent la validité de ce modèle sur la perception des voyelles et observent que les auditeurs sont capables de discriminer entre différentes réalisations de la même catégorie. Ces résultats mènent à la naissance de la ‘théorie motrice’ (Liberman *et al.*, 1963; Studdert-Kennedy *et al.*, 1970; Liberman & Mattingly, 1985). Les auteurs suggèrent que les consonnes sont perçues de façon catégorielle à cause de leur nature : la transition de /ba/ à /da/, par exemple, est discontinue parce qu’un articulatoire différent est activé. Dans le continuum /i/-/e/, en revanche, un déplacement graduel du même articulatoire se produit. Les auditeurs récréent inconsciemment les gestes articulatoires produits par le locuteur. La désactivation d’un articulatoire et l’activation d’un autre pour une consonne est plus saillant perceptivement.

Alors que la perception catégorielle a été initialement considérée comme spécifique aux phonèmes, une série de travaux soulignent que d’autres catégories sémantiques sont perçues de façon catégorielle : les couleurs (Bornstein & Korda, 1984; Ozgen & Davies, 2002; Roberson & Davidoff, 2000), les accords musicaux (Howard *et al.*, 1992; Locke & Kellar, 1973), les expressions faciales (Beale & Keil, 1995; Etcoff & Magee, 1992), les tons lexicaux (Chan *et al.*, 1975; Hallé *et al.*, 2004) et les configurations manuelles en Langue des Signes Américaine (Baker *et al.*, 2005; Emmorey *et al.*, 2003)². De même, des travaux sur la cognition animale³ démontrent que la PC est un principe relevant du traitement d’information auditive qui influence la perception des sons (May *et al.*, 1989, p.837).

Il est difficile de décider si une classe d’objets est perçue de façon catégorielle et des critères stricts doivent être définis. Les quatre critères (Studdert-Kennedy *et al.*, 1970) les plus récurrents dans la littérature sont les suivants :

- des catégories distinctes séparées par une frontière relativement abrupte ;
- performance autour du seuil de hasard pour les paires intra-catégorielles ;
- un pic de discrimination à la frontière catégorielle ;
- une relation étroite entre la fonction de discrimination et le patron de catégorisation.

Schouten *et al.* (2003) abordent l’influence de la qualité des stimuli et de la tâche expérimentale sur le degré de PC. En citant un travail antérieur (Schouten & van Hessen, 1992), ils suggèrent que le réalisme des stimuli joue un rôle sur le degré de PC. Pourtant, la différence entre les deux groupes d’auditeurs (naïfs *vs.* sujets entraînés en phonétique) affaiblit l’importance du modèle de PC pour l’étude de la communication. Quand un protocole considéré libre de biais (dans le sens de la détection du signal) est utilisé, la relation entre catégorisation et discrimination est largement réduite.

Pisoni (1976) teste des sujets anglophones sur une opposition de voisement à trois niveaux : non-voisé aspiré, non-voisé et voisé. Le continuum explore l’indice de VOT (Voice Onset Time). Les patrons de réponse sont similaires à ceux obtenus dans le cadre de la PC bien que seulement deux niveaux de l’opposition soient

2. Pour une critique de ces deux travaux et pour des résultats en faveur d’une perception non-catégorielle en Langue des Signes Française, voir Boutora & Karypidis (2009a,b)

3. Kluender *et al.* (1987) sur les cailles japonaises, Burdick & Miller (1975); Kuhl & Miller (1978) sur les chinchillas, Kuhl & Padden (1982) sur les macaques ; voir Ehret (1987) pour une revue de la littérature sur les études animales

phonologiquement contrastifs en anglais. Les auteurs concluent qu'il est possible de modifier les mécanismes de perception des sujets dans des conditions de laboratoire.

Le modèle de PC ne fait aucune supposition quant à la structure interne des catégories et tous les exemplaires semblent être équivalents. En même temps, la nature des représentations par rapport auxquelles les stimuli sont évalués n'est pas explicite. L'explication évidente serait qu'après le dépôt temporaire, le stimulus arrive à la MCT. Il est décomposé en éléments et puis le résultat de cette analyse est utilisé pour chercher la MLT, où des informations concernant les propriétés physiques des exemplaires de chaque catégorie phonémique sont stockées. Une fois que les propriétés du stimulus en ligne sont comparées à celles des différents phonèmes et qu'un meilleur candidat est localisé, le stimulus reçoit l'étiquette de cette catégorie. Ce mécanisme concerne plutôt la version originale du modèle de PC.

Quelques années après l'article original, Liberman *et al.* (1963), Studdert-Kennedy *et al.* (1970) et Liberman & Mattingly (1985) formulent l'hypothèse selon laquelle l'interaction entre MCT et MLT n'implique pas les propriétés acoustiques du stimulus entrant mais plutôt des configurations de l'appareil phonatoire. Au moyen d'une inversion acoustique → articulatoire, la représentation mentale que le sujet se fait du stimulus entrant est comparée aux patrons emmagasinés dans la MLT. Nous rappelons ici que la théorie de PC n'utilise pas une approche basée sur la psychologie cognitive et n'explique pas les différentes étapes du décodage et encodage à travers les modules mémoriels. Les lignes précédentes sont alors notre interprétation du modèle en question.

1.3.2 Le rôle des prototypes

Rosch (1973), en suivant l'idée de Wertheimer (1912) et l'école de Gestalt, aborde la notion de 'prototype' et fait la distinction entre le prototype naturel et la tendance centrale d'une catégorie. Le prototype naturel contient les propriétés les plus récurrentes de la catégorie respective et les sujets le considèrent comme le plus typique. Les autres exemplaires s'organisent autour de ce prototype. La tendance centrale correspond à l'occurrence moyenne. Les deux exemplaires peuvent se chevaucher, en fonction de la nature de la catégorie en question. En résumant la littérature, Rosch (p.330) suggèrent que :

- il est plus facile d'apprendre une catégorie dans laquelle le prototype est situé au centre d'un ensemble de variations qu'une catégorie dans laquelle la distorsion d'un prototype est centrale et le prototype naturel constitue un membre périphérique
- les sujets ont tendance à apprendre le prototype en premier, qu'il possède une position centrale ou non dans l'ensemble des stimuli
- les sujets organisent un ensemble de variations autour du prototype naturel, qu'il soit ou non central à la distribution des stimuli utilisés

Afin de vérifier ses hypothèses, elle étudie les membres d'une population de l'Âge de Pierre qui ne possède pas de concept de teinte de couleur ni de forme géométrique. Les trois postulats sont confirmés par cette étude.

Rosch approfondit l'idée de Wertheimer selon laquelle, à l'intérieur de certains types de catégories sémantiques, il se trouve des sous-catégories qui servent de points

d’ancrage cognitif (Rosch, 1975). Menant une recherche sur la catégorisation des couleurs, des nombres et des lignes, elle suggère que les stimuli déviant légèrement de ces points d’ancrage sont plus facilement assimilés, et jugés plus proches de ces points. La distance psychologique autour des stimuli de référence n’est pas symétrique. Par exemple, la phrase ‘996 est essentiellement 1000’ est plus facilement acceptée par les sujets que la phrase ‘1000 est essentiellement 996’. De même, une ligne de 85° est presque verticale mais l’inverse n’est pas vrai. Ici, ‘1000’ et la ligne verticale sont les points de référence cognitifs auxquels ‘996’ et la ligne de 85° sont comparés. Par conséquent, ce procédé de ‘généralisation’ pendant laquelle le point de référence cognitif assimile le stimulus adjacent est activé quand le point d’ancrage est en seconde position dans une paire. La théorie de Rosch est soutenue par les patrons des réponses des participants pour les trois types de catégories étudiées.

Le concept des prototypes est intégré dans les modèles de reconnaissance de la parole et y joue un rôle primordial. Dans le modèle de traitement d’information auditive général de Oden & Massaro (1978), une détection des traits phonétiques est produite et le résultat de cette analyse est stocké dans le dépôt auditif préperceptif (DAP). Les traits acoustiques, qui sont censés être traités séparément, sont ensuite stockés dans le DAP et comparés à des unités perceptives emmagasinées dans la mémoire à long terme. Les auteurs suggèrent que les prototypes sont spécifiés en termes de traits acoustiques qui définissent l’information acoustique idéale du résultat de la détection des traits qui est stocké dans le DAP (p.174). Le modèle de Oden & Massaro (1978) revendique explicitement le caractère continu de la perception et propose que les différentes étapes du modèle consistent en des quantifications de la probabilité pour que certaines conditions — présence d’un trait acoustique, appartenance à une catégorie — soient vraies. Cette idée est inspirée par la théorie des ensembles flous de Zadeh (1965).

Le modèle de Samuel (1982) explique le phénomène d’adaptation sélective privilégiant le concept de ‘prototype’. Samuel localise les prototypes pour /ka/ et /ga/ sur un continuum synthétique de 20 stimuli. Basé sur ces résultats, il choisit trois adaptateurs pour chaque sujet : un stimulus proche du prototype, un deuxième sur la frontière phonémique et un troisième situé loin du prototype. Les résultats indiquent que le prototype est un meilleur adaptateur, générant un plus grand déplacement dans le patron de réponses.

1.3.3 Mode de perception unique pour la parole ?

Alors que Liberman (1970) postule que la parole implique un mode de perception unique au langage, de nombreux travaux apportent des preuves contre cette hypothèse. Lane (1965) est parmi les premiers à contredire la théorie motrice et l’hypothèse selon laquelle la parole est perçue au moyen de l’articulation. En comparant des résultats d’expériences utilisant des stimuli de parole et de non-parole, il soutient que les mêmes relations entre identification et discrimination sont retrouvées pour les deux types de stimuli.

Cutting *et al.* (1976) et Cutting (1978) utilisent des ondes en dents de scie (angl. *sawtooth waves*) et soutiennent que la perception catégorielle et l’adaptation sélective ne sont pas uniques aux stimuli linguistiques. De même, Pisoni & Tash (1975)

proposent une approche auditive aux dépens d'une approche abstraite qui utilise des unités linguistiques.

Dans sa discussion sur les différences censées exister entre la perception de la parole et celle de la non-parole, Schouten (1980) revisite six aspects de la perception :

- avantage de l'oreille droite
- perception catégorielle
- patrons d'interférences mutuelles, différentes pour la parole et la non-parole
- adaptation sélective
- avec des stimuli linguistiques, une moindre durée est nécessaire pour obtenir des jugements d'ordre corrects
- différence d'efficacité des indices acoustiques des consonnes dans les contextes CV et VC

Ces travaux ont comme objet de remettre en cause la distinction bien établie dans la littérature entre un niveau auditif et un autre phonémique. Ils ont alors comme objet de démontrer qu'une interaction entre la MCT et la MLT n'est pas nécessaire pendant une tâche de discrimination. Les sujets n'utilisent pas d'étiquettes catégorielles pour distinguer des stimuli mais simplement leurs traces auditives.

1.3.4 Le modèle du processus duel

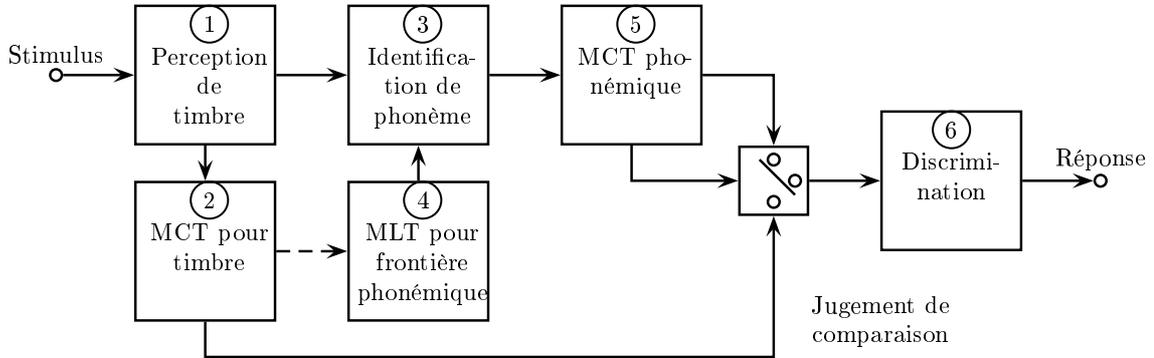
Fujisaki & Kawashima (1969, 1970, 1971); Fujisaki (1979) proposent un modèle de discriminabilité pour le paradigme ABX, qui est illustré dans la figure 1.5. Si la sensation d'un stimulus dépasse le seuil du bruit interne, un processus auditif établit une correspondance entre le continuum physique des stimuli et le continuum perceptif de timbre (case 1). Pour ceci, un ensemble de paramètres adéquats doit être adopté pour le traitement des stimuli. Le résultat de ce traitement est stocké sous forme analogique dans une mémoire à court terme pour le timbre (case 2) et en même temps procède à l'étape d'identification phonémique (case 3). Cette catégorisation est effectuée à l'aide d'informations qui concernent les frontières phonémiques et qui sont stockées dans une mémoire à long terme (case 4). Le résultat de la case 3 est un symbole phonémique et est stocké dans une mémoire à court terme séparée (case 5). Quand cette chaîne d'évènements est exécutée pour les trois stimuli, les résultats sont analysés dans une étape de prise de décision. Si les stimuli A et B sont jugés appartenir à des catégories phonémiques distinctes, seuls les résultats de la case 5 sont utilisés. Dans le cas contraire, le contenu de la mémoire à court terme pour le timbre (case 2) est réanalysé.

Il faut noter ici la différence entre les deux mémoires. Le contenu de la MCT pour timbre (case 2) est sous forme analogique, se dégrade très rapidement et se caractérise par des fluctuations durant les différentes procédures. La MCT phonémique contient des unités discrètes qui restent relativement stables pendant plusieurs secondes. En regardant la figure 1.3, nous constatons que la MCT pour timbre correspond au dépôt auditif temporaire et la MCT phonémique à la mémoire à court terme.

Fujisaki (1979) soutient que de l'incertitude peut naître de la génération de bruit pendant certaines étapes du processus :

- incertitude dans la correspondance entre le continuum des stimuli et le continuum perceptif. Le bruit gaussien aléatoire généré est similaire pour tous les

FIGURE 1.5: Le modèle de discriminabilité de Fujisaki (1979) pour un paradigme ABX. MCT : mémoire à court terme; MLT : mémoire à long terme.



stimuli

- incertitude dans la frontière phonémique. Des fluctuations dans la MLT peuvent être mises en place à cause de l'influence des stimuli précédents sur le stimulus en ligne. Ces variations des frontières phonémiques sont appelées des 'effets de contexte', qui consistent en un déplacement temporaire de la localisation des frontières
- incertitude dans la MCT pour timbre, due au bruit généré pendant les processus de rétention et de récupération dans la MCT pour timbre.

1.3.5 La théorie trace-contexte

Macmillan (1977) adapte le modèle de résolution de Durlach & Braida (1969) de l'intensité pour la perception des sons. Après la publication de Braida *et al.* (1984), ils ont implémenté le concept de 'point d'ancrage perceptif' (Macmillan *et al.*, 1987, 1988) dans leur modèle trace-contexte (MTC).

Le MTC comporte des sources de fluctuation qui empêchent le traitement d'un stimulus. La variance sensorielle concerne la transduction neuronale et correspond au bruit interne (voir la section suivante sur la théorie de la détection). La variance contextuelle correspond à la MLT pour frontière phonémique du modèle duel et engendre de l'incertitude chez le sujet. Plus la dispersion des ensembles des stimuli utilisés dans une tâche est large et l'incertitude est élevée, plus la résolution est basse. La variance de trace limite la capacité de l'auditeur à comparer les stimuli et correspond au bruit généré dans la MCT pour timbre pendant les processus de rétention et de récupération dans la MCT pour timbre. La variance de trace est positivement corrélée à la durée de l'Intervalle Interstimulus. Plus l'IIS est long, plus la trace mémorielle des stimuli se dégrade et plus il est difficile à l'auditeur d'utiliser les détails plus fins du signal.

La différence principale entre le modèle duel et la théorie trace-contexte repose sur la nature des catégories mémorielles. Le modèle duel suppose que l'interaction entre la MCT et la MLT ne concerne que la récupération d'informations sur les frontières phonémiques. Le point d'ancrage perceptif, en revanche, est un concept plus souple et sa nature peut varier en fonction de la tâche et des stimuli utilisés.

Braida *et al.* (1984) soutiennent que les points d’ancrage peuvent être dérivés :

- de standards explicitement présentés, comme dans les tâches d’adaptation sélective
- de régions quantiques du système sensoriel, à savoir de frontières naturelles comme le seuil de détection, le seuil de confort, etc.
- de transitions abruptes dans les valeurs des paramètres décrivant l’ensemble des stimuli
- de transitions abruptes dans les probabilités de présentation des stimuli.

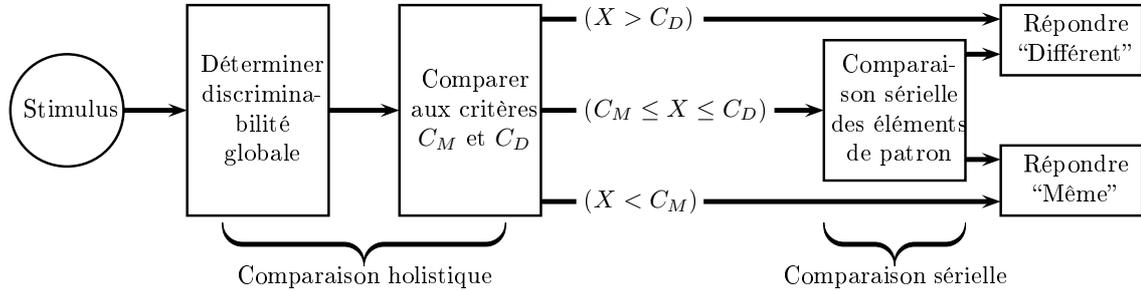
Une fois que l’auditeur a défini un point d’ancrage durant la tâche, il code l’excitation dérivée par les stimuli en estimant leur position par rapport à ce point. Quand plusieurs points d’ancrage sont disponibles, les mesures de distance sont combinées. La détermination de points d’ancrage est achevée par une reconstruction mentale du continuum physique. Par conséquent, plus l’étendue des stimuli est large, plus l’incertitude du sujet est importante. Et plus la distance est importante entre un stimulus et le point d’ancrage, moindre est la résolution du stimulus.

1.3.6 Le modèle à deux étapes

Le modèle de Kroll & Hershenson (1980), similaire à ceux de Lockhead (1972) et de Atkinson & Juola (1974), est basé sur la perception visuelle et comporte deux étapes. Dans la première étape, le sujet détermine la discriminabilité globale des stimuli (X) qu’il compare aux critères de similarité et de dissimilarité qu’il s’est fixés. Si la valeur de discriminabilité est inférieure au seuil de similarité (C_M), il répondra qu’il s’agit d’occurrences du même stimulus. Si la valeur est supérieure au seuil de dissimilarité (C_D), le sujet répondra qu’il s’agit de stimuli ‘différents’. Au cas où la valeur ne satisfait aucun des deux critères et qu’aucune décision ne peut être prise, une deuxième étape est initiée. Le sujet procède à une analyse sérielle des paramètres. Les valeurs des stimuli pour le premier paramètre seront extraites. Si la différence n’est pas suffisamment importante, le sujet procèdera à l’analyse du deuxième paramètre et ainsi de suite, jusqu’à ce qu’une différence soit trouvée.

Une différence cruciale entre le modèle de Kroll & Hershenson (1980) et les modèles décrits ci-avant regarde les prédictions qu’il fait concernant les temps de réaction. Dans le modèle duel, par exemple, le sujet compare d’abord les étiquettes des stimuli analysés. Si elles sont différentes, il répond ‘différent’. Si leur distance perceptive n’est pas suffisante, il procèdera à une deuxième analyse plus fine. Ainsi, ce mécanisme prédit des temps de réaction plus rapides pour les réponses ‘différent’. Néanmoins, les réponses correctes aux stimuli identiques sont généralement plus rapides que les réponses correctes aux stimuli différents. Nickerson (1972, 1973, 1978) discute certains des paradoxes émergeant dans les tâches ‘même-différent’ et qui contredisent les prédictions du modèle duel.

FIGURE 1.6: Le modèle de discriminabilité à deux étapes de Kroll & Hershenson (1980).



1.4 Théorie de la détection du signal⁴

La théorie de détection (du signal), introduite initialement par Tanner & Swets (1954), a été appliquée en psychophysique par Green & Swets (1966). L'idée principale de ce modèle est que toute décision se produit dans des conditions d'incertitude. La théorie de détection propose donc des outils pour analyser et quantifier notre capacité à discriminer entre le signal et le bruit.

Prenons l'exemple classique qui sert à illustrer le cœur de ce modèle. Soit un radiologue qui étudie une image de TDM (tomodensitométrie) en quête d'une tumeur. Même après un entraînement intensif, l'interprétation de ce type d'images est compliquée et toute décision est accompagnée d'un degré d'incertitude. Quatre possibilités sont envisageables (figure 1.7) : une tumeur existe et le médecin la détecte (détection correcte), une tumeur existe et le médecin ne la détecte pas (omission), le patient n'a pas de tumeur et le médecin croit en voir une (fausse alarme) ou le patient n'a pas de tumeur et le médecin n'en voit pas (rejet correct). Les détections correctes et les rejets corrects sont les réponses (comportements) désirables. Le patron de réponse optimal serait : 0% de fausses alarmes et 0% d'omissions. Deux facteurs, pourtant, peuvent compromettre ces résultats optimaux : le bruit et le biais.

Le mécanisme de la prise de décision implique deux concepts : l'acquisition de l'information et le biais.

La capacité d'un sujet à détecter le signal ou à discriminer deux stimuli dépend de la qualité de l'information qu'il va en obtenir. Il lui sera donc plus facile de reconnaître un objet si celui-ci lui est présenté pendant une durée suffisamment importante. De même, plus il a d'informations sur un objet, plus il lui est facile de le reconnaître. Mémoriser 40 items au lieu de 10 compliquera la tâche et le sujet aura plus de mal à les distinguer.

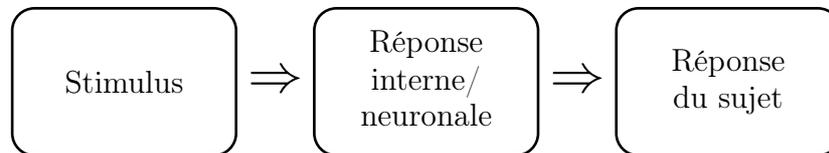
Nonobstant la qualité et quantité des informations, un sujet peut éventuellement opter pour une certaine réponse. La détection d'une tumeur est compliquée et toute réponse pèse différemment. Une fausse alarme peut mener à des examens supplémentaires et avoir comme résultat une exposition plus importante du patient à la radiation. Une omission, en revanche, augmentera le risque de mortalité du patient. Le médecin doit alors trouver un équilibre et décider si le risque d'une omission est

4. Pour la rédaction de cette section, nous avons tiré des informations à partir du site de David Heeger (<http://www.cns.nyu.edu/~david/handouts/sdt/sdt.html>), ainsi qu'à partir de l'entrée correspondante de Wikipedia (http://en.wikipedia.org/wiki/Detection_theory/)

FIGURE 1.7: Les quatre possibilités dans la théorie de la détection.

		Tumeur	
		Présent	Absent
Réponse	Oui	Détection Correcte (DC)	Fausse Alarme (FA)
	Non	Omission	Rejet Correct (RC)

FIGURE 1.8: Les étapes exécutées pendant une tâche de détection.



plus dangereux qu'une dose supplémentaire de radiation. Différents médecins auront, par conséquent, différents critères de détection de tumeurs.

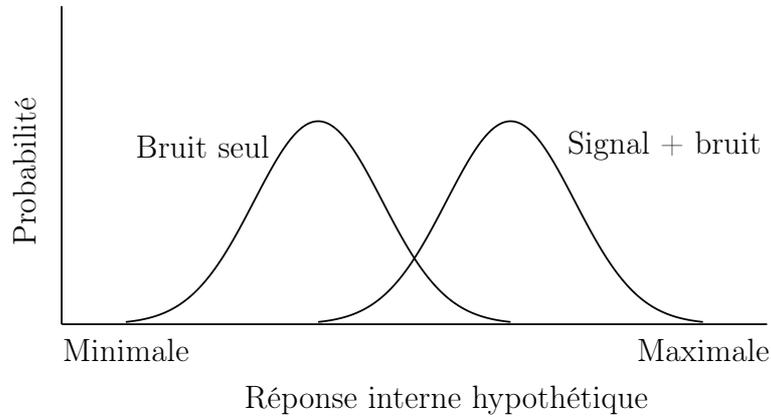
Nous avons écrit plus haut que toute décision se produit dans des conditions d'incertitude due à la présence de bruit, qui peut être externe ou interne.

Le bruit externe peut provenir de plusieurs sources. Dans le cas des TDM, une tâche, un mauvais contraste de l'image peuvent s'avérer de vrais obstacles pour le médecin. En psychoacoustique, la mauvaise qualité des casques ou des enceintes, un mauvais appareil d'enregistrement, le mauvais état mental ou physique du sujet (fatigue, malaise par rapport aux outils expérimentaux, ennui) peuvent introduire de la variance dans les réponses.

Le bruit interne est lié à l'activation des neurones dont les réponses sont 'bruyantes'. Il peut y avoir de la variance dans les réponses neuronales pour le même stimulus.

La figure 1.8 présente les deux étapes mises en route pendant une tâche de reconnaissance. Le stimulus, une fois traité par l'oreille, activera certains neurones dans le cerveau. Des théories ont été formulées sur la nature des détecteurs de traits mais des preuves solides n'ont pas encore été fournies en faveur de l'un ou de l'autre modèle. Étant donné que le traitement des informations ne nous est pas

FIGURE 1.9: Proportion des items ayant concédé le niveau d'excitation indiqué en abscisse.



explicite, nous ne pouvons pas faire d'inférences sur la localisation et la nature de ces neurones. Ainsi parlerons-nous de la réponse interne du sujet (ou de la sensation d'un stimulus) sans prendre de décision concernant la nature du traitement auditif ou visuel du stimulus. Comme nous venons de le préciser, cette réponse est sujette à du bruit inhérent. Ensuite, le sujet doit prendre une décision concernant la présence ou absence du stimulus en se basant sur des critères et des seuils subjectifs. Cette décision sera exprimée par la réponse effective.

Il est à noter ici qu'il est difficile de mesurer l'intensité du bruit interne et de le séparer du signal. Kontsevich *et al.* (2002) proposent une méthode psychophysique pour séparer le bruit du signal visuel. Pour une revue des différentes méthodes utilisées dans la littérature, le lecteur pourra se référer au même article.

Soit une tâche de détection avec 200 stimuli dont la moitié sont blancs (pas de signal) alors que l'autre moitié contiennent du bruit accompagné d'un signal. Même les stimuli blancs laissent une sensation, une réponse interne, qui est censée être caractérisée par une distribution normale (voir figure 1.9). Parfois ces excitations seront plus fortes, parfois plus faibles. Les stimuli contenant un signal engendreront généralement des sensations plus fortes que les stimuli blancs. Pourtant, dans certains cas, ces deux types de stimuli auront une réponse équivalente. Vu que le sujet ne connaît pas d'emblée la distribution à laquelle le stimulus présenté appartient, il doit fixer un seuil, à savoir un critère, qui séparera les deux distributions. Si le niveau d'excitation d'un stimulus dépasse ce seuil, le sujet répondra 'oui', sinon il répondra qu'il n'a pas détecté un signal.

La figure 1.10 présente les distributions des conditions 'bruit' (B) et 'signal + bruit' ($S\&B$), ainsi que le critère de sélection du sujet. La partie de la distribution de $S\&B$ au-dessus — à droite — contient les détections correctes alors que la partie de la même distribution en dessous du critère concerne les omissions. Quant à la condition B , les réponses en dessous du critère sont des rejets corrects, alors que la partie au-dessus du critère sont des fausses alarmes.

Nous pouvons estimer la propension du sujet à donner une réponse plus souvent que l'autre avec la formule 1.1 :

$$c = -\frac{1}{2}[z(DC) + z(FA)] \quad (1.1)$$

FIGURE 1.10: Proportion des items ayant concédé le niveau d'excitation indiqué en abscisse et le critère de décision. DC : détection correcte; FA : fausse alarme; RC : rejet correct.

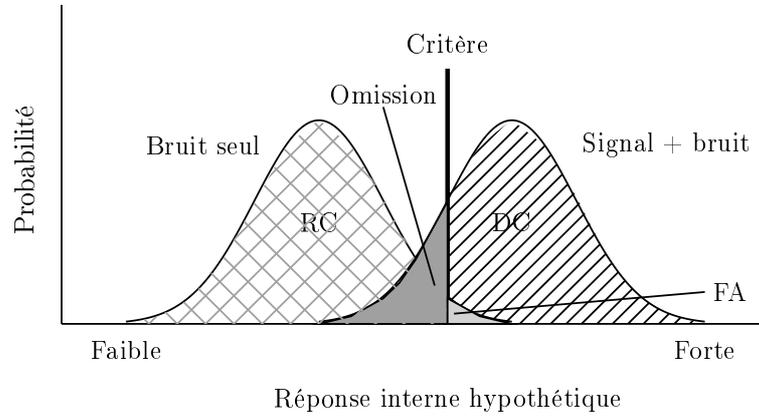
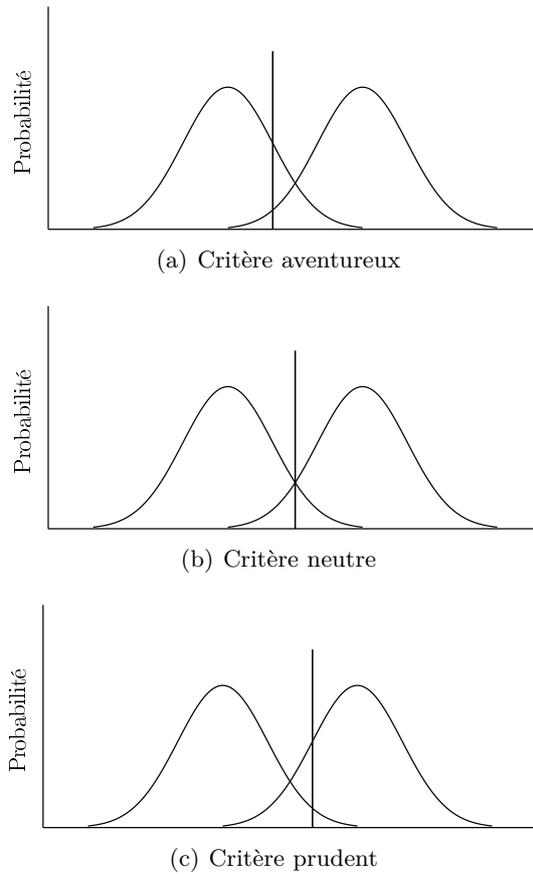


FIGURE 1.11: Distributions du bruit (à gauche) et du signal avec bruit (à droite) dans trois conditions différant dans l'emplacement du critère.



où c est la mesure du biais, et $z(DC)$ et $z(FA)$ sont les valeurs normales inverses des proportions des détections correctes et des fausses alarmes respectivement.

Le deuxième paramètre important des deux distributions — le premier étant le critère — est la sensibilité, qui représente le degré de séparation des deux distributions. Plus le signal est fort, plus les deux fonctions sont séparées. De même, la variance de chaque courbe décidera si le chevauchement sera considérable. Nous pouvons estimer la sensibilité avec la formule 1.2 :

$$d' = z(DC) - z(FA) \quad (1.2)$$

Nous trouvons dans la figure 1.11 les distributions hypothétiques du bruit et du signal avec bruit dans trois conditions différentes. Dans 1.12(a), le sujet fixe un critère aventureux et a tendance à détecter des signaux. Une stimulation relativement faible amène le sujet à considérer qu'il s'agit d'un signal. Quand le critère est neutre (figure 1.12(b)), une stimulation moyenne peut être classée tantôt comme un signal, tantôt comme du bruit. Quand le sujet est prudent (figure 1.12(c)), il ne détecte un signal que quand la stimulation est relativement forte. Il a donc tendance à donner moins souvent la réponse 'oui'.

1.5 Conclusion

Dans ce chapitre, nous avons présenté le cadre théorique dans lequel ce travail s'inscrit. Nous avons commencé par présenter quelques concepts primaires de la psychophysique et de la perception pour continuer avec une description des différents niveaux de la mémoire (précatégorielle, à court terme et à long terme), de leur nature et de leur fonction. Ce modèle nous a aidé dans l'illustration des différentes théories qui tentent d'expliquer les mécanismes implémentés dans les tâches d'identification et de discrimination. Dans la dernière section, nous avons fait une courte introduction à la théorie de la détection du signal.

Dans le chapitre suivant, nous passons en revue les hypothèses émises sur l'origine des asymétries en perception et nous les expliquons à travers les modèles abordés dans ce chapitre.

Chapitre 2

Asymétries et hypothèses

"Make everything as simple as possible, but not simpler."
Albert Einstein

Nous avons mentionné au début du chapitre précédent que l'effet d'ordre de la présentation des stimuli jouait un rôle sur leur discriminabilité et que ce phénomène avait été attesté par Fechner il y a 150 ans. Ce phénomène de résolution différentielle de stimuli consécutifs a été négligé par les chercheurs qui, très souvent, le considèrent comme une source de variabilité, bien qu'il ait été démontré par une série de travaux que nous allons déployer ci-après. Nous allons réinterpréter chaque hypothèse formulée sur la provenance des asymétries en perception vocalique à travers des modèles de discriminabilité présentés dans le premier chapitre.

2.1 Asymétries et stimuli non-linguistiques

2.1.1 Formes géométriques

Dans son travail sur les prototypes, Rosch (1975) envisage la similarité de lignes droites. Elle choisit trois directions comme point de référence : horizontale (0°), verticale (90°) et diagonale (45° et 135°). Les résultats montrent que la distance psychologique entre deux lignes est moins importante quand un de ces points de référence est placé en seconde position dans une paire.

Op de Beeck *et al.* (2003) comparent les performances de singes et d'humains afin de déterminer si le mécanisme d'asymétrie est phylogénétique, c'est-à-dire lié à des mécanismes neurophysiologiques communs aux deux espèces. Les auteurs trouvent une corrélation positive entre typicalité et ordre de présentation au niveau du groupe. La similarité entre deux stimuli est jugée plus basse quand le prototype est en première position. Or, les analyses au niveau inter-sujet montrent une variabilité importante aussi bien pour les sujets humains que pour les singes.

2.1.2 Couleurs

Rosch (1975) examine les quatre couleurs basiques 'rouge', 'jaune', 'vert' et 'bleu'. Elle choisit comme prototype des nuances de ces couleurs qui ont été jugées typiques par des sujets venant de plusieurs langues (Berlin & Kay, 1969). Les stimuli dévient

des prototypes et des stimuli-contrôle différent en teinte ou en degré de saturation. Rosch rapporte que dans les paires comprenant un prototype, la similarité entre les deux membres est plus basse quand le prototype est présenté en seconde position.

Polk *et al.* (2002) travaillent avec les couleurs ‘bleu’ et ‘vert’ et construisent cinq stimuli pour chaque catégorie en modifiant les valeurs RVB (Rouge-Vert-Bleu). Le but de l’étude est d’apporter des indications en faveur de l’impact de l’exposition sur le degré d’asymétrie dans les jugements de similarité. Les résultats montrent que les asymétries sont significativement plus prononcées après l’entraînement qu’avant. Les auteurs suggèrent que l’exposition peut mener à des états stables dans un réseau récurrent.

2.1.3 Tons modulés

Dooley & Moore (1988) étudient les seuils de discrimination de fréquence. En utilisant des tons modulés (angl. *glides*) montants ou descendants, ils détectent des asymétries directionnelles. Ils rapportent que le seuil fréquentiel est significativement plus élevé pour les tons modulés montants dans les fréquences jusqu’à 2000 Hz, et pour des durées comprises entre 50 et 750 ms à 2000 Hz. Tsumura *et al.* (1973) signalent des effets d’ordre de la même direction, alors que les asymétries attestées par Gardner & Wilson (1979) et Schouten (1985, 1986) sont de la direction opposée. Arlinger *et al.* (1977) ne trouvent pas d’effet d’ordre. Dooley & Moore (1988) prennent en considération les différences — en matière de durée, d’expérience des sujets et de patron fréquentiel — entre les stimuli utilisés dans les études ci-mentionnées. Ils concluent qu’ils n’arrivent pas à déceler la raison pour laquelle une différence de directionnalité entre leurs propres résultats et ceux de Gardner & Wilson (1979) a été retrouvée. La seule explication proposée concerne des différences individuelles.

2.1.4 Accords musicaux

Bharucha & Krumhansl (1983) utilisent des paires d’accords musicaux pour étudier l’asymétrie contextuelle (angl. *Contextual Asymmetry*) et l’asymétrie intratonale (angl. *Intrakey Asymmetry*). Les patrons de réponses indiquent qu’il est plus facile de détecter le changement d’un accord juste (angl. *on-key*) vers un accord faux (angl. *off-key*) que dans l’ordre inverse. Bharucha & Pryor (1986) trouvent que dans une tâche ‘même-différent’, les performances sont plus élevées quand, dans un paire de séquences tonales, la séquence rythmique est présentée en première position et la variante discontinue en seconde position. Des asymétries dans la perception de stimuli musicaux sont aussi attestées dans Schellenberg & Trehub (1999), Schellenberg (2001, 2002) et Trainor (1997).

2.2 Asymétries et stimuli linguistiques

2.2.1 Matrices de confusion et asymétries en reconnaissance

La représentation des résultats de catégorisation sous la forme de matrices de confusion est bien répandue dans la littérature. La structure des matrices est toujours la même : les lignes représentent les segments cibles et les colonnes sont les réponses

Tableau 2.1: Matrice de confusion des voyelles anglaises présentées avec un ratio signal/ bruit à -14 dB (Benkí, 2003).

	a	æ	ai	ε	e	o	ɪ	i	u	Λ
a	227	17	13	0	1	4	1	0	2	11
æ	26	68	0	6	4	0	2	1	0	2
ai	7	2	92	4	5	0	2	3	1	4
ε	11	7	2	57	17	8	1	1	1	24
e	2	5	4	2	91	12	13	1	0	1
o	9	5	11	7	45	34	10	0	0	5
ɪ	1	2	0	0	22	1	87	1	2	3
i	1	1	1	4	1	0	11	94	20	2
u	0	0	2	3	6	0	13	60	41	0

des auditeurs. Le contenu des matrices est souvent analysé dans le cadre de la théorie de l'information (Miller & Nicely, 1955; Bell *et al.*, 1989). Pourtant, en lisant une matrice, nous nous rendons immédiatement compte d'asymétries prononcées entre des paires des segments.

Benkí (2003) enregistre des non-mots CVC contenant dix consonnes et dix voyelles de l'anglais et les présente à des auditeurs anglophones sous quatre ratio de signal/bruit, -14, -11, -8 et -5 dB. Le tableau 2.1 présente la matrice qui concerne la reconnaissance des voyelles dans la condition S/B -14 dB. Les cadres 2x2 indiquent des asymétries. Nous observons que /ε/ est souvent perçu comme /e/ mais l'inverse est très rare. La même tendance est attestée pour d'autres paires de voyelles. Le degré d'asymétrie dépend de la condition et diminue quand le signal est assez fort par rapport au bruit (effets de seuils). Les auteurs parviennent à la conclusion qu'il y a un biais vers les voyelles antérieures. Karypidis *et al.* (2006a, 2009) utilisent un ensemble de 33 stimuli synthétiques qui sont dispersés dans l'espace vocalique. Ils observent un biais vers les voyelles antérieures pour les grecophones, vers les voyelles postérieures pour les hispanophones et une symétrie pour les griekophones et les auditeurs de l'italien salentin du sud. Benkí (2003) remarque aussi que la hauteur vocalique est plus robuste que la place d'articulation des voyelles, que l'opposition de voisement chez les consonnes est plus robuste dans la position d'attaque syllabique qu'en coda, et qu'il y a moins d'erreurs de catégorisation pour les consonnes en position initiale que pour les consonnes en fin de mot.

Selon Plauché *et al.* (1997), lorsqu'un son X possède un indice que le son Y ne possède pas, il est plus probable que les auditeurs ratent cet indice au lieu de l'introduire. Prenons l'exemple plus concret des voyelles /i/ et /y/. Les auditeurs vont donc omettre l'indice de labialité¹, plutôt que de l'ajouter au segment entendu. Par conséquent, les auditeurs percevront [y] comme /i/ mais ils ne confondront pas [i] avec /y/. Pour une hypothèse similaire, voir le postulat de *mismatch* de Lipski *et al.* (2007) plus bas dans ce chapitre.

Le tableau 2.2 illustre des résultats de reconnaissance des consonnes de l'anglais américain extraits de Miller & Nicely (1955). De nouveau, des asymétries sont

1. Il faut noter que la labialité n'a pas d'indice spécifique et a un effet global sur les propriétés spectrales d'un segment. Notre exemple ne sert qu'à démontrer le raisonnement initial.

Tableau 2.2: Matrice de confusion des voyelles anglaises présentées avec un ratio signal/ bruit à -12 dB (Miller & Nicely, 1955).

	p	t	k	f	θ	s	ʃ	b	d	g	v	ð	z	ʒ	m	n
p	51	53	65	22	19	6	11	2		2	3	3	1	5	8	5
t	64	57	74	20	24	22	14	2	3	1	1	2	1	1	5	1
k	50	42	62	22	18	16	11	4	1	1	1	2			4	2
f	31	22	28	85	34	15	11	3	5		8	8	3		3	
θ	26	22	25	63	45	27	12	6	9	3	11	9	3	2		1
s	16	15	16	33	24	53	48	3	5	6	3	1	6	2		1
ʃ	23	32	20	14	27	25	115	1	4	5	3		6	3	4	2
b	4	2	2	18	7	7	1	60	18	18	44	25	14	6	20	10
d	3		1	4	7	4	11	18	48	35	16	24	26	14	9	12
g	3	1	1	1	4	5	7	20	38	29	16	29	29	38	10	9
v		1	1	12	5	4	5	37	20	23	71	16	14	4	14	9
ð		1	4	17	2	3	2	53	31	25	50	33	23	5	13	6
z	6	1	2	2	6	14	8	23	29	27	24	19	40	26	3	6
ʒ	3	2	2	1		6	7	7	30	23	9	7	39	77	5	14
m		1			1	1		11	3	6	8	11		1	109	60
n	1			1	1	1		2	2	6	7	1	1	9	84	145

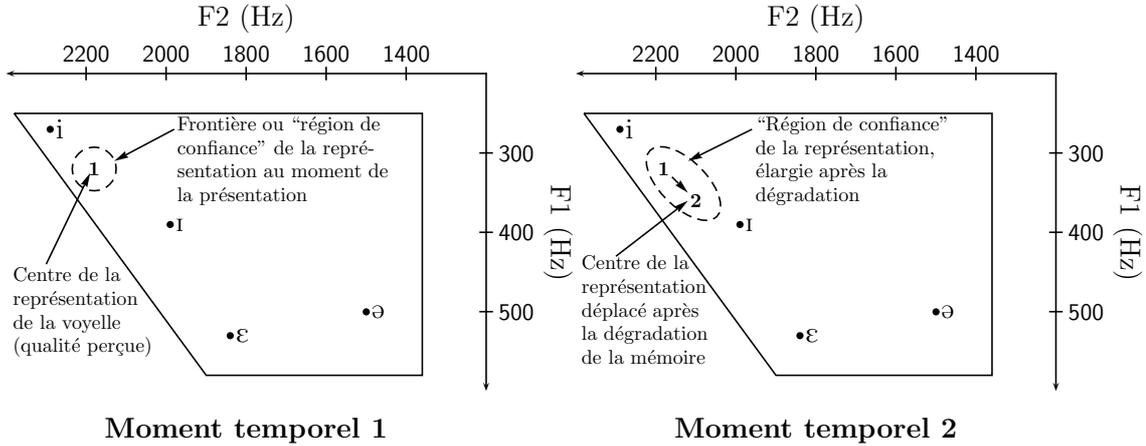
retrouvées dans toute la matrice.

2.2.2 L'hypothèse de la neutralisation

Repp *et al.* (1979) sont probablement les premiers à faire état des effets d'ordre dans la perception des voyelles. En menant des expériences d'identification et de discrimination sur le continuum /i/-/ɪ/-/e/ de Pisoni (1971), ils manipulent l'Intervalle Interstimulus (IIS) qui est rempli, soit de silence, soit de répétitions de [y]. Les résultats révèlent des effets d'ordre dans la région de /i/ et de /ɪ/, où les scores de discrimination sont plus élevés quand la voyelle la plus 'fermée', c'est-à-dire celle avec la valeur de F1 la plus basse, est présentée en seconde position dans une paire de stimuli. Un effet inverse mais de moindre degré est retrouvé à l'extrémité /e/. Les auteurs observent ces tendances dans l'expérience 2 de leur travail. Dans leur conclusion, ils postulent que les asymétries ne sont pas affectées par les interférences induites pendant l'IIS.

Quelques années plus tard, Cowan & Morse (1986) utilisent le continuum /i/-/ɪ/ de Pisoni (1973) et confirment que la discrimination est plus facile quand la deuxième voyelle dans une paire est plus proche de l'extrémité /i/. Ils avancent que cette asymétrie dans la distance perçue entre deux stimuli est motivée par un changement de qualité vocalique vers le centre de l'espace vocalique, qui correspond à la région occupée par la voyelle /ə/. Le mécanisme de ce changement est présenté dans la figure 2.1. À gauche, nous voyons la trace auditive que la voyelle a laissée 250 ms après sa présentation. La ligne discontinue délimite la 'région de confiance' de la voyelle. Au centre de cette région est localisé le centre de la représentation mentale que le locuteur se fait du stimulus. Au moment temporel 2, soit 2 secondes après la présentation du stimulus, la 'région de confiance' est élargie et le centre de

FIGURE 2.1: Le modèle de changement de qualité vocalique, adapté de Cowan & Morse (1986). Au moment temporel 1, la représentation mentale du stimulus est précise et étroite; au moment 2, la dégradation de la trace mémorielle du stimulus déplace le centre de la ‘région de confiance’, qui est élargie vers un point central dans l’espace.



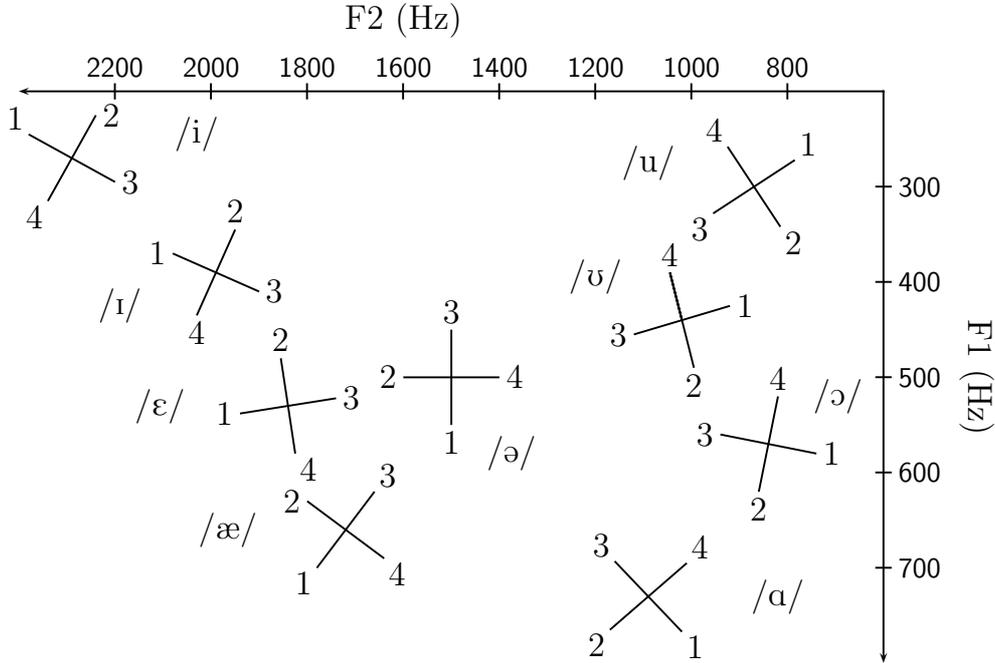
la représentation est déplacé vers le centre de l’espace vocalique.

Ce mouvement d’expansion dépend de la position du stimulus dans l’espace. Dans la figure 2.1, le stimulus se trouve très près de la périphérie, et ‘sa région de confiance’ ne pouvant s’élargir vers un endroit encore plus extrême dans l’espace, elle s’élargit obligatoirement vers l’intérieur de l’espace vocalique. Dans le cas d’une voyelle non périphérique, de telles contraintes acoustiques ne s’imposent pas et l’expansion de la ‘région de confiance’ peut se réaliser dans une direction quelconque. L’introduction du modèle de changement vocalique (‘quality shift model’) est aussi motivée par la corrélation positive entre les effets d’ordre et l’IIS : plus l’IIS est long, plus la trace mémorielle se déplace vers un endroit cible dans l’espace et, par conséquent, plus les effets d’ordre sont importants.

En construisant un protocole *ad hoc*, Repp & Crowder (1990) (R&C dans la suite du texte) mettent à l’épreuve l’hypothèse de dégradation qu’ils ont nommée ‘hypothèse de la neutralisation’. Dans l’expérience principale de leur article, l’ensemble des stimuli comprend neuf stimuli ‘prototypes’ dont les valeurs formantiques de F1 et F2 correspondent aux valeurs moyennes pour les voyelles /i, ɪ, e, æ, a, ɔ, u, u/ de l’anglais américain, telles qu’elles ont été produites par des locuteurs masculins (Peterson & Barney, 1952). Les valeurs d’un tube uniforme de 17 cm, fermé au niveau de la source et ouvert à l’autre extrémité (Chiba & Kajiyama, 1941), et de longueur égale à celle de l’appareil phonatoire d’un homme adulte ($F1 = 500$ Hz, $F2 = 1500$ Hz), sont attribuées à la voyelle /ə/, qui n’est pas incluse dans le travail de Peterson et Barney, et qui est ajoutée à l’ensemble des stimuli.

Autour de chacun des neuf prototypes, quatre satellites (S1-S4) sont positionnés en forme de croix (figure 2.2), avec l’axe S1-S3 pointant vers [ə] et S1 étant toujours le plus éloigné de [ə]. Les satellites sont numérotés dans le sens des aiguilles d’une montre en commençant par le stimulus le plus éloigné de [ə], à savoir S1. Quant à la catégorie /ə/, les axes de S1-S3 et de S2-S4 sont parallèles à ceux de F1 et F2

FIGURE 2.2: Les 45 stimuli de Repp & Crowder (1990). Au centre de chaque croix se trouve le prototype de la catégorie respective.



respectivement, avec $F1_{S1} > F1_{S3}$ et $F2_{S2} > F2_{S4}$.

La longueur des branches de chaque croix est décidée d'après des pré-tests, de sorte qu'il soit raisonnablement difficile de distinguer les satellites de leurs prototypes respectifs. F3 est fixe pour toutes les voyelles, soit à 2440 Hz, sauf pour la catégorie /i/ pour laquelle F3 est égal à 3010 Hz.

Au sein de chaque catégorie, tous les stimuli (P, S1-S4) sont appariés avec eux-mêmes (P/P, S_i/S_i); les quatre satellites sont également groupés avec le prototype dans les deux ordres possibles (P/ S_i , S_i /P). L'IIS est fixé à 500 ms entre les deux membres de chaque paire, à 2 s entre les paires et à 5 s après chaque groupe de 13 paires. Sept blocs, chacun contenant toutes les paires en ordre aléatoire — mais identique pour tous les auditeurs — sont préparés. 24 locuteurs natifs (anglais américain) participent à cette expérience.

Suivant l'hypothèse de la neutralisation, nous nous attendons à des effets d'ordre systématiques pour les satellites S1 et S3, et à des effets mineurs ou inexistantes pour S2 et S4. Dans le cas de S1, la discrimination est plus facile dans l'ordre P/S1, étant donné que P se dégrade vers [ə] (voir figure 2.1), donc loin de S1. Dans l'ordre inverse, l'expansion de la région de confiance se fait en direction de [ə], donc vers P, réduisant ainsi la distance perceptive entre les deux stimuli. Dans le cas de S3, la discrimination est plus facile dans l'ordre S3/P : S3 se déplace vers [ə] donc loin de P. Dès lors, les effets sont positifs pour S1 [$(P/S1) - (S1 - P) > 0$] et négatifs pour S3 [$(P/S3) - (S3 - P) < 0$]. Dans le cas de S2 et S4, l'ordre ne joue pas de rôle important dans la discriminabilité des stimuli.

Afin d'étudier ces prédictions, Repp et Crowder (R&B) mènent des analyses de variance (ANOVA) séparées pour S1 et S3, d'une part, et pour S2 et S4 de l'autre.

Un effet d'ordre issu d'une analyse commune pour les quatre satellites ne saurait être expliqué de façon simple. Les résultats de la catégorie /ə/ ne sont pas inclus, puisque les satellites ne sont pas numérotés de la même façon que ceux des autres catégories. Il faut remarquer ici que les effets d'ordre au sein de la catégorie /ə/ sont relativement bas, à savoir autour de 7%.

Pour S1 et S3, l'ANOVA montre un effet d'Ordre significatif. L'interaction Satellite* Ordre confirme l'hypothèse initiale prédisant des polarités inverses pour les deux satellites. Une triple interaction Voyelle* Satellite* Ordre révèle que le degré d'asymétrie varie à travers l'espace vocalique. Dans le cadre de l'hypothèse, ceci s'explique par le fait que la distance entre chaque catégorie et [ə] n'est pas fixe. Théoriquement, plus une voyelle est proche de [ə], moins la représentation mentale de la première voyelle dans une paire se déplace. Dans le cas de [ɛ], le déplacement du premier stimulus est restreint, vu qu'il se situe très près de [ə]. Bien que les auteurs n'effectuent pas des analyses de corrélation entre le degré d'asymétrie et la distance de chaque catégorie par rapport à [ə], les scores bruts ne révèlent pas une telle tendance : les effets pour S1 et S3 de /ɛ/, par exemple, sont plus importants que pour ceux de /i/.

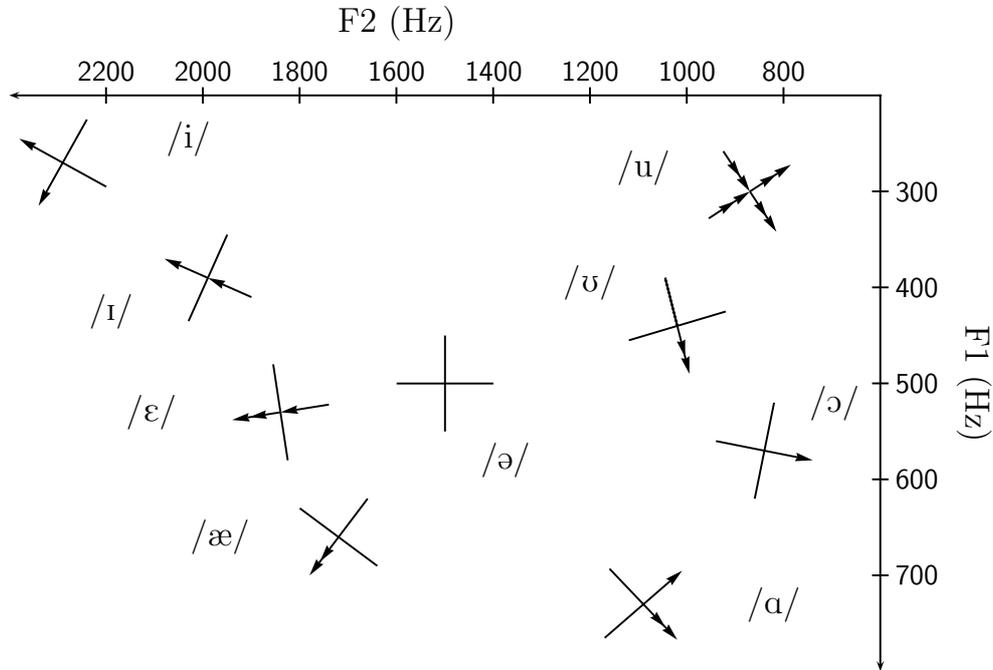
Pour S2 et S4, l'effet significatif d'Ordre et la triple interaction Voyelle* Satellite* Ordre affaiblissent l'hypothèse de départ. Or, les scores bruts indiquent qu'un nombre limité de catégories présente des asymétries considérables : /u/ et /ʊ/ pour S2, et /i/, /ɑ/ et /u/ pour S4. Ce nombre relativement bas de catégories n'affaiblit pas nécessairement l'hypothèse de la neutralisation, qui prédit peu ou pas d'asymétries pour S2 et S4. L'interaction Satellite* Ordre n'atteint pas le niveau de significativité.

Les analyses statistiques susmentionnées apportent des indications en faveur de l'hypothèse de la neutralisation. En revanche, en regardant les asymétries pour chaque paire de stimuli (figure 2.3), les auteurs constatent que, alors que pour les voyelles antérieures le point cible semble être [ə], les voyelles postérieures auraient plutôt comme destination la région autour de S3 de /ɔ/. Cette divergence, ajoutée aux résultats d'un travail antérieur, les amène à suggérer que la cible dépend probablement de l'ensemble des stimuli utilisés dans une tâche. Il est à noter que les seuils que nous avons fixés pour les asymétries dans la figure 2.3 sont subjectifs et ne servent qu'à déchiffrer les données en question.

Dans l'expérience 2 de leur article, Repp et Crowder examinent la relation entre le point de dégradation et l'étendue de l'ensemble des stimuli. Les données ne confirment pas l'hypothèse de départ. Une relation entre le degré d'asymétrie et l'IIS est constatée, ce qui mène au postulat que l'appartenance catégorielle - soit l'étiquette assignée à un stimulus, telle que /i/ ou /e/ - joue un rôle. La possibilité d'un changement de qualité vers la tendance centrale ou le meilleur exemplaire d'une catégorie contredit les résultats de l'expérience 1.

En résumant les résultats des trois expériences (l'expérience 3 n'est pas décrite ici), les auteurs concluent que seule l'expérience 1, qui inclut la plus grande variété de stimuli et qui est désignée *ad hoc*, corrobore l'hypothèse de la neutralisation. Or, le point faible de cette démonstration est que le point neutre dans l'espace vocalique correspond approximativement au centre de l'ensemble des stimuli. Ceci tend plutôt à montrer qu'il existe une cible de dégradation, qui serait propre à la distribution des

FIGURE 2.3: Représentation graphique des asymétries retrouvées dans les résultats de Repp & Crowder (1990). Les flèches indiquent l'ordre dans lequel la discrimination a été plus facile. Flèches simples : effet d'ordre supérieur à 10% ; flèches doubles : effet supérieur à 20%.



stimuli utilisés dans une tâche. Le rôle de l'IIS est interprété dans le cadre du modèle de code duel de Fujisaki & Kawashima (1970) : la première voyelle d'une paire est encodée sous deux formes : (a) une forme catégorielle qui résulte de l'attribution d'une étiquette telle que /i/ ou /u/, et (b) une forme continue qui consiste en la forme analogique du signal avec des détails fins. Avec des IIS courts, les auditeurs utilisent la trace auditive des stimuli pour les discriminer, alors que l'introduction d'un IIS relativement long oblige les sujets à recourir aux résultats de catégorisation.

2.2.3 Périphéricité et focalisation

Au cours des dix dernières années, le centre d'intérêt s'est déplacé des prototypes vers les extrémités de l'espace vocalique qui, à leur tour, jouent le rôle d'aimants perceptifs. Ce glissement d'intérêt peut être dû au fait que la définition des prototypes est très difficile : les voyelles sont multidimensionnelles et un nombre assez important de combinaisons de fréquences est nécessaire afin qu'un exemplaire ressemblant le plus aux représentations mentales hypothétiques emmagasinées en MLT puisse être localisé. En même temps, la variabilité entre les sujets dans leur choix de prototypes exige que des protocoles personnalisés à chaque sujet soient utilisés. De même, l'utilisation de la notion de prototype en discrimination exige de poser *a priori* que les sujets utilisent un seul prototype pour tout contexte. Ce postulat ne peut être défendu étant donné que la quête des indices invariants des sons de la parole n'a jamais abouti à des résultats satisfaisants. Une dernière contre-indication ne permettant pas l'utilisation des prototypes pour expliquer les asymétries est le fait

que très souvent dans les continua utilisés, seule une catégorie de celles exploitées dans un travail donné se comportait comme un point d’ancrage.

Polka & Bohn (2003), en résumant les travaux sur la perception des voyelles par les enfants, postulent qu’il est plus facile de discriminer quand la voyelle la plus périphérique dans une paire est présentée en seconde position. Une voyelle périphérique est caractérisée par une ou plusieurs valeurs formantiques extrêmes et se localise donc sur les bords de l’espace acoustique des voyelles (Vallée, 1994; Schwartz *et al.*, 1997). La figure 2.4 présente les oppositions étudiées dans les différents travaux de la littérature. La figure 2.5(a) montre que pour les données d’enfants, il est en effet plus facile de discriminer deux voyelles quand le changement vocalique se fait vers une position plus extrême dans l’espace. La seule exception est l’opposition /u/-/y/. Cet effet d’ordre plutôt inattendu — les deux voyelles possèdent également des positions périphériques — amène Best & Faber (2000) à postuler que la position sur le plan F1/F2 ne peut peut-être pas expliquer certains résultats et que la périphéricité sur le plan F2/F3 peut aussi jouer un rôle, du moins pour le /y/ norvégien.

Les résultats des études animales (figure 2.5(b)) sont similaires à ceux des enfants. Cette fois-ci, le changement vocalique est plus facilement perçu quand nous avons une élévation de F1 et F2. La seule exception est l’opposition /ε/-/æ/ avec un passage vers une voyelle plus haute.

Polka & Bohn (2003) essaient d’éliminer certains facteurs qui pourraient déclencher des effets d’ordre. Une première hypothèse non vérifiée dans la littérature concerne le statut phonologique des voyelles. Le tableau 2.3 présente les différents travaux de la littérature où des asymétries ont été attestées. Polka & Werker (1994) observent une asymétrie $\gamma \rightarrow \upsilon$, ce qui suggère que la voyelle non-native /y/ est assimilé à la voyelle /u/ via un effet d’aimant perceptif (Kuhl *et al.*, 1992). Cette hypothèse est étudiée dans un travail ultérieur Polka & Bohn (1996). Cette fois, les chercheurs trouvent une asymétrie $\varepsilon \rightarrow \text{æ}$ pour des enfants anglophones et germanophones. La voyelle /æ/ n’appartenant pas au système de l’allemand, l’impact du facteur de familiarité est abandonné. Deux travaux ultérieurs (Best *et al.*, 1997; Best & Faber, 2000) attestent des asymétries chez les enfants anglophones en faveur de la voyelle non-native /y/ dans les paires /u/-/y/ et /i/-/y/.

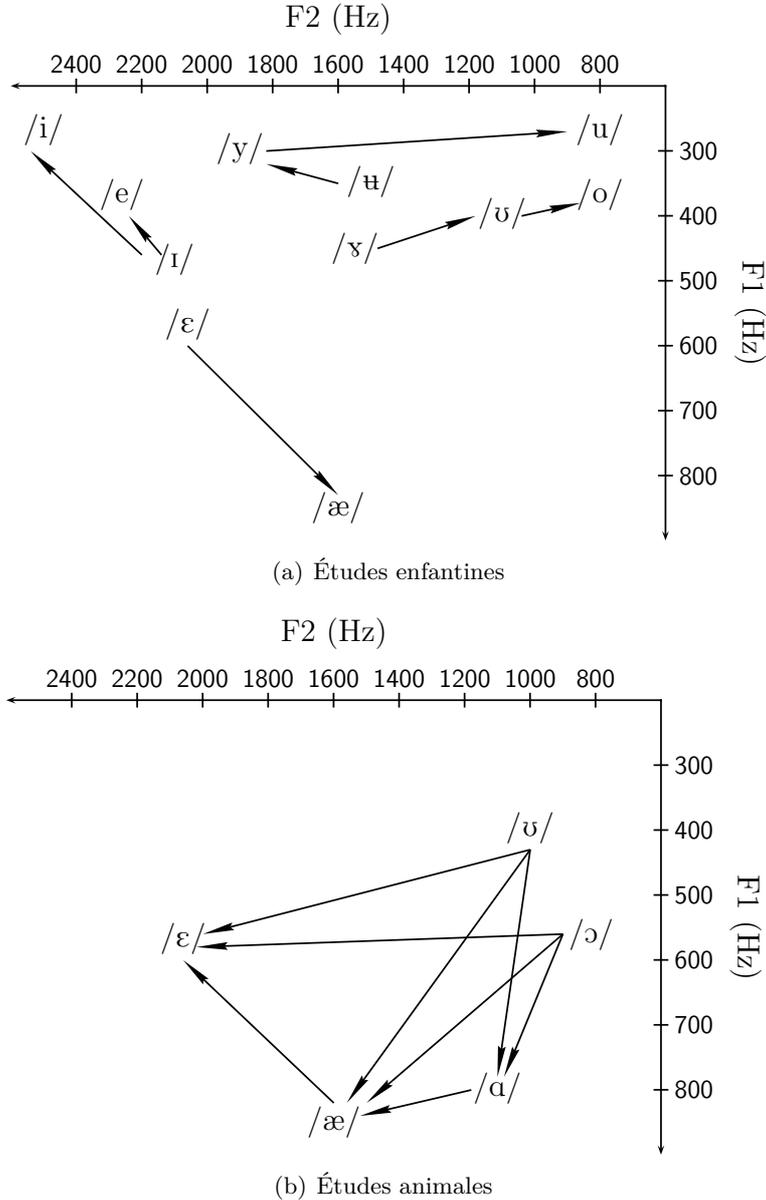
Un deuxième facteur qui ne semble pas être valide est la tendance vers les voyelles favorisées dans les langues du monde. Il s’agirait d’un biais perceptif. La figure 2.5 présente la fréquence des 37 voyelles ‘basiques’ apparaissant dans la base de données de Maddieson (1991). Nous observons donc que l’hypothèse du biais vers les voyelles les plus fréquentes est valide pour la plupart des oppositions. Une nette exception est pourtant retrouvée, la paire /ε/-/æ/.

Polka & Bohn (2003) éliminent également les facteurs d’amplitude (intensité) et de durée. Les oppositions /u/-/y/, /γ/-/υ/ et /ε/-/æ/ ont été choisies de façon à ce que leurs valeurs sur ces deux dimensions se chevauchent. De même, la différence de durée pour les membres de la paire /ε/-/æ/ est neutralisée.

Le dernier facteur n’offrant pas de solution robuste est la direction du changement sur le plan F1 ou F2. Les flèches dans la figure 2.5(a) pointent vers toutes les directions.

Les auteurs reviennent alors au postulat originel de la périphéricité : la discrimination est plus facile quand le changement de qualité se fait vers une voyelle

FIGURE 2.4: Les oppositions phonémiques étudiées dans les travaux ci-indiqués et la direction de changement dans laquelle la discrimination a été plus facile. (a) : études d'enfants; (b) études animales. Les valeurs formantiques sont indicatives.



plus extrême. Il s'agit d'un biais perceptif universel et non-spécifique à la langue d'exposition du sujet.

Schwartz *et al.* (2005) placent le phénomène d'asymétrie dans le cadre de la théorie de Dispersion-Focalisation (Schwartz *et al.*, 1997) et proposent que :

“more *peripheral* vowels in Polka & Bohn's experiments are in fact more *focal* vowels, and (...) more focal vowels within a contrast would serve as a reference or perceptual anchor for discrimination, thanks to

Tableau 2.3: Études en perception enfantine et animale ayant indiqué des effets d'ordre. Les flèches indiquent la direction dans laquelle la discrimination a été plus facile (Polka & Bohn, 2003). Entre parenthèses, les oppositions pour lesquelles l'effet d'ordre n'a pas été significatif.

Référence	Langue d'exposition	Opposition
<i>(a) Enfants</i>		
Polka & Werker (1994)	Anglais	y→u y→ʊ
Polka & Bohn (1996)	Anglais	y→u ε→æ
	Allemand	y→u ε→æ
Bohn & Polka (2001)	Allemand	i→e (e→i) (ʊ→o)
Swoboda <i>et al.</i> (1976)	Anglais	(i→i)
Swoboda <i>et al.</i> (1978)	Anglais	i→i
Dejardins & Trainor (1998)	Anglais	i→i
Best <i>et al.</i> (1997)	Anglais	ʌ→y
Best & Faber (2000)	Anglais	i→y
	Espèce	
<i>(b) Animaux</i>		
Hienz <i>et al.</i> (1981)	Merle	ɔ→æ ɔ→ɑ ɔ→ε ɑ→æ
Hienz <i>et al.</i> (1996)	Chat	ɔ→æ ɔ→ɑ ɔ→ε ɑ→æ ʊ→ε ʊ→æ ʊ→ɑ

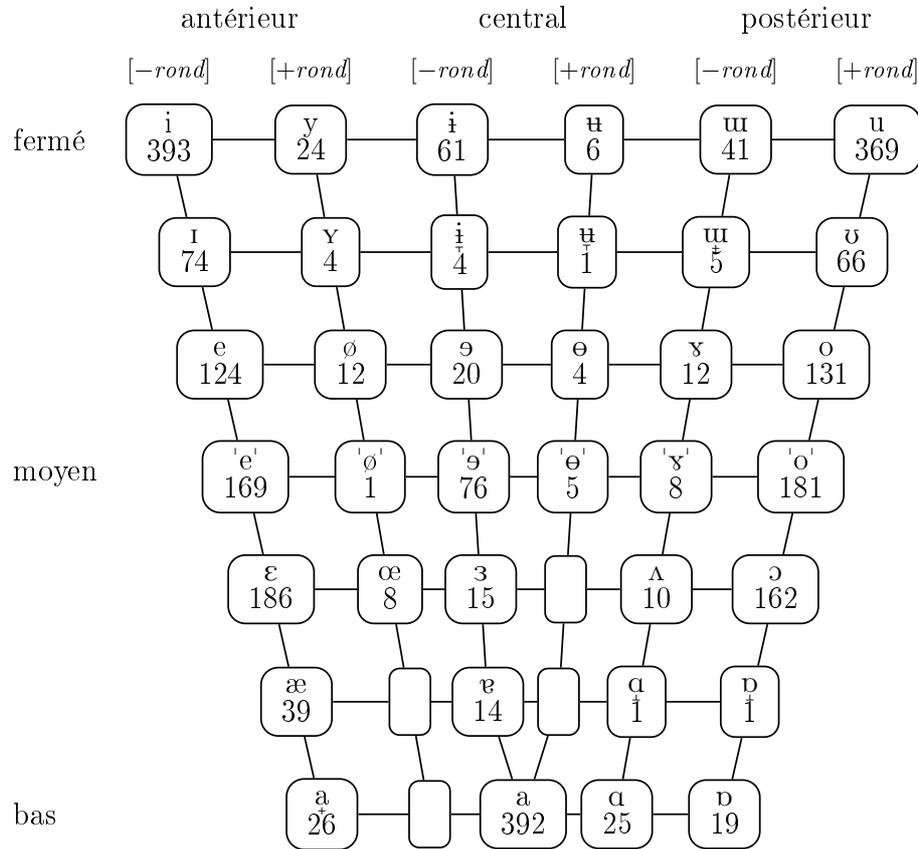
their increased auditory salience” (p. 431).

Le rôle de la focalisation dans la perception des voyelles est démontré dans Schwartz & Escudier (1989). Dans leur continuum synthétisé, F0, F1, F2 et F4 sont fixés à 100 Hz, 450 Hz, 2000 Hz et 3350 Hz respectivement. Le troisième formant varie de 2300 Hz à 3000 Hz. Les deux extrémités forment des stimuli focaux. Les résultats révèlent de fausses alarmes moins importantes pour ces stimuli focaux, indiquant la stabilité de la trace mémorielle de ces derniers. Ceci pointe vers un comportement psychophysique lié à des processus auditifs tels que l'intégration spectrale.

2.2.4 L'effet des aimants perceptifs

Kuhl (1991) est la première à mettre le concept de 'prototype' à l'épreuve de façon

FIGURE 2.5: Les 37 qualités vocaliques basiques (des segments oraux, non-longs, monophthongaux, à voix normale) retrouvées dans UPSID (Maddieson, 1991). Dans chaque ovale, le symbole API de la voyelle et le nombre de langues (sur 451) qui possèdent la voyelle en question. Adapté de Karypidis & Montagu (2009).



systematique. En construisant deux ensembles de stimuli à l'intérieur de la catégorie /i/, elle mène d'abord une expérience de 'goodness rating'. Les deux ensembles sont organisés sous forme d'étoile, au centre desquels sont localisés un /i/ prototype et un /i/ non-typique. Les résultats montrent une structuration interne de /i/, avec certains stimuli recevant de meilleures notes que d'autres. Les expériences 2 à 4 du même travail démontrent que la typicalité a un impact sur la capacité des auditeurs à discriminer les sons. Ceci est le cas pour les enfants et les adultes, mais pas pour les singes. Le critère adopté par Kuhl est la différence en scores bruts (réponses correctes) de discrimination entre la condition Prototype et la condition Non-prototype. Elle postule alors que la généralisation est positivement liée à la typicalité et que cette corrélation est un phénomène ontogénétique ayant lieu pendant les premiers mois d'un enfant. Pendant cette période, les enfants se mettent en accord avec leur langue maternelle plutôt que d'adopter des contraintes locales universelles.

Lively & Pisoni (1997) réutilisent les stimuli de Kuhl. Alors qu'ils trouvent une gradation dans les résultats de 'goodness rating', certains stimuli étant évalués comme meilleurs exemplaires de la catégorie /i/, ce phénomène n'est pas robuste. Dans leur expérience de discrimination, ils utilisent deux ensembles, celui de Kuhl et

des ensembles personnalisés pour chaque auditeur. Un effet de généralisation n'est retrouvé pour aucun des deux groupes de stimuli. Les auteurs suggèrent que le critère d'interaction Condition*Distance n'est pas suffisant. Ils proposent deux critères supplémentaires : a. les scores des jugements de 'goodness rating' doivent être plus élevés pour certains membres de la catégorie que pour d'autres, et b. une corrélation négative entre typicalité et sensibilité. La troisième tâche consiste en une catégorisation des stimuli de Kuhl. Les réponses des sujets indiquent que les stimuli de Kuhl couvrent plusieurs catégories vocaliques et non pas seulement /i/.

Lotto *et al.* (1998) remettent en question le fait que dans les travaux sur l'effet des aimants perceptifs, les stimuli sont présentés en contexte isolé pour les tâches d'identification et d'évaluation, mais en paires pour la discrimination. Ils mènent alors des tâches d'identification et de discrimination en présentant les stimuli en paires. Le contexte a un impact important sur les scores d'identification. Les résultats de discrimination montrent que la discrimination est hautement prévisible par les scores d'identification et qu'un effet d'aimant n'est pas attesté. Les auteurs postulent que la capacité des sujets à discriminer mieux autour du non-prototype consiste simplement en des effets de frontières : il est plus facile de discriminer des stimuli appartenant à des catégories différentes qu'à la même catégorie.

Frieda *et al.* (1999) utilisent des ensembles personnalisés pour les prototypes de chaque auditeur. Alors que l'effet de Distance est significatif, ceci n'est pas le cas pour l'interaction Condition*Distance.

Iverson & Kuhl (2000) défendent l'hypothèse de généralisation en étudiant la relation entre l'effet d'aimant et les frontières phonémiques. En menant une expérience de 'même-différent', ils comparent les résultats d'une tâche de discrimination 'fixed' à ceux d'une tâche 'roving'. Ils trouvent que l'étendue des stimuli n'a pas le même effet sur les pics de sensibilité autour des frontières que sur l'effet d'aimant. Ceci les amène à postuler que l'effet d'aimant ne provient pas du même mécanisme.

Un grand nombre de travaux sur les aimants perceptifs est publié (Iverson & Kuhl, 1995, 2000; Sussman & Lauckner-Morano, 1995; Thyer *et al.*, 2000, parmi autres). Ici, nous avons présenté un petit nombre d'articles afin d'explicitier le modèle et de formuler des hypothèses.

Tandis que Rosch explicite le rôle de l'ordre de présentation, en expliquant que la généralisation en faveur du prototype s'effectue quand celui-ci est présenté en second, Kuhl n'inclut pas le facteur Ordre dans ses travaux (Kuhl, 1991; Iverson & Kuhl, 1995, 2000). Nous supposons que la généralisation dans son modèle se met en place indépendamment de l'ordre de présentation.

2.2.5 Sous-spécification et *mismatch negativity*

Lipski *et al.* (2007) utilisent la méthode de *mismatch negativity* (MMN), qui enregistre les réponses de détections automatiques du cerveau. En classifiant les voyelles du turc et de l'allemand dans le cadre du modèle FUL ('Featurally Underspecified Lexicon', Lahiri & Reetz, 2002), ils prédisent qu'un changement vocalique est plus prononcé quand les deux segments comprennent des traits de surface et lexicaux mutuellement exclusifs. Le tableau 2.4 illustre les traits non-communs pour les paires de voyelles /i/-/y/ et /i/-/e/. Nous observons que les deux langues spécifient diffé-

Tableau 2.4: Les traits lexicaux et de surface des voyelles /i e y/ de l’allemand et du turc dans le cadre du modèle FUL (Lipski *et al.*, 2007).

Direction	Langue	Traits de surface et lexicaux	Résultat de <i>mismatch</i>
[e]→/i/	turc	[bas]-[_____]	pas de <i>mismatch</i>
	allemand	[_____]-[haut]	pas de <i>mismatch</i>
[i]→/e/	turc	[haut]-[bas]	<i>mismatch</i>
	allemand	[haut]-[_____]	pas de <i>mismatch</i>
[y]→/i/	turc	[lab, haut]-[_____]	pas de <i>mismatch</i>
	allemand	[lab, haut]-[haut]	pas de <i>mismatch</i>
[i]→/y/	turc	[haut]-[lab]	pas de <i>mismatch</i>
	allemand	[haut]-[lab, haut]	pas de <i>mismatch</i>

remment les voyelles moyennes : le turc les spécifie pour [bas] alors qu’en allemand, elles sont sous-spécifiées. Ainsi, un conflit sera-t-il produit dans la direction [i]→/e/ pour le turc mais pas pour l’allemand.

Des germanophones et des bilingues turc-allemand ayant le turc comme langue maternelle participent à cette expérience. Les auteurs indiquent une interaction significative de Langue*Direction, ainsi qu’un effet significatif d’Ordre pour la paire /i/-/e/ chez les bilingues. Il faut noter, pourtant, que les valeurs F des ANOVA citées ne sont pas très élevées ($F = 6.28$ et $F = 5.39$, respectivement) et que l’étude souffre apparemment d’un faible pouvoir statistique.

2.3 Modèles génériques

2.3.1 Tversky (1977) et les asymétries de similarité

Le travail de Tversky (1977) sur la similarité est devenu classique et sert d’exemple pour les travaux qui ont suivi. L’auteur remet en question les modèles géométriques qui représentent les objets comme des points sur un espace de coordonnées de façon à ce que les dissimilarités entre les objets correspondent à des mesures de distance. Il reprend les trois axiomes qui caractérisent les distances métriques : la minimalité, la symétrie et l’inégalité du triangle. Soit δ la distance entre deux objets. Il est traditionnellement postulé que :

$$\delta(a, b) \geq \delta(a, a) = 0 \quad (2.1)$$

$$\delta(a, b) = \delta(b, a) \quad (2.2)$$

$$\delta(a, b) + \delta(b, c) \geq \delta(a, c) \quad (2.3)$$

L’axiome de minimalité (2.1) présuppose que les rejets corrects sont stables pour tous les stimuli et que les DC sont plus élevées que les (FA) rejets corrects. Ceci ne paraît pas être le cas. Comme nous l’avons mentionné, Schwartz & Escudier (1989) citent que les FA — et, par conséquent, les RC — dépendent de la saillance des stimuli : plus un stimulus est saillant, plus sa trace est stable au niveau de la mémoire à court terme et moins il déclenche de fausses alarmes. Quant à la supériorité des

DC, les figures 2.1 et 2.2 indiquent que ceci n'est pas validé par les données. Si la reconnaissance est basée sur un processus de similarité, l'axiome de minimalité est infirmé.

L'axiome de symétrie est celui qui paraît le plus erroné. Soit deux objets, a et b. Quand on dit "a est comme b", a est le sujet et b le référent (le point de référence). Le rôle de chaque objet est souvent défini par leur saillance. Un locuteur préfère la phrase "le fils ressemble à son père" que "le père ressemble à son fils". Ainsi, l'objet le plus saillant sert de référent.

Dans d'autres cas, les deux directions sont naturelles mais apportent un sens différent. "Cette pièce de théâtre est comme la vie" signifie que la pièce est réaliste. "La vie est comme une pièce de théâtre" signifie que les gens jouent des rôles. Cette différence de sens entre les deux phrases est due à l'implémentation de différents traits. Vu que, dans ces exemples, nous n'exploitons pas les mêmes traits pour évaluer la similarité entre les deux objets, des asymétries sont probables.

Tversky cite un exemple de William James pour démontrer que le principe d'inégalité peut être falsifié. Considérons la similarité de pays : la Jamaïque est similaire à Cuba à cause de leur proximité géographique ; Cuba est similaire à la Russie à cause de leur affinité politique ; mais la Jamaïque et la Russie ne sont pas du tout similaires. La similarité n'est donc pas transitive. Dans le cas de la perception auditive, le principe d'inégalité est valide. Soient trois sons purs a, b et c, à 300, 320 et 340 Hz respectivement. Si on compare ces stimuli, nous observons que l'inégalité est valable pour ce type de stimuli.

Un point intrigant que l'auteur aborde concerne la relation entre la similarité et la différence. Il est généralement considéré que les deux mécanismes sont négativement corrélés : plus deux objets sont similaires, moins ils sont différents. Tversky remet ce postulat en question et soutient que les deux tâches ne vont pas attribuer le même poids aux traits. Les sujets vont privilégier les traits communs dans une tâche de similarité, et les traits discordants dans une tâche de discrimination.

Un dernier point de l'article de Tversky qui est lié à notre travail est l'idée de saillance. L'auteur décompose la saillance d'un trait en deux éléments : l'intensité et la diagnosticité. L'intensité est déterminée par des facteurs perceptifs et cognitifs qui sont indépendants du contexte et qui accroissent le ratio signal/bruit : la fraîcheur d'une lumière, l'amplitude d'une tonalité, la saturation d'une couleur. La diagnosticité d'un trait est déterminée par l'importance des classifications qui sont basées sur ces traits, et est liée au contexte. Le trait 'réel' n'a pas de valeur diagnostique pour l'ensemble des animaux existants puisqu'il est partagé par tous les membres de la catégorie. Les sujets ne peuvent donc pas l'utiliser pour catégoriser ces animaux. Si, en revanche, l'ensemble contient des animaux existants et mythiques, le trait 'réel' obtient la valeur diagnostique.

Une transposition des éléments de la saillance sur la perception de voyelles est possible. L'intensité concerne des mécanismes auditifs de bas niveau (sensoriels) et fait partie de la structure inhérente du stimulus. Stevens (1972, 1989) postule que les langues privilégient les régions articulatoires quantiques afin de garantir la stabilité acoustique et articulatoire des segments. Dans une région quantique, une précision articulatoire n'est pas nécessaire puisque des perturbations des organes ne produisent pas d'effets acoustiques significatifs. Cette stabilité acoustique aura comme effet des

déviations standard relativement basses pour ces segments, ce qui peut amener à une meilleure reconnaissance. De même, les segments quantiques sont à la fois focaux parce qu'ils se localisent sur des endroits de changement d'affiliation de formants. Ceci conduit au rapprochement de deux formants consécutifs et, par conséquent, à une focalisation. Comme nous venons de le voir, les segments focaux sont considérés comme plus stables et produisent moins de FA.

L'expérience 2 de Repp & Crowder (1990) apporte des éléments en faveur de la valeur diagnostique des traits. Ils mènent trois tâches séparées. Chaque tâche comprend les stimuli de la catégorie /ε/, ainsi que les stimuli d'une des trois catégories /ɪ/, /æ/ ou /ə/. Les résultats indiquent que la discriminabilité des membres de la catégorie /ε/ dépend de la dispersion de l'ensemble des stimuli. Les effets d'ordre montrent aussi que la localisation du point de dégradation est déplacé en fonction des autres stimuli inclus dans la même tâche.

2.3.2 L'erreur de l'ordre temporel

Étant donné la nature différente des catégories sémantiques mentionnées ci-dessus nous avons besoin d'un modèle qui soit suffisamment générique afin de justifier les asymétries attestées. Hellström (1979, 1985, 1989, 2003), travaillant sur l'erreur de l'ordre temporel, propose un modèle utilisant le poids de la sensation pour calculer la différence entre deux stimuli :

$$d = [s_1\psi_1 + (1 - s_1)\psi_{r1}] - [s_2\psi_2 + (1 - s_2)\psi_{r2}] \quad (2.4)$$

où d est la différence, ψ_1 et ψ_2 le degré de sensation des stimuli, et ψ_{r1} et ψ_{r2} les niveaux de référence. Si $\psi_{r1} = \psi_{r2} = \psi_r$, la formule est réduite à :

$$d = s_1(\psi_1 - \psi_r) - s_2(\psi_2 - \psi_r) \quad (2.5)$$

Hellström (2008) étend son modèle pour rendre compte de l'effet des aimants perceptifs. Il propose alors une généralisation de son modèle (Hellström, 2008) aux stimuli multidimensionnels :

$$d_{12} = \left\{ \sum_{i=1}^p [s_1(\psi_{1i} - \psi_{ri}) - s_2(\psi_{2i} - \psi_{ri})]^m \right\}^{\frac{1}{m}} \quad (2.6)$$

avec p pour le nombre des dimensions perceptives prises en compte (comme les formants) et m la constante de Minkowski.

Malgré la plausibilité apparente du modèle de 'sensation-weighting' et son application réussie avec des stimuli de différentes natures, il serait très difficile à l'heure actuelle de l'appliquer sur les résultats de perception vocalique. Ceci est dû au fait qu'il nous serait impossible de proposer un nombre spécifique de paramètres servant à discriminer des sons vocaliques. Comme nous le verrons plus tard, la distance euclidienne seule sur le plan F1/F2 ne saurait justifier la sensibilité. Pour utiliser un modèle paramétrique, nous croyons qu'un travail approfondi sur la structure des voyelles en tant qu'objets auditifs doit être mené. Dans la conclusion de cette thèse, nous proposons de nouvelles voies que nous comptons suivre pour mieux cerner le phénomène des asymétries.

Une deuxième raison pour laquelle nous n'utilisons pas de formule générique est que la plupart des hypothèses sur la provenance des effets d'ordre présupposent l'utilisation d'informations *top-down*. Nous avons jugé nécessaire de recadrer ces hypothèses et de les réinterpréter afin d'évaluer leur capacité à prédire l'existence, le degré et la localisation d'asymétries.

2.4 Recadrage des théories et hypothèses formulées

Dans cette section, nous allons reprendre les modèles de discriminabilité présentés dans le chapitre précédent et les confronter aux diverses hypothèses sur la source des asymétries. Ceci nous permettra de formuler des hypothèses qui peuvent être formulées dans le cadre de chaque modèle sans que nous violions les principes de ce dernier.

2.4.1 Perception catégorielle

Le modèle de perception catégorielle postule que les sujets procèdent automatiquement à un étiquetage des stimuli, indépendamment de la tâche expérimentale. Cette étiquette leur permettra de catégoriser ou de discriminer les stimuli en ligne. Le modèle présuppose que chaque stimulus est traité séparément et qu'aucune interaction entre les stimuli n'est mise en action. Ce postulat signifie que pour une paire de stimuli :

1. la trace du premier stimulus dans le dépôt temporaire n'interfère pas avec le traitement du deuxième stimulus
2. le sujet ne construit pas d'image auditive pour les stimuli. Dans le cas contraire, cette image se dégrade très rapidement
3. une fois que le sujet analyse les indices acoustiques du premier stimulus et qu'il catégorise ce dernier, le contenu de la mémoire à court terme se dégrade à un rythme très rapide (un effacement quasi-total doit avoir lieu avant la présentation du deuxième stimulus, à savoir en moins d'une seconde) pour éviter des interférences avec le traitement du deuxième stimulus

Nous avons vu plus haut que les trois postulats ne sont pas plausibles. De même, le modèle ne prédit pas d'asymétries. À titre d'exemple, nous allons considérer ici que des interférences sont possibles mais que les étiquettes sont tout de même le seul critère de discrimination de deux stimuli. Nous attendons alors une relation étroite entre les effets de contexte en identification et les effets d'ordre en discrimination.

Une voyelle périphérique dans l'espace acoustique est également extrême au niveau articulatoire. La voyelle /i/, par exemple, implique une élévation et un avancement extrêmes de la langue. Nous imaginons qu'il est plus facile au locuteur d'identifier une voyelle périphérique parce que la configuration de l'appareil phonatoire qui lui est assignée dans la mémoire à long terme est plus facile à repérer. Cette facilité produira moins de fautes d'identification, à savoir plus de RC et moins de FA. Une voyelle extrême est théoriquement plus stable au niveau articulatoire. Une perturbation importante dans la région de [i] produira la semi-voyelle [j]. Mais dans la région de [e], une telle précision n'est pas nécessaire. La précision articulatoire

renvoie à la théorie quantique de Stevens. Et puisqu'un segment quantique est aussi focal, les segments focaux auront un statut particulier.

Les prototypes seront privilégiés puisque leur reconnaissance est plus facile. Une corrélation négative entre la typicalité et les temps de réaction est attendue. De nouveau, les FA seront corrélés négativement à la typicalité.

L'Intervalle Interstimulus ne joue probablement pas de rôle. Étant donné le taux très rapide de dégradation de la mémoire à court terme et de la trace mémorielle, des différences pour des IIS supérieurs à 500 ms ne seront pas attestées.

Les asymétries pouvant s'exprimer uniquement à travers des stratégies d'identification et étant donné que les scores d'identification sont stables à l'intérieur des catégories phonémiques, la seule région présentant des fluctuations dans un continuum est la frontière phonémique.

2.4.2 Modèle unifié pour parole et non-parole

Ce modèle avance que des représentations linguistiques, emmagasinées dans la MLT, ne sont pas nécessaires dans le traitement des stimuli. Ceci signifie que le facteur de l'expérience linguistique ne joue pas de rôle dans les tâches d'identification et de discrimination. Si des divergences sont trouvées parmi différents groupes linguistiques, elles seront dues à un traitement différentiel des stimuli au niveau de la mémoire à court terme.

Toute différence entre les stimuli n'est due qu'à leur nature inhérente. Les voyelles périphériques sont des candidats pour des effets d'ordre. Les prototypes ne servent pas de points d'ancrage.

2.4.3 Le modèle du processus duel

Ce modèle postule que l'identification des segments se fait automatiquement. Quand les stimuli à discriminer sont jugés appartenir à des catégories différentes, le sujet fait recours à la mémoire de timbre. Nous nous attendons à ce que les réponses 'identique' soient plus rapides que les réponses 'différent'.

Les asymétries peuvent être dues :

- au déplacement de la frontière phonémique : plus le premier stimulus est saillant, plus la frontière sera déplacée, ce qui aura comme effet des asymétries en identification.
- au taux différentiel de la dégradation de la mémoire pour timbre : puisque la mémoire à court terme phonémique est considérée comme relativement stable et que son contenu peut être retenu pendant plusieurs secondes, c'est la dégradation de la mémoire pour timbre qui produira des effets d'ordre.

Dans le premier cas, la durée de l'IIS n'aura pas d'impact sur les effets d'ordre. Une corrélation entre les effets de contexte et les effets d'ordre est prédite. Dans le second cas, plus l'IIS est long, plus la dégradation du premier stimulus sera importante et plus le sujet sera obligé de baser son jugement sur le deuxième stimulus. Ainsi, plus le deuxième stimulus sera saillant, plus les asymétries seront prononcées.

2.4.4 La théorie trace-contexte

Le modèle trace-contexte implémente le concept de variance de trace. Nous avons expliqué plus haut que la variance de trace est positivement corrélée à la durée de l'Intervalle Interstimulus. Nous supposons que les voyelles focales résisteront davantage à cette variance.

Le point de référence peut être lié à différents processus. Ce point peut être localisé sur les extrémités d'un continuum comme Durlach et Braida le proposent, sur la frontière phonémique ou sur des régions quantiques. Il peut aussi correspondre aux prototypes catégoriels. Des asymétries peuvent se trouver autour de n'importe laquelle des régions mentionnées. Plus un stimulus est proche du point d'ancrage, plus il sera facile à l'auditeur de mieux stocker sa position dans l'espace perceptif qu'il a reconstruit.

2.4.5 Le modèle à deux étapes de Kroll & Hershenson (1980)

La seule hypothèse que nous pouvons formuler ici concerne la stabilité des traces auditives qui serviront à la discrimination des stimuli. Les mêmes contraintes mémorielles s'appliqueront ici.

Tableau 2.5: Les hypothèses formulées dans le cadre des différents modèles de discriminabilité.

Hypothèses	Modèles				
	CP	Modèle unifié	Duel	Trace	2 étapes
Impact de l'IIS sur les effets d'ordre		x	x	x	x
L'IIS ne joue pas de rôle	x				
Corrélation entre effets de contexte et effets d'ordre	x		x	x	
Moins de FA pour les voyelles focales	x	x	x	x	x
Moins de FA pour les voyelles périphériques	x				
Corrélation négative entre typicalité et FA	x			x	
Effets d'ordre quand le point de référence est en seconde position	x	x	x	x	x
Statut privilégié de la périphérie	x			x	
Point de référence : focalisation		x	x	x	x
Point de référence : prototype	x				
Point de référence : frontières phonémiques				x	
Point de référence : région quantique	x			x	

2.4.6 Récapitulatif des hypothèses

Le tableau 2.5 présente les hypothèses que nous venons de formuler dans le cadre de chaque modèle de discriminabilité. En analysant chaque théorie, nous constatons que les prédictions plausibles ou interdites dépendent des contraintes du modèle. Nous nous apercevons vite que l'adoption d'un cadre théorique dans un travail est d'une importance primordiale. Nos expériences portent sur les facteurs que nous

avons décrits et nous essayons de fournir des arguments pour valider l'un ou l'autre modèle. Nous allons donc tenter de vérifier certaines de ces hypothèses dans les chapitres suivants.

Deuxième partie

Partie expérimentale

Chapitre 3

Asymétries et expérience linguistique

"One can judge from experiment, or one can blindly accept authority. To the scientific mind, experimental proof is all important and theory is merely a convenience in description, to be junked when it no longer fits. To the academic mind, authority is everything and facts are junked when they do not fit theory laid down by authority."
"Doctor Pinero" in *Life-Line* (1939)

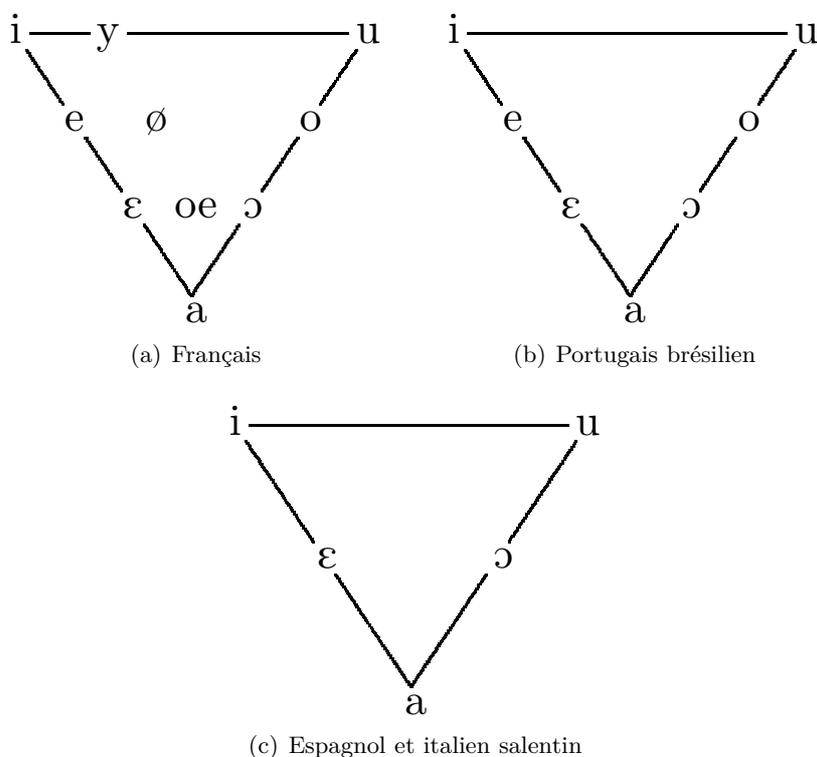
Dans le Chapitre 2, nous avons présenté les différentes hypothèses sur l'origine des asymétries. Or, au lieu d'examiner chacun des facteurs implémentés au sein de chaque théorie, nous avons décidé de commencer par un facteur qui pourrait déclencher des effets d'ordre : l'expérience linguistique de l'auditeur. Notre choix a suivi la logique ci-après. Si des asymétries robustes sont retrouvées chez les locuteurs de différentes langues, ceci indique que le déclencheur est un facteur lié aux processus cognitifs communs à tous les humains (procédures de la mémoire à court ou à long terme ou propre au système auditif). La périphéricité, la focalisation ou la neutralisation seraient des candidats plausibles qui mériteraient une étude plus approfondie. Si, en revanche, des asymétries sont retrouvées chez des locuteurs venant d'un système linguistique donné mais pas d'un autre, le déclencheur est possiblement lié à la façon dont les locuteurs de chaque groupe perçoivent les voyelles et aux indices qu'ils attribuent à celles-ci.

Afin de déterminer le rôle de l'expérience linguistique, nous avons mené deux études séparées. Dans la première, nous avons étudié le patron des réponses des locuteurs de quatre langues¹ : le français et le portugais brésilien d'une part (avec quatre hauteurs distinctives), l'espagnol et l'italien salentin d'autre part (avec trois hauteurs). La configuration des quatre systèmes est disponible dans la Figure 3.1.

Ce choix n'est certainement pas aléatoire. Étant donné que les deux groupes de langues spécifient différemment la hauteur de la langue (voir page 53) pour /i/ et /e/, qui font l'objet principal de cette thèse, nous nous attendons à ce que des asymétries soient attestées seulement pour l'espagnol et l'italien.

1. Nous utilisons ici le terme 'langue' de façon abusive et pour des raisons pratiques. L'italien salentin est plutôt un dialecte de l'italien et non pas une langue à part entière

FIGURE 3.1: Les systèmes vocaliques des quatre langues étudiées.



L'expérience linguistique peut avoir un impact sur les indices liés à la perception des voyelles d'une langue. Dans un système dont l'axe des voyelles antérieures est très chargé (plus de trois hauteurs), la différence entre les valeurs des deux premiers formants de deux voyelles adjacentes n'est parfois pas suffisante pour discriminer celles-ci, ce qui inciterait les auditeurs à utiliser des indices additionnels, tels que $F2'$, pendant des tâches d'identification et de discrimination. Selon cette logique, des effets d'ordre seraient retrouvés seulement pour le français et le portugais. Il est à rappeler ici que notre thèse n'a pas comme objet l'examen des indices permettant la perception de sons vocaliques. Aussi n'allons-nous pas vérifier si $F2'$ est, en effet, un indice perceptif robuste.

Deux expériences ont été menées pour cette première étude. Dans l'expérience 1, les participants ont catégorisé les dix stimuli d'un continuum $[i]-[e]$, préparés avec un modèle articulatoire et présentés en contexte isolé. L'expérience 2 a consisté en une tâche 'même-différent'. Les auditeurs ont discriminé des voyelles présentées en paires et qui différaient d'un ou de deux pas sur le continuum. Des effets d'ordre systématiques ont été détectés pour le français et l'italien, ainsi que des effets mineurs pour l'espagnol.

Pour notre deuxième étude, nous avons utilisé un autre ensemble de stimuli. Nous avons synthétisé les trente-trois stimuli 'prototypes' de Schwartz *et al.* (1997) et avons mené un test d'identification AXB avec des locuteurs de quatre langues à cinq voyelles $[i \text{ a } u \text{ e } o]$: espagnol, grec, italien salentin et griko. Les stimuli A et B (qui ont servi de standard) de chaque paire étaient les deux voyelles périphériques

Tableau 3.1: Nombre (N) d’auditeurs par langue, étendue d’âge, moyenne d’âge (M) et déviation standard (ds).

Langue	Région	N	Étendue	M	ds
Portugais Brésilien (PBR)	Rio	20	20-27	23.60	1.80
Français (FR)	Paris	34	18-51	30.88	8.99
Italien Salentin du Sud (ISS)	Salento	23	20-33	28.00	3.09
Espagnol (ES)	Andalousie	30	20-53	29.03	8.54

sur une hauteur donnée, tandis que X était une voyelle non-périphérique de la même hauteur, par exemple i-y-u, i-i-u. Les voyelles périphériques ont été présentées dans les deux ordres : AXB et BXA. Nous avons ainsi essayé de vérifier si l’un des deux axes, ‘F2-F1 maximum’ ou ‘F2-F1 minimum’, exerçait un pouvoir plus important sur le stimulus à classifier.

3.1 Expérience 1 : Catégorisation ²

Ce test de catégorisation à un intervalle nous a permis de : a. filtrer les auditeurs afin de ne garder que ceux dont la courbe d’identification avait une forme S, b. délimiter les deux catégories phonétiques afin de pouvoir les traiter séparément dans la section suivante, et c. détecter des différences éventuelles au niveau de l’emplacement de la frontière inter-catégorielle.

3.1.1 Participants

107 auditeurs au total ont participé à cette expérience. Tous ont indiqué être locuteurs natifs et ne pas avoir de trouble langagier. Les informations concernant les quatre groupes sont disponibles dans le tableau 3.1.

3.1.2 Stimuli

Les stimuli ont été synthétisés avec le modèle articulatoire de Maeda (1990), VTCALCs. Le modèle propose, pour chacune des voyelles du français, une configuration de l’appareil phonatoire que le créateur du logiciel a choisie parmi plusieurs, en se basant sur la ressemblance articulatoire, acoustique et perceptive avec un bon exemplaire de leurs catégories respectives. La simplicité de modélisation a également été un critère.

Nous avons commencé par les configurations de [i] et de [e] (Figure 3.2) et avons utilisé les valeurs par défaut des sept paramètres articulatoires implémentés dans le modèle. Les lignes 1 et 10 du tableau 3.2 nous informent que [i] et [e], dorénavant stimuli 1 et 10 respectivement, ne diffèrent que par la hauteur de la mâchoire et par la position de la langue.

En interpolant de façon linéaire les paramètres présentés dans le tableau 3.2, nous avons préparé trois continua : un où la hauteur de la mâchoire et la position

2. Des résultats partiels des expériences 1 et 2 ont été publiés dans Karypidis *et al.* (2006b), Karypidis *et al.* (2007) et Karypidis *et al.* (2008).

FIGURE 3.2: Les configurations de l'appareil phonatoire proposées par VTCALCs pour [i] et [e] respectivement. À gauche : la cavité labiale.

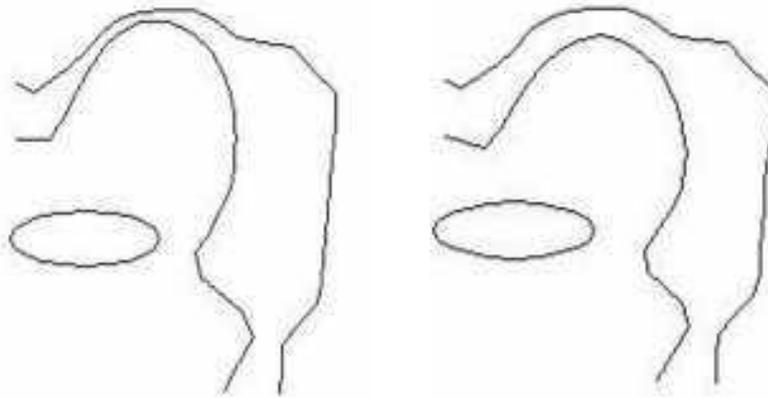


Tableau 3.2: Les paramètres articulatoires des dix stimuli synthétisés. Les chiffres correspondent à des valeurs standardisées des paramètres articulatoires. HM : Hauteur de la mâchoire; PL : Position de la langue; FL : Forme de la langue; PA : Position de l'apex; HLe : Hauteur des lèvres; PLe : Protrusion des lèvres; HLa : Hauteur du larynx

No. du stimulus	HM	PL	FL	PA	HLe	PLe	HLa
1	0.500	-2.00	1	-2	1	-1	0
2	0.444	-1.88	1	-2	1	-1	0
3	0.388	-1.77	1	-2	1	-1	0
4	0.332	-1.66	1	-2	1	-1	0
5	0.276	-1.55	1	-2	1	-1	0
6	0.220	-1.44	1	-2	1	-1	0
7	0.164	-1.33	1	-2	1	-1	0
8	0.108	-1.22	1	-2	1	-1	0
9	0.052	-1.11	1	-2	1	-1	0
10	0	-1.00	1	-2	1	-1	0

de la langue ont été modifiées simultanément, et deux autres où chacun des deux paramètres a été modifié séparément. Dans chaque continuum, le stimulus 1 était le plus périphérique (avec le F1 le plus bas et le F2 le plus élevé) et le stimulus 10 ressemblant le plus à un /e/. Pour tous les stimuli, la fréquence fondamentale (F0) croissait de 121 à 130 Hz pendant le premier tiers pour ensuite chuter à 100 Hz pendant les deux derniers tiers. La durée totale des stimuli a été fixée à 350 ms. Le taux d'échantillonnage était de 11025 Hz et la profondeur à 16 bit. Les stimuli ont été sauvegardés en fichiers .wav mono et en format PCM (Pulse Code Modulation).

Nous avons produit des estimations des deux premiers formants de tous les stimuli (tableau 3.3) afin que le continuum le plus apte à nos besoins soit défini. L'extraction des formants a été effectuée avec le logiciel Praat (Boersma & Weenink, 2001). La fréquence fondamentale n'étant pas stable pendant le stimulus, une estimation des formants sur tout le stimulus (longueur de la fenêtre : 350 ms) et non pas au point médian a été choisie. Ensuite, la différence moyenne en F1 (Δ_{F1}) et F2

Tableau 3.3: Valeurs de F1 et F2 (en Hertz) et différence moyenne entre des stimuli consécutifs (dernière ligne) pour les trois continua. Les paramètres modifiés à chaque fois ont été : la hauteur de la mâchoire (colonnes 2 et 3), la position de la langue (colonnes 4 et 5), ou les deux (colonnes 6 et 7).

Stimulus	Mâchoire		Langue		Mâchoire + Langue	
	F1	F2	F1	F2	F1	F2
1	249	2266	249	2265	249	2265
2	252	2264	260	2251	262	2248
3	255	2261	270	2236	276	2229
4	257	2258	281	2219	291	2206
5	260	2256	293	2200	306	2182
6	263	2251	304	2180	322	2156
7	265	2248	316	2158	337	2126
8	268	2245	329	2135	352	2095
9	272	2241	339	2109	365	2061
10	274	2238	350	2082	378	2024
Différ. moyenne	2.8	3.1	11.2	20.3	14.3	26.8

Tableau 3.4: Valeurs formantiques (F1–F4) en Hertz (colonnes 2–5) et en Bark (colonnes 6–9) des dix stimuli synthétisés.

Stimulus	Hertz				Bark			
	F1	F2	F3	F4	F1	F2	F3	F4
1	249	2265	3122	3618	2.49	13.84	15.94	16.86
2	262	2248	3046	3635	2.63	13.79	15.78	16.89
3	276	2229	2937	3667	2.78	13.74	15.55	16.94
4	291	2206	2831	3698	2.94	13.67	15.31	16.99
5	306	2182	2752	3720	3.09	13.59	15.13	17.03
6	322	2156	2686	3728	3.25	13.51	14.97	17.04
7	337	2126	2638	3738	3.40	13.42	14.85	17.06
8	352	2095	2596	3738	3.55	13.32	14.75	17.06
9	365	2061	2563	3743	3.68	13.21	14.66	17.07
10	378	2024	2529	3733	3.80	13.09	14.57	17.05

(Δ_{F_2}) pour chaque continuum a été calculée. Le tableau 3.3 nous informe que quand un seul paramètre articulatoire (hauteur de la mâchoire ou position de la langue) est modifié à la fois, la différence moyenne entre des stimuli consécutifs (dernière ligne du tableau) est bien inférieure aux seuils de discrimination proposés par Flanagan (1957) : ± 20 Hz pour F1 et ± 50 Hz pour F2. Ainsi, nous avons retenu le troisième continuum pour nos expériences (colonnes 6 et 7 du même tableau).

Il est à noter que même à l'intérieur de ce troisième continuum, $\Delta_{F_1} < 20$ Hz et $\Delta_{F_2} < 50$ Hz. Par conséquent, les paires différant de deux pas sur le continuum ont été incluses (voir Expérience 3.2).

Les paramètres articulatoires des dix stimuli retenus sont présentés dans le tableau 3.2, et leurs valeurs formantiques en Hertz et en Bark dans le tableau 3.4. Avant d'être présentés aux auditeurs, ils ont été appariés en énergie RMS à -10 dB à l'aide de Sound Forge 6.0.

Il ne fait aucun doute que les deux extrémités du continuum ne constituent pas

Tableau 3.5: Les étapes du test de catégorisation.

Étape	Durée (en ms)
1. Présentation du stimulus	350
2. Affichage sur l'écran des deux réponses possibles	—
3. Le sujet répond par un clic	—
4. Intervalle inter-trial	500
5. Pause (tous les 20 items)	—

des exemplaires typiques des catégories /i/ et /e/ du français, étant donné que les auditeurs d'une langue ne partagent pas nécessairement les mêmes représentations typiques [voir Frieda *et al.* (1999) pour la dispersion des prototypes de /i/ pour des locuteurs de l'anglais américain]. Il en va de même que ce n'est pas la seule façon de construire un continuum franchissant les deux catégories /i/ et /e/. Il est possible d'élever F1 en ouvrant uniquement la mâchoire. Cependant, les valeurs articulatoires attribuées aux stimuli seraient très similaires à celles d'un [a] et par là même, non-réalistes.

3.1.3 Procédure

Les stimuli ont été présentés en stéréo via des écouteurs dans de petites salles calmes et l'expérience a été menée sur différents ordinateurs portables par Angelica V. Costagliola pour l'italien, Gilles Guglielmi pour le français, Antonia Colazo-Simon pour l'espagnol et Cirineu Cicote Stein pour le portugais. Nous avons choisi de manière intuitive un niveau d'intensité approximatif (autour des 70 dB SPL) qui a été identique pour les locuteurs de la même langue. Cependant, l'intensité a pu être légèrement différente, compte-tenu de la variabilité des cartes son utilisées dans les quatre groupes.

Praat a servi d'interface. Les stimuli ont été présentés en 7 blocs avec la fonction de randomisation 'PermuteBalancedNoDoublets' de Praat, qui permet 1) d'éviter les répétitions consécutives du même stimulus et 2) de garantir l'équilibrage la présentation des stimuli entre les blocs.

Deux réponses ont été proposées, <i> et <é> pour FR, <i> et <e> pour ES, ISS et PBR. Pour que les auditeurs comprennent que le deuxième symbole correspondait à la voyelle antérieure mi-fermée /e/ et non pas à la mi-ouverte /ɛ/, nous leur avons fourni des explications et des exemples au début du test. Les deux réponses possibles ont été présentées à l'écran à l'intérieur de deux larges carrés jaunes, sur lesquels les auditeurs étaient censés cliquer pour catégoriser le stimulus. Après chaque réponse, le stimulus suivant a été présenté avec un délai de 500 ms. Une pause a été introduite tous les deux blocs — 20 stimuli — pendant laquelle les participants étaient censés cliquer sur l'écran pour reprendre le test. L'expérience, qui a duré approximativement 5 minutes au total, a été précédée d'une session courte d'entraînement, où les dix stimuli ont été présentés deux fois avec la même méthode de randomisation. Une description des différentes étapes de la procédure est disponible dans le tableau 3.5.

FIGURE 3.3: Fonctions d'identification pour les quatre groupes. La ligne horizontale désigne le point de hasard.

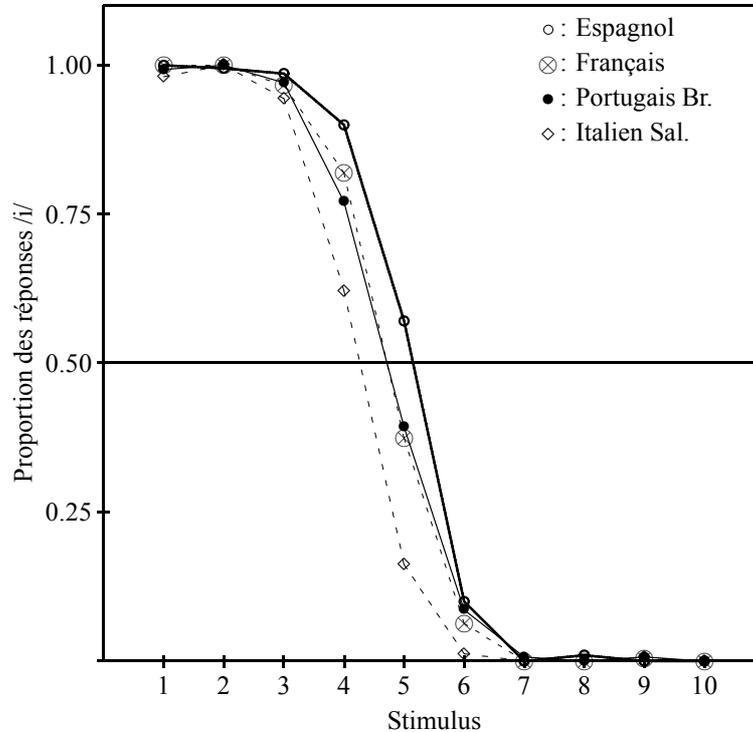


Tableau 3.6: Tableau des moyennes de la frontière sur le continuum pour les quatre groupes.

Langue	N	M	sd
PBR	20	4.78	.73
FR	34	4.72	.64
ISS	23	4.19	.52
ES	30	5.09	.49

3.1.4 Résultats

Un test binomial (Uitenbroek, 1997) pour chaque groupe a montré qu'aucun des stimuli (à part le stimulus 5 pour ESP) n'a été localisé dans la zone du hasard ($\alpha = 0.01$). Ceci a comme résultat la discontinuité des courbes d'identification dans la figure 3.3. Cette transition brusque a été retrouvée chez tous les auditeurs qui ont, par conséquent, été retenus pour l'expérience 2.

Nous avons ensuite estimé l'emplacement de la frontière catégorielle (angl. 'cut-off point'), qui correspond au croisement de la fonction d'identification avec la ligne horizontale indiquant le hasard, conçu comme un point précis, dans la figure 3.3. La méthode choisie a été celle de l'interpolation linéaire des scores se trouvant des deux côtés de la frontière. Nous trouvons les statistiques descriptives de ces données dans le tableau 3.6.

Les valeurs de frontière ont été utilisées comme input dans une analyse de

Tableau 3.7: Résultats de PLSD ($\alpha = 0.01$) pour comparaison de frontière entre les quatre groupes.

Groupes	Différence moyenne	Différence critique	p
es, fr	.370	.392	.015
es, it	.899	.433	<.01
es, pt	.313	.452	.072
fr, it	.529	.422	<.01
fr, pt	-.058	.441	.733
it, pt	-.586	.478	<.01

Kruskal-Wallis, qui a révélé un effet significatif pour le facteur ‘Groupe’ [$K(3 = 23.369, p < .01)$]³. Un test PLSD de Fisher (tableau 3.7) a indiqué que la frontière pour ISS a été significativement différente de celui pour FR ($p < .01$), PBR ($p < .01$) et ES ($p < .01$).

3.2 Expérience 2 : Discrimination

Cette expérience consiste en une tâche ‘même-différent’, où les stimuli de l’expérience 1 ont été présentés en paires et dans les deux ordres, tout en différant d’un ou de deux pas sur le continuum. Le but de cette étude a été d’examiner si l’ordre de présentation des stimuli incitait les auditeurs à donner des réponses ‘différent’ plus souvent dans une direction de changement vocalique que dans l’autre.

3.2.1 Participants

Au total, 99 auditeurs de l’échantillon général ont participé. Deux auditeurs espagnols et un français ont ensuite été exclus à force d’avoir répondu ‘différent’ dans plus de 90% des cas.

3.2.2 Stimuli

Les stimuli ont été les mêmes que pour l’expérience 1.

3.2.3 Procédure

Les conditions externes ont été les mêmes que pour l’expérience 1. Aux auditeurs ont été présentés cinq blocs de 18 paires de voyelles différant d’un pas (9 combinaisons x 2 ordres) et de 16 paires de deux pas (8 combinaisons x 2 ordres), soumis en ordre aléatoire et sans doublets. Dans l’ordre ‘*Progressif*’, le premier stimulus présenté a été plus extrême (F1 plus bas, F2 plus élevé), tandis que dans l’ordre ‘*Régressif*’, le stimulus le plus extrême a été présenté en seconde position. Les paires de stimuli identiques n’ont pas été incluses, ne faisant pas partie des hypothèses initiales. Les différentes combinaisons de stimuli sont présentées dans le tableau 3.8.

3. Nous proposons ici la valeur K corrigée pour la présence de rangs ex-aequo, puisque des valeurs identiques ont été fréquemment rencontrées dans nos données.

Tableau 3.8: Les quatre groupes de paires de stimuli et leurs caractéristiques : ordre, différence en pas sur le continuum, étiquette (la façon dont le groupe est référencé dans le texte) et nombre de paires dans chaque groupe.

Groupe de paires	Ordre	Différence en pas	Étiquette	N de paires
1-2, ..., 9-10	Progressif	un	Pro ₁	9
2-1, ..., 10-9	Régressif	un	Reg ₁	9
1-3, ..., 8-10	Progressif	deux	Pro ₂	8
3-1, ..., 10-8	Régressif	deux	Reg ₂	8

L'Intervalle Interstimulus (ISI) a été fixé à 250 ms. Selon Cowan & Morse (1986), on observe des pics de discrimination quand l'ISI est fixé autour de cette valeur. Les sujets ont dû indiquer dans chaque cas si les voyelles dans chaque paire étaient 'absolument identiques' ou 'même légèrement différentes'. Après chaque réponse, la première voyelle de la paire suivante a été présentée avec un délai de 500 ms. Des pauses ont été introduites toutes les 15 paires. Les auditeurs devaient cliquer sur l'écran afin de reprendre la tâche. Les différentes étapes de ce test sont affichées dans le tableau 3.9.

Tableau 3.9: Les étapes du test de discrimination.

Étape	Durée (en ms)
1. Présentation du stimulus A	350
2. Intervalle Interstimulus	250
3. Présentation du stimulus B	350
4. Le sujet répond par un clic	—
5. Intervalle inter-trial	500
6. Pause (toutes les 15 paires)	—

L'expérience a été précédée d'une courte session d'entraînement pendant laquelle les 34 paires ont été présentées une fois et dans un ordre aléatoire.

3.2.4 Résultats

Une première hypothèse issue des différentes théories présentées dans le Chapitre 2 prédit que les scores de discrimination pour les paires où l'élément le plus périphérique est présenté en premier (groupes Pro₁ et Pro₂) seront inférieurs à ceux des paires dans l'ordre Régressive (groupes Reg₁ et Reg₂). En d'autres termes, les hypothèses principales sont :

$$H_1 : \text{Pro}_1 < \text{Reg}_1$$

$$H_2 : \text{Pro}_2 < \text{Reg}_2$$

Une ANOVA à mesures répétées sur les deux facteurs, Paire et Ordre, a révélé un effet significatif d'Ordre pour FR et ISS dans les deux conditions (un pas et deux pas) et pas d'effet pour PBR ou ES dans aucune des deux conditions. L'effet de Paire a été significatif pour les quatre groupes et dans les deux conditions. Les résultats d'ANOVA sont disponibles dans les tableaux 3.10 (condition d'un pas) et 3.11 (deux pas).

Tableau 3.10: Résultats d'ANOVA (mesures répétées sur deux facteurs : Ordre et Paire) pour la condition d'un pas ; ddl= degré de liberté.

Facteur	Langue	ddl	F	p
Ordre	FR	1,32	11.801	< .01
	ISS	1,14	25.895	< .01
	PBR	1,19	.199	.660
	ES	1,27	3.173	.086
Paire	FR	8,256	9.038	< .01
	ISS	8,112	11.817	< .01
	PBR	8,152	20.781	< .01
	ES	8,216	15.495	< .01
Paire*Ordre	FR	8,256	4.578	< .01
	ISS	8,112	5.068	< .01
	PBR	8,152	.458	.884
	ES	8,216	1.174	.091

Tableau 3.11: Résultats d'ANOVA (mesures répétées sur deux facteurs : Ordre et Paire) pour la condition de deux pas ; ddl= degré de liberté.

Facteur	Langue	ddl	F	p
Ordre	FR	1,32	14.484	< .01
	ISS	1,14	7.259	.05
	PBR	1,19	.353	.559
	ES	1,27	.273	.605
Paire	FR	7,224	20.468	< .01
	ISS	7,98	20.161	< .01
	PBR	7,133	35.981	< .01
	ES	7,189	24.765	< .01
Paire*Ordre	FR	7,224	9.124	< .01
	ISS	7,98	1.215	.302
	PBR	7,133	1.469	.184
	ES	7,189	3.497	< .01

Les ANOVA nous informent que H_1 est validée pour FR et ISS tandis que H_2 est validée pour FR et marginalement pour ISS (dépendant du seuil α). Un effet de Paire signifie que les fonctions de discrimination n'ont été plates pour aucune langue ni condition. Les interactions Paire*Ordre indiquent que les asymétries retrouvées pour une paire dépendent de l'emplacement de la paire sur le continuum.

La figure 3.4 présente les scores pour chaque paire dans les deux conditions et essaie d'éclaircir les interactions Paire*Ordre. Or, afin de mieux étudier le rôle de l'emplacement de chaque paire sur le continuum, nous avons fait des tests de rang de Wilcoxon, où nous avons comparé les scores des deux ordres pour chaque paire séparément. Les effets les plus robustes se trouvent dans l'extrémité 'gauche' du continuum et pour les trois ou quatre paires contenant des stimuli ressemblant le plus au /i/, et semblent décroître vers l'extrémité /e/. Ceci est très évident pour FR dans les deux conditions, ainsi que pour ES et ISS dans certains cas. Les résultats du PBR ne révèlent pas de perception asymétrique. Une deuxième tendance moins solide est l'effet d'ordre négatif (Reg < Pro) au milieu du continuum pour l'espagnol.

FIGURE 3.4: Fonctions de discrimination pour les quatre groupes. Pro : ordre progressif; Reg : ordre régressif.

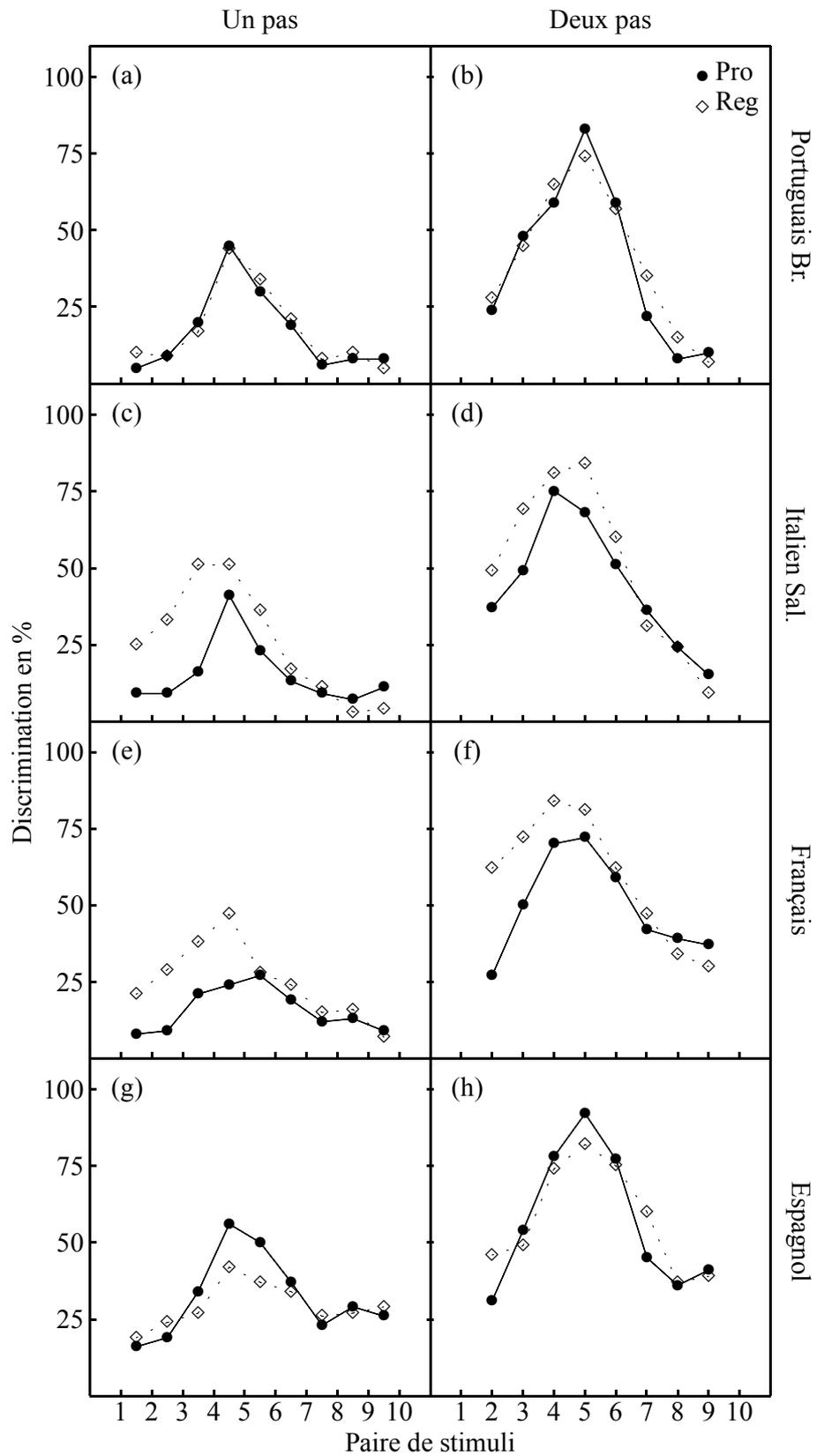


Tableau 3.12: Résultats des tests χ^2 2x2 pour les quatre langues : nombre de réponses ‘différent’, valeur χ^2 et valeur p .

	n	χ^2	p
FR	2038	8.17	<.005
ISS	853	5.33	<.05
PRB	947	0	1
ES	2056	1.42	0.233

Ceci est retrouvé dans les deux conditions.

Nous avons ensuite conduit des tests χ^2 2x2 pour déterminer la relation entre l’ordre de présentation (progressif ou régressif) et la distance acoustique (un pas ou deux pas). Le nombre de réponses ‘différent’ a servi de variable dépendante. Le tableau 3.12 indique qu’une relation est retrouvée pour le français et l’italien. En effet, le degré d’asymétrie diminue quand la distance acoustique est plus grande. Ceci est probablement dû au fait que dans la condition de deux pas, un effet de plafond est plus facilement atteint chez certains auditeurs dans les deux ordres.

Afin de généraliser sur les deux catégories vocaliques, nous avons groupé les paires contenant des membres de la même catégorie et avons refait des tests Wilcoxon, en considérant les variables Paire et Auditeur comme facteurs aléatoires. Seuls les stimuli identifiés à plus de 90% dans leur catégorie respective ont été inclus dans l’analyse. Dans la figure 3.3, les stimuli identifiés comme /i/ étaient les stimuli 1-3 pour le PBR, FR et ISS et les stimuli 1-4 pour ES. Les stimuli dans la catégorie /e/ étaient les stimuli 6-10 pour les quatre groupes. La figure 3.6 révèle que les seules asymétries attestées sont retrouvées à l’intérieur de la catégorie /i/ et ceci pour FR et ISS.

Il est possible que ces asymétries soient liées au postulat de Schwartz *et al.* (2005) selon lequel la focalisation, combinée avec la périphéricité, peut engendrer des asymétries. Nous avons alors estimé les valeurs de $F2'$ pour les dix stimuli. Ce moyen a été choisi parce que la focalisation est liée à la concentration d’énergie spectrale et que les formules d’estimation de $F2'$ utilisent la distance entre les formants pour trouver le centre de gravité du spectre.

Nous avons converti en Bark les valeurs brutes (en Hertz) des stimuli avant de les utiliser dans la formule de $F2'$ (Mantakas *et al.*, 1986). Nous avons utilisé cinq formules (3.1-3.5) de conversion de Hertz en Bark.

$$z = 6 \cdot \operatorname{arcsinh} \frac{f}{600} \quad (3.1)$$

$$z = 7 \cdot \operatorname{arcsinh} \frac{f}{650} \quad (3.2)$$

$$z = \frac{26.81}{1+1960} - 0.53 \quad (3.3)$$

$$z = 6.7 \cdot \operatorname{arcsinh} \frac{f-20}{600} \quad (3.4)$$

$$z = 8.7 + 14.2 \cdot \lg 10 \cdot \frac{f}{100} \quad (3.5)$$

où f représente la fréquence en Hertz et z la valeur de bande critique en Bark. Les résultats de la conversion sont disponibles dans le tableau 3.13.

FIGURE 3.5: Les effets d'ordre en % pour chaque paire avec une erreur standard. Les astérisques indiquent une différence significative telle qu'indiquée avec un test de rang de Wilcoxon. * : $\alpha = 0.05$, ** : $\alpha = 0.01$. À gauche : un pas ; à droite : deux pas. Du haut au bas : portugais brésilien, italien, français et espagnol.

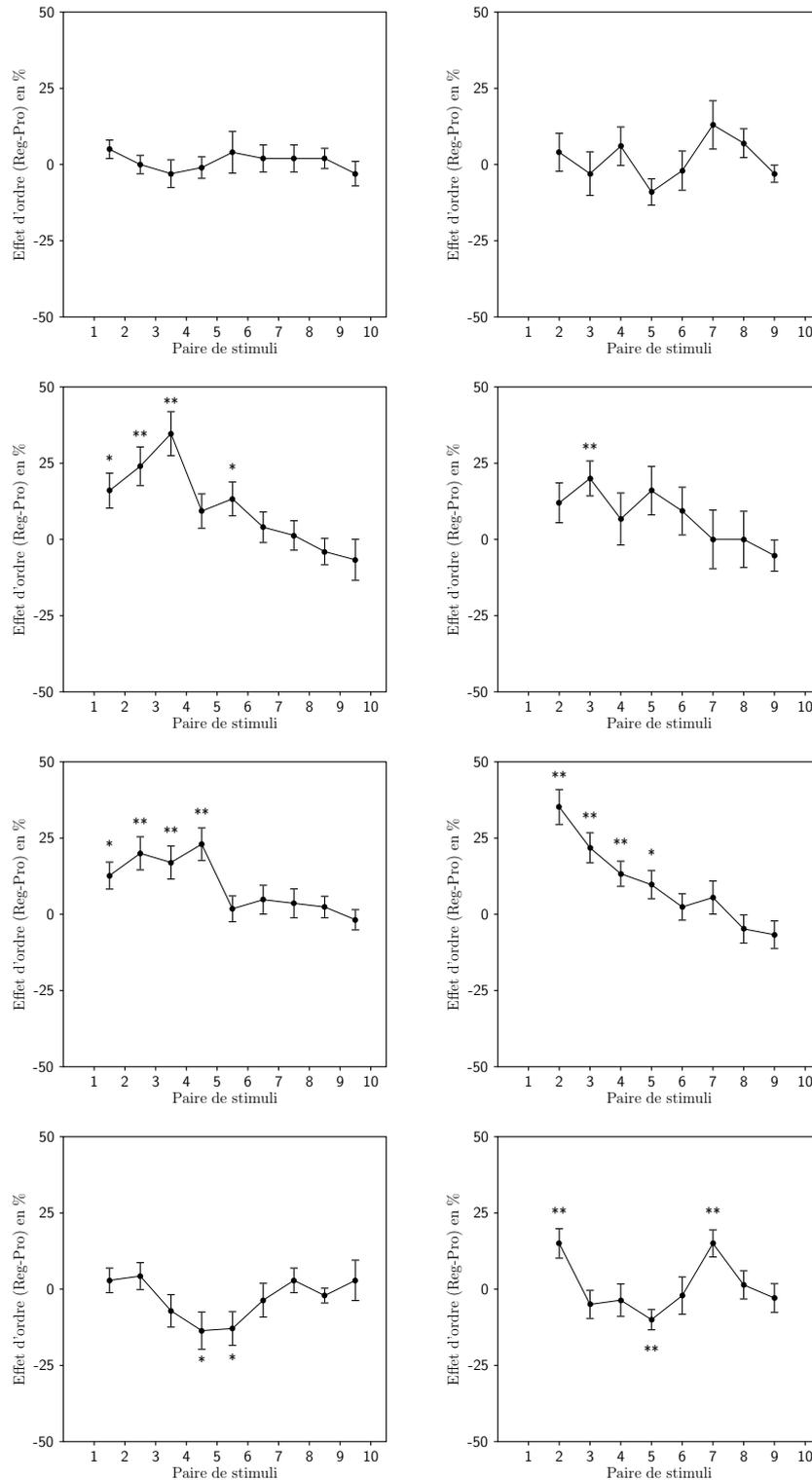
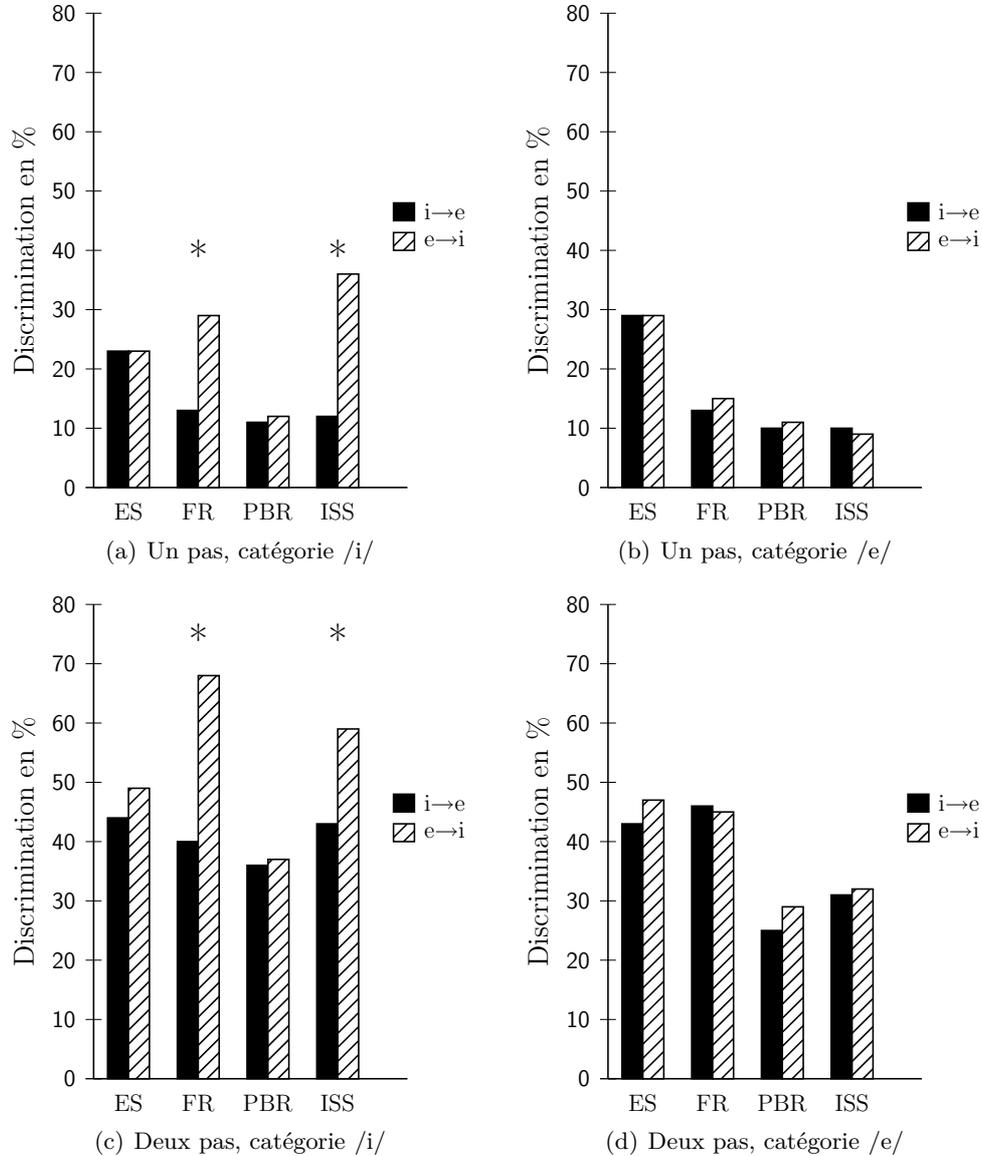


FIGURE 3.6: Comparaison des scores de discrimination pour les deux directions de changement vocalique. * : $\alpha = 0.01$ avec un test Wilcoxon.



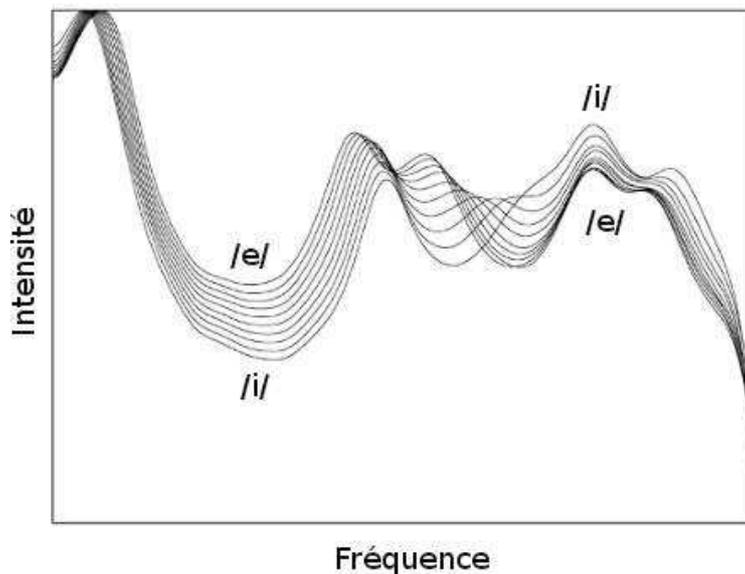
La figure 3.8 est une visualisation du tableau 3.13. Elle révèle que les cinq formules prédisent le même patron : l'évolution de $F2'$ est discontinue, avec une rupture considérable entre les stimuli 3 et 4. Cette non-linéarité est due au déplacement de $F3$ qui, à partir du stimulus 4, est plus proche de $F2$ que de $F4$, créant une prééminence spectrale avec le premier (figure 3.7). Nous constatons alors que la focalisation ne déclenche pas d'asymétries pour les locuteurs des quatre groupes.

Les faibles asymétries pour ES, coïncidant avec la frontière catégorielle, ont potentiellement comme source la spécification de la hauteur de la langue dans le système linguistique en question. Lipski *et al.* (2007), basés sur le modèle FUL (Lahiri &

Tableau 3.13: Les valeurs de $F2'$ estimées pour les dix stimuli en utilisant différentes formules de conversion Hertz-Bark. Les en-têtes des colonnes font référence aux formules 3.1-3.5 présentées plus haut, correspondant à celles proposées par Fourcin *et al.* (1977), Schroeder *et al.* (1979), Traummüller (1990), Tjmov (1971) et Zwicker & Terhardt (1980) respectivement.

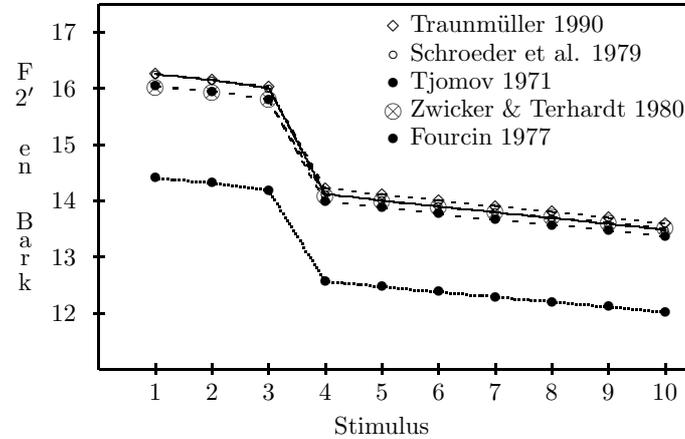
Stimulus	Formule				
	3.1	3.2	3.3	3.4	3.5
1	14.40	16.25	16.25	16.04	16.02
2	14.31	16.15	16.15	15.94	15.93
3	14.19	16.00	16.01	15.80	15.80
4	12.56	14.12	14.22	13.98	14.09
5	12.47	14.00	14.11	13.87	13.99
6	12.37	13.89	14.00	13.76	13.90
7	12.28	13.79	13.90	13.66	13.80
8	12.20	13.69	13.80	13.56	13.70
9	12.11	13.58	13.70	13.46	13.61
10	12.01	13.47	13.59	13.35	13.51

FIGURE 3.7: Représentation spectrale des dix stimuli.



Reetz, 2002), soutiennent que la transition de /i/ à /e/ consisterait en une exclusion mutuelle entre les traits de surface et les traits spécifiés pour le turc (à trois hauteurs vocaliques, comme l'italien et l'espagnol). Nous nous attendrions alors à des effets d'ordre négatifs ($Reg < Pro \Rightarrow /e/ \rightarrow /i/ > /i/ \rightarrow /e/$). Dans la direction inverse, les traits de surface ne sont pas en conflit avec les traits sous-jacents. En allemand, le non-conflit concerne les deux directions de changement vocalique.

En suivant la logique de Lipski *et al.* (2007), nous obtiendrions des effets d'ordre autour des frontières pour l'espagnol et l'italien, mais pas pour le français et le portugais. Un contre-exemple vient de l'italien. Or, il est possible que cette langue,

FIGURE 3.8: Représentation graphique des valeurs de $F2'$ estimées en utilisant différentes formules de conversion Hertz-Bark.

provenant de l'italien standard, un système à 7 voyelles et à quatre hauteurs, encode la hauteur de la langue différemment de l'espagnol.

3.3 Expérience 3 : Catégorisation AXB

Pour cette expérience, nous avons utilisé un nouvel ensemble de 33 stimuli dispersés dans tout l'espace vocalique. Des locuteurs de quatre langues à cinq voyelles /i a u e o/ : espagnol, grec, italien salentin et griko, ont été soumis à une tâche AXB. Le but a été d'examiner si l'un des deux axes, 'F2-F1 maximum' ou 'F2-F1 minimum', exerçait une force plus importante sur le stimulus à classifier.

3.3.1 Participants

85 auditeurs au total ont participé à cette expérience. Tous ont indiqué être locuteurs natifs et ne pas avoir de trouble langagier. Les informations concernant les quatre groupes sont disponibles dans le tableau 3.14.

Tableau 3.14: Groupe linguistique, étiquette sous laquelle le groupe apparaît dans le texte, nombre (N) d'auditeurs par langue, étendue d'âge, moyenne d'âge (M) et déviation standard (ds).

Langue	Étiquette	N	Étendue	M	ds
Espagnol	ES	30	20-53	29.0	8.7
Griko	GK	14	29-71	52.9	12.6
Grec	GR	15	18-30	23.6	3.7
Italien Salentin du Sud	ISS	26	20-31	26.5	3.1

3.3.2 Stimuli

Dans leur tentative pour expliquer des patrons inter-langue de configuration de systèmes vocaliques dans le cadre de la théorie de Dispersion-Focalisation, Schwartz *et al.* (1997) proposent un ensemble de trente-trois ‘prototypes’ représentant une approximation adéquate des qualités vocaliques majeures, attestées dans UPSID (Maddieson, 1991). Ces prototypes couvrent l’intégralité de l’espace vocalique et leurs distances $F1-F2'$ sont raisonnablement symétriques. Leurs valeurs formantiques ont servi d’input à un synthétiseur de formants parallèle (Carré, 2004). Nous avons sélectionné ce type de synthèse afin d’examiner directement l’effet des formants. Dans la synthèse cascade, l’amplitude et la largeur des bandes formantiques dépendent de la distance relative des formants. Ceci aurait donné des voyelles focalisées/quantiques (Schwartz *et al.*, 1997; Stevens, 1989), avec une concentration d’énergie significative dans les hautes fréquences. Notre but était d’éliminer, autant que possible, le rôle des formants supérieurs sans appliquer de filtre coupe-bande. Le tableau 3.15 présente les valeurs formantiques des stimuli.

La largeur des bandes a été fixée à 50, 100, 200 et 200 pour les quatre premiers formants respectivement. La durée des stimuli, synthétisés à un échantillonnage de 16 KHz et à un débit binaire de 16 octets, a été fixée à 200 ms. La fréquence fondamentale croissait de 103 à 110 Hz pendant le premier tiers des stimuli pour ensuite chuter à 94 Hz durant les deux derniers tiers. Avant d’être soumis aux auditeurs, les stimuli ont été appariés en énergie RMS (−10 dB) avec Sound Forge 6.0.

3.3.3 Procédure

Des ordinateurs portables fonctionnant sous Windows XP ont été utilisés. Les stimuli ont été présentés en stéréo par casque dans des salles calmes. Un niveau d’intensité approximatif (autour des 70 dB SPL) a été choisi intuitivement afin que les stimuli donnent une impression quasi-réaliste. Ce procédé a été repris à l’identique pour les participants de chaque langue. Les cartes son étant différentes d’un ordinateur à un autre, l’intensité a pu varier d’un groupe à un autre. Praat (Boersma & Weenink, 2001) a servi d’interface pour le test.

Les stimuli ont été présentés par triades dans une tâche AXB. La figure 3.9 présente les huit hauteurs vocaliques. Pour les sept premières (en commençant par le haut), la méthode suivie pour la disposition des triades a été la suivante : le premier (A) et le dernier (B) stimulus de chaque triade ont été les deux voyelles extérieures d’un degré d’aperture donné et X était une des voyelles intérieures de la même hauteur. Nous avons par exemple <i-y-u>. Les deux ordres de présentation — AXB et BXA — ont été proposés. La voyelle [a], étant la seule sur sa propre hauteur, a été combinée avec les voyelles périphériques de la septième hauteur. Nous avons donc 19 stimuli intérieurs * cinq répétitions * deux ordres = 190 triades par auditeur. La fonction de randomisation ‘PermuteBalancedNoDoublets’ de Praat a été choisie.

Cinq réponses ont été proposées, <i>, <a>, <u>, <e> et <o>, qui correspondent aux voyelles natives des auditeurs. Les auditeurs devaient cliquer sur de grands carrés jaunes qui présentaient les cinq choix possibles. Ils devaient juger si le deuxième stimulus ressemblait plus au premier ou au troisième. Après chaque réponse, la triade suivante était présentée avec un délai de 500 ms. Toutes les 15

Tableau 3.15: Les valeurs des quatre formants pour les trente-trois stimuli synthétisés (Schwartz *et al.*, 1997).

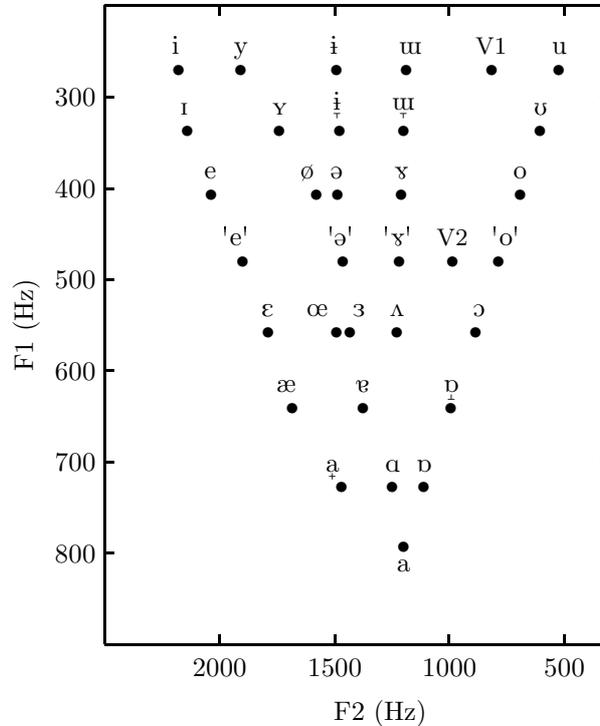
Voyelle	Stimulus	F1	F2	F3	F4
i	01	277	2208	3079	3560
y	02	277	1937	2232	3560
ï	03	277	1520	2310	3560
u	04	277	1218	2500	3560
V1	05	277	845	2460	3560
u	06	277	553	2420	3560
ɪ	07	344	2170	2660	3560
ɤ	08	344	1770	2230	3560
ĩ	09	344	1507	2390	3560
ɥ	10	344	1228	2500	3560
ʊ	11	344	635	2413	3560
e	12	414	2065	2570	3560
ø	13	414	1608	2250	3560
ə	14	414	1516	2500	3560
ɤ	15	414	1238	2500	3560
o	16	414	721	2406	3560
'e'	17	487	1928	2580	3560
'ə'	18	487	1492	2505	3560
'ɤ'	19	487	1248	2500	3560
V2	20	487	1015	2450	3560
'o'	21	487	815	2393	3560
ɛ	22	565	1819	2528	3560
œ	23	565	1520	2500	3560
ɜ	24	565	1462	2500	3560
ʌ	25	565	1258	2500	3560
ɔ	26	565	915	2373	3560
æ	27	648	1712	2490	3560
ɐ	28	648	1405	2500	3560
ɹ	29	648	1023	2500	3560
ɑ	30	735	1498	2537	3560
ɑ	31	735	1278	2500	3560
ɒ	32	735	1141	2280	3560
a	33	800	1228	2500	3560

triades, une pause permettait aux auditeurs de se détendre. Ils devaient ensuite cliquer sur l'écran pour reprendre l'expérience. Le scénario *a priori* optimal aurait été d'introduire une pause à la fin de chaque bloc (38 triades) mais des pré-tests ont montré que cette méthode était fatigante et désagréable pour les auditeurs.

L'expérience a été précédée d'un entraînement, pendant lequel 15 triades concernant les trois premières hauteurs vocaliques (F1 bas), ont été présentées.

3.3.4 Résultats

La figure 3.10 présente les résultats pour les deux ordres de présentation. Les chiffres représentent le taux auquel les sujets ont assimilé les stimuli à la voyelle antérieure de leur série. Dans chaque fraction, le numérateur correspond aux scores pour les triades où la voyelle antérieure était placée en première position. Nous

FIGURE 3.9: Représentation sur le plan F1/F2 des trente-trois stimuli prototypes (Schwartz *et al.*, 1997).

n'allons pas focaliser sur les scores *per se* de cette figure mais seulement sur les différences retrouvées pour des deux ordres.

Karypidis *et al.* (2009) présentent une analyse adéquate des scores bruts.

L'analyse du graphe montre que les asymétries significatives (en italique) ne sont pas régulières, à savoir concentrées dans une région de l'espace acoustique ou déclenchées uniquement par les voyelles antérieures ou postérieures. Pour l'espagnol, griko et grec, les effets d'ordres se situent dans les moyennes et basses voyelles, avec l'exception de [y] pour le griko. L'italien salentin, en revanche, révèle des asymétries à tous les hauteurs vocaliques et pour les voyelles antérieures, centrales et postérieures. Nous observons alors qu'il est difficile d'associer les effets d'ordre à une hauteur ou à une trajectoire vocalique spécifiques.

Nous observons, néanmoins, des différences entre les quatre groupes. Afin de généraliser sur le mécanisme de prise de décision utilisé par les sujets dans cette tâche AXB, nous avons refait une analyse Wilcoxon. Nous avons d'abord établi la moyenne pour les performances de chaque sujet par triade (ce qui donne 19 paires de valeurs par sujet) et puis traité la variable Sujet comme une variable aléatoire. L'analyse a montré que pour les quatre groupes de locuteurs, il est significativement plus probable qu'un stimulus soit assimilé à sa voyelle antérieure correspondante s'il est précédé par celle-ci. Prenons comme exemple les résultats de [y] pour les griko-phones. La figure 3.11 présente les deux triades <i-y-u> et <u-y-i> et le score pour chaque triade. Nous pouvons interpréter l'interaction de [i] et [y] de deux façons :

FIGURE 3.10: Pourcentage auquel les 19 stimuli ont été jugés plus similaires à la voyelle antérieure de leur série. Numérateur : ordre AXB, dénominateur : ordre BXA avec A= voyelle antérieure et B= voyelle postérieure. En italiques : les différences significatives à $\alpha = .01$ avec un test Wilcoxon.

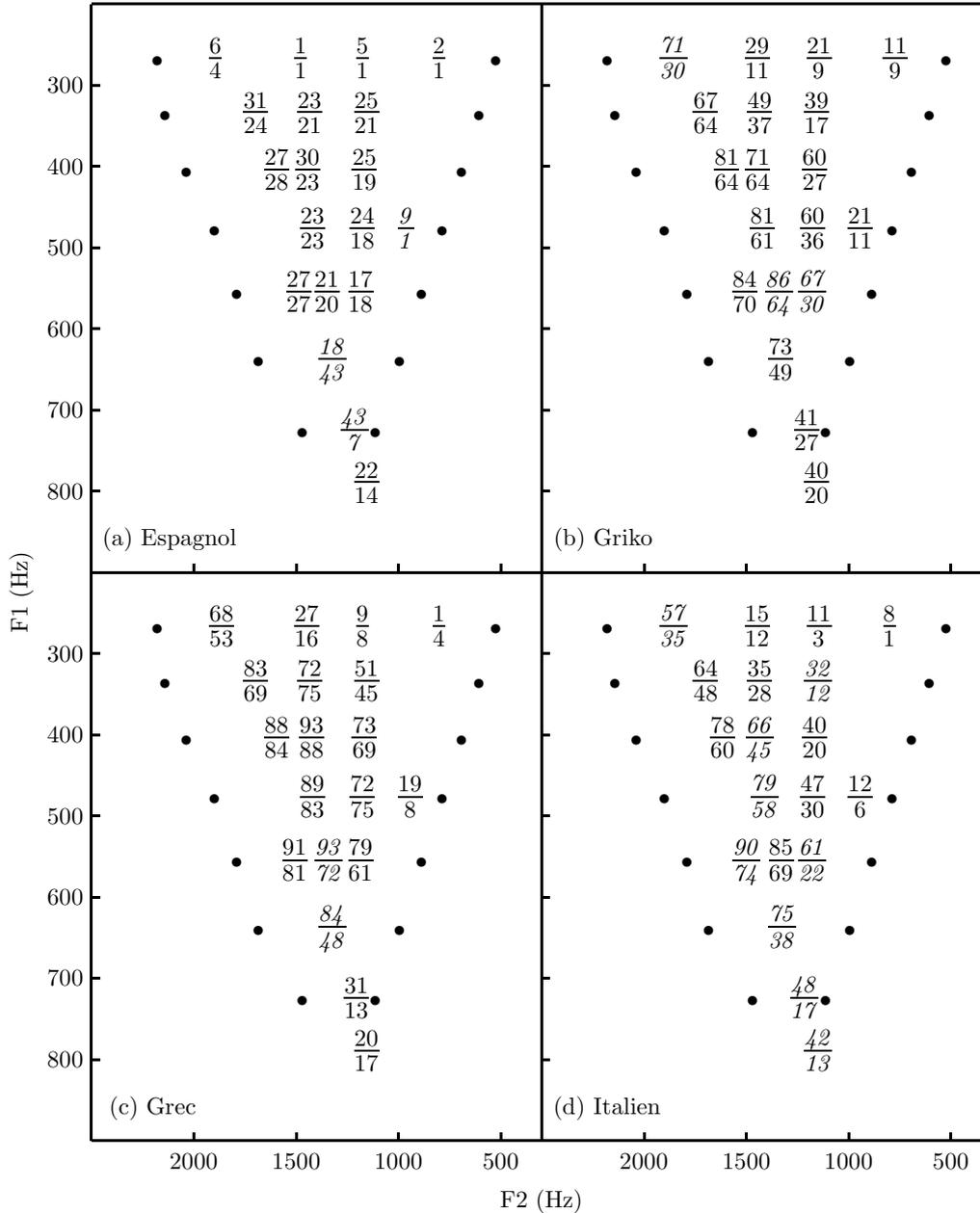
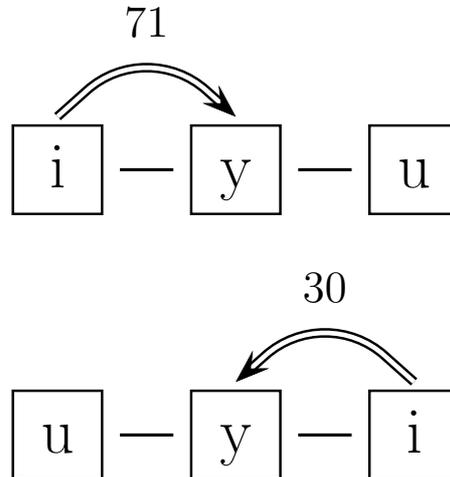


FIGURE 3.11: Les deux triades dans lesquelles le stimulus [y] peut être présenté. Les flèches démarquent la position de la voyelle antérieure standard à laquelle [y] est comparée et les chiffres correspondent aux scores pour [y], comme il a été jugé similaire à /i/ par les grikoophones.



- [i] influence [y] via un processus généralisation : les propriétés de [i] sont stockées en mémoire et créent des interférences pendant le traitement de [y]. L'effet d'ordre est alors expliqué par une supériorité de l'effet de primauté par rapport à l'effet de récence ;
- [i] influence [y] via un effet de contraste : [y] est comparé au standard [i] et l'étiquette inverse ['postérieur'] lui est attribuée. L'effet d'ordre est alors expliqué par une supériorité de l'effet de récence.

Afin de mieux saisir le rôle de l'ordre sur la catégorisation, nous avons mené des tests χ^2 2x2 sur les variables Ordre (AXB et BXA) et Réponse (stimulus assimilé à la voyelle antérieure ou la voyelle postérieure de la triade). Le tableau 3.16 présente les résultats. Nous observons une relation significative entre les deux variables pour les quatre langues.

Tableau 3.16: Résultats des tests χ^2 pour les quatre langues : valeur χ^2 de Yates corrigée pour continuité, degrés de liberté, taille de l'échantillon et valeur p.

Langue	χ^2	ddl	n	p
ES	11.32	1	5700	<.001
GK	91.53	1	2660	<.001
GR	23.29	1	2850	<.001
ISS	174.84	1	4940	<.001

3.4 Conclusions générales

Dans ce chapitre, nous avons examiné le rôle que l'expérience linguistique joue sur le mécanisme de prise de décision, en menant deux études.

Notre premier travail (expérience 1 et 2) a exploré un continuum [i]-[e] et a inclus les locuteurs de quatre langues. Aussi avons-nous constaté que les stimuli localisés à l'extrémité /i/ du continuum déclenchent des effets d'ordre, incitant les auditeurs à donner plus de réponses 'différent' quand le stimulus le plus proche à l'extrémité /i/ est présenté en deuxième position dans une paire. Une telle tendance n'a pas été retrouvée au sein de la catégorie /e/.

Nous avons démontré l'impact du facteur linguistique par le fait que des asymétries ont été attestées seulement pour les auditeurs français et italiens. Par conséquent, nous avons pu formuler l'hypothèse que le déclencheur des asymétries n'est pas localisé uniquement au niveau de la mémoire auditive précatégorielle, étant donnée l'interférence de l'expérience linguistique.

La deuxième étude (expérience 3) a fait intervenir des locuteurs de quatre langues dont les systèmes vocaliques partagent la même structure. Des stimuli dispersés dans l'espace acoustique de façon symétrique ont été synthétisés et utilisés dans une tâche d'identification AXB. Les résultats ont révélé que l'ordre a un effet significatif sur la catégorisation pour les auditeurs des quatre langues. Les résultats au niveau du stimulus ont montré, pourtant, que les effets d'ordre ne dépendent pas de la hauteur vocalique ni de la trajectoire des voyelles antérieures ou postérieures. L'italien présente davantage d'effets que les autres langues.

Tandis que ces deux travaux corroborent l'hypothèse selon laquelle l'exposition aux sons d'une langue façonne le traitement des stimuli, il n'est pas possible de formuler une hypothèse forte sur l'implémentation de la mémoire à long terme. Ceci est dû au fait que les protocoles utilisés dans les expériences 1 et 2 n'ont pas été les mêmes — contexte isolé et tâche 'même-différent' respectivement. Ils ne nous permettent donc pas de vérifier si les effets d'ordre en discrimination présupposent une catégorisation implicite asymétrique.

Dans les chapitres suivants, nous aurons l'occasion de comparer les mécanismes d'identification et de discrimination dans le même contexte et consécutivement, d'examiner le rôle du contexte et du paradigme expérimental.

Chapitre 4

Asymétries et configurations manuelles

"Try again, fail again. Fail better."
Samuel Beckett, *Worstward Ho* (1983)

Dans le chapitre précédent, nous avons examiné le rôle de l'expérience linguistique dans la perception de stimuli auditifs. Pour la suite, nous avons voulu étudié le rôle du facteur linguistique en perception visuelle. Nous avons sélectionné deux groupes de sujets : des signeurs sourds, natifs de la Langue des Signes Française (LSF) et des non-signeurs (NS) entendants n'ayant pas d'expérience en LSF. Nous avons synthétisé deux continua explorant trois configurations manuelles en dactylographie de la LSF. Ces configurations ont un statut linguistique pour les participants sourds mais pas pour les entendants.

Nous avons effectué des tâches d'identification et de discrimination. Deux aspects des résultats nous ont intéressé : l'effet de contexte et les asymétries. L'effet de contexte en identification se réfère à l'influence qu'un stimulus a sur l'identification d'un stimulus adjacent. Notre hypothèse est que plus la trace mémorielle d'un stimulus est robuste, moindre est l'impact du stimulus adjacent sur son étiquetage. La robustesse de la représentation mentale peut dépendre de plusieurs facteurs — saillance perceptive, typicalité par rapport à la sur-catégorie, parmi d'autres ; voir le chapitre 3 — mais cette étude ne focalisera pas directement sur ces facteurs. Ce qui nous concerne ici est de savoir s'il existe des différences entre les deux groupes. Si des divergences entre les deux groupes sont attestées, nous pourrions faire des inférences quant à leur nature et à leur origine.

La tâche de discrimination pourra révéler des asymétries. Leur localisation sur le continuum peut nous mener à des hypothèses concernant leur origine. Si des effets d'ordre sont retrouvés autour des extrémités et que seuls les résultats des sourds révèlent des asymétries, nous parlerons de points de référence cognitifs (linguistiques), à savoir des traces mémorielles emmagasinées dans la mémoire à long terme et liées à l'exposition des sujets à la LSF. Si, en revanche, des asymétries sont attestées pour les deux groupes au même endroit, il s'agira plutôt de prototypes temporaires, issus de la reconstruction mentale du continuum par les sujets et de la construction de catégories temporaires *ad hoc* (Barsalou, 1983) à partir des extrémités du continuum (Durlach & Braida, 1969). Si les effets d'ordre sont localisés autour des frontières,

Tableau 4.1: Moyenne d'âge (M) des seize participants de chaque groupe, déviation standard (ds) et étendue d'âge.

	UV		VX	
	LSF	NS	LSF	NS
M	45.5	37.3	45.4	37.3
ds	14.0	14.2	13.9	14.2
Étendue	26–66	25–62	26–66	25–62

nous expliquerons ces résultats avec le modèle *FUL* de Lahiri & Reetz (2002). De même, en combinant les résultats d'identification avec ceux de discrimination, nous pourrions vérifier si les effets d'ordre en discrimination sont déclenchés par des effets de contexte issus d'un étiquetage implicite.

Les deux tests présentés ici font partie d'un projet plus large sur la perception catégorielle en LSF. Nous n'allons par conséquent pas présenter les résultats qui ne sont pas liés directement à nos hypothèses. Le lecteur pourra se référer à Boutora & Karypidis (2007, 2009a,b) et Boutora (2008) pour des analyses plus générales.

4.1 Expérience 1 : Catégorisation

Ce test de catégorisation nous permet de : a. examiner le rôle du contexte pendant la catégorisation de stimuli, b. délimiter la localisation des effets sur les continua, et c. étudier les différences entre les deux groupes linguistiques.

4.1.1 Participants

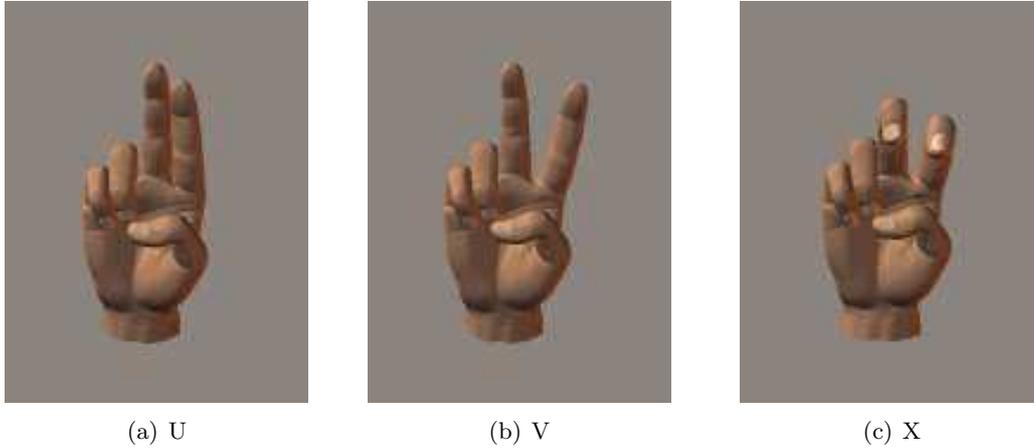
Trente-deux sujets ont participé à chaque séance, seize signeurs de la Langue des Signes Française (LSF) et seize entendants non-signeurs (NS). Nous avons utilisé les mêmes entendants pour les continua U–V et V–X tandis qu'avec le groupe des signeurs, nous avons dû remplacer une personne pour le continuum V–X. Des détails sur les quatre groupes des participants sont disponibles dans le tableau 4.1.

4.1.2 Stimuli

À l'aide du logiciel de modélisation 3D 'Poser 6.0', nous avons préparé trois configurations manuelles correspondant aux configurations U, V et X utilisées dans la dactylologie de la LSF (figure 4.1). Ensuite, en interpolant de façon linéaire les configurations d'origine, nous avons créé deux continua s'étendant : a. de U à V, et b. de V à X. Les critères de sélection des trois configurations sont discutés dans Boutora (2007).

La taille physique des images sur l'écran a été de 9.2 cm de largeur et 12.3 cm de hauteur. Étant donné que différents ordinateurs ont été utilisés, nous avons redimensionné les fichiers image par rapport à la résolution native et à la taille physique de chaque écran, ceci afin que la dimension de l'image projetée soit stable.

FIGURE 4.1: Les trois configurations manuelles localisées sur les extrémités des continua U-V et V-X.



4.1.3 Procédure

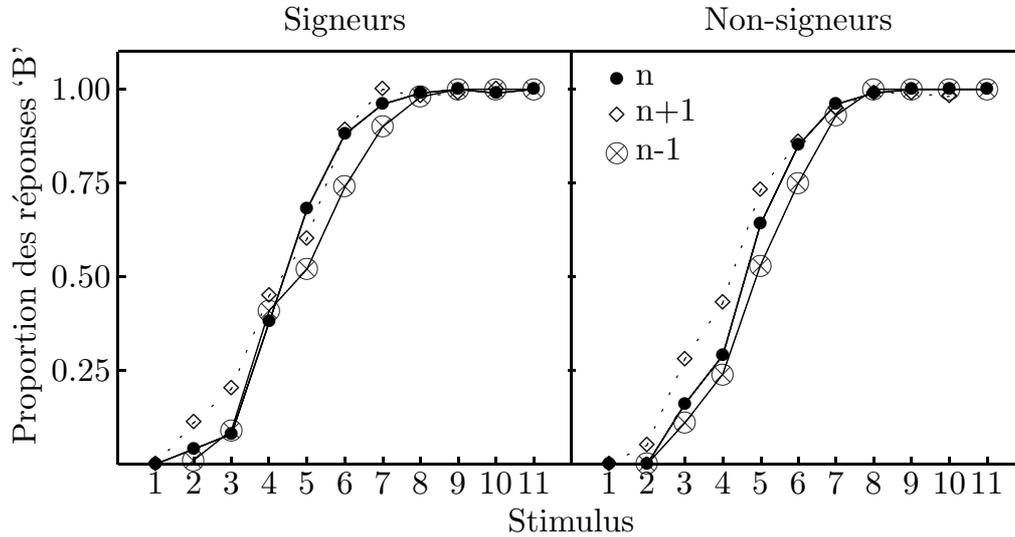
Les deux extrémités de chaque continuum ont été présentées à l'écran simultanément pendant trois secondes au début de chaque bloc expérimental (mais pendant six secondes au début de l'entraînement). Après la disparition de l'inscription à l'écran, des paires d'images ont été présentées, dans toutes les combinaisons possibles — AA, BB, AB, BA, où A et B étaient deux stimuli consécutifs sur le continuum. La première image apparaissait au centre de l'écran et la seconde un peu plus en bas et à droite par rapport à la première. Nous avons introduit cette particularité pour rendre les images relativement plus difficiles à discriminer. Chaque stimulus a été présenté pendant 300 ms. L'écran a ensuite été effacé. Nous avons fixé l'intervalle interstimulus à 500 ms. Les sujets avaient six secondes pour répondre et le premier membre de la paire suivante était présenté avec un délai d'une seconde.

Les participants devaient ignorer la première image et décider si la seconde image de chaque paire ressemblait plus à l'“Image A” — l'extrémité de chaque continuum présentée à gauche au début de chaque bloc — ou à l'“Image B” — présentée à droite. Ils devaient cliquer sur les lettres <F> et <L> de leur clavier respectivement. Nous avons choisi de présenter les images en paires afin que le contexte soit identique pour l'identification et pour la discrimination. Selon Repp *et al.* (1979), la relation entre l'identification et la discrimination est plus étroite quand le même protocole est utilisé pour les deux tâches.

Nous avons choisi un étiquetage avec les lettres <F> et <L> afin que les deux groupes, signeurs et non-signeurs, puissent avoir des stratégies d'identification similaires, sans que les signeurs utilisent des informations linguistiques de façon explicite. Les consignes ont été présentées en français pour les non-signeurs et dans les deux langues pour les signeurs.

Nous avons utilisé six blocs expérimentaux, chacun incluant une répétition de toutes les paires. Le premier bloc a servi de mise en route, sans que les données soient recueillies. À l'issue de cet entraînement, un message a averti les participants que dès lors, leurs réponses seraient enregistrées. Les deux continua ont été étudiés

FIGURE 4.2: Proportions des réponses ‘V’ pour le continuum U–V en fonction du contexte. n , $n+1$, $n-1$: position relative sur le continuum du stimulus précédant la voyelle à catégoriser. $N = 16$.



séparément.

FLXLab 1.9 a servi d'interface pour l'expérience. Nous avons également enregistré les temps de réaction.

4.1.4 Résultats

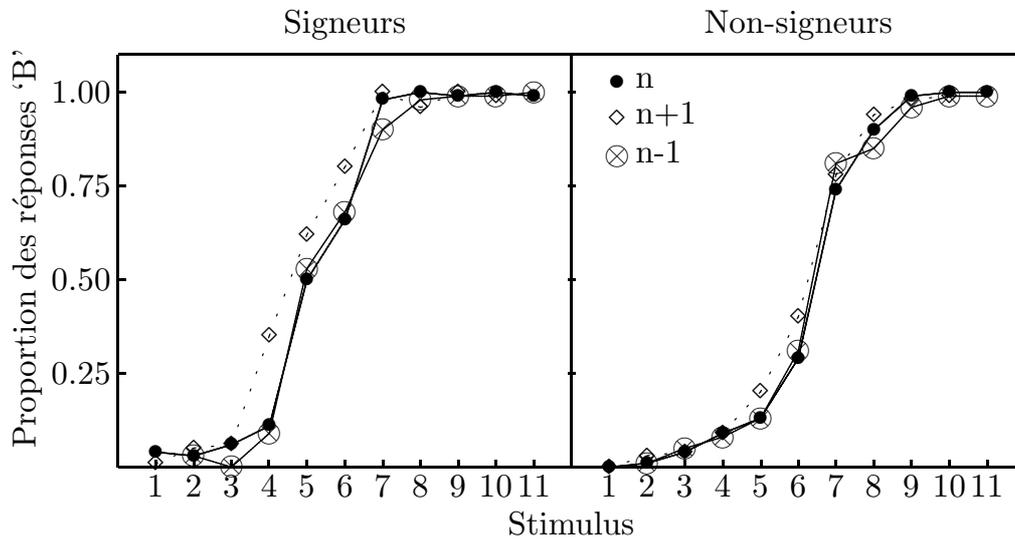
Les figures 4.2 et 4.3 présentent les résultats de la catégorisation des onze stimuli en fonction de leur contexte. Trois courbes sont alors tracées pour chaque continuum, une pour chaque contexte. n signifie qu'un stimulus est précédé par lui-même. Dans le cas du stimulus 2 par exemple, n se réfère à la paire 02–02, $n+1$ à 03–02 et $n-1$ à 01–02.

Nous avons précisé dans l'introduction de cette expérience que trois aspects des résultats nous intéressaient : le rôle du contexte, les différences entre les deux groupes linguistiques et la localisation des effets de contexte. Afin d'aborder les deux premiers aspects, nous avons effectué des analyses de Kruskal-Wallis sur le facteur Contexte. Les résultats détaillés sont disponibles dans l'annexe

Il est à noter ici que nous avons dû compléter une partie des données puisque les paires $(n-1)$ –01 et 11– $(n+1)$ ne sont pas possibles. Nous avons alors attribué à ces paires des scores virtuels : '0' à $(n-1)$ –01 et 100 à 11– $(n+1)$.

Pour les deux continua et les deux groupes, l'effet de Contexte n'est significatif. Ceci est dû au fait que pour les stimuli localisés sur les extrémités, l'effet est très faible et non-existant, ce qui baisse la valeur K de Kruskal. De même, les figures 4.2 et 4.3 nous montrent que certaines comparaisons de contexte — notamment $n+1$ vs. $n-1$ — semblent plus robustes que d'autres. Nous devons donc étudier chaque paire séparément en menant des tests de rang de Wilcoxon.

FIGURE 4.3: Proportions des réponses ‘X’ pour le continuum V–X en fonction du contexte. n , $n+1$, $n-1$: position relative sur le continuum du stimulus précédant la voyelle à catégoriser. $N = 16$.



Le test de Wilcoxon ($N = 352$) montre que quand les trois contextes sont comparés deux à deux ($n/n+1$, $n/n-1$, $n-1/n+1$), l’effet de contexte est significatif pour les deux continua pour LSF, mais seulement pour le continuum U–V pour NS. Cette similarité entre les deux groupes de locuteurs pour U–V peut s’expliquer par le fait que les paramètres des stimuli de ce continuum peuvent être réduits à un seul : la distance entre les deux doigts. Or, V–X est plus complexe, avec davantage de paramètres implémentés, tels que les ongles et l’ombre des doigts qui entrent en jeu. En effet, Boutora & Karypidis (2009b) et Mathur & Best (2007) montrent que la discrimination des stimuli du continuum U–V est régie par la loi de Weber : plus la distance entre les deux doigts augmente, moins les sujets y sont sensibles.

Tableau 4.2: Les comparaisons de contexte par paire pour U–V, significatives à $\alpha = .05$ avec un test de rang de Wilcoxon ($N = 16$).

Stimulus	n/n+1		n/n-1		n-1/n+1	
	LSF	NS	LSF	NS	LSF	NS
1						
2		*				*
3	*				*	*
4		*				*
5				*		*
6			*	*		
7					*	
8						
9						
10						
11						

Tableau 4.3: Les comparaisons de contexte par paire pour V-X, significatives à $\alpha = .05$ avec un test de rang de Wilcoxon ($N = 16$).

Stimulus	n/n+1		n/n-1		n-1/n+1	
	LSF	NS	LSF	NS	LSF	NS
1						
2						
3					*	
4	*				*	
5						
6						
7			*		*	
8						*
9						
10						
11						

Les tableaux 4.2 et 4.3 illustrent les trois contrastes de contexte significatifs pour chaque paire tels qu'ils ont été retrouvés avec un test de Wilcoxon. Nous observons que pour le continuum U-V, les effets de contexte se localisent entre les stimuli 3 et 7 pour LSF et entre 2 et 6 pour NS. Le nombre de contrastes significativement différents est plus important pour NS que pour LSF. De même, le contraste $n-1/n+1$ révèle le plus d'effets, ce qui indique que les deux contextes exercent des forces opposées sur les stimuli à catégoriser. En regardant la figure 4.2, nous constatons que quand le premier stimulus d'une paire se rapproche de l'extrémité 'V', les sujets assimilent le deuxième stimulus plus souvent à cette extrémité. Il s'agit donc, soit d'un contraste rétroactif, soit d'une généralisation proactive. Déterminer lequel explique au mieux nos données est impossible sans que nous exploitions le rôle de l'intervalle interstimulus de façon directe. À notre connaissance, de telles études systématiques sur le rôle de l'intervalle interstimulus dans des tâches de catégorisation n'ont pas été effectuées.

Quant à V-X, le tableau 4.3 indique que le seul contraste significatif pour NS est localisé au stimulus 8. De nouveau, la plupart des effets sont attestés pour la comparaison $n-1/n+1$ et la même tendance est retrouvée : les sujets répondent plus souvent 'X' quand le premier stimulus se rapproche de cette extrémité.

4.2 Expérience 2 : Discrimination

Cette tâche de discrimination nous permet de : a. détecter des effets d'ordre à travers les deux continua, b. délimiter la localisation des effets sur les continua, et c. examiner si les effets d'ordre sont déclenchés par des effets de contexte issus d'un étiquetage implicite.

4.2.1 Participants

Les sujets ayant participé à l'expérience 1 se sont également prêtés à cette tâche.

4.2.2 Stimuli

Les stimuli sont les mêmes que pour l'expérience 1.

4.2.3 Procédure

Les mêmes stimuli que dans l'expérience 1 ont été présentés aux participants, qui ont dû indiquer dans chaque cas si les voyelles dans chaque paire étaient "absolument identiques" ou "même légèrement différentes". Nous avons donné cette instruction spécifique afin d'induire un rôle plus actif de la mémoire à court terme. Nous avons souhaité que les sujets se fient moins à l'étiquette du stimulus et davantage à ses aspects physiques. Les mêmes touches ont été utilisées pour valider les réponses, <F> pour 'même' et <L> pour 'différent'.

L'intervalle interstimulus, la durée de la présentation des stimuli et la pause après chaque réponse (angl. '*intertrial interval*') ont été les mêmes que pour l'expérience 1.

Six blocs, dont le premier a servi d'entraînement, ont été présentés. Tandis que la tâche d'identification portait sur des différences d'un pas sur le continuum, nous avons ajouté une tâche portant sur des différences de deux pas. Les sujets ont alors participé à deux tests séparés pour chaque continuum.

4.2.4 Résultats

Nous allons présenter les résultats des deux continua séparément afin de rendre la lecture plus claire.

4.2.4.1 Continuum U-V

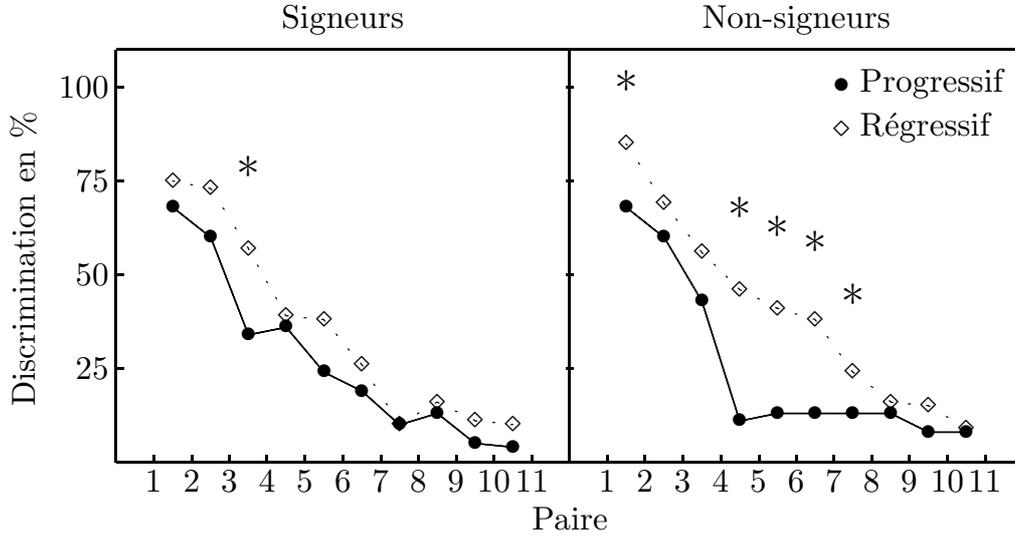
Nous avons mené des analyses de Wilcoxon ($N = 160$ et $N = 144$ pour les conditions d'un pas et de deux pas, respectivement) sur chaque groupe séparément. Dans les quatre cas, les sujets ont répondu 'différent' beaucoup plus souvent dans l'ordre Régressif ($U \leftarrow V$) que dans l'ordre Progressif ($U \rightarrow V$), à savoir quand la distance entre les deux doigts sélectionnés diminue. Dans les figures 4.4 et 4.5, nous observons que les asymétries sont localisées généralement entre les paires 3-4 et 7-8. La condition à un pas révèle davantage d'asymétries pour les NS.

Pour vérifier si les effets d'ordre sont déclenchés par un mécanisme de catégorisation asymétrique, nous avons comparé les résultats issus des deux tâches. Prenons l'exemple de la paire incluant les stimuli 1 et 2. Nous calculons d'abord la différence entre le pourcentage des réponses 'V' pour les paires 1-2 et 2-1. Ensuite nous calculons la différence de réponses 'différent' pour les mêmes paires. Ces valeurs servent d'input pour une analyse de corrélation. Nous avons effectué cette procédure seulement pour la condition d'un pas.

L'analyse de Pearson ($N = 176$) montre qu'aucune corrélation n'existe entre les effets de contexte et les asymétries. Ceci indique que ces dernières ne sont pas en relation directe avec une stratégie d'étiquetage différentiel.

Comme le titre de cette thèse le signale, nous nous sommes particulièrement intéressé à la notion de robustesse des représentations. Pour évaluer la stabilité de traces mémorielles de nos stimuli, nous avons travaillé sur les fausses alarmes. Notre

FIGURE 4.4: Scores de discrimination pour les deux directions sur le continuum U-V à un pas. * : $\alpha = 0.05$ avec un test Wilcoxon ($N = 16$).



hypothèse est qu'un stimulus saillant, après son stockage dans la mémoire, est traité de façon à ce qu'il soit plus facilement détecté comme similaire à un exemplaire identique.

Le test de Kruskal ($N = 176$) montre que la position d'un stimulus sur le continuum n'a pas d'effet significatif sur le taux de fausses alarmes. Ceci est le cas pour les deux conditions et groupes.

FIGURE 4.5: Scores de discrimination pour les deux directions sur le continuum U-V à deux pas. * : $\alpha = 0.05$ avec un test Wilcoxon ($N = 16$).

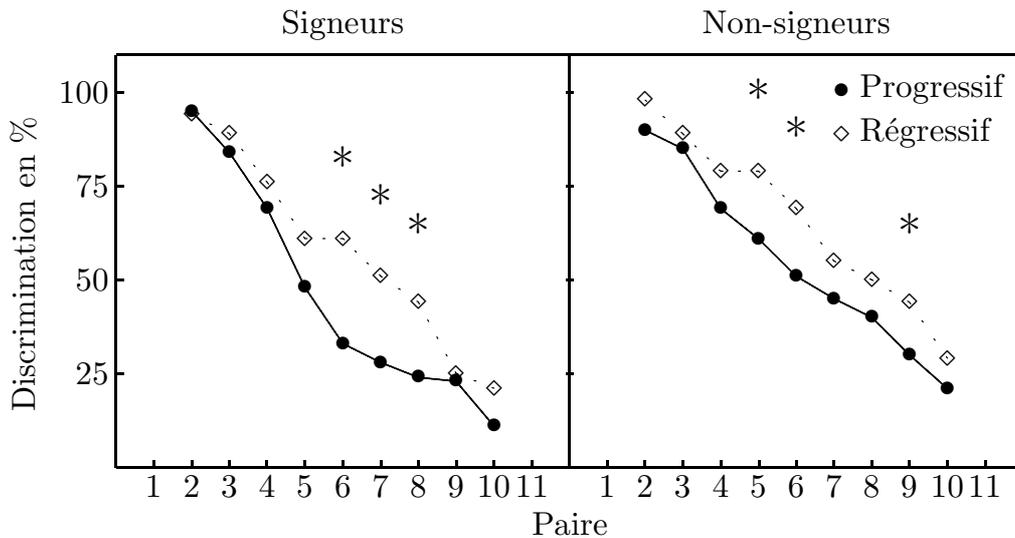


FIGURE 4.6: Scores de discrimination pour les deux directions sur le continuum V-X à un pas. * : $\alpha = 0.05$ avec un test Wilcoxon ($N = 16$).

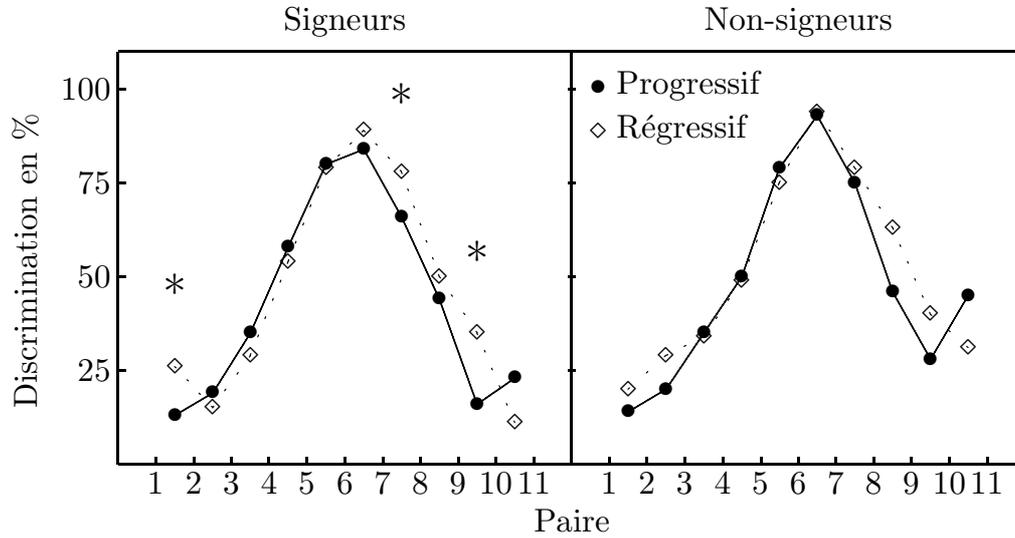
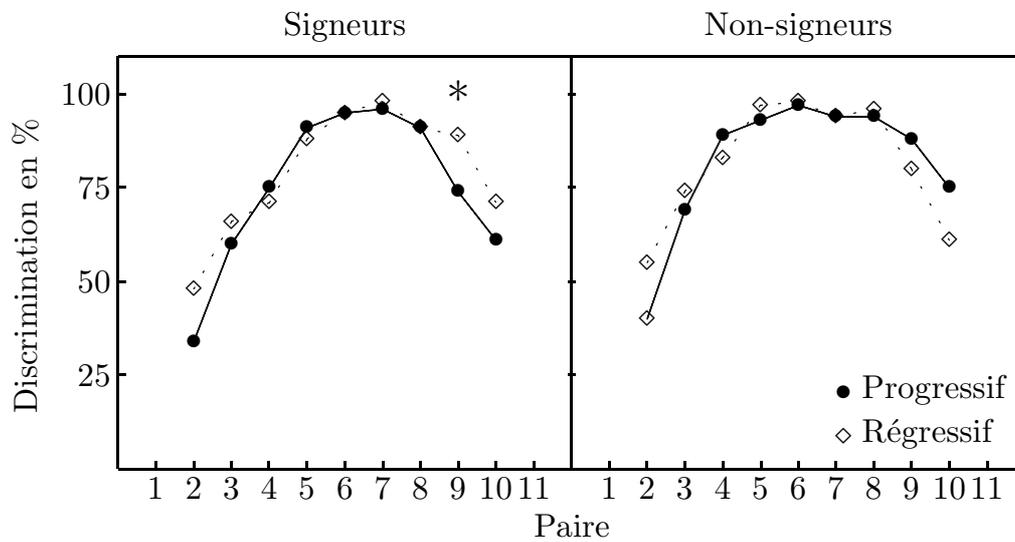


FIGURE 4.7: Scores de discrimination pour les deux directions sur le continuum V-X à deux pas. * : $\alpha = 0.05$ avec un test Wilcoxon ($N = 16$).



4.2.4.2 Continuum V-X

Les résultats de Wilcoxon indiquent que la différence entre les deux ordres de présentation n'est significative pour aucun des deux groupes. Nous pouvons détecter quelques effets non-systématiques pour les LSF dans les figures 4.6 et 4.7.

Comme pour U-V, l'analyse de Pearson montre qu'aucune corrélation n'existe entre les effets de contexte et les asymétries pour le continuum V-X. Ceci indique que ces dernières ne sont pas en relation directe avec une stratégie d'étiquetage

différentiel.

Le test de Kruskal ($N = 176$) montre que la position d'un stimulus sur le continuum n'a pas d'effet significatif sur le taux de fausses alarmes. Ceci est le cas pour les deux conditions et groupes.

4.3 Conclusions générales

Dans cette partie de notre travail, nous avons envisagé les différences entre deux groupes linguistiques dans leur façon de percevoir des stimuli visuels. Le critère de comparaison que nous avons fixé a été le degré auquel des effets de contexte et d'ordre sont attestés.

Nous avons considéré deux continua constitués d'unités de la dactylogogie de la Langue des Signes Française, U-V et V-X. Le continuum U-V peut être réduit à un continuum à une dimension, à savoir la distance entre les deux doigts. Ceci infirme l'hypothèse selon laquelle les deux groupes linguistiques utilisent des points de référence cognitifs de nature différente. Ainsi avons-nous trouvé des effets de contexte pour U-V, à part égale pour les deux groupes. Les tâches de discrimination ont révélé de fortes asymétries pour les deux conditions et pour les deux langues.

Nous avons eu des indications en faveur du rôle de l'expérience linguistique à partir des résultats pour le continuum V-X, pour lequel des effets de contexte ont été démontrés. Nous constatons donc que la nature et la complexité des stimuli d'un continuum peuvent avoir un impact sur la façon dont les sujets décodent le signal.

Quant à nos hypothèses de départ sur la localisation des effets sur le continuum, nous avons retrouvé une tendance plutôt forte à l'intérieur du continuum et pas autour des extrémités. Ceci fait penser que ces effets sont liés aux frontières et, par conséquent, à l'étiquetage des stimuli. Nos analyses de corrélation indiquent que les asymétries et les effets de contexte ne sont interdépendants pour aucune condition, groupe ou continuum.

Finalement, nous avons mené des analyses de Kruskal sur l'effet de l'emplacement d'un stimulus par rapport aux fausses alarmes. De nouveau, la position d'un stimulus n'a pas d'impact sur le taux de fausses alarmes ou bien sur la robustesse de la représentation mentale du stimulus lui-même.

Chapitre 5

L'hypothèse de la neutralisation

*"A belief is not merely an idea the mind possesses.
It is an idea that possesses the mind."
Robert Oxton Bolton*

Ce chapitre consiste à répliquer l'étude de Repp & Crowder (1990) tout en évitant les inconvénients que le protocole original a introduits (voir section 2.2.2, page 44) et qui n'ont pas pu amener les chercheurs à des résultats concluants. Aussi avons-nous mené une expérience avec dix voyelles du français que nous avons séparées en trois groupes, de telle sorte que le centre de chaque distribution ne se situe pas dans la région de [ə]. Ceci a permis de vérifier à la fois trois hypothèses : a. dégradation vers [ə] ; b. dégradation vers le centre de la dispersion des stimuli ; c. hypothèse de périphéricité.

5.1 Expérience 1 : Discrimination¹

5.1.1 Sujets

Au total, 35 sujets francophones ont participé à cette expérience. Certains n'ont pas participé aux trois tâches prévues ; nous avons alors retenu 23 auditeurs pour chaque séance. Tous les auditeurs ont indiqué être des locuteurs natifs du français, monolingues et ne pas avoir de pertes auditives connues. Le tableau 5.1 présente les détails des auditeurs des trois séances.

1. Des résultats partiels de l'expérience 1 ont été publiés dans Karypidis (2007c, 2009) et présentés dans plusieurs conférences (Karypidis, 2007a,b).

Tableau 5.1: Moyenne d'âge, déviation standard (ds) et étendue d'âge des 23 auditeurs de chaque séance.

	u o ɔ a	y ø ə	i e ε
Moyenne	26.5	25.5	25.9
ds	6.0	5.1	5.3
Étendue	21–47	21–47	21–47

5.1.2 Stimuli

Neuf stimuli correspondant aux valeurs acoustiques moyennes des neuf voyelles orales françaises /i, e, ε, a, ɔ, o, u, y, ø/ produites par des locuteurs adultes masculins (Calliope, 1989) ont été synthétisés avec le synthétiseur en série de Klatt (1980). La voyelle /œ/ n'a pas été incluse, étant donné sa proximité avec [ə] sur le plan F1/F2. Un tel rapprochement aurait comme résultat des chevauchements entre les satellites des deux catégories ainsi que des différences de catégorisation entre deux stimuli. Dans ce cas, les auditeurs utiliseraient probablement comme critère de discrimination la différence d'étiquette — deux stimuli jugés appartenir à deux catégories distinctes — et non pas les détails fins du signal. Le but de notre expérience était l'examen de la dégradation acoustique au niveau de la mémoire sensorielle.

Tableau 5.2 – Les valeurs formantiques en Hertz des 50 stimuli synthétisés. P : le prototype de chaque catégorie ; S1–S4 : les quatre satellites d'un prototype.

		F1	F2	F3	F4
	P	308	2064	2976	3407
	S1	283	2160	2976	3407
i	S2	279	1986	2976	3407
	S3	334	1972	2976	3407
	S4	338	2145	2976	3407
	P	365	1961	2644	3362
	S1	342	2060	2644	3362
e	S2	332	1895	2644	3362
	S3	389	1867	2644	3362
	S4	399	2029	2644	3362
	P	530	1718	2558	3300
	S1	543	1820	2558	3300
ε	S2	485	1746	2558	3300
	S3	517	1622	2558	3300
	S4	576	1690	2558	3300
	P	684	1256	2503	3262
	S1	727	1209	2503	3262
a	S2	717	1319	2503	3262
	S3	643	1305	2503	3262
	S4	652	1195	2503	3262
	P	531	998	2399	3278
	S1	536	933	2399	3278
ɔ	S2	579	1005	2399	3278
	S3	526	1066	2399	3278
	S4	485	991	2399	3278
	P	383	793	2283	3256
	S1	372	738	2283	3256
o	S2	424	778	2283	3256
	S3	394	850	2283	3256

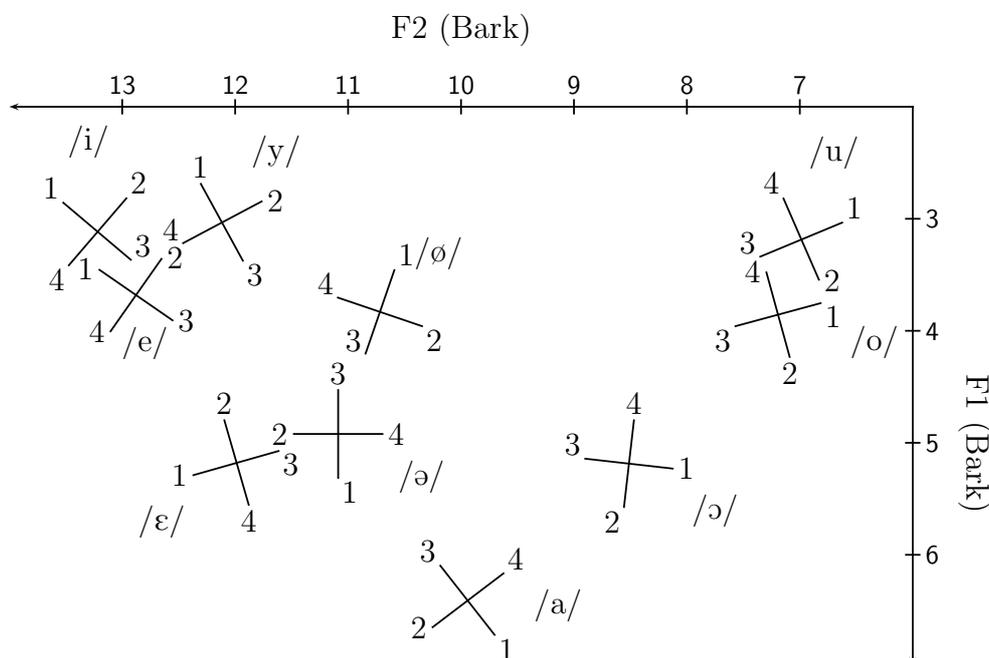
Tableau 5.2 – (cont.)

		F1	F2	F3	F4
	S4	343	808	2283	3256
	P	315	764	2027	3118
	S1	300	713	2027	3118
u	S2	352	742	2027	3118
	S3	330	817	2027	3118
	S4	279	786	2027	3118
	P	300	1750	2120	3145
	S1	266	1801	2120	3145
y	S2	282	1660	2120	3145
	S3	335	1701	2120	3145
	S4	319	1844	2120	3145
	P	381	1417	2235	3215
	S1	342	1389	2235	3215
ø	S2	395	1337	2235	3215
	S3	421	1445	2235	3215
	S4	368	1501	2235	3215
	P	500	1499	2501	3502
	S1	547	1499	2501	3502
ə	S2	500	1593	2501	3502
	S3	455	1499	2501	3502
	S4	500	1411	2501	3502

Les propriétés de [ə] ont été celles d'un tube uniforme d'une longueur égale à celle de l'appareil phonatoire d'un homme adulte (Chiba & Kajiyama, 1941) : F1= 500, F2= 1500, F3= 2500, F4= 3500 Hz, et identiques à celles de [ə] de Repp & Crowder (R&C) . Les valeurs formantiques en Hertz (tableau 5.2) des dix voyelles ont été converties en Bark pour que la différence en résolution fréquentielle de l'oreille soit compensée (Traunmüller, 1990). Autour de chacune de ces dix voyelles prototypes (P)², quatre stimuli satellites (S1-S4) ont été arrangés sous forme de croix, dont le centre coïncidait avec le prototype. S1-S4 ont été placés sur les quatre extrémités de chaque croix, avec un axe pointé vers [ə]. Chaque branche de la croix équivalait à la distance euclidienne de 0.4 Bark sur le plan F1/F2 entre le prototype et chacun des satellites. S1 était placé sur l'axe pointé vers [ə] et se trouvait le plus loin de celui-ci, contrairement à S3 qui se trouvait le plus près de [ə] sur le même axe. Les autres stimuli ont été étiquetés dans le sens des aiguilles d'une montre. Quant à la catégorie /ə/, les axes de S1-S3 et de S2-S4 ont été placés en parallèle avec ceux de F1 et F2 respectivement, avec $F1_{S1} > F1_{S3}$ et $F2_{S2} > F2_{S4}$. Une représentation sur le plan F1/F2 des 50 stimuli synthétisés est disponible dans la figure 5.1. Leurs valeurs formantiques sont disponibles dans le tableau 5.2.

2. Dans notre protocole et pour le reste de ce chapitre, nous utilisons le terme 'prototype' à l'instar de Repp & Crowder pour désigner le stimulus central à l'intérieur de chaque catégorie vocalique de notre ensemble de stimuli. Il ne s'agit pas d'une représentation mentale correspondant au meilleur exemplaire ou à la tendance moyenne d'une catégorie.

FIGURE 5.1: Les 50 stimuli synthétisés sur le plan F1/F2 (en Bark).



Contrairement à R&C qui ont assigné une valeur unique de F3 à tous les stimuli - sauf à ceux appartenant à la catégorie /i/ - nous avons fixé F3 et F4 séparément pour chaque catégorie. Ainsi, la valeur de F3 et F4 de chaque prototype a été utilisée pour les quatre satellites respectifs. Nous avons fixé la durée de chaque stimulus à 250 ms et nous avons choisi un contour descendant pour F0 (100→80 Hz). Les stimuli ont été appariés en énergie RMS à -10 dB avec Sound Forge 6.0. La largeur de bande et l'amplitude des formants ont été définies par le synthétiseur de Klatt en fonction de la distance relative entre les formants successifs de chaque stimulus.

5.1.3 Procédure

À l'instar de R&C, nous avons choisi une tâche de 'même-différent'. Au sein de chaque catégorie phonétique, tous les stimuli (P, S1-S4) ont été appariés avec eux-mêmes (S_i/S_i); les quatre satellites ont également été appariés avec le prototype dans les deux ordres possibles (P/S_i, S_i/P). L'intervalle interstimulus a été fixé à 500 ms entre les deux membres de chaque paire, à 1 s entre les paires et à 3 s après chaque groupe de 13 paires. Nous avons diminué les intervalles par rapport au protocole de R&C, qui ont utilisé un intervalle à 2 s entre les paires et à 5 s après chaque groupe de paires, parce que pendant les pré-tests, les auditeurs ont estimé le rythme de l'expérience relativement lent. Les différentes étapes de cette tâche sont illustrées dans le tableau 5.3.

Nous avons utilisé six blocs expérimentaux, chacun incluant une répétition de toutes les paires en ordre aléatoire — mais identique pour tous les auditeurs. Le premier bloc a servi d'entraînement et les résultats n'ont pas été pris en compte. Les participants devaient juger si les deux stimuli appariés étaient 'absolument identi-

Tableau 5.3: Les étapes du test de discrimination.

Étape	Durée (en ms)
1. Présentation du stimulus A	250
2. Intervalle Interstimulus	500
3. Présentation du stimulus B	250
4. Le sujet répond par un clic	—
5. Intervalle inter-trial	1000
6. Pause (toutes les 13 paires)	3000

ques’ ou ‘même légèrement différents’ en tapant sur la touche ‘m’ de leur clavier pour ‘mêmes’ ou ‘d’ pour ‘différents’ respectivement. Aucune information concernant la compatibilité stimulus-réponse n’a été donnée après chaque réponse.

Dans R&C (1990), le centre de l’ensemble des stimuli et le point [ə] se chevauchaient. Afin d’éviter cet inconvénient qui nous empêcherait de bien distinguer les deux points cibles d’une éventuelle dégradation, soit le point central de la distribution des stimuli et le point central [ə] de l’espace acoustique, nous avons regroupé les dix catégories vocaliques en trois ensembles :

- /i e ε/
- /y ø ə/
- /u o ɔ a/

Ce regroupement a été motivé par le fait que les membres de chacun des trois ensembles partagent un espace commun dans l’espace vocalique : (a) /i e ε/ sont des voyelles périphériques avec F2+F1 maximal ; (b) /u o ɔ/ sont périphériques avec F2-F1 minimum ; (c) /y ø ə/ sont des voyelles non périphériques. La catégorie /a/ a été regroupée avec /u o ɔ/ pour que le centre de cet ensemble de stimuli se localise entre /o/ et /ɔ/. Un différent classement de /a/, c’est-à-dire son regroupement avec l’un des deux autres ensembles, aurait comme effet le chevauchement du centre de l’ensemble donné avec [ə].

5.1.4 Résultats

L’hypothèse de la neutralisation permet quatre prédictions :

- [Pr1a] Effet d’ordre positif pour S1 : $\checkmark(P-S1)-(S1-P) > 0$
- [Pr1b] Effet d’ordre négatif pour S3 : $(P-S3)-(S3-P) < 0$
- [Pr2] Pas d’effet ou effet mineur pour S2 et S4
- [Pr3] Pas d’effet ou effet mineur pour la catégorie /ə/

Il est à noter ici que Repp et Crowder n’avancent que les trois premières hypothèses, soutenant qu’à l’intérieur de la catégorie /ə/, la dégradation sera minime, puisque la proximité des satellites par rapport à [ə] est considérable. Il est tout de même possible d’avancer l’hypothèse que pour cette catégorie, nous obtiendrons des effets positifs :

- dans l’ordre P/S_i, la distance perceptive sera approximativement égale à la distance acoustique originale, le prototype [ə] étant la cible finale de toute dégradation ;

- dans l'ordre S_i/P , la distance entre S_i et P sera moins importante, quelle que soit la direction de la dégradation.

Le tableau 2 présente les scores de discrimination pour les paires identiques, les paires non identiques, ainsi que les effets d'ordre calculés sur la différence entre les scores pour l'ordre P/S_i et ceux pour l'ordre S_i/P . La première remarque que nous pouvons faire est que les fausses alarmes sont bien inférieures aux 'détections correctes' de façon uniforme, ce qui indique que les auditeurs n'ont pas eu de problèmes avec cette tâche de discrimination.

Tableau 5.4: Les scores de discrimination pour les paires identiques (premier bloc de valeurs) et non identiques dans l'ordre $P-S_i$ (deuxième bloc) et S_i-P pour les dix catégories. Le dernier bloc de valeurs représente la différence entre les deux ordres.

	i	e	ɛ	y	ø	ə	u	o	ɔ	a
Paires identiques										
P-P	16.52	13.04	15.65	14.78	13.04	20.00	7.83	7.83	12.17	19.13
S1-S1	11.30	19.13	12.17	12.17	11.30	13.91	13.04	8.70	8.70	13.91
S2-S2	11.30	14.78	14.78	13.04	6.96	24.35	21.74	8.70	11.30	20.87
S3-S3	15.65	11.30	5.22	19.13	8.70	12.17	9.57	9.57	15.65	13.04
S4-S4	13.04	15.65	12.17	12.17	9.57	8.70	2.61	9.57	14.78	21.74
Paires différentes : prototype en premier										
P-S1	58.26	73.91	80.00	65.22	64.35	74.78	43.48	31.30	54.78	73.91
P-S2	57.39	72.17	42.61	59.13	56.52	66.96	61.74	43.48	51.30	62.61
P-S3	38.26	62.61	60.87	31.30	50.43	49.57	22.61	22.61	29.57	73.91
P-S4	64.35	74.78	58.26	82.61	44.35	63.48	45.22	70.43	53.91	45.22
Paires différentes : prototype en second										
S1-P	33.04	45.22	78.26	23.48	35.65	48.70	12.17	11.30	30.43	60.00
S2-P	34.78	63.48	39.13	56.52	25.22	59.13	71.30	28.70	44.35	44.35
S3-P	46.96	62.61	46.96	44.35	56.52	62.61	25.22	31.30	35.65	59.13
S4-P	55.65	72.17	40.00	67.83	44.35	46.96	13.04	73.04	35.65	45.22
Effets d'ordre : (P-S1)-(S1-P)										
S1	25.22	28.70	1.74	41.74	28.70	26.09	31.30	20.00	24.35	13.91
S2	22.61	8.70	3.48	2.61	31.30	7.83	-9.57	14.78	6.96	18.26
S3	-8.70	0.00	13.91	-13.04	-6.09	-13.04	-2.61	-8.70	-6.09	14.78
S4	8.70	2.61	18.26	14.78	0.00	16.52	32.17	-2.61	18.26	0.00

Afin de vérifier les prédictions Pr1a, Pr1b et Pr2, nous avons mené des analyses de variance séparées pour S1 et S3 d'un côté et pour S2 et S4 de l'autre. Les analyses exploitaient trois variables : Voyelle(9), Satellite(2) et Ordre(2). Nous avons exclu les paires de la catégorie /ə/ puisque leur numérotation ne suivait pas le protocole général. Ces paires ont été examinées séparément.

L'analyse commune pour S1 et S3 a révélé que l'effet d'Ordre et l'interaction Ordre* Satellite ont été significatifs [$F(1, 792) = 26.900$, $p < 0.01$ et $F(1, 792) = 36.569$, $p < 0.01$] ($p < 0.01$ pour les deux). L'absence d'interaction significative pour interaction Ordre* Satellite* Voyelle [$F(8, 792) = 2.476$, $p = 0.012$] pourtant, indique que l'ordre n'a pas le même effet sur les satellites des voyelles. En effet, le tableau 5.4 montre que pour certaines voyelles telles que /a/, /e/ et /ɛ/, les prédictions Pr1a et Pr1b ne sont pas confirmées. Des tests *post-hoc* sont alors nécessaires.

Les résultats des analyses *post-hoc* montrent que l'effet d'Ordre est significatif pour S1 [$F(1, 396) = 68.979$, $p < 0.01$], ce qui confirme Pr1a. L'absence d'un tel effet pour S3 [$F(1, 396) = 0.341$, $p < 0.559$] infirme Pr1b.

Quant à S2 et S4, nous avons trouvé un effet d'Ordre [$F(1, 792) = 25.056$, $p < 0.01$] mais l'absence d'interaction significative pour Ordre* Satellite [$F(1, 792) = 0.033$, $p = 0.856$] confirme Pr2. La triple interaction Ordre* Satellite* Voyelle significative [$F(8, 792) = 3.141$, $p < 0.01$] nous oblige à mener des analyses *post-hoc*, comme nous avons fait pour S1 et S3.

Quant à l'analyse séparée pour S2 et S4, nous avons trouvé un effet d'Ordre significatif pour les deux satellites [$F(1, 396) = 12.507$, $p < 0.01$ et $F(1, 396) = 12.589$, $p < 0.01$ pour S2 et S4 respectivement]. Ces résultats infirment alors Pr2.

Nous avons utilisé les données de la catégorie /ə/ dans une analyse séparée. L'effet d'Ordre et l'interaction Ordre*Satellite étaient significatifs pour S1 et S3 [$F(1, 180) = 7.753$, $p < 0.01$ et $F(1, 180) = 23.628$, $p < 0.01$ respectivement], ainsi infirmant Pr3. Ceci n'a pas été le cas pour S2 et S4 [$F(1, 180) = 6.060$, $p = 0.015$ et $F(1, 180) = 0.935$, $p = 0.335$ respectivement], confirmant Pr3. L'absence d'effet d'Ordre pour les quatre satellites séparément a confirmé Pr3. Il est à noter, pourtant, que cette absence pourrait être due au degré de liberté relativement bas (1,44) de chaque ANOVA.

5.1.5 Discussion

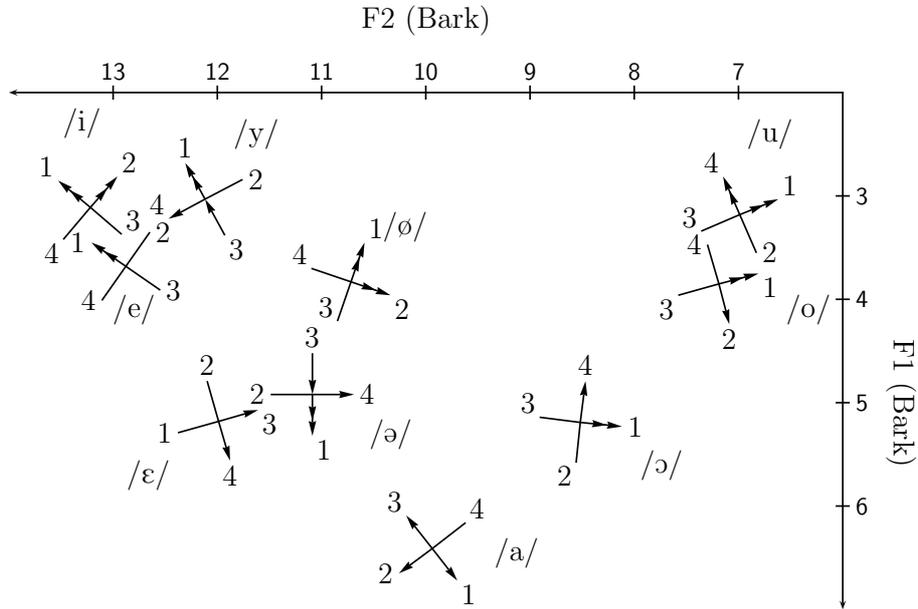
Nous avons vu que parmi les quatre prédictions faites par l'hypothèse de la neutralisation, certaines n'ont pas été confirmées. Aussi avons-nous constaté que dans les paires contenant S1, il était plus facile de discriminer les deux stimuli quand le prototype est en première position dans la paire. L'effet inverse n'a pas été attesté pour S3, ce qui remet en question la validité du modèle de la neutralisation. Les effets d'ordre significatifs pour S2 et S4, ainsi que les effets pour la catégorie /ə/ ont remis en cause l'hypothèse de la neutralisation. Une première conclusion est qu'une partie des données ne peut pas être expliquée par le modèle.

En commentant la figure 2.1, nous avons expliqué que plus une voyelle est localisée dans la périphérie de l'espace, plus le déplacement du centre de sa région de confiance est restreint et doit se faire dans la direction de centre de l'espace vocalique. Une prédiction basée sur ce postulat est que le degré de ce déplacement est corrélé à la distance entre un stimulus et /ə/. Des analyses de corrélation de Spearman séparées pour S1 et S3 n'ont détecté une corrélation significative ni pour S1 ($\rho = -0.286$) ni pour S3 ($\rho = -0.141$). Les valeurs que nous avons utilisées pour ces analyses sont celles apparaissant dans le dernier bloc du tableau 5.4. Notre réanalyse des données de Repp et Crowder (voir tableau II de leur article) a également indiqué qu'une telle corrélation n'était pas significative pour S1 ($\rho = -0.021$, $p = 0.953$) ou S3 ($\rho = -0.500$, $p = 0.157$).

La figure 5.2 est une représentation graphique des effets d'ordre du tableau 5.4. Ici, nous n'avons retenu que les scores supérieurs à 10%. Ce seuil étant subjectif, cette figure ne sert qu'à déceler les tendances centrales des scores bruts.

Nous allons examiner deux possibilités pour interpréter ces résultats. La première approche est de partir de *l'a priori* que le premier stimulus dans une paire se dégrade

FIGURE 5.2: Les effets d'ordre attestés dans nos données. Les flèches marquent la direction dans laquelle la discrimination a été plus facile. Flèche simple : effet d'ordre supérieur à 10% ; flèche double : effet supérieur à 20%.

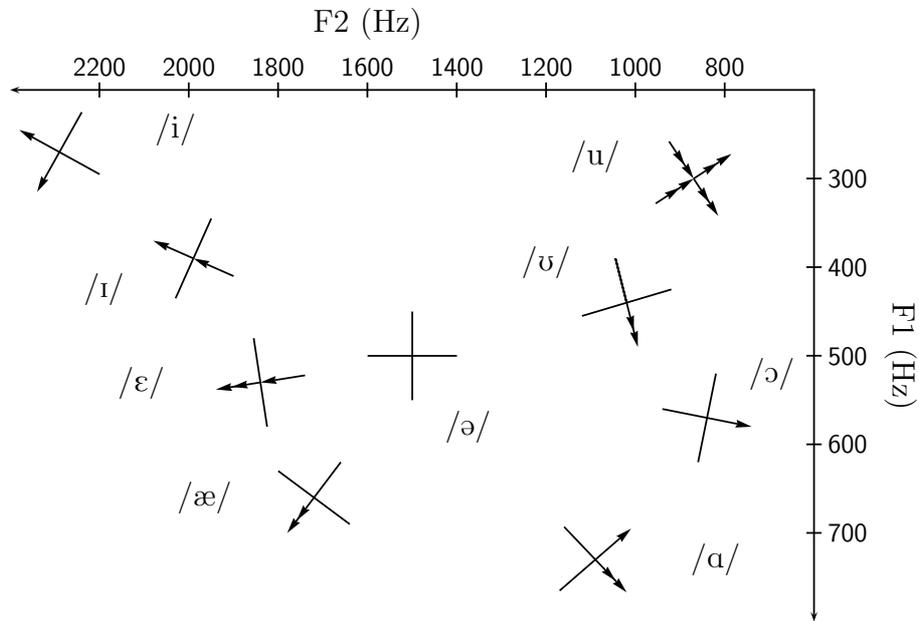


vers un certain point dans l'espace. Dans ce cas, les flèches de la figure 5.2 pointent loin de la cible de dégradation. Nous observons donc que les flèches pour S1 sont en conformité avec une dégradation vers [ə], mais les asymétries pour S2 et S4 indiquent que cette hypothèse n'est pas valide.

En regardant les effets au sein de chaque catégorie, nous remarquons que le 'prototype' de chaque voyelle apparaît comme une cible possible de la dégradation des stimuli. Les flèches à l'intérieur de /i/, /e/, /ɛ/, /a/, /ø/, /o/ et /ɔ/ sont en concordance entre elles, toutes pointant loin du centre de leur catégorie respective. Ce fait pourrait être vu comme une tendance des satellites à se dégrader vers leurs prototypes.

Il est très important de noter que par 'prototype' nous nous référons toujours aux stimuli situés au centre de chaque catégorie et non pas aux représentations mentales que les auditeurs se font des unités linguistiques et qui sont forgées par l'expérience langagière des auditeurs. Afin de faire la distinction entre les dix stimuli prototypes dans notre protocole et les représentations abstraites des catégories linguistiques, nous allons appeler les premiers 'prototypes' et les secondes 'prototypes linguistiques'. Une hypothèse concernant les prototypes linguistiques serait hasardeuse si l'on considère leur variabilité inter- et intra-sujet (Frieda *et al.*, 1999). Elle exigerait une étude de cas qui focaliserait sur la structuration interne des catégories chez chaque auditeur séparément. Dans notre protocole, les prototypes pourraient constituer des points de référence car ce sont les stimuli qui ont été présentés le plus fréquemment : parmi les 13 combinaisons des cinq stimuli de chaque voyelle, les neuf paires comprenaient le prototype donc leur statut particulier pourrait découler de leur fréquence. Or, ce statut n'est pas confirmé dans nos données pour deux raisons :

FIGURE 5.3: Les effets d'ordre attestés dans les données de Repp & Crowder (1990). Les flèches marquent la direction dans laquelle la discrimination a été plus facile. Flèche simple : effet d'ordre supérieur à 10% ; flèche double : effet supérieur à 20%.



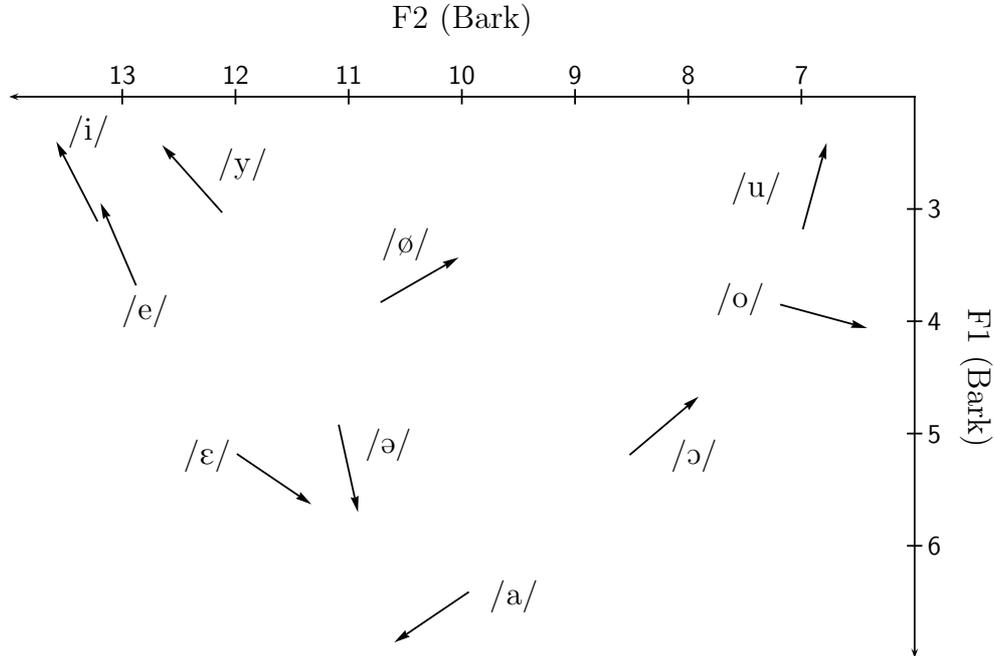
a. des asymétries positives devraient être constatées pour tous les satellites d'une catégorie, ce qui n'a pas été le cas pour aucune des dix voyelles ; b. le degré des effets devrait être plutôt homogène parmi les satellites.

Quant au postulat de R&B selon lequel le point cible de dégradation est plutôt le centre de la dispersion des stimuli, une telle tendance ne trouve pas de soutien particulier pour les groupes /i, e, ɛ/ et /u, o, ɔ, a/. Le groupe /y, ø, ə/, en revanche, indique qu'une telle possibilité, quoique faible, serait envisageable. Le point cible dans ce cas semble être situé autour de S3 de /ə/. Or, les effets pour S3 de /ə/ et pour S4 de /y/ nous empêchent de tirer une conclusion définitive.

Une deuxième approche des asymétries de la figure 5.2 se base sur l'*a priori* selon lequel les effets d'ordre sont dus, non pas à la dégradation du premier stimulus, mais à un effet de contraste déclenché par le deuxième stimulus d'une paire. Ainsi, la discrimination dépendrait de la saillance perceptive de la trace auditive de ce stimulus, la trace du premier stimulus étant plus faible à cause de la dégradation rapide du contenu de la mémoire à court terme. Quand on passe d'un stimulus moins saillant vers un stimulus plus saillant, la distance perceptive des deux stimuli est plus importante. Dans cette logique, quand une flèche pointe vers un des deux membres d'une paire, c'est ce stimulus qui déclenche l'asymétrie et qui est perceptivement plus saillant. Il faut noter ici que par 'saillance perceptive' nous entendons la probabilité qu'un stimulus se comporte comme un point de référence auquel un autre stimulus se compare.

Deux points de référence potentiels sont la périphéricité, c'est-à-dire la proximité d'un stimulus des limites de l'espace acoustique, et la typicalité, soit le degré auquel un stimulus représente la tendance moyenne de sa catégorie linguistique. Dans les

FIGURE 5.4: Représentation vectorielle des effets d'ordre retrouvés dans nos données.



deux cas, les flèches de la figure 5.2 sont interprétées de la façon suivante : la discrimination est plus facile quand le deuxième stimulus est plus saillant au niveau perceptif. Ceci signifie que les flèches de la figure pointent vers un point de référence.

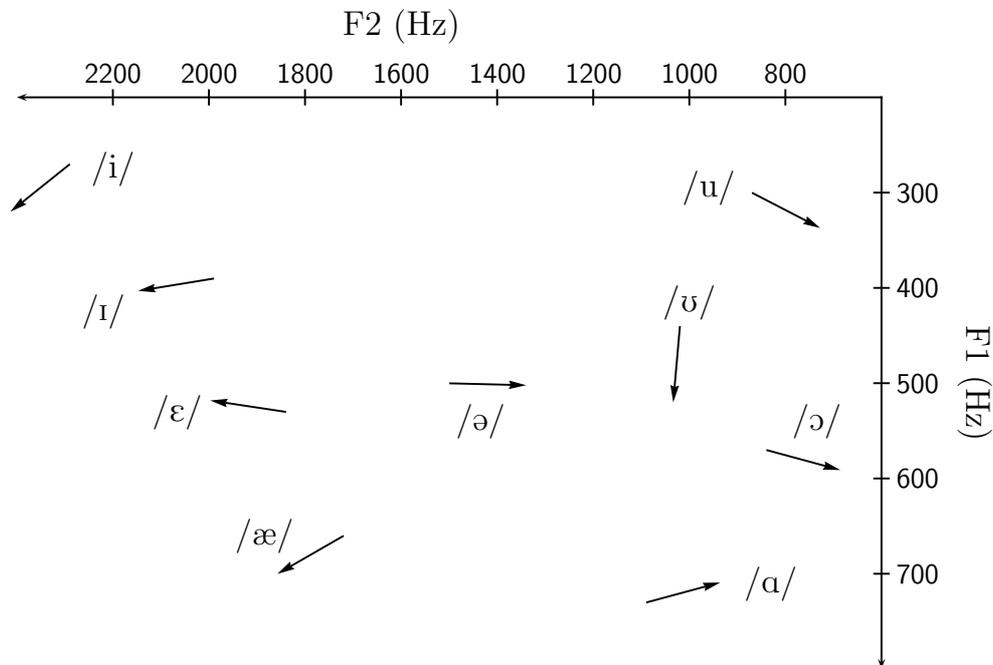
Nous observons que l'hypothèse de la saillance perceptive des prototypes linguistiques est en congruence avec la direction des asymétries : hormis certaines exceptions peu nombreuses, les flèches pointent loin des prototypes. Néanmoins, nous remarquons une tendance relativement constante concernant la direction des asymétries : la grande majorité des flèches pointent vers la périphérie de l'espace vocalique. Ceci est particulièrement évident pour les catégories /i, y, u, o/.

La figure 5.3 présente les asymétries retrouvées dans les données de Repp & Crowder (1990). Nous observons, de nouveau, que la majeure partie des flèches pointe vers les bords de l'espace, fournissant des indications supplémentaires en faveur du rôle de la périphérie.

Vu la subjectivité des seuils sélectionnés pour les figures 5.2 et 5.3, nous avons transformé les effets d'ordre pour les quatre satellites de chaque catégorie en vecteurs. Cette méthode offre l'avantage considérable de prendre en compte tous les effets sont pris en compte afin que la tendance générale de chaque catégorie soit décelée. La transformation a été réalisée de la façon suivante : 1. nous avons additionné les effets pour les satellites situés sur le même axe, soit S1+S3 et S2+S4 ; 2. nous avons additionné les deux 'super-vecteurs' en prenant en considération leur angle et leur longueur. Une calculatrice d'addition de vecteurs a été utilisée. Les résultats de ces calculs sont présentés dans les figures 5.4 et 5.5. Nous avons maintenu la longueur de tous les vecteurs stable, puisqu'il n'est pas possible de quantifier les vecteurs en Hz.

En reprenant les deux approches ci-dessus — dégradation *vs.* effet de contraste

FIGURE 5.5: Représentation vectorielle des effets d'ordre retrouvés dans Repp & Crowder (1990).



— nous constatons que l'hypothèse de dégradation semble être corroborée seulement par le groupe /y, ø, ə/. Néanmoins, la cible de la dégradation ne semble pas être [ə] mais plutôt le point central de cet ensemble. Pour les deux autres groupes, /i, e, ε/ et /u, o, ɔ, a/, une telle convergence n'est pas observée. Les données de R&B révèlent une absence similaire d'indication en faveur de la dégradation vers un point spécifique dans l'espace.

En examinant les vecteurs sous l'optique de la saillance perceptive et de l'effet de contraste, nous observons que la plupart des asymétries se dirigent vers la périphérie de l'espace. La direction des flèches dans les résultats de R&B confirme cette tendance en pointant vers la périphérie. La seule exception est la catégorie /ʊ/.

5.2 Conclusions générales

Dans le présent chapitre, nous avons revisité l'hypothèse de la neutralisation selon laquelle les asymétries dans la perception des voyelles sont provoquées par la dégradation du premier stimulus dans une paire vers le centre de l'espace vocalique. À l'appui d'un ensemble de stimuli construits à partir du système vocalique du français, nous n'avons pas obtenu d'indications suffisamment convaincantes en faveur de l'hypothèse de départ. Les stimuli ne semblent donc pas se dégrader vers le centre de l'espace ni vers le centre de la dispersion des stimuli utilisés dans une tâche.

Or, en interprétant les asymétries dans nos résultats ainsi que dans ceux de Repp & Crowder (1990), nous jugeons que la périphéricité d'une voyelle (sa position marginale dans l'espace vocalique) est un indice plausible expliquant pourquoi l'ordre

de la présentation des stimuli a un effet sur leur discriminabilité. Le postulat que nous formulons alors est que les asymétries ne sont pas dues à la dégradation au niveau de la mémoire précatégorielle mais qu'elles sont plutôt provoquées par le stimulus présenté en dernier dans une paire de voyelles. Ceci est probablement dû à la saillance perceptive des voyelles périphériques, qui semblent agir comme des points d'ancrage auxquels le premier stimulus dans une paire est comparé. Ainsi, quand on passe d'une voyelle vers une autre plus périphérique, la discrimination des deux sons est facilitée.

Malgré les nombreuses études sur ce phénomène, les effets d'ordre restent toujours dissociés des stratégies de production de la parole. Il est possible que les asymétries en perception aient un effet sur le degré de coarticulation des segments. Notre idée est que quand l'ordre de la réalisation de deux voyelles est favorisé par le système auditif, comme par exemple dans le cas d'une séquence /e-/i/, la coarticulation serait moins importante que dans l'ordre inverse, soit /i-/e/. Ceci résulterait de la connaissance implicite du locuteur qu'un moindre effort articulatoire est nécessaire pendant la transition vers une voyelle plus périphérique. Dans le chapitre 7, nous étudions les asymétries en production pour la paire /i-/e/. De même, nous nous penchons sur le concept de la stabilité des cibles acoustiques en fonction de la position d'une voyelle dans l'espace vocalique, afin d'établir un lien entre les mécanismes de la production et ceux de la perception.

Nous allons consacrer le chapitre 6 à l'étude du rôle du paradigme expérimental, ainsi qu'à la stabilité des points de référence dans chaque tâche. Ceci nous permettra de mieux comprendre les effets de contraste et de généralisation. Nous allons aussi essayer de savoir si les asymétries sont déclenchées par des formants spécifiques.

Chapitre 6

La variabilité inter-sujet

*"Le savant n'est pas l'homme qui fournit les vraies réponses;
c'est celui qui pose les vraies questions."
Claude Lévi-Strauss, *Le Cru et le Cuit* (1964)*

Dans les chapitres précédents, nous avons abordé certaines hypothèses principales sur l'origine des asymétries. Dans le chapitre 3, nous avons observé que l'ordre de présentation jouait un rôle sur les patrons de catégorisation (expérience 3) pour les auditeurs des quatre langues étudiées. L'expérience 2 du même chapitre a démontré, pourtant, que dans certains cas, l'expérience linguistique pouvait jouer un rôle. Nous avons retrouvé des asymétries pour le français et pour l'italien, mais pas pour l'espagnol et pour le portugais. La même tâche a indiqué que ce patron asymétrique de discrimination était déclenché pour les stimuli situés autour de l'extrémité /i/ du continuum utilisé. Afin de mieux comprendre pourquoi nous n'avons pas retrouvé d'effets d'ordre autour des deux extrémités, nous avons mené une batterie de tests qui nous ont permis d'étudier les effets de contexte et de contraste que ces deux extrémités exerçaient sur les autres stimuli. Nous avons aussi envisagé ces effets à travers tout le continuum, ainsi que l'impact que l'Intervalle Interstimulus avait sur les effets d'ordre.

Un autre aspect des stimuli que nous avons envisagé sont les fréquences formantiques. Partant de l'hypothèse que les effets d'ordre peuvent être déclenchés par des fréquences localisées dans une région spécifique, nous avons construit 3 continua contenant des tons purs dont les fréquences correspondaient aux fréquences des trois premiers formants des stimuli du continuum originel. Dans une tâche similaire, nous avons filtré les voyelles originelles en appliquant un filtre passe-bande qui n'a retenu que les deux premiers formants de chaque stimulus.

Cette section est consacrée à la variabilité intra-sujet. Nous ne sommes donc pas intéressé par le comportement du groupe. Néanmoins, il ne s'agit que d'une étude pilote qui sert à exploiter quelques pistes sur les diverses sources d'asymétries. Il est donc impossible de tirer des conclusions robustes de ces expériences. Nous essaierons de démontrer que la variabilité intra-sujet est considérable et que la méthode optimale pour étudier les asymétries serait une étude de cas extensive qui ne se baserait pas sur des réponses catégorielles ('même' ou 'différent').

Tableau 6.1: Identifiant des auditeurs, langue maternelle et âge.

Code	Langue	Âge
AR_F1	arabe	30
AR_M1	arabe	32
AR_M2	arabe	30
AR_M3	arabe	26
BB_M1	berbère	36
BP_F2	portugais brésilien	25
BP_F3	portugais brésilien	28
BP_F4	portugais brésilien	24
ES_F1	espagnol	24
FR_F1	français	24
FR_F2	français	47
FR_F3	français	24
FR_F6	français	30
FR_F7	français	24
FR_F8	français	26
FR_F9	français	25
FR_F10	français	23
FR_F11	français	25
FR_F12	français	22
FR_F13	français	30
FR_F14	français	37
FR_F15	français	28
FR_F16	français	26
FR_M4	français	25
FR_M5	français	25
FR_M6	français	45
FS_M1	farsi	35
IT_F1	italien	28
NP_M1	nepali	35
PL_F1	polonais	35
RO_F1	roumain	25
RO_F2	roumain	24
RU_F1	russe	27
RU_M1	russe	43
SB_F1	serbe	25
UK_F1	ukrainien	28

6.1 Sujets

Neuf sujets ont participé à toutes les tâches proposées. Afin de mieux cerner certains aspects de notre travail, nous avons fait appel à d'autres auditeurs pour certaines combinaisons de tâches. Le tableau 6.1 présente les détails des auditeurs. L'identifiant de chaque auditeur est encodé de la façon suivante : 'code de la langue' + 'sexe' (F= femme, M= homme) + 'numéro de série'.

6.2 Stimuli

Les stimuli vocaliques ont été les mêmes que pour l'expérience 1. Pour mieux étudier le rôle que certaines fréquences pouvaient avoir dans le phénomène d'asy-

métrie, nous avons construit trois continua de tons purs. La fréquence des tons de chaque continuum suivait l'évolution de l'un des trois premiers formants des stimuli vocaliques. Les fréquences des stimuli vocaliques et sinusoïdaux sont présentées dans le tableau 3.4, page 3.4.

6.3 Procédure

Nous avons proposé 19 tâches aux auditeurs. Les stimuli ont été présentés en stéréo via des écouteurs dans de petites salles calmes. Nous avons mené la plupart (80%) des expériences dans la chambre sourde de l'ILPGA. Le reste des expériences a été mené dans une salle calme. Nous avons calibré le casque (Sennheiser HD 205) à 70 dB, en utilisant un ton pur à 1000 Hz et le sonomètre Velleman AVM2050. FLXLab 2.0 a servi d'interface. Ces conditions ont été les mêmes pour toutes les expériences.

6.3.1 Identification X (iX)

Deux réponses ont été proposées, <i> et <é>. Après chaque réponse, le stimulus suivant a été présenté avec un délai de 500 ms. Les dix blocs proposés contenaient chacun une répétition des dix stimuli. Une pause a été introduite à la fin de chaque bloc pendant laquelle les participants étaient censés cliquer sur l'écran pour reprendre le test. L'expérience a été précédée d'une session courte d'entraînement, où deux blocs ont été proposés. Les stimuli ont été présentés en ordre aléatoire, déterminé par l'expérimentateur. Par conséquent, l'ordre a été le même pour tous les auditeurs.

6.3.2 Identification avec présentation des deux extrémités (iAXi et iAXe)

Dans ces deux tâches, les dix stimuli du continuum ont été présentés en contexte. Dans iAXi, chaque stimulus a été précédé par le stimulus 1, qui correspondait à l'extrémité /i/. Dans iAXe, chaque stimulus a été précédé par le stimulus 10, qui correspondait à l'extrémité /e/. L'Intervalle Interstimulus a été fixé à 500 ms. Nous avons demandé aux auditeurs d'ignorer le premier stimulus de chaque paire et de catégoriser le deuxième membre de la paire.

6.3.3 Identification avec appariement des stimuli (iAX1r et iAX2r)

Dans ces tâches, nous avons sept blocs de toutes les combinaisons de stimuli. Pour iAX1r, il s'agissait de 18 paires de voyelles différant d'un pas (9 combinaisons x 2 ordres) et 10 paires de voyelles identiques. Pour iAX2r, il s'agissait de 16 paires de voyelles différant d'un pas (8 combinaisons x 2 ordres) et 10 paires de voyelles identiques. Dans l'ordre '*Progressif*', le premier stimulus présenté a été le plus extrême (F1 plus bas, F2 plus élevé), tandis que dans l'ordre '*Régressif*', le stimulus le plus extrême a été présenté en seconde position.

Les expériences ont été précédées d'une courte session d'entraînement contenant un bloc (28 et 26 paires pour iAX1r et iAX2r respectivement).

6.3.4 Discrimination avec présentation des deux extrémités (dAXi et dAXe)

Le protocole de ces expériences a été le même que pour iAXi et iAXe. Les participants ont dû juger si les deux stimuli appariés étaient ‘absolument identiques’ ou ‘même légèrement différents’ en tapant sur la touche ‘m’ de leur clavier pour ‘mêmes’ ou ‘d’ pour ‘différents’ respectivement.

6.3.5 Discrimination avec appariement des stimuli (dAX1r et dAX2r)

Le protocole de ces expériences a été le même que pour iAX1r et iAX2r.

6.3.6 Discrimination avec appariement des stimuli filtrés (dAX2r_fl)

Le protocole de cette expérience a été le même que pour dAX2r. La seule différence concerne les stimuli. Nous avons appliqué un filtre passe-bande sur les dix stimuli. La plage du filtre a été de 0 Hz au deuxième formant du stimulus, avec une fenêtre de glissement de 100 Hz.

6.3.7 Discrimination et Intervalle Interstimulus (dAX2r_250, dAX2r_500 et dAX2r_1000)

Le protocole de ces expériences a été similaire à celui de dAX2r. La première différence concernait l’Intervalle Interstimulus. Nous avons proposé trois tâches avec IIS= 250 (dAX2r_250), 500 (dAX2r_500) ou 1000 (dAX2r_1000) ms. Nous avons aussi réduit le nombre de paires. Nous n’avons inclus que les paires contenant les stimuli 1, 3, 5, 7 et 9.

Six blocs expérimentaux ont été proposés, dont l’un a servi d’entraînement.

6.3.8 Discrimination et protocole ‘fixed’ (dAX2f)

Au lieu d’utiliser un protocole ‘roving’ comme dans dAX2r, où toutes les paires sont présentées en un seul bloc, nous avons utilisé un paradigme ‘fixed’, où chaque bloc ne contient que les paires incluant deux stimuli. Le premier bloc a compris 8 répétitions des paires 01-01, 02-02, 01-02 et 02-01. Le deuxième bloc a compris les paires 02-02, 03-03, 02-03 et 03-02, et ainsi de suite.

6.3.9 Discrimination de tons purs (f1i, f1ii, f2i, f2ii, f3i, f3ii)

Ces six tâches suivaient les protocoles de dAX1r et dAX2r. Les tons purs ayant servi de stimuli (voir plus haut pour une description) suivaient l’évolution des formants des dix stimuli originels. Dans f1i, f2i, et f3i, nous avons intégré les combinaisons de stimuli différant d’un pas. Dans f1ii, f2ii et f3ii, les stimuli appariés différaient de deux pas. Comme dans les autres tâches ‘roving’, les paires contenant des stimuli identiques ont été incluses.

6.4 Résultats

Dans les sous-sections suivantes, nous allons envisager les diverses hypothèses une à une et présenter les graphes et les statistiques qui en découlent.

6.4.1 Effets de contexte – ordre et les extrémités /i/ et /e/

Dans le chapitre 3, nous n'avons retrouvé d'effet d'ordre qu'autour de l'extrémité /i/. Nous avons alors formulé l'hypothèse selon laquelle ces stimuli périphériques et focaux servent de points de référence. Quand ces stimuli sont en seconde position dans une paire (ordre progressif), les auditeurs ont tendance à donner davantage de réponses 'différent' que dans l'ordre régressif.

Dans les tâches suivantes, nous avons utilisé les deux extrémités comme standards auxquels les autres stimuli sont comparés. Nous avons ensuite conduit des tests χ^2 2x2 pour chaque auditeur, où les lignes correspondaient aux deux tâches et les colonnes au nombre de réponses ('i' et 'é'). Le tableau 6.2 illustre les résultats de ces tests. Nous observons que la relation entre le standard (stimulus 01 ou 10) et la proportion des réponses est significative seulement pour 8/20 auditeurs en identification et pour 5/22 auditeurs en discrimination. Il paraît donc difficile de défendre le postulat que les stimuli focaux déclenchent des effets de contexte ou de contraste de façon universelle. Dans les tâches d'identification, nous obtenons des effets de contexte pour le sujet italien et pour la plupart des auditeurs francophones. Il s'agit des deux groupes pour lesquels des asymétries ont été retrouvées dans l'étude inter-langue. Néanmoins, les tâches de discrimination nous amènent à la conclusion que la focalisation ou la périphéricité déclenchent davantage d'effets de contexte que de contraste. L'absence de cohérence dans les tests χ^2 pour l'identification et pour la discrimination indique que les effets de contraste et de contexte ne sont pas interdépendants.

Les figures 6.1 illustrent les résultats des quatre tâches. Ici, les stimuli dont l'identification est le plus affecté par la présence d'un standard sont ceux qui sont situés au milieu du continuum (stimuli no. 04–07). Les stimuli sur les extrémités paraissent très stables. Il est possible que les extrémités agissent comme des points d'ancrage, comme Durlach & Braida (1969) et Macmillan *et al.* (1988) soutiennent. Le modèle duel (page 34) accepte comme points de repère les frontières phonémiques. Ceci n'explique pas pourquoi les extrémités sont plus résistantes à l'influence du standard. Le modèle de trace-contexte paraît plus adapté ici. Le sujet estime la distance de chaque stimulus par rapport aux deux extrémités. Ensuite, des informations sur la frontière catégorielle sont récupérées de la MLT. La présentation du standard aura un effet sur la localisation de la frontière et la déplacera de façon temporaire.

6.4.2 Effets de contexte et distance acoustique (iAX1r et iAX2r)

Nous avons continué l'étude des effets de contexte en utilisant un paradigme 'roving'. Nous avons de nouveau mené des tests χ^2 3x2 pour déterminer la relation entre le contexte (le stimulus précédent, à savoir le standard) et la proportion des réponses 'i' et 'é'. Prenons le stimulus 03. Les trois contextes ont été n (03-03), $n+1$ (04-03) et $n-1$ (02-03). Pour les conditions $n+1$ et $n-1$, le stimulus servant de

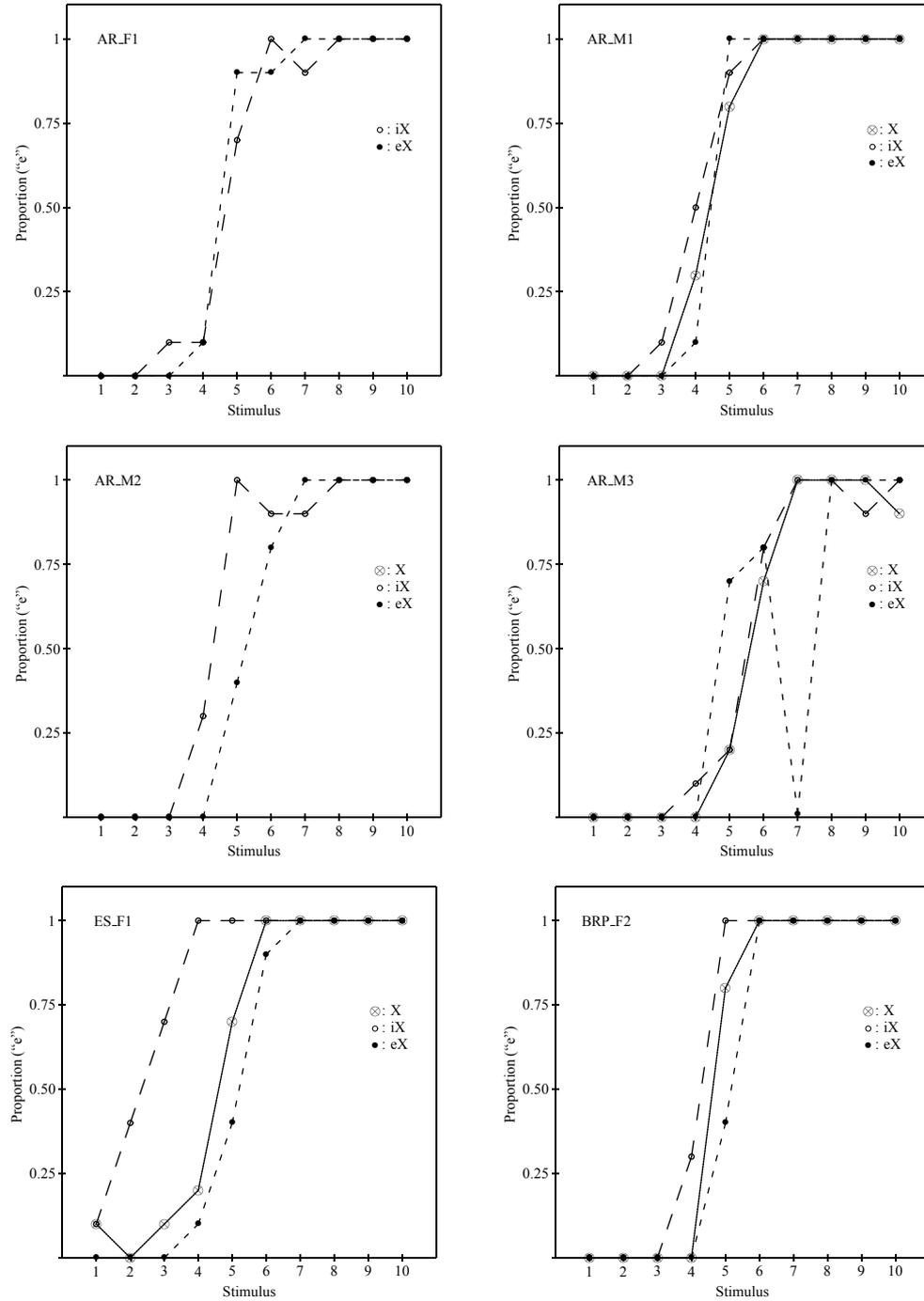
FIGURE 6.1: Les résultats d'identification et de discrimination pour les tâches iAXi-iAXe et dAXi-dAXe. $n = 10$.

FIGURE 6.1: Les résultats d'identification et de discrimination pour les tâches iAXi-iAXe et dAXi-dAXe. (cont.)

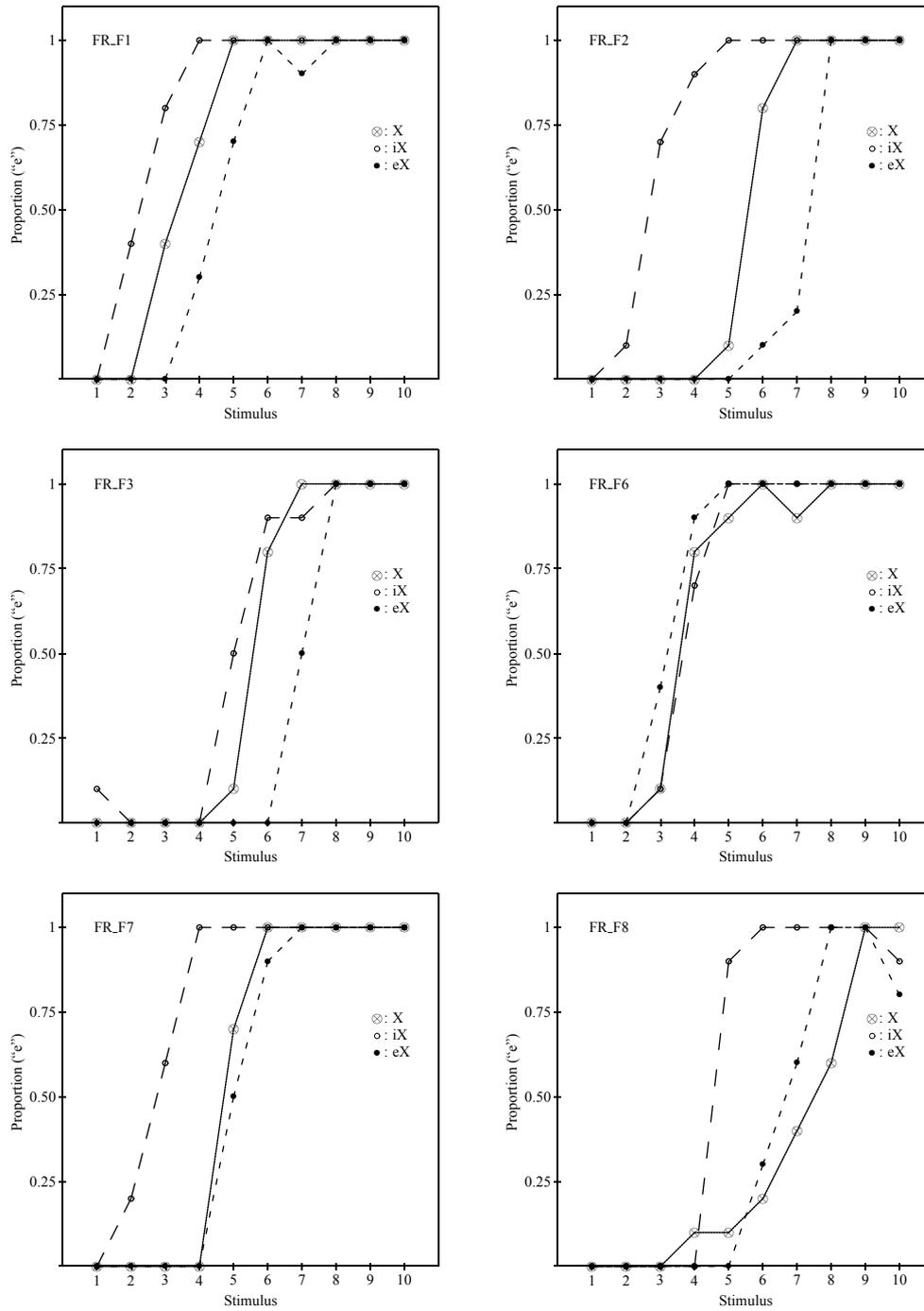


FIGURE 6.1: Les résultats d'identification et de discrimination pour les tâches iAXi-iAXe et dAXi-dAXe. (cont.)

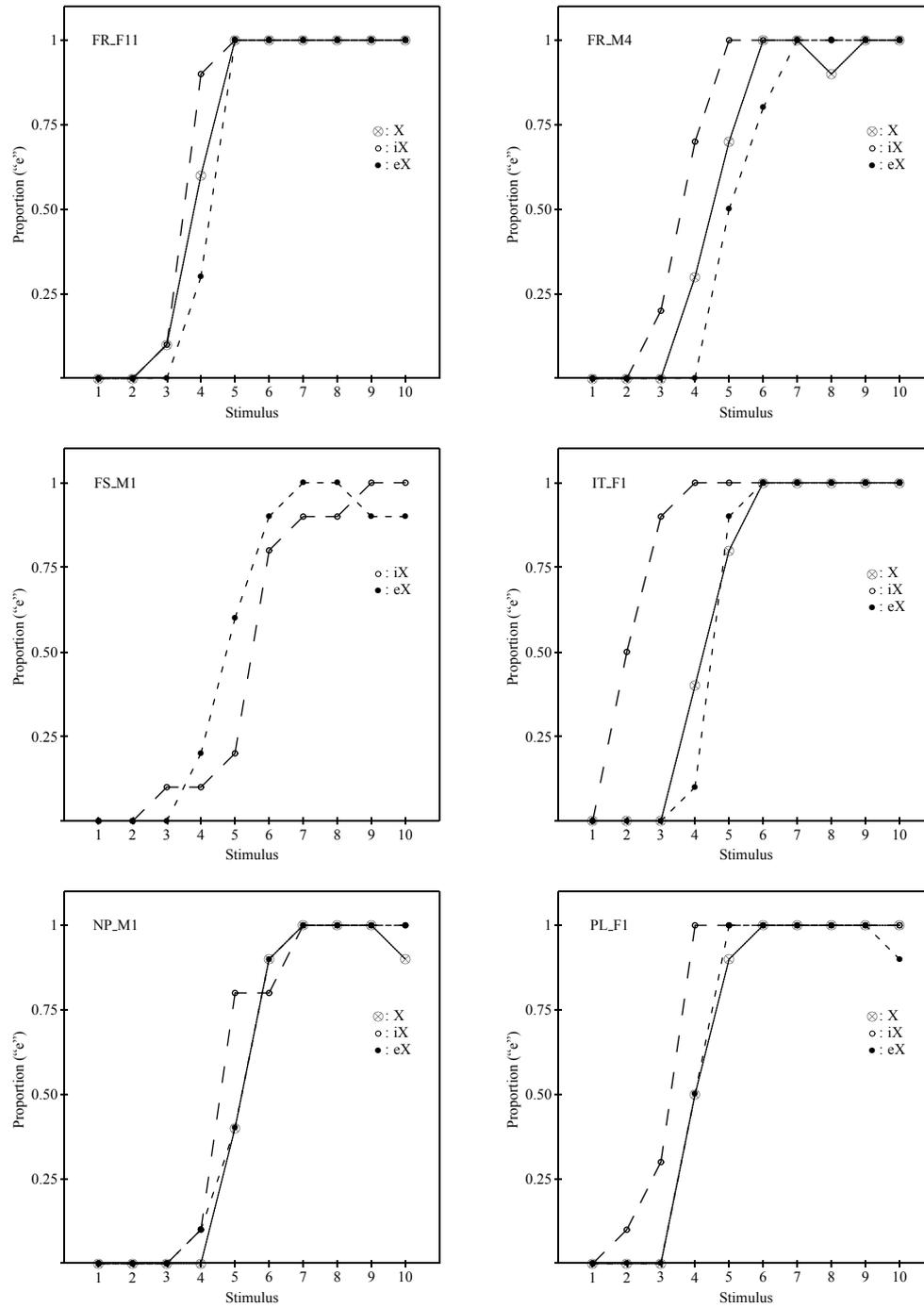


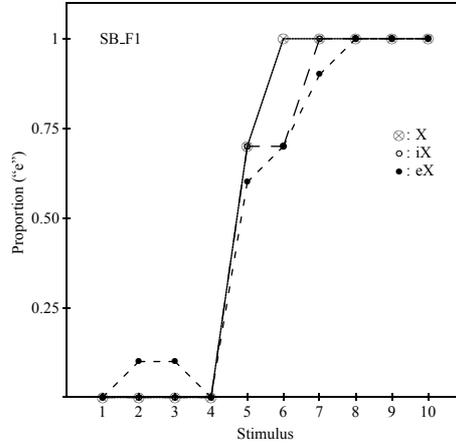
Tableau 6.2: Résultats des tests χ^2 ($ddl = 1$, $n = 200$) pour les tests iAXi-iAXe (colonnes 2 et 3) et dAXi-dAXe (colonnes 4 et 5). * : le chiffre dans au-moins une cellule de l'analyse était inférieur à 5.

	iAXi-iAXe		dAXi-dAXe	
	χ^2	p	χ^2	p
AR_F1	0.00	1.000	7.25	<.01
AR_M1	0.19	0.663	3.11	0.078
AR_M2	1.30	0.254	0.02	0.888
AR_M3	0.32	0.572	2.42	0.120
BRP_F2	1.32	0.251	1.48	0.224
ES_F1	16.75	<.001	32.00	<.001
FR_F1	11.64	<.001	0.03	0.863
FR_F2	37.35	<.001	2.66	0.103
FR_F3	6.56	<.05	0.53	0.467
FR_F6	0.38	0.538	0.03	0.863
FR_F7	11.79	<.001	0.65	0.420
FR_F8	8.02	<.01	4.19	<.05
FR_F10			1.64	0.200
FR_F11	0.81	0.368	4.01	<.05
FR_F17	1.33	0.249	7.51	<.01
FR_M4	4.73	<.05	1.84	0.175
FS_M1	0.32	0.572		
IT_F1	13.12	<.001	0.02	0.888
NP_M1	0.08	0.777	0.99	0.320
PL_F1	1.89	0.169	10.5	<.005
RO_F1			1.48	0.224
RU_F1			3.07	0.080
SB_F1	0.02	0.888	1.35	0.245

contexte (point de repère) différait d'un (tâche iAX1r) ou de deux pas (tâche iAX2r) sur le continuum. Les résultats des analyses sont disponibles dans le tableau 6.3.

Nous observons que les différences entre les trois contextes ne sont significatives que dans la condition de deux pas (iAX2r). Ceci pourrait être dû au fait que dans la condition d'un pas, la différence acoustique en termes de formants entre les stimuli consécutifs est inférieure au seuil de discriminabilité. En revanche, nous avons vu dans l'expérience 2 du chapitre 3 que des effets d'ordre prononcés se manifestent dans la condition d'un pas. Ceci, en combinaison avec l'absence de corrélation entre les effets de contexte et d'ordre pour la Langue des Signes Française (chapitre 4), nous amène au postulat que les effets d'ordre ne sont pas dus à des stratégies différentielles d'étiquetage ou au déplacement temporaire des frontières phonémiques. Un argument en faveur de ce postulat est le fait que les effets d'ordre sont concentrés à l'extrémité /i/ et non pas autour des frontières phonémiques. Dans la section précédente, nous avons démontré que les stimuli aux extrémités étaient résistants aux effets de contexte et de contraste.

FIGURE 6.1: Les résultats d'identification et de discrimination pour les tâches iAXi-iAXe et dAXi-dAXe. (cont.)

Tableau 6.3: Résultats des tests χ^2 pour les tâches iAX1r ($ddl = 2$, $n = 196$) et iAX2r ($ddl = 2$, $n = 182$). * : le contenu d'au-moins une cellule était inférieur à 5.

	iAX1r		iAX2r	
	χ^2	p	χ^2	p
AR_F1	4.13	0.127	10.68	<.005
AR_M1	0.04	0.980	6.49	<.05
FR_F1	0.03	0.985	21.75	<.001
FR_F3	0.71	0.701	10.54	<.01
FR_F6	0.69	0.708	16.23*	<.001
FR_F7	1.16	0.560	10.28	<.01
FR_F8	0.53	0.767	15.87	<.001
FR_F17	1.28	0.527	14.55	<.001
FR_M4	0.23	0.891	7.01	<.05
IT_F1			16.63	<.001
NP_M1	0.85	0.654	15.29	<.001
PL_F1	1.42	0.492	2.14	0.343

6.4.3 Effets d'ordre et distance acoustique (dAX1r et dAX2r)

Nous avons vu dans le chapitre 3 (voir le tableau 3.12, page 3.12) que les asymétries étaient en relation avec la distance acoustique. Nous avons effectué un test χ^2 2x2, où les deux variables indépendantes étaient la tâche (dAX1r et dAX2r) et l'ordre (Progressif ou Régressif) et la variable dépendante, les nombres de bonnes détections. Les résultats sont disponibles dans le tableau 6.4. Nous observons que la relation entre les deux variables n'est significative que pour deux auditeurs sur 19. Étant donné la taille de l'échantillon (de 48 à 186) plutôt limitée, il est difficile de conclure sur l'impact de la distance acoustique sur les asymétries.

Dans la figure 6.2, nous illustrons les effets d'ordre pour les deux tâches. Pour la plupart des 18 sujets, les effets sont soit inexistantes, soit irréguliers. Les résultats

Tableau 6.4: Résultats des tests χ^2 2x2 sur les tâches dAX1r et dAX2r : nombre de détections correctes, valeur χ^2 et valeur p .

	n	χ^2	p
AR_F1	48	0	1.000
AR_M1	154	0.1	0.752
BRP_F2	84	1.81	0.179
BRP_F3	141	0.54	0.462
BRP_F4	148	0	1.000
FR_F1	146	18.3	<.001
FR_F3	185	0.11	0.740
FR_F6	186	0.01	0.920
FR_F7	127	3.02	0.082
FR_F8	116	0.05	0.823
FR_F10	74	0.12	0.729
FR_F11	59	0.42	0.517
FR_F12	78	0.50	0.480
FR_F17	162	0.54	0.462
FR_M4	122	7.18*	<.01
FR_M6	172	0.03	0.863
IT_F1	133	0.16	0.689
NP_M1	112	0	1.000
PL_F1	117	0.14	0.708

des cinq sujets confirment les tendances de l'expérience 2 du chapitre 3 : FR_F3, FR_F11, IT_F1 et PL_F1. Nous avons des effets positifs autour de l'extrémité /i/ ($02 - 01 > 01 - 02$) et des effets négatifs autour de l'extrémité /e/ ($09 - 10 > 10 - 09$). Dans les deux cas, la discrimination est plus élevée quand le deuxième stimulus est plus proche de l'extrémité respective, ce qui apporte des indications en faveur du rôle des extrémités en tant que points d'ancrage. Le deuxième stimulus semble exercer ici un effet de contraste.

6.4.4 Effets d'ordre et apport des formants supérieurs (dAX2r et dAX2r_fil)

Nous avons postulé qu'une source possible d'asymétries pouvait être le caractère focal des voyelles. Pour vérifier cette hypothèse, nous avons filtré les dix stimuli afin d'enlever les informations concernant F3 et F4, qui sont la source de cette focalisation dans notre continuum. Les résultats apparaissent dans la figure 6.3. Nous observons deux tendances générales. Pour les sujets AR_M1, FR_F11, IT_F1 et PL_F1, les effets d'ordre sont positifs autour de /i/, ils se neutralisent autour de la frontière pour ensuite s'inverser vers l'extrémité /e/. Ce patron renforce l'hypothèse des points d'ancrage. Pour la plupart des autres sujets, les asymétries ne sont pas régulières. Ce qui est fréquent, pourtant, est un effet négatif ($Pro > Rég$) pour la paire 04-06, qui rejoint l'hypothèse de Lipski *et al.* (2007) que le changement de /i/ à /e/ provoque un contraste plus prononcé. Il est donc possible que cet effet soit lié à la frontière catégorielle.

Le tableau 6.5 démontre que la relation entre la nature des stimuli et l'ordre n'est

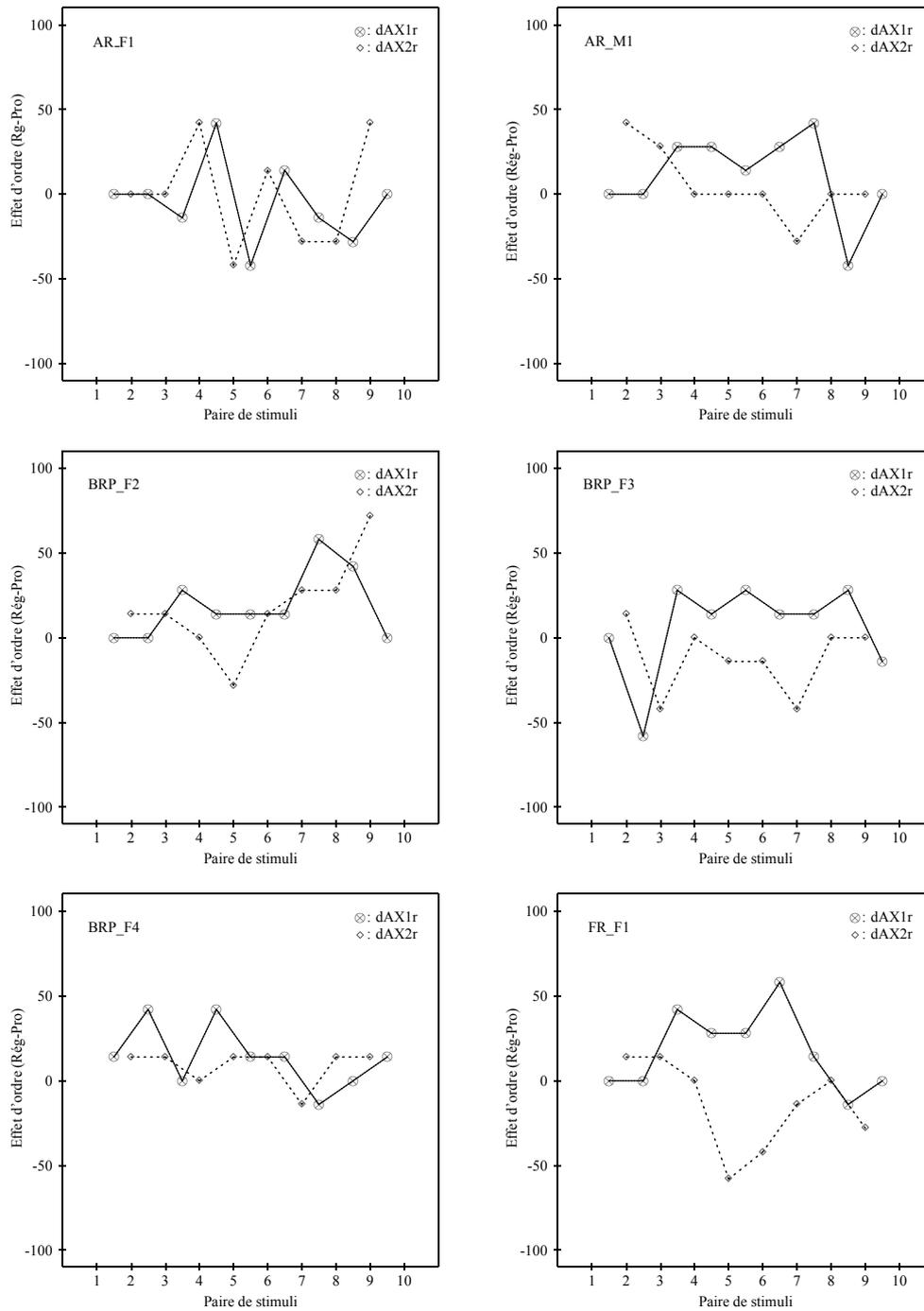
FIGURE 6.2: Les effets d'ordre pour les tâches dAX1r et dAX2r. $n = 7$.

FIGURE 6.2: Les effets d'ordre pour les tâches dAX1r et dAX2r. (cont.)

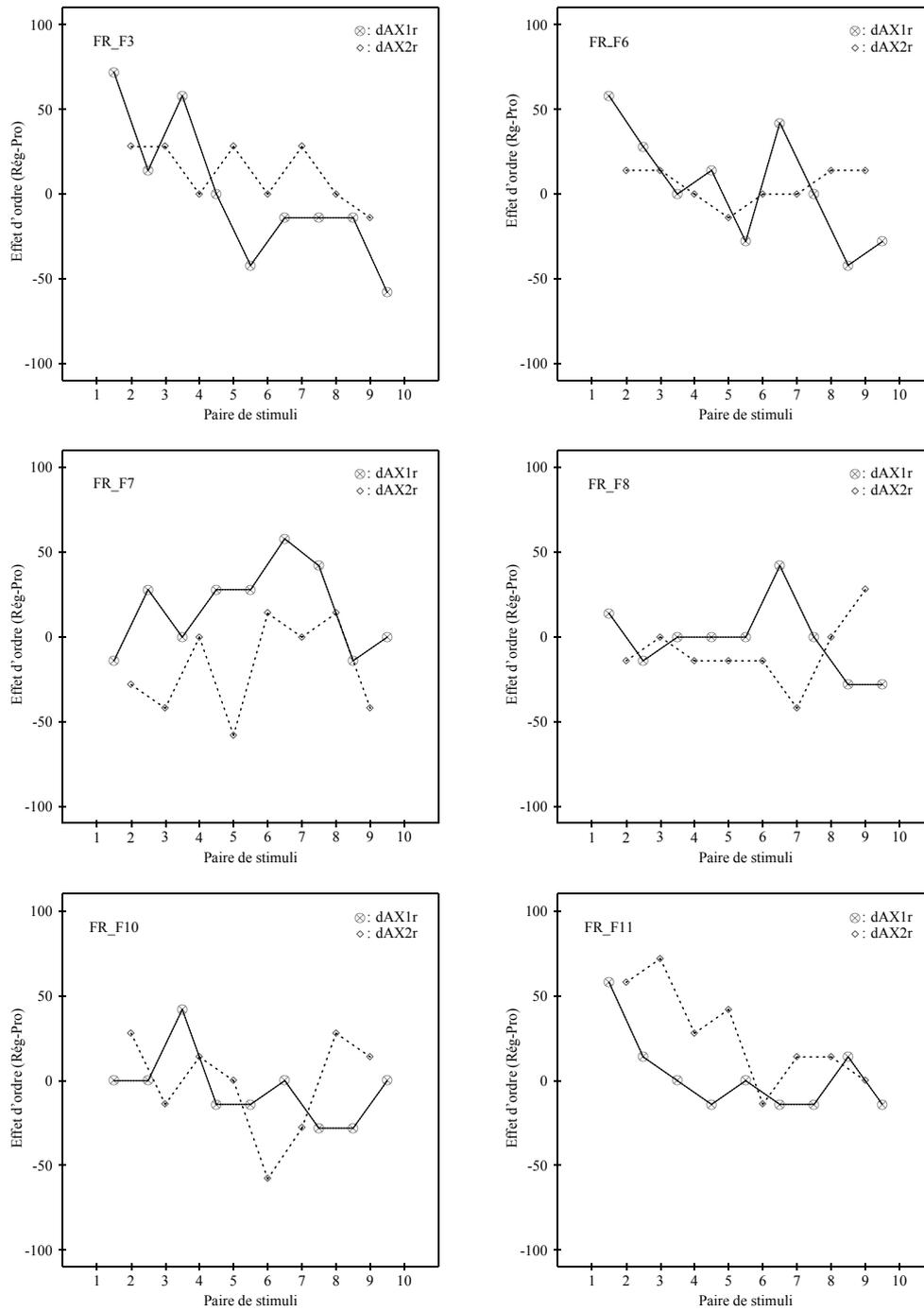


FIGURE 6.2: Les effets d'ordre pour les tâches dAX1r et dAX2r. (cont.)

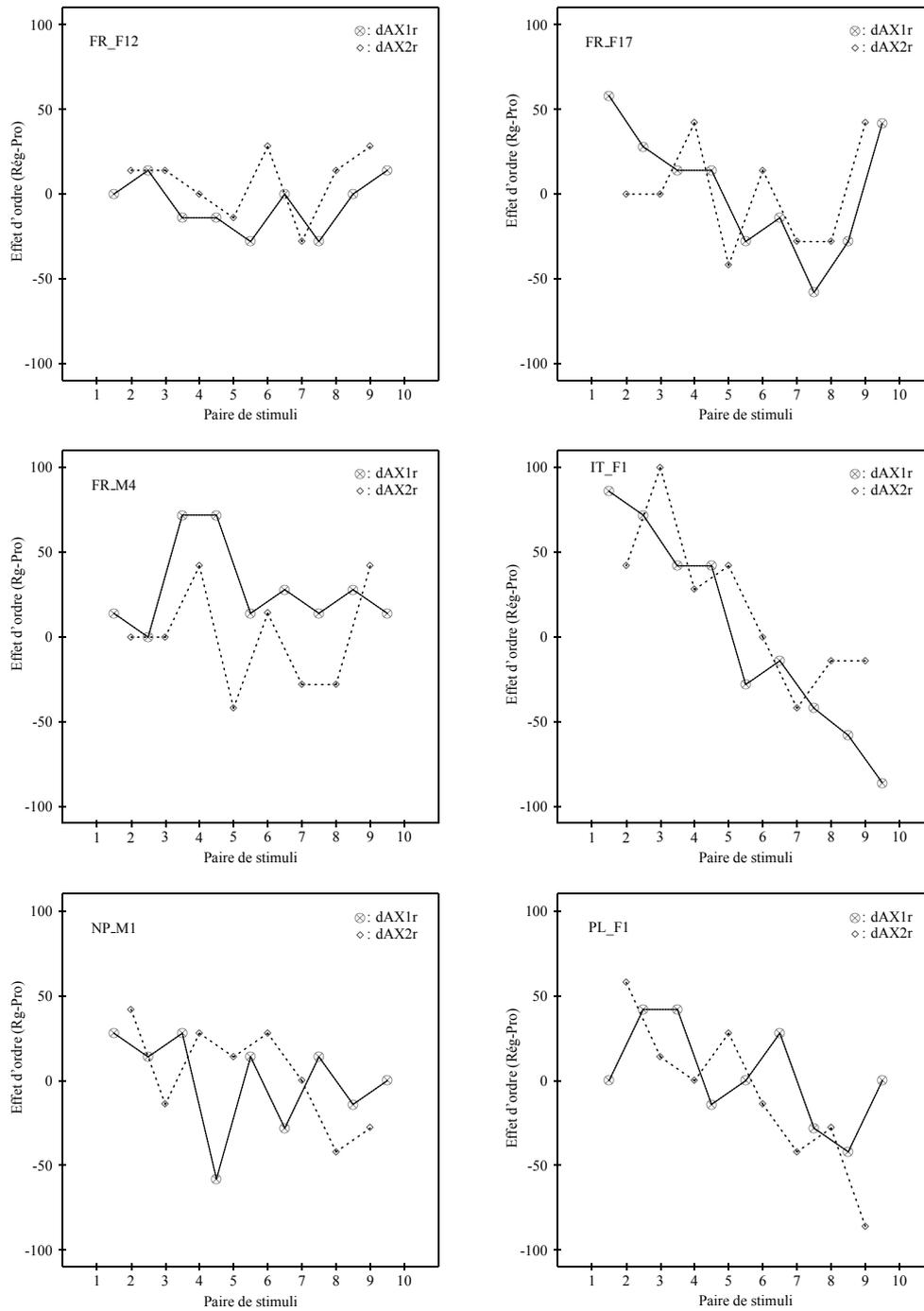


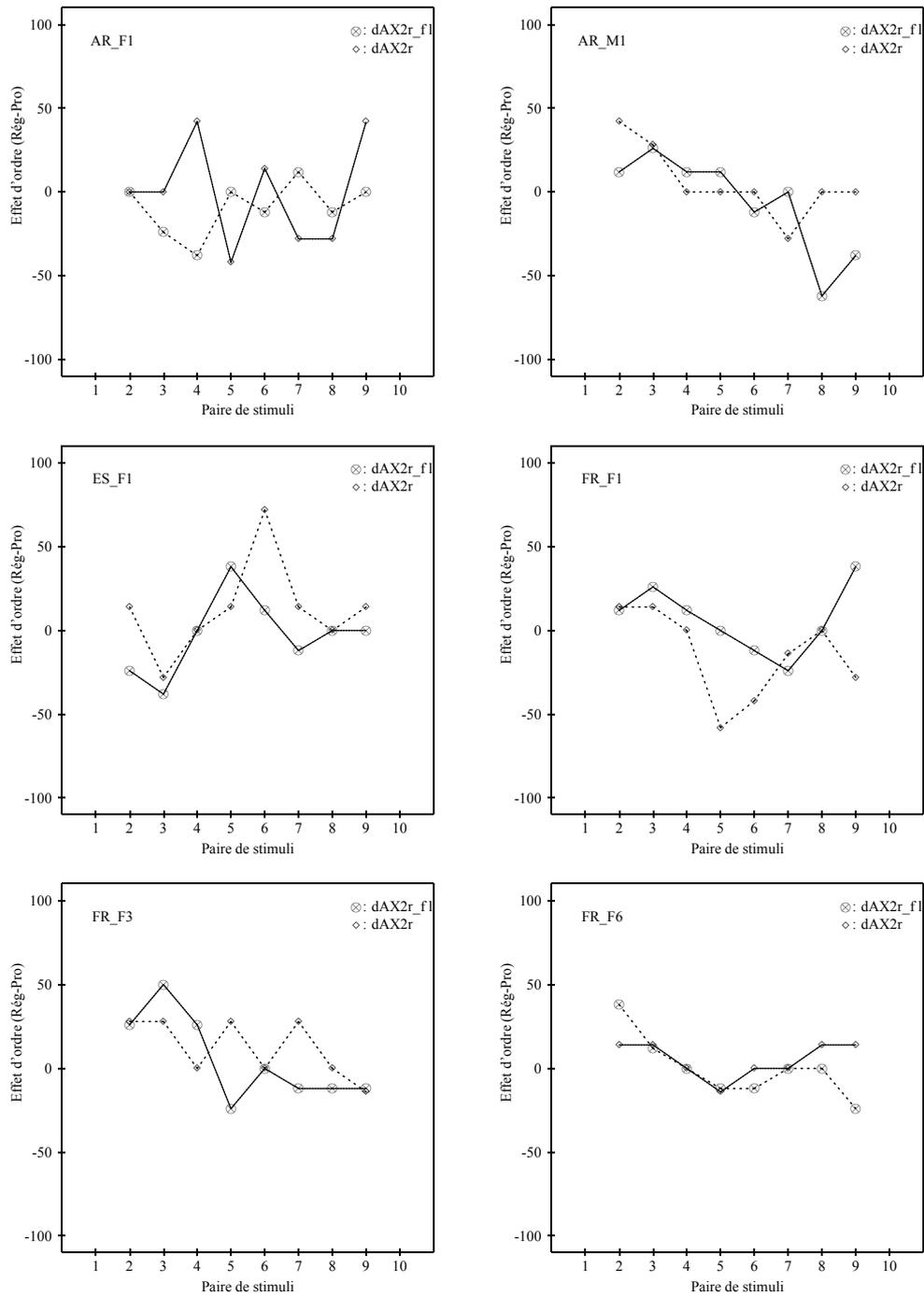
FIGURE 6.3: Les effets d'ordre pour les tâches dAX2r et dAX2r_fil. $n = 7$.

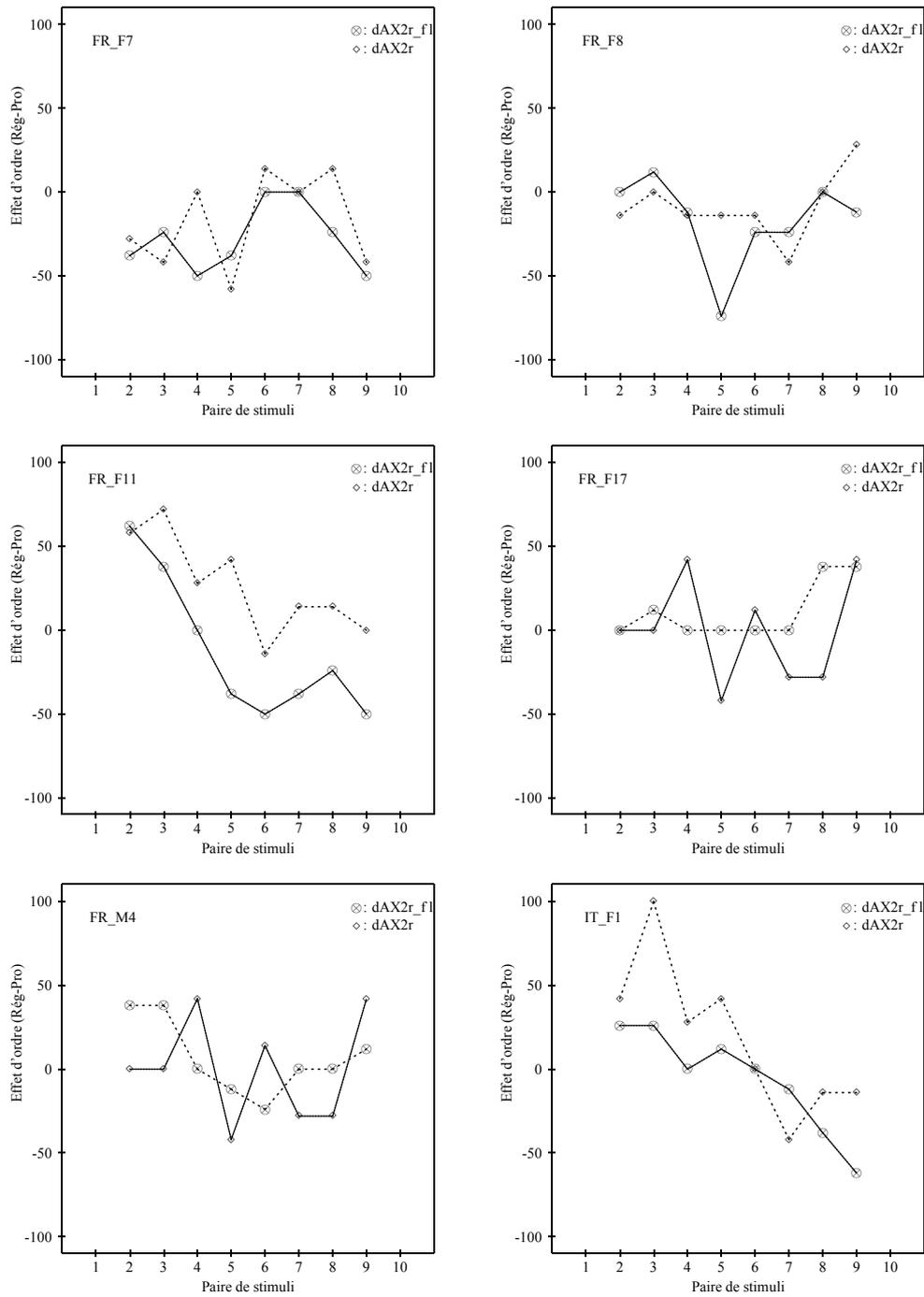
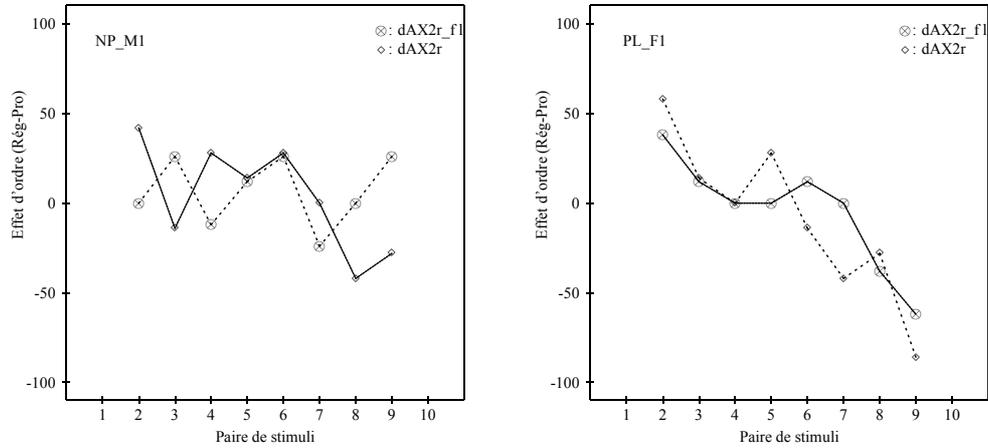
FIGURE 6.3: Les effets d'ordre pour les tâches dAX2r et dAX2r_fil. $n = 7$. (cont.)

FIGURE 6.3: Les effets d'ordre pour les tâches dAX2r et dAX2r_fil. $n = 7$. (cont.)Tableau 6.5: Résultats des tests χ^2 2x2 sur les tâches dAX2r et dAX2r_fil : nombre de détections correctes, valeur χ^2 et valeur p .

	n	χ^2	p
AR_F1	53	0.05	0.823
AR_M1	169	0.05	0.823
ES_F1	86	0.71	0.399
FR_F1	142	0.63	0.427
FR_F3	191	0.04	0.842
FR_F6	195	0.00	1.000
FR_F7	143	0.15	0.699
FR_F8	108	0.05	0.823
FR_F11	122	4.72	<.05
FR_F17	188	1.09	0.297
FR_M4	181	0.00	1.000
IT_F1	177	0.75	0.387
NP_M1	100	0.14	0.708
PL_F1	173	0.00	1.000

significative que pour FR_F11. Pour ce sujet, il y a eu plus de détections correctes dans l'ordre Régressif que dans l'ordre Progressif.

6.4.5 Effets d'ordre et différence entre paradigmes (dAX2r et dAX2f)

Dans une expérience 'même-différent' qui utilise un paradigme 'roving', chaque bloc contient toutes les paires possibles. Dans une tâche 'fixed', un bloc ne comprend que deux stimuli du continuum, à savoir les quatre combinaisons possibles : XX, XY, YX et YY. Dans le cadre de la théorie de trace-contexte, cette différence de protocole aura une conséquence sur la résolution des stimuli. Plus la dispersion des stimuli est importante, plus la variance contextuelle est élevée et plus le sujet aura

Tableau 6.6: Résultats des tests χ^2 2x2 sur les tâches dAX2r et dAX2f : nombre de détections correctes, valeur χ^2 et valeur p .

	n	χ^2	p
AR_F1	64	0.08	0.777
AR_M1	205	0.35	0.554
BRP_F2	175	0.49	0.484
ES_F1	153	0.15	0.699
FR_F1	209	0.23	0.632
FR_F2	183	1.18	0.277
FR_F3	218	0.1	0.752
FR_F6	233	0.00	1.000
FR_F7	185	0.63	0.427
FR_F8	147	0.24	0.624
FR_F10	166	0.01	0.920
FR_F11	145	2.91	0.088
FR_F17	213	0.26	0.610
FR_M4	223	0.07	0.791
IT_F1	206	0.51	0.475
NP_M1	177	0.00	1.000
PL_F1	204	0.01	0.920

du mal à détecter des différences entre les stimuli. En revanche, nous ne pouvons pas prédire quel sera l'effet de cette variance sur les asymétries. Le tableau indique que le patron des effets d'ordre n'est pas significativement différent pour les deux tâches. Ceci est dû au fait que dans la tâche dAX2f, des effets de plafond ont été très facilement atteints. Les effets de plafond ont eu comme conséquence des fonctions plates (figure 6.4).

6.4.6 Effets d'ordre et l'Intervalle Interstimulus

Les hypothèses que nous pouvons formuler concernant le rôle de l'Intervalle Interstimulus dépendent de notre présupposition sur la nature des points d'ancrage :

- périphérie : des effets positifs sont retrouvés à travers tout le continuum. L'impact d'IIS renforce l'effet de contraste et, par là, les effets d'ordre
- focalisation : des effets positifs sont retrouvés à l'extrémité /i/. L'impact d'ISI renforce l'effet de contraste et, par là, les effets d'ordre
- extrémités : les extrémités servent de points d'ancrage. Des effets positifs autour de /i/ et des effets négatifs autour de /e/ sont prédits

En regardant les 16 sous-figures, nous ne trouvons pas d'exemple typique pour les deux premiers scénarios. La troisième hypothèse, pourtant, semble être partiellement confirmée par les résultats des sujets FR_F15, NP_M1, PL_M1 et IT_F1, où nous observons une inversion des asymétries autour des extrémités. Or, la différence entre les trois IIS ne semble pas favoriser le postulat d'effet de contraste. Nous nous attendions à des effets d'ordre plus prononcés pour des IIS plus grands, étant donné la dégradation plus importante du premier stimulus. Cette dégradation incite le sujet à se baser plus sur le deuxième stimulus. Et plus le deuxième stimulus est saillant, plus il doit déclencher d'asymétries. Malgré quelques différences entre les trois IIS,

FIGURE 6.4: Les effets d'ordre pour les tâches dAX2r et dAX2f. $n = 7$.

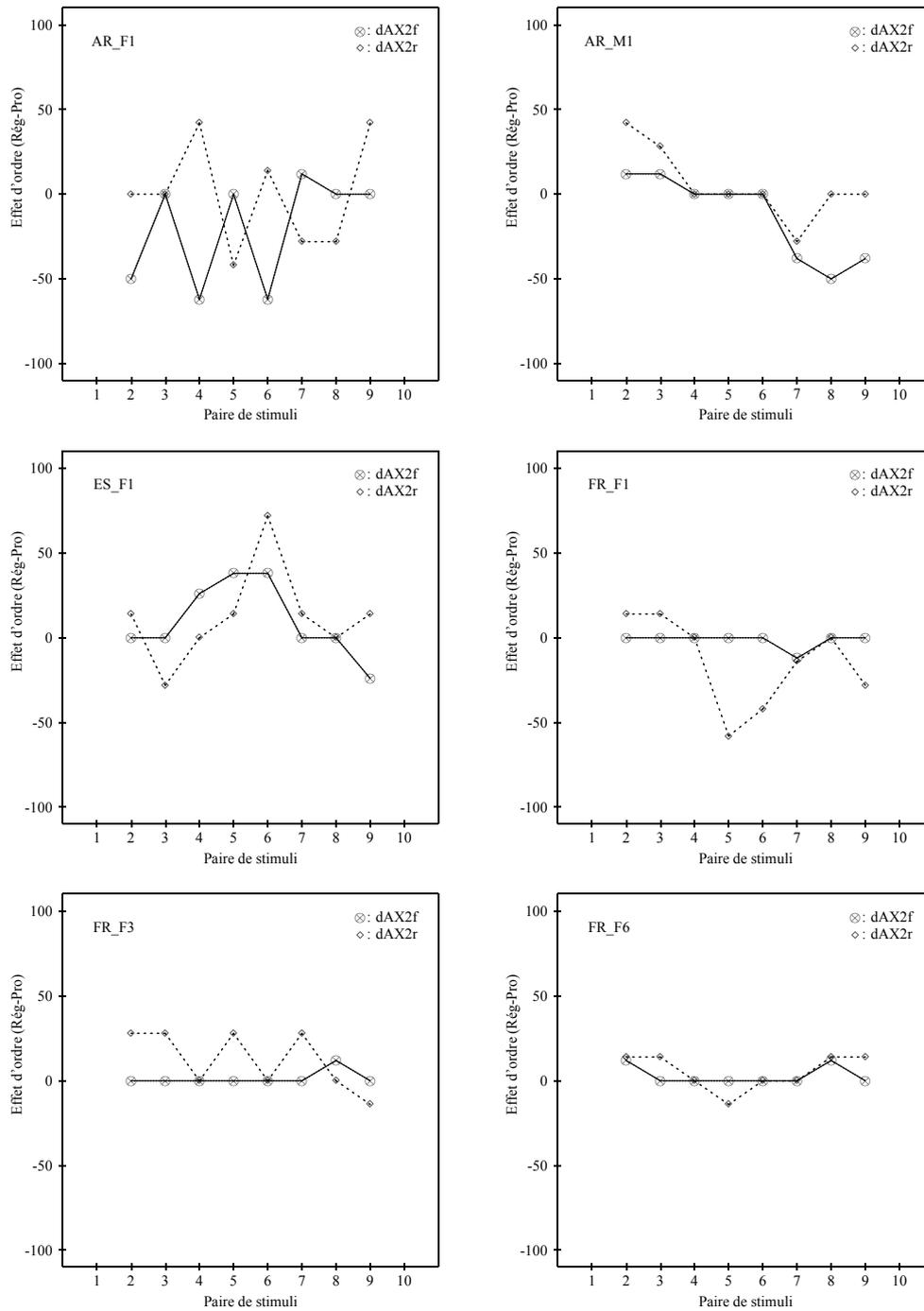


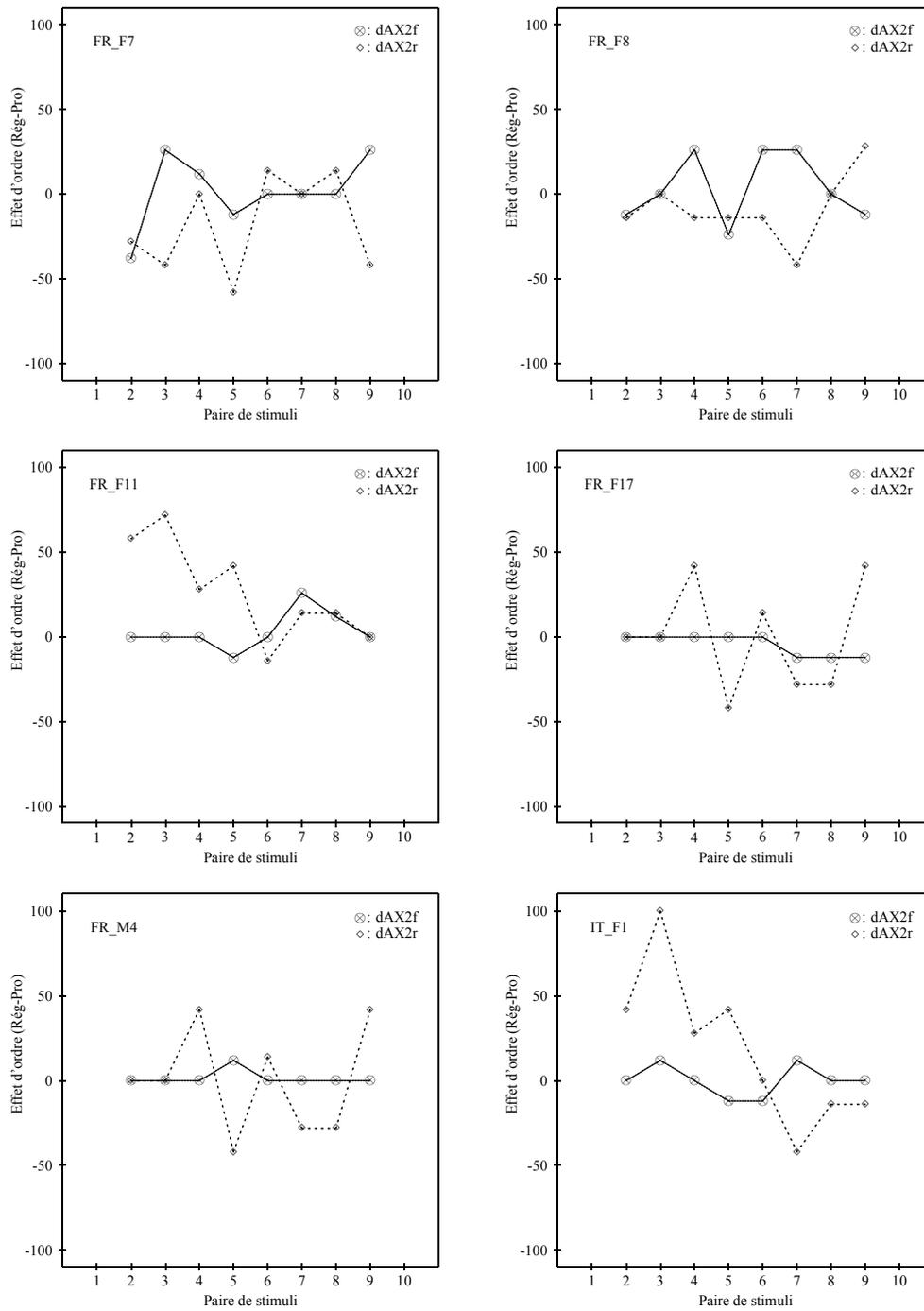
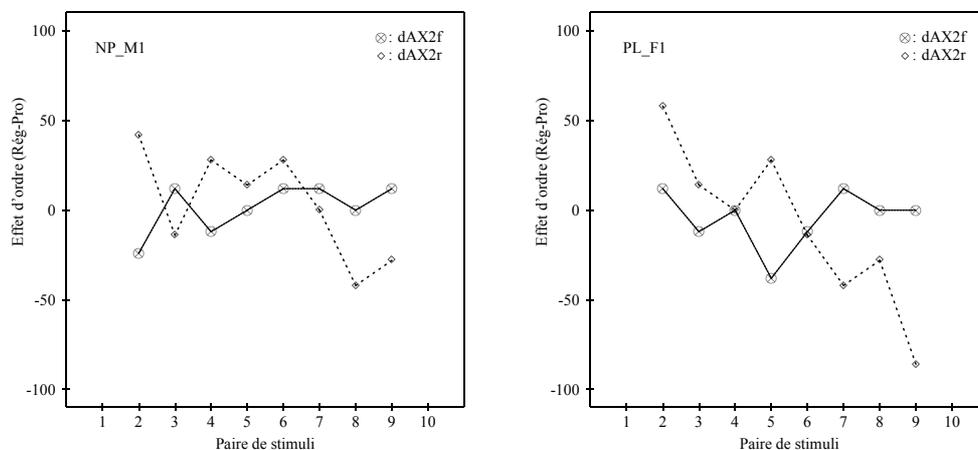
FIGURE 6.4: Les effets d'ordre pour les tâches dAX2r et dAX2f. $n = 7$. (cont.)

FIGURE 6.4: Les effets d'ordre pour les tâches dAX2r et dAX2f. $n = 7$. (cont.)

celles-ci ne sont pas régulières : paire 05-07 pour RO_F2, 07-09 pour FR_F8, 03-05 pour FR_F3, 01-03 pour FR_M4. Il est donc difficile de conclure sur le rôle de l'IIS.

6.4.7 Effets d'ordre et tons purs

Les six prochaines tâches contenant des tons purs nous ont permis de vérifier deux hypothèses principales :

- les extrémités comme points d’ancrage : le patron décrit ci-avant est attendu, à savoir des effets positifs autour du stimulus 01, et des effets négatifs autour du stimulus 10
- tons et asymétries directionnelles : nous avons vu dans les sections 2.1.3 et 2.1.4 (page 42) que la perception des tons modulés est plus facile dans un ordre que dans l’autre. Nous nous attendons alors à ce que des effets de la même polarité soient retrouvés à travers tout le continuum
- périphérie et fréquences préférées : l’hypothèse de périphérie se base sur la présupposition que des informations sur les fréquences de la périphérie de l’espace vocalique sont soit stockées dans la mémoire à long terme et sont récupérées pendant le traitement auditif, soit implémentées dans les règles du traitement auditif. Dans ce cas-là, nous nous attendons à ce que des effets d’ordre soient positifs à travers tout le continuum.

Dans la figure 6.6, nous n’observons pas d’effets d’ordre à cause des effets de plafond. Seule exception est le sujet FR_F7, où la fonction d’asymétrie est positivement liée à la distance avec le stimulus 10. Cette extrémité joue éventuellement un rôle de point d’ancrage. La figure 6.7 démontre que dans la condition de deux pas, les asymétries sont absentes.

Dans la tâche f2i, nous trouvons quelques effets négatifs au milieu du continuum, entre les stimuli 03 et 08. Dans f2ii, les effets de plafond nous empêchent de détecter des asymétries régulières.

Dans la tâche f3i, nous trouvons plusieurs fluctuations à travers le continuum, surtout pour les sujets PL_F1, FR_F8 et FR_F1. Dans f3ii, les fonctions ont été

FIGURE 6.5: Les effets d'ordre pour les tâches dAX2r_250, dAX2r_500 et dAX2r_1000. $n = 5$.

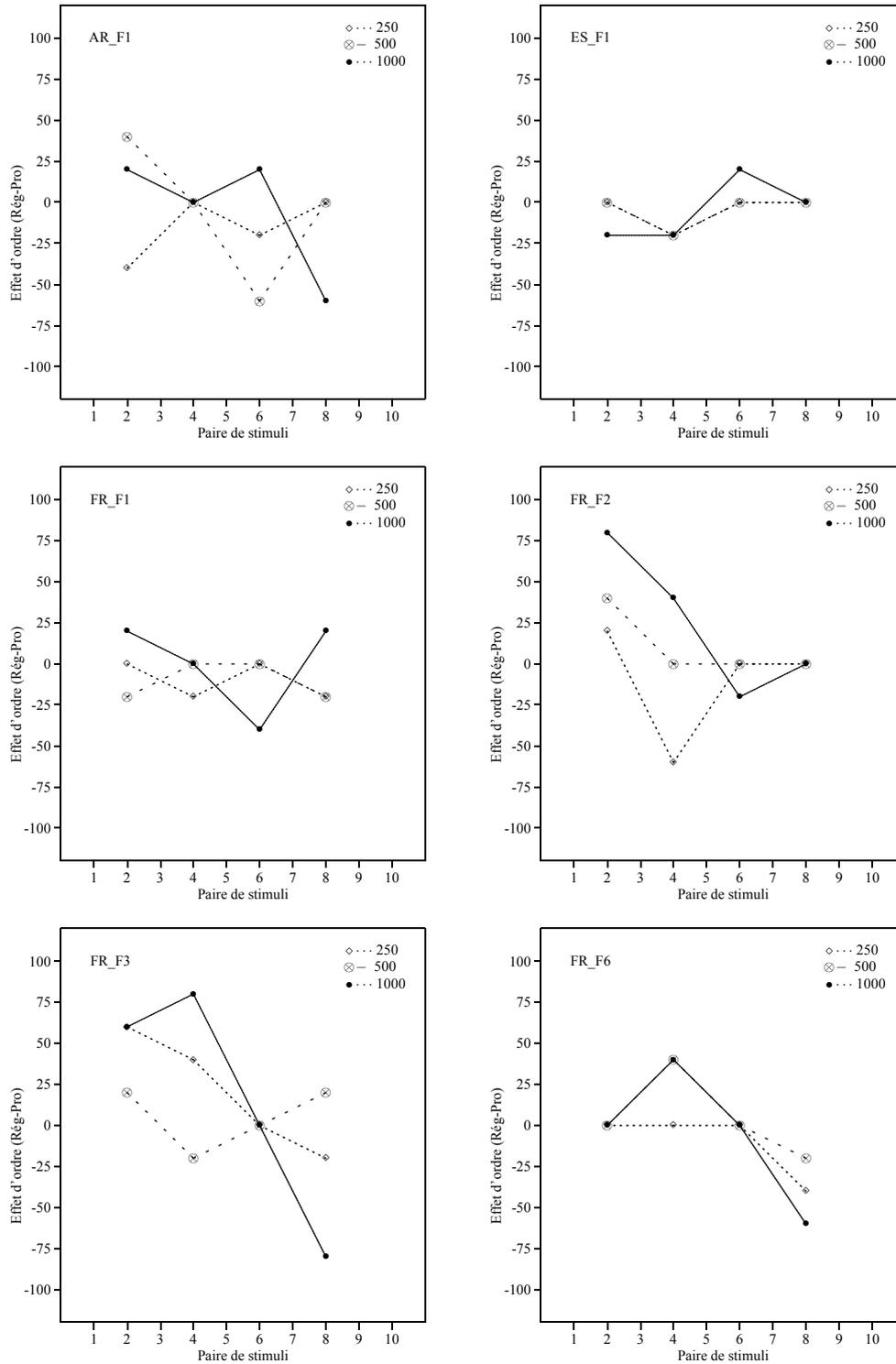


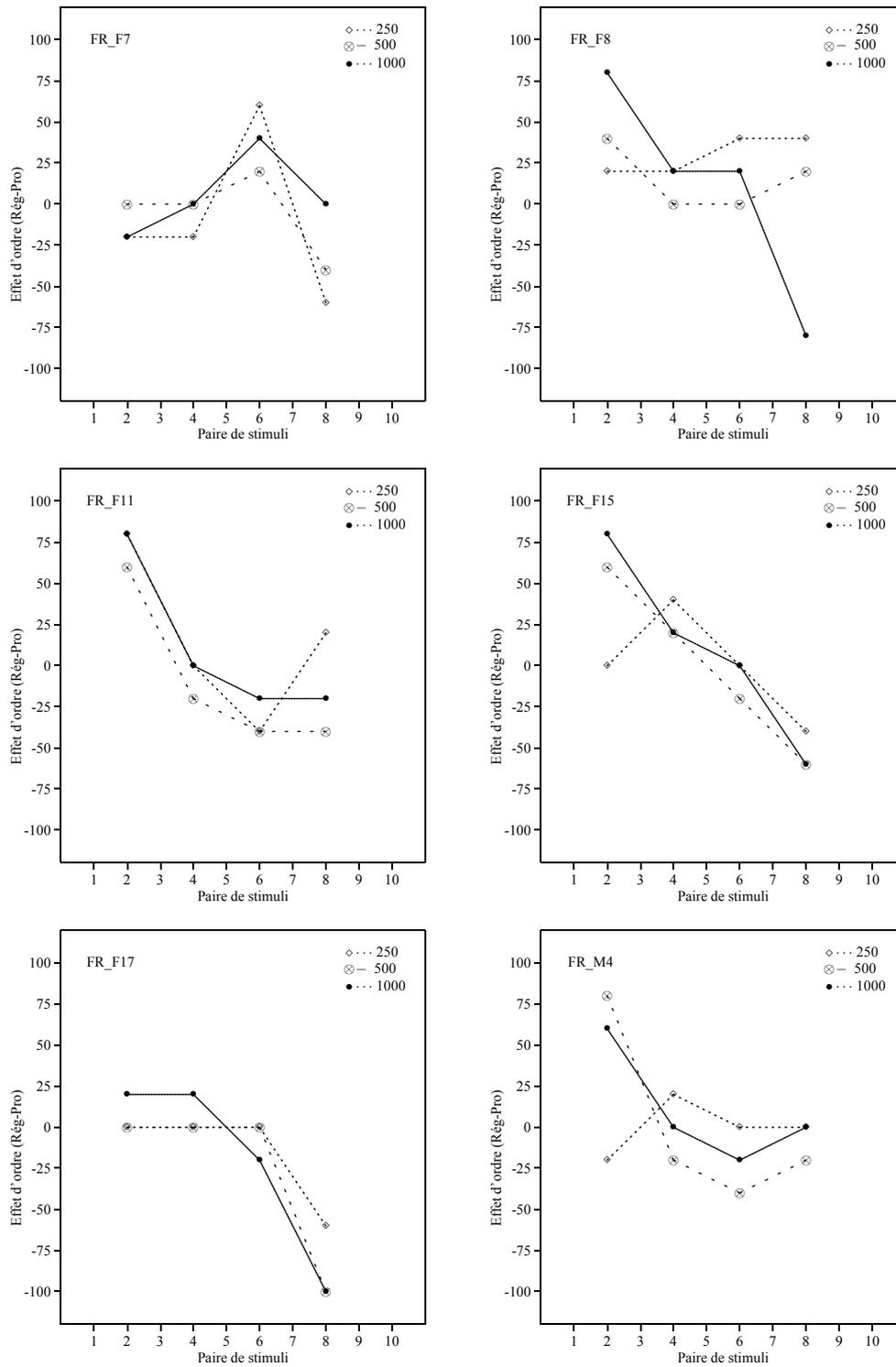
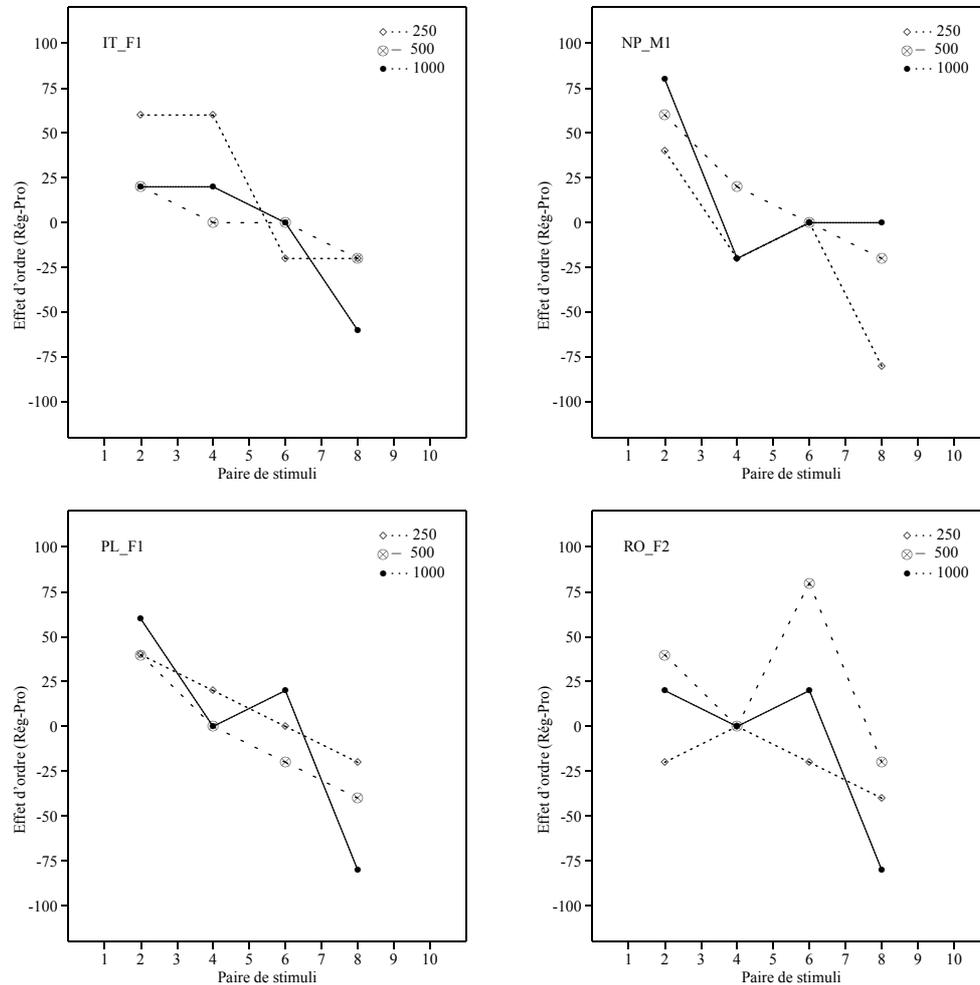
FIGURE 6.5: Les effets d'ordre pour les tâches dAX2r_250, dAX2r_500 et dAX2r_1000. $n = 5$. (cont.)

FIGURE 6.5: Les effets d'ordre pour les tâches dAX2r_250, dAX2r_500 et dAX2r_1000. $n = 5$. (cont.)

relativement plates.

6.5 Conclusions générales

Dans ce chapitre, nous avons mené une étude restreinte sur l'impact que divers facteurs pouvaient avoir sur les effets d'ordre.

Les résultats des tâches iAXi, iAXe, dAXi et dAXe ont révélé que les deux extrémités /i/ et /e/ n'avaient pas le même effet sur les stimuli adjacents. Nous avons retrouvé des différences dans les effets de contexte pour les sujets francophones et pour l'auditrice italienne, confirmant les résultats du chapitre 3 sur le rôle de la focalisation. Les effets de contraste, pourtant, nous ont amené à la conclusion que les effets de contexte et de contraste ne sont pas étroitement liés, ce qui a confirmé les résultats de l'étude sur la perception des configurations manuelles (chapitre 4).

FIGURE 6.6: Les effets d'ordre pour les tâches dAX1r et fli. $n = 7$.

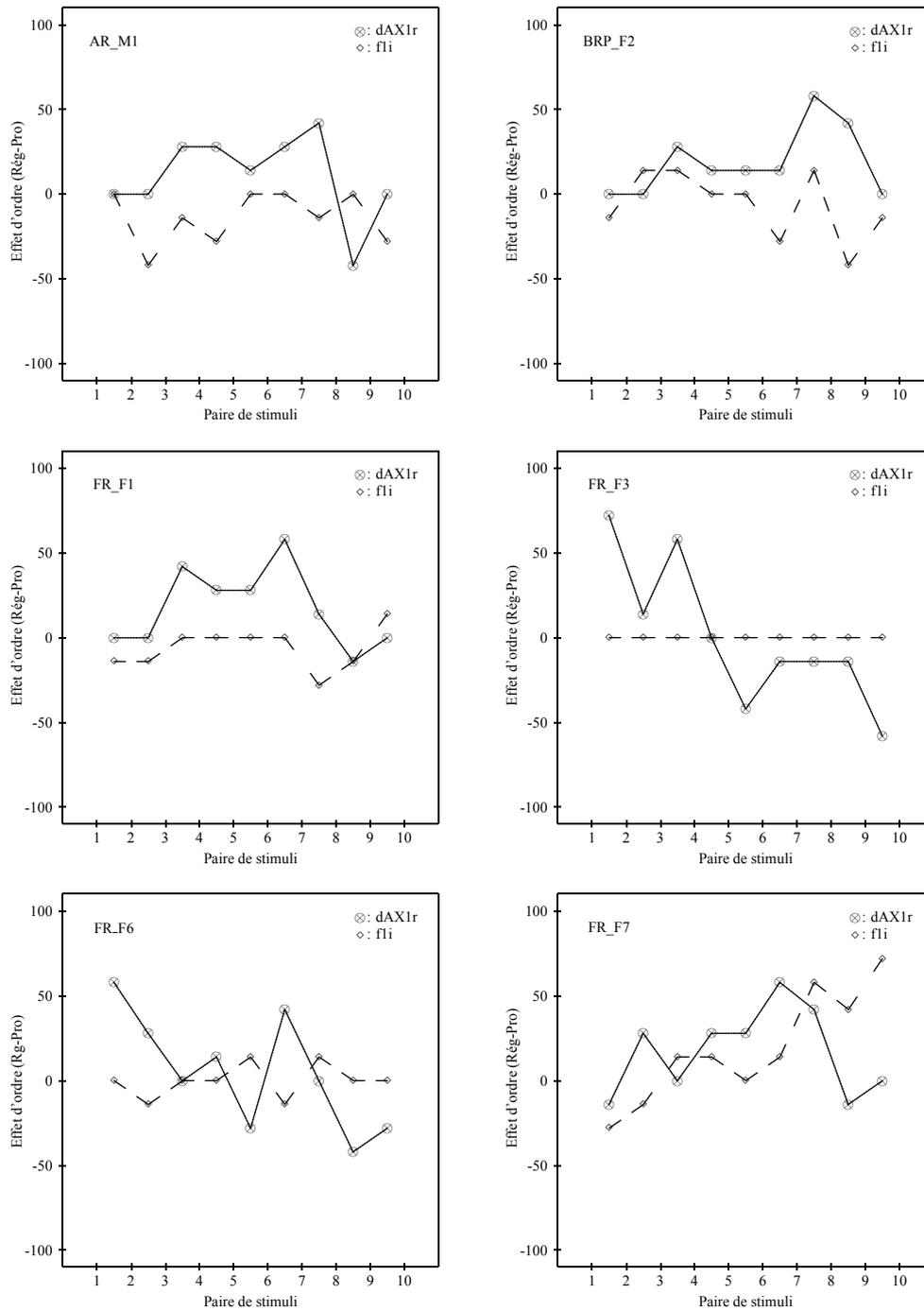


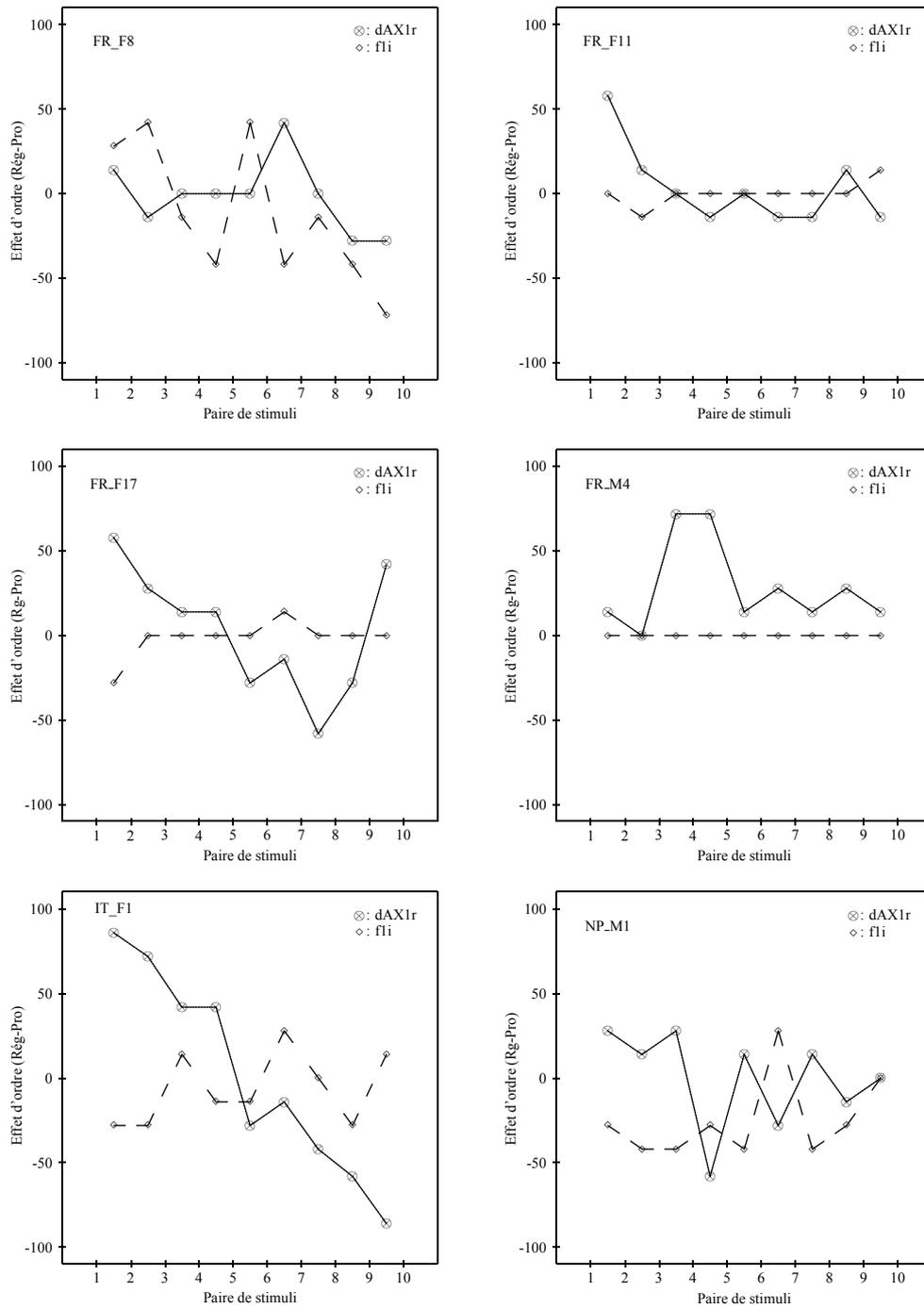
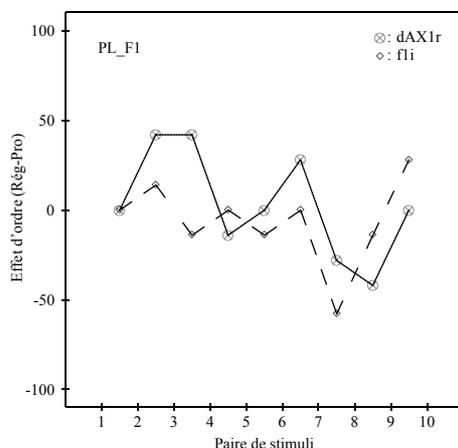
FIGURE 6.6: Les effets d'ordre pour les tâches dAX1r et fli. $n = 7$. (cont.)

FIGURE 6.6: Les effets d'ordre pour les tâches dAX1r et fli. $n = 7$. (cont.)

Dans les tâches iAX1r et iAX2r, le contexte a joué un rôle sur l'étiquetage des stimuli. Pour que cet effet de contexte soit significatif, une distance acoustique minimale doit exister entre le standard et le stimulus à évaluer. Cette contrainte n'est pas appliquée aux tâches dAX1r et dAX2r, où la distance acoustique n'a pas eu d'effet significatif sur le nombre de détections correctes.

La relation non-significative entre le nombre de détections correctes et la nature des stimuli (filtrés ou non-filtrés), en combinaison avec les effets d'ordre, indique que la focalisation n'est pas le facteur principal déclenchant des asymétries.

Dans la tâche dAX2f, des seuils de plafond ont été atteints. Les effets d'ordre ont par conséquent été relativement inexistantes.

Les résultats des tâches exploitant le rôle de l'Intervalle Interstimulus ne sont pas facilement cernables. L'hypothèse de périphéricité et de focalisation ne semblent pas s'appliquer sur les données. Les patrons de réponses des quatre sujets apportent des preuves en faveur du rôle des extrémités en tant que points d'ancrage.

Les six tâches incluant des tons purs ont eu comme but de nous indiquer si les effets d'ordres pouvaient être attribués à la perception différentielle des fréquences formantiques des dix stimuli. Les effets de plafond, ainsi que les effets d'ordre irréguliers, indiquent que les asymétries dans la perception des voyelles ne sont pas expliquées par les données sur les tons purs.

La conclusion que nous pouvons tirer ici est qu'il est difficile de détecter les facteurs provoquant des effets d'ordre. Ceci est dû dans une certaine mesure à la nature des données recueillies. Les données issues des tâches 'même-différent' ne sont pas facilement exploitables dans nos types d'analyse. Afin de déterminer si un effet d'ordre est significatif pour une paire de stimuli, la seule analyse applicable semble être un test χ . Pour que cette analyse soit fiable, nous avons besoin d'un grand nombre de détections correctes. Le problème vient du fait que la tâche et la distance acoustique peuvent très facilement mener à des effets de plafond qui cacheront les effets d'ordre. Il nous faut donc une autre façon de mesurer la discriminabilité des stimuli. Nous y reviendrons dans la conclusion de cette thèse.

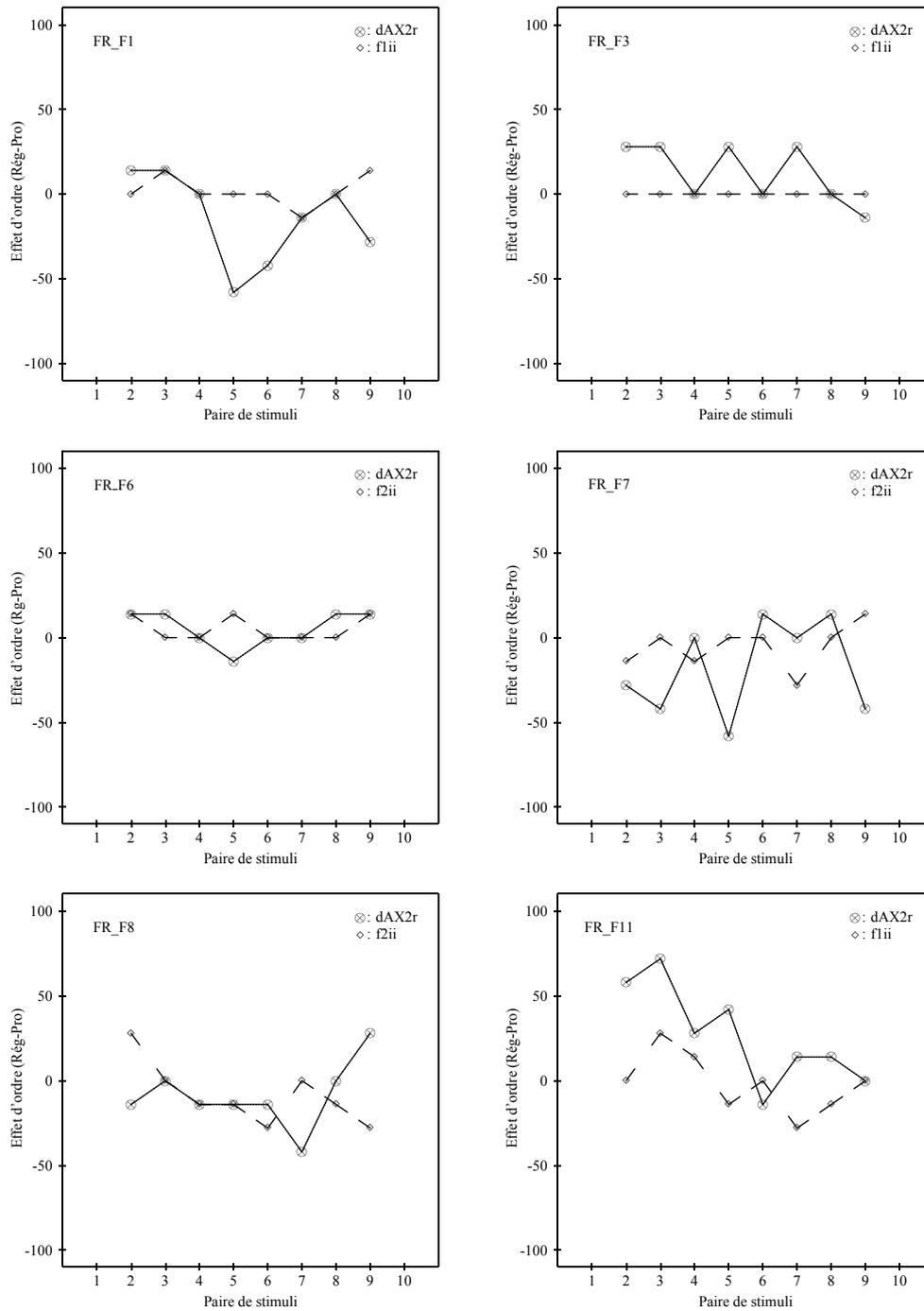
FIGURE 6.7: Les effets d'ordre pour les tâches dAX2r et flii. $n = 7$.

FIGURE 6.7: Les effets d'ordre pour les tâches dAX2r et flii. $n = 7$. (cont.)

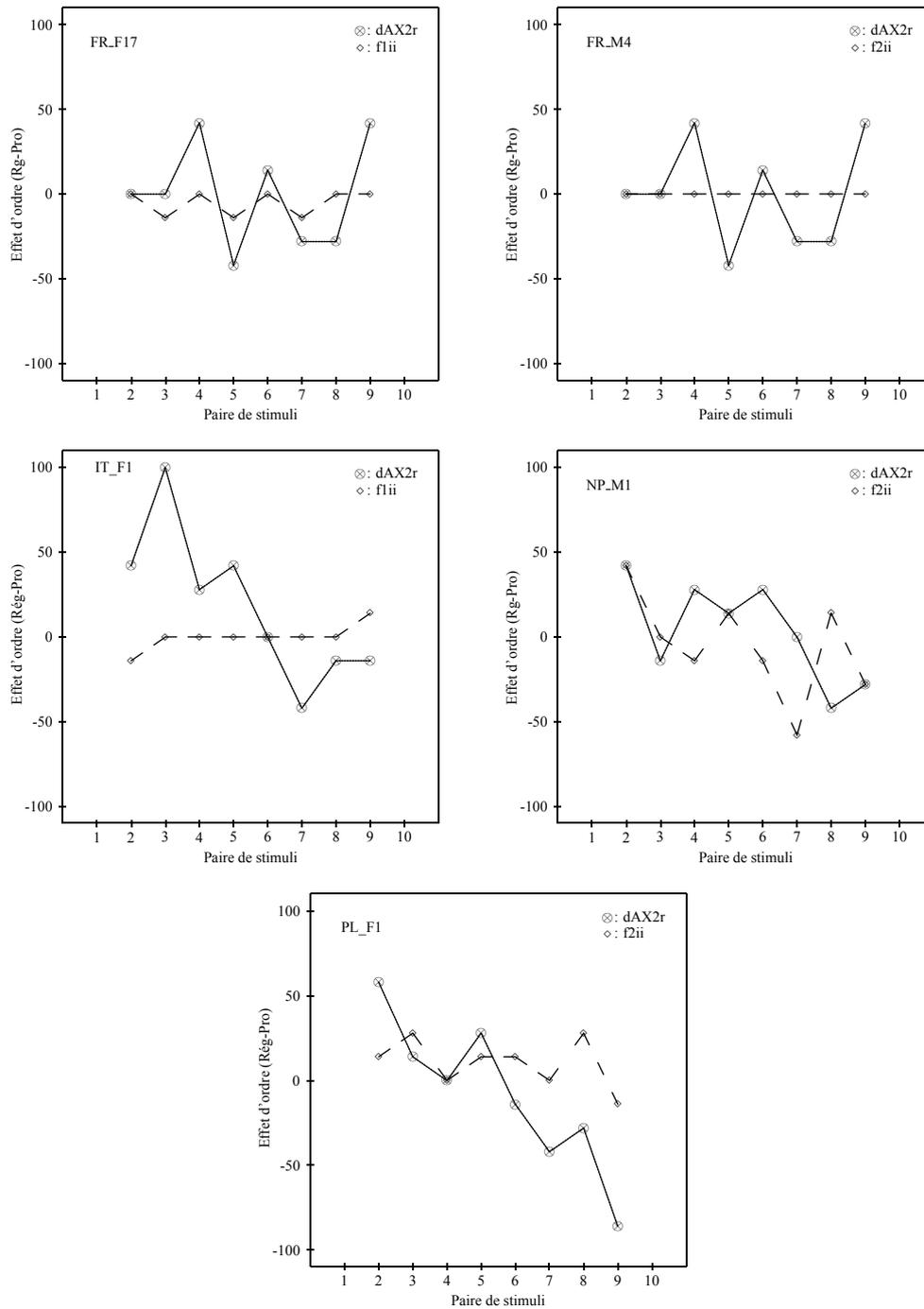


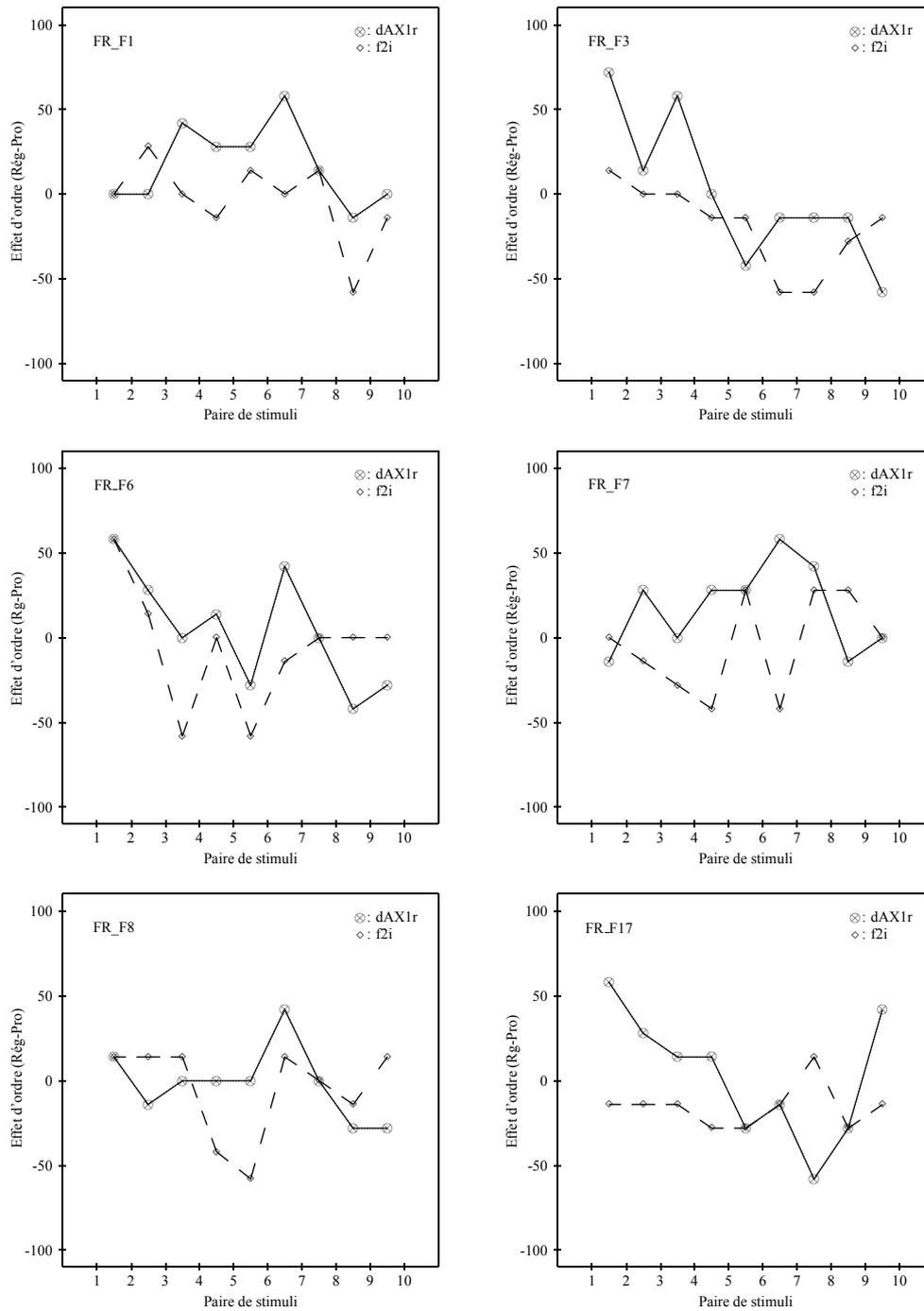
FIGURE 6.8: Les effets d'ordre pour les tâches dAX1r et f2i. $n = 7$.

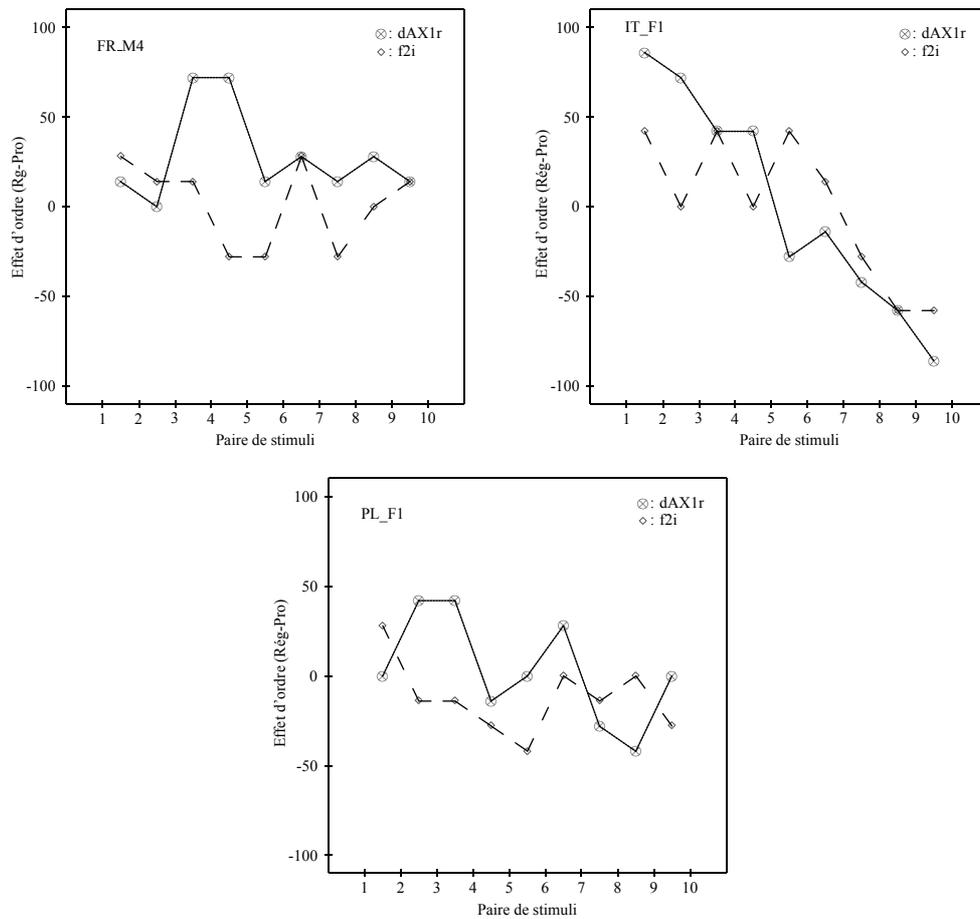
FIGURE 6.8: Les effets d'ordre pour les tâches dAX1r et f2i. $n = 7$. (cont.)

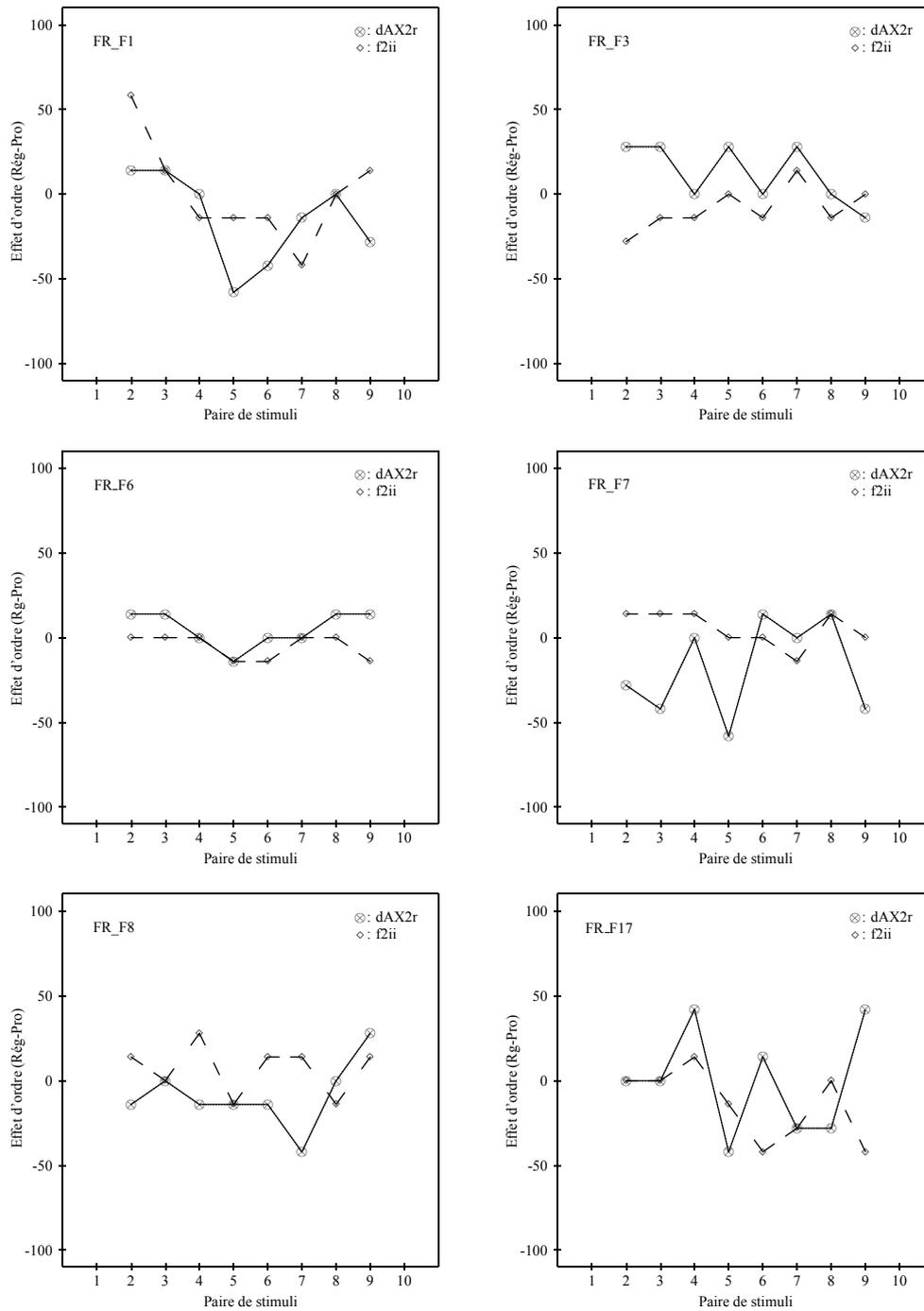
FIGURE 6.9: Les effets d'ordre pour les tâches dAX2r et f2ii. $n = 7$.

FIGURE 6.9: Les effets d'ordre pour les tâches dAX2r et f2ii. $n = 7$. (cont.)

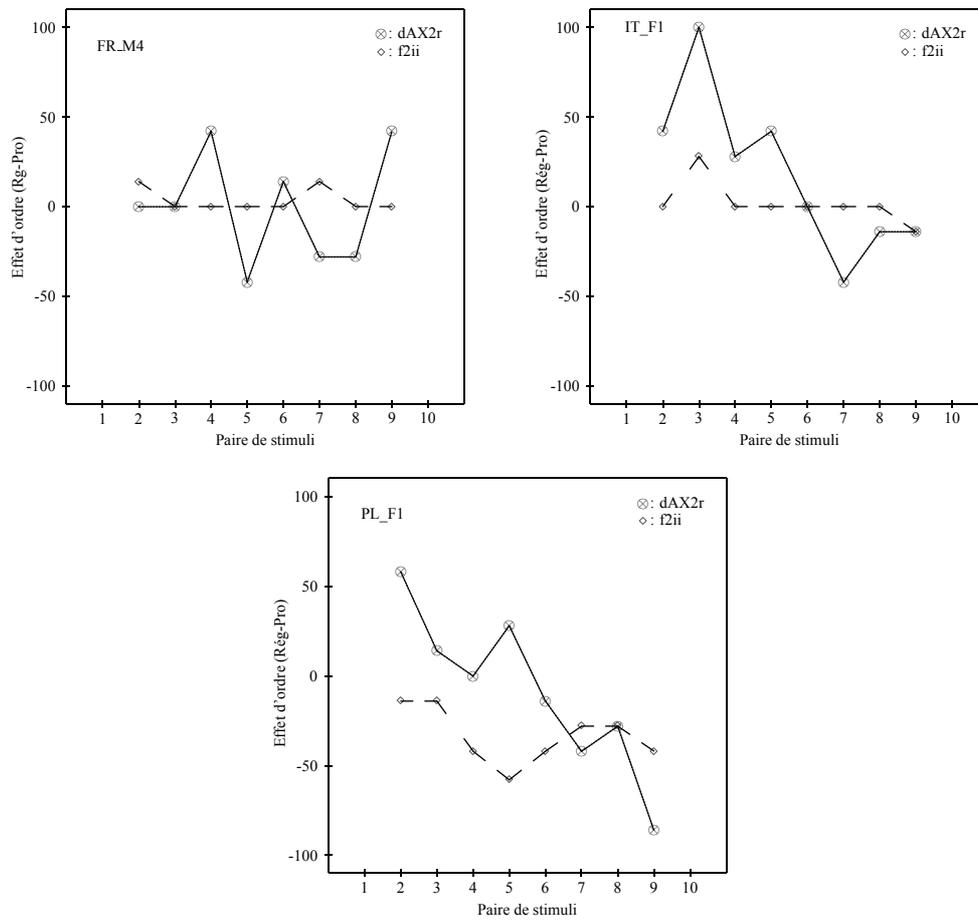


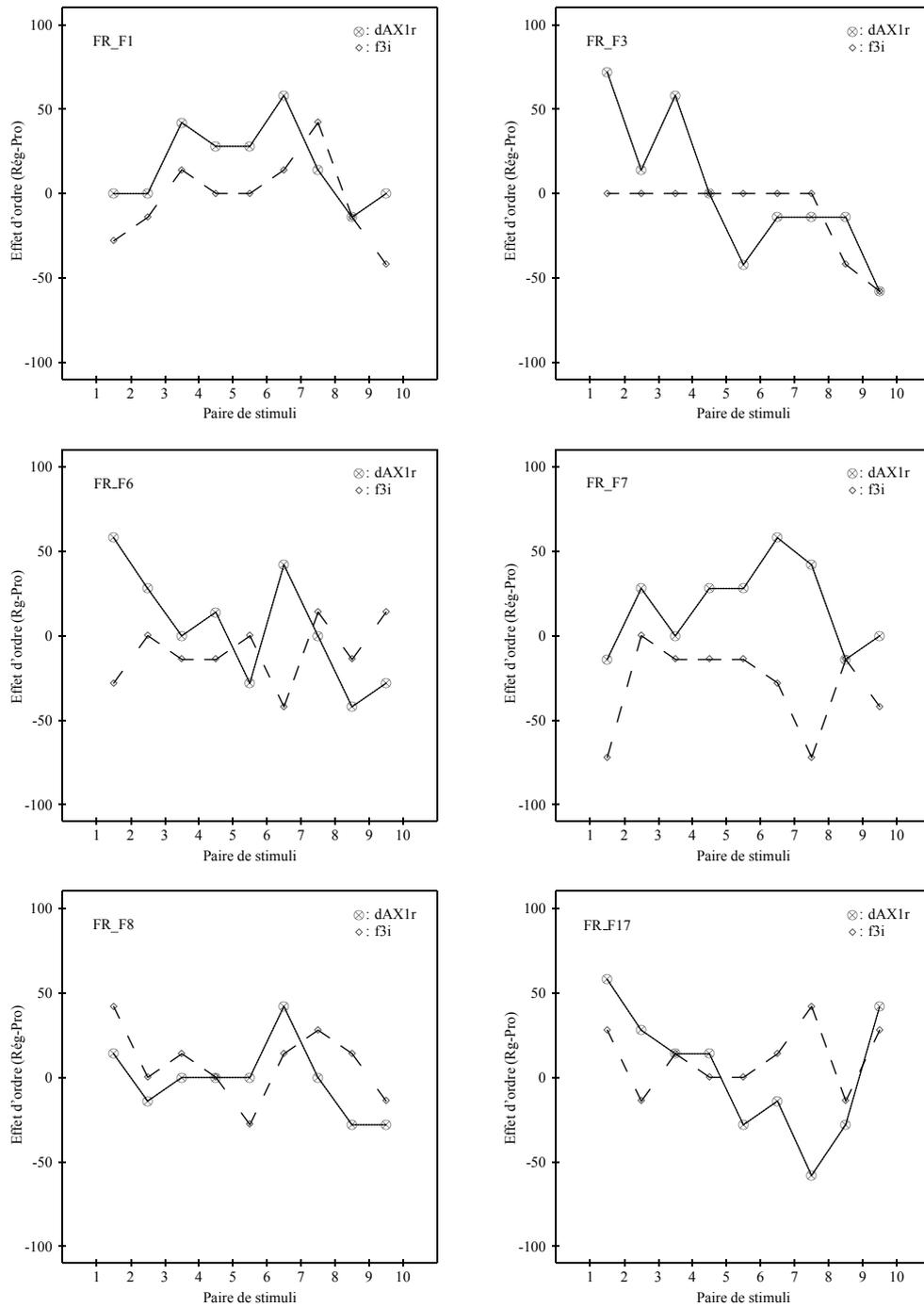
FIGURE 6.10: Les effets d'ordre pour les tâches dAX1r et f3i. $n = 7$.

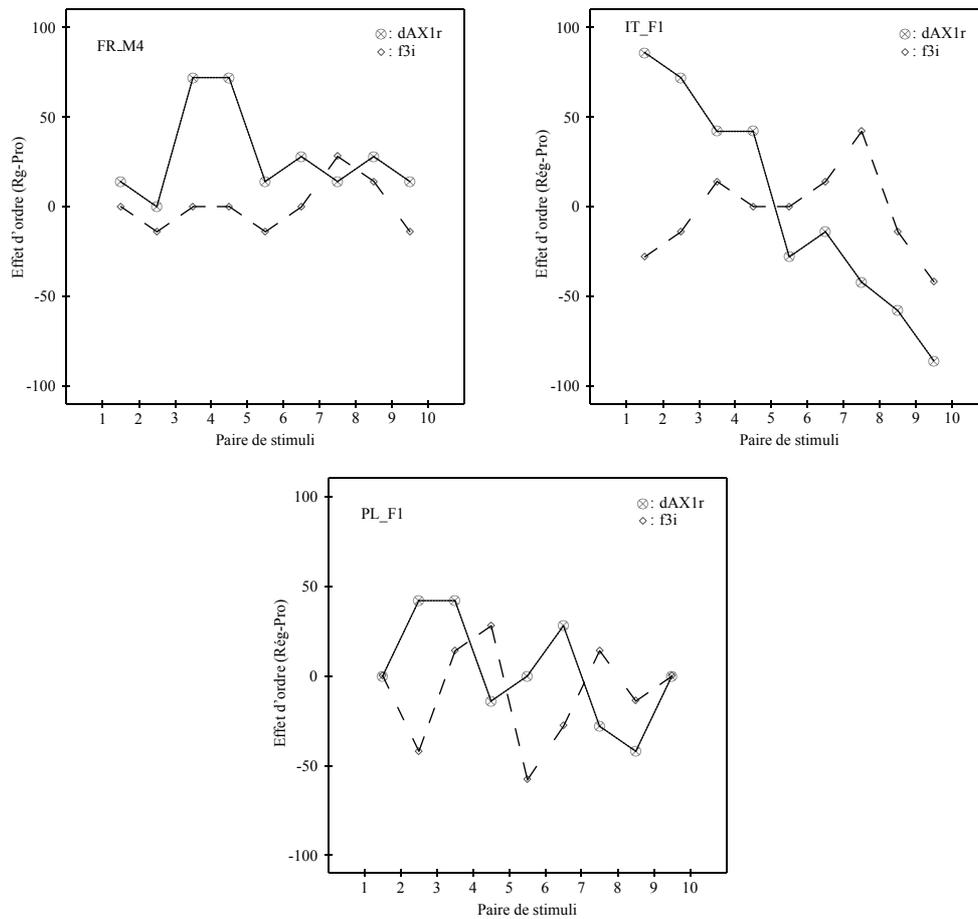
FIGURE 6.10: Les effets d'ordre pour les tâches dAX1r et f3i. $n = 7$. (cont.)

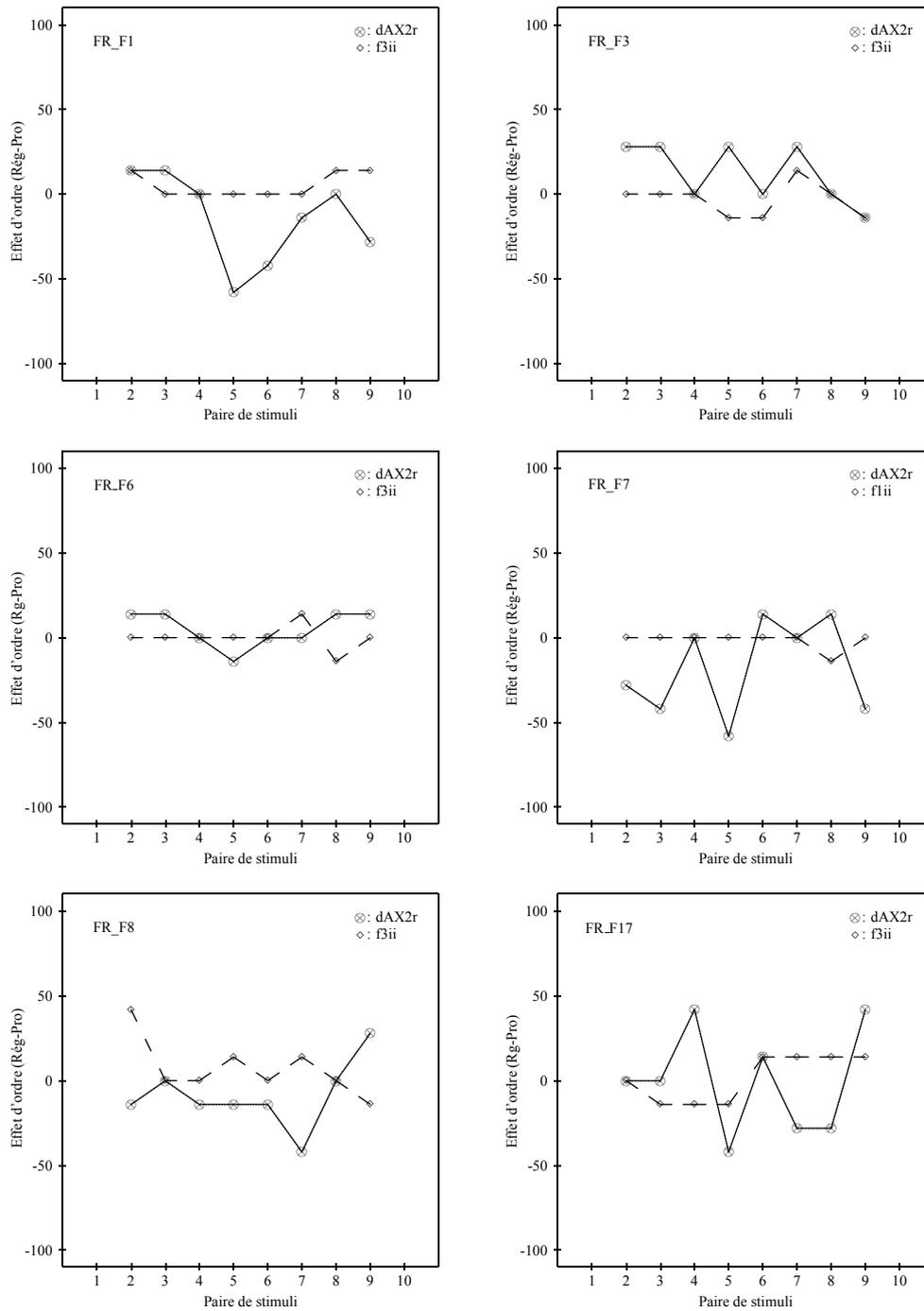
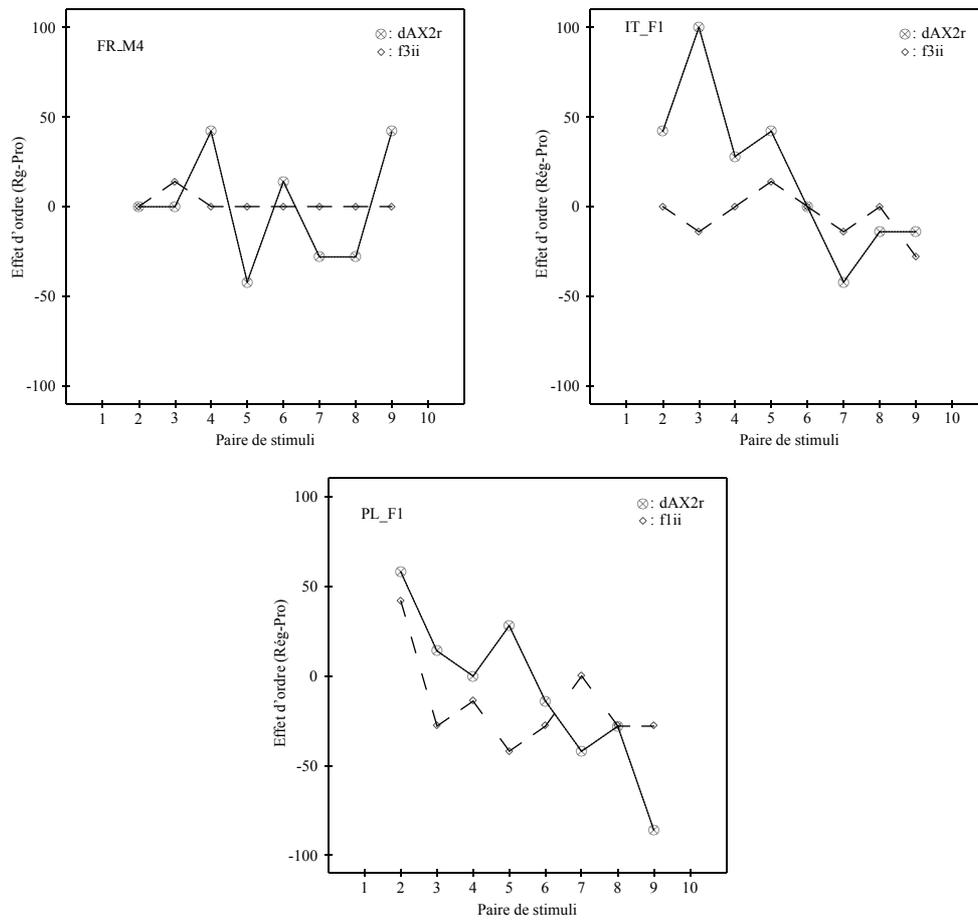
FIGURE 6.11: Les effets d'ordre pour les tâches dAX2r et f3ii. $n = 7$.

FIGURE 6.11: Les effets d'ordre pour les tâches dAX2r et f3ii. $n = 7$. (cont.)



Chapitre 7

Asymétries et cibles acoustiques

*"It is a capital mistake to theorize before one has data.
Insensibly one begins to twist facts to suit theories,
instead of theories to suit facts."
Sherlock Holmes, A Scandal in Bohemia*

Les hypothèses formulées jusqu'à présent concernent la place que les voyelles occupent dans l'espace perceptif ou bien leur typicalité. Ici nous avons tenté d'établir un lien entre la production et la perception. Dans le chapitre 5, nous avons postulé que les effets d'ordre en perception pouvaient éventuellement avoir un effet sur les stratégies de production des locuteurs. Nous avons alors formulé l'hypothèse selon laquelle les locuteurs utiliseraient leurs connaissances implicites sur la discriminabilité des voyelles afin de rendre des voyelles consécutives plus dissimilaires. Reprenons l'exemple de la paire /i-/e/. Afin d'atteindre le même degré de dissimilarité des deux voyelles, le locuteur a besoin d'un mouvement articuloire plus marqué dans l'ordre [i]-[e] que dans l'ordre [i]-[e]. Un moindre effort articuloire est donc nécessaire pendant la transition vers une voyelle plus périphérique.

Une deuxième approche des asymétries en production consiste en ce que dans l'ordre [e]-[i], des frontières physiologiques se mettent en place, puisqu'une élévation trop importante de la langue résultera en une semi-voyelle ou une consonne fricative. Dans l'ordre inverse, un déplacement articuloire plus ample est possible.

De même, nous avons corrélé le concept de stabilité d'une cible acoustique à la position d'une voyelle dans l'espace vocalique. Nous allons vérifier s'il est plus facile d'attaquer la cible acoustique d'une voyelle extrême (périphérique) que celle d'une voyelle moins périphérique. Cette probable facilité serait due au fait que la périphérie est une région acoustiquement plus stable et qui agit comme point d'ancrage. Il serait plus facile de 'retrouver' un point extrême dans l'espace vocalique qu'un point médian.

Nous avons mené deux expériences. Pour la première, nous avons enregistré dix-neuf locuteurs d'origines linguistiques diverses. Nous leur avons demandé de prononcer des logatomes bisyllabiques CVCV, où C=[p t k] et V=[i e]. Ceci nous a donné l'occasion de comparer la différence de valeurs formantiques entre [CiCe] et [CeCi]. De même, les séquences [CiCi] et [CeCe] nous ont aidé à vérifier notre hypothèse sur la stabilité des cibles acoustiques.

Tableau 7.1: Identifiant des auditeurs, langue maternelle et âge.

ID	Langue	Age
FR_F1	français	25
FR_F2	français	47
FR_F3	français	26
BRP_F1	portugais brésilien	32
FR_M1	français	26
IT_F1	italien	31
FR_M2	français	29
FR_F18	français	30
FR_M3	français	30
BRP_F2	portugais brésilien	27
AR_M1	arabe	32
FR_F19	français	25
RO_F1	roumain	26
CZ_F1	tchèque	26
FR_M4	français	27
AR_F1	arabe	30
FR_F6	français	30
FR_F7	français	27
PL_F1	polonais	35

Pour la deuxième expérience, nous avons fait appel à dix locuteurs qui ont prononcé 20 triphongues $V_1V_2V_1$ où V_1 et V_2 sont des voyelles françaises, occupant des positions adjacentes dans l'espace acoustique. Nous avons étudié exclusivement l'hypothèse de stabilité acoustique.

7.1 Expérience 1 : Production de logatomes CVCV

Cette expérience nous a servi pour obtenir une réponse aux deux questions suivantes :

- Le déplacement acoustique est-il plus important dans la séquence [CiCe] que dans [CeCi] ?
- La transition vocalique [CiCi] est-elle plus stable que la transition [CeCe] ?

7.1.1 Participants

Dix-neuf locuteurs de diverses langues maternelles ont participé à cette expérience. Le tableau 7.1 présente des détails concernant les locuteurs. L'identifiant de chaque auditeur est encodé de la façon suivante : 'code de la langue' + 'sexe' (F= femme, M= homme) + 'numéro de série'. Nous avons attribué aux sujets ayant participé également à l'étude de cas du chapitre 6 le même identifiant. Tous les locuteurs habitent dans la région parisienne depuis plusieurs années.

7.1.2 Corpus

Nous avons enregistré la réalisation de logatomes bisyllabiques d'une structure CVCV, où $C=[p \ t \ k]$ et $V=[i \ e]$. Nous avons alors 12 logatomes : [pipi pipe pepi

pepe], [titi tite teti tete] et [kiki kike keki keke].

7.1.3 Procédure

Nous avons présenté les 40 items — 4 mots x 10 répétitions par mot — de chacun des trois groupes de logatomes sous la forme d’une liste et en ordre aléatoire, mais identique pour tous les locuteurs. L’ordre des groupes a été [p]-[t]-[k]. Cela fait 120 mots au total par locuteur. Le débit de parole proposé a été d’un mot toutes les 2–2.5 secondes. Les locuteurs pouvaient faire une pause quand ils le désiraient.

Nous avons effectué les enregistrements dans la chambre sourde de l’ILPGA. Nous avons utilisé un micro AudioTechnica ATM8033, qui était placé sur un pied de micro de table à 30 cm du locuteur. Cette distance, pourtant, n’a pas été constante pour tous les locuteurs, puisque certains d’entre eux parlaient à voix basse. Dans ce cas, le micro était placé plus près du locuteur. Nous avons souvent utilisé l’amplificateur de la table mixeur pour obtenir un signal plus fort. Afin d’éviter toute saturation pour les consonnes occlusives, nous avons fixé un écran anti-pop devant le micro.

7.1.4 Résultats

Nous avons estimé les valeurs formantiques des voyelles à l’aide d’un script sur Praat. Le paramétrage de l’analyse originelle a été le suivant :

- Plafond de détection formantique : 5000 Hz (hommes), 5500 Hz (femmes)
- Nombre de formants à détecter : 5
- Largeur de la fenêtre : 0.010 ms
- Étendue dynamique : 30.0 dB

Les valeurs 5000 Hz et 5500 Hz sont traditionnellement utilisées dans la littérature. Cette analyse a révélé un certain nombre de valeurs extrêmes (angl. ‘*outliers*’) suite à une mauvaise détection des formants. Nous avons alors décidé d’appliquer la méthode d’optimisation du plafond formantique de Escudero *et al.* (2009). Celle-ci consiste en l’adaptation du plafond formantique pour chaque locuteur et voyelle de la façon suivante :

- pour les 120 occurrences de chaque catégorie vocalique (/i/ et /e/) énoncées par un locuteur, nous estimons les deux premiers formants 201 fois, à savoir pour tous les plafonds entre 4000 Hz et 6000 Hz par pas de 10 Hz pour les hommes, et pour tous les plafonds entre 4500 Hz et 6500 Hz par pas de 10 Hz pour les femmes ;
- nous convertissons les 120 valeurs en valeurs logarithmiques ;
- nous calculons la variance des 120 valeurs ;
- nous additionnons la variance de F1 et de F2 pour chaque occurrence ;
- nous détectons le plafond ‘optimal’ parmi les 201 plafonds, à savoir celui qui présente la variance la plus faible parmi les 120 valeurs.

Tableau 7.2: Plafonds optimaux pour chaque voyelle prononcée par nos locuteurs. Colonnes 2–3 : résultats obtenus avec la méthodologie de Escudero *et al.* (2009). Colonnes 4–5 : résultats obtenus avec notre protocole adapté.

	Analyse initiale		Deuxième analyse	
	i	e	i	e
FR_F1	5100	5720	5200	5700
FR_F2	4820	4860	5200	6400
FR_F3	4930	4870	5300	5200
BRP_F1	5540	4730	5500	5200
FR_M1	4040	5460	4900	5500
IT_F1	6140	6380	6100	6500
FR_M2	4020	5430	4600	5400
FR_F18	4550	4740	5200	5200
FR_M3	4590	4850	4600	5000
BRP_F2	6500	5830	6500	5800
AR_M1	4220	4380	4600	5200
FR_F19	4550	4710	5600	5200
RO_F1	6120	5460	6500	5400
CZ_F1	4690	4840	5300	5200
FR_M4	4000	4820	4600	4600
AR_F1	4600	4750	5200	5700
FR_F6	5930	4590	5900	5200
FR_F7	5620	5780	5600	5800
PL_F1	6370	4570	6300	6500

Tableau 7.3 – Les valeurs des bandes critiques et des largeurs de bande critiques en fonction de la fréquence (Zwicker & Terhardt, 1980).

Bande critique (Bark)	Fréquence (Hz)	Largeur de bande critique (Hz)	Fréquence centrale (Hz)
0	0		
		100	50
1	100		
		100	150
2	200		
		100	250
3	300		
		100	650
4	400		
		110	450
5	510		
		120	570
6	630		
		140	700
7	770		
		150	840

Tableau 7.3 – (cont.)

Bande critique (Bark)	Fréquence (Hz)	Largeur de bande cri- tique (Hz)	Fréquence cen- trale (Hz)
8	920		
		160	1000
9	1080		
		190	1170
10	1270		
		210	1370
11	1480		
		240	1600
12	1720		
		280	1850
13	2000		
		320	2150
14	2320		
		380	2500
15	2700		
		450	2900
16	3150		
		550	3400
17	3700		
		700	4000
18	4400		
		900	4800
19	5300		
		1100	5800
20	6400		
		1300	7000
21	7700		
		1800	8500
22	9500		
		2500	10500
23	12000		
		3500	13500
24	15500		

Les résultats sont disponibles dans le tableau 7.2 (colonnes 2 et 3). En regardant les valeurs brutes des données, nous avons vu que cette méthode a donné des valeurs non-réalistes, particulièrement pour F2 de /i/. Ainsi, avons-nous modifié cette méthode en restreignant la plage des fréquences pour les plafonds possibles (5200–6500 pour les femmes et 4500–6000 pour les hommes). Nous avons également modifié le degré des pas à 100 Hz. Les résultats sont disponibles dans le tableau 7.2 (colonnes 4 et 5). Cette méthode a aussi donné des valeurs problématiques. Nous avons ainsi opté pour l'estimation de l'énergie des bandes critiques (Fletcher, 1940).

Pour chaque logatome, nous avons extrait les deux voyelles que nous avons normalisées à 75 dB au niveau de l'amplitude RMS. Nous avons ensuite produit une coupe spectrale pour chaque voyelle. Un filtre passe-bande a été appliqué sur le spectre des voyelles pour l'étendue des fréquences 199–8000 Hz. Nous avons exclu les deux premières bandes critiques (fréquences 0–199 Hz) pour éviter la différenciation des voyelles sur la base de la fréquence fondamentale (voir Hillenbrand *et al.*, 1995, pour la F0 intrinsèque des voyelles de l'anglais). Le tableau 7.3 présente les fréquences minimale et maximale de chacune des 24 bandes critiques, la largeur de bande et la fréquence centrale. Nous avons calculé l'énergie de chacune des bandes 3 à 20 pour chaque voyelle et avons extrait la différence absolue de ces 18 paires de valeurs.

Les résultats de l'analyse en bandes critiques ont servi pour des analyses multivariées (MANOVA) à une variable indépendante (paire de voyelles) et 18 variables dépendantes (les bandes critiques). Nous avons comparé les paires de voyelles deux à deux afin de mieux envisager les différentes hypothèses.

Les résultats des MANOVA sont présentés dans le tableau 7.4. La première hypothèse à valider est celle de la stabilité des cibles. Basé sur la théorie de périphéricité et sur la théorie quantique, nous avons postulé que les cibles acoustiques de /i/ sont plus stables que celles de /e/. La différence donc entre deux occurrences de /i/, comme dans [titi] serait inférieure à la distance spectrale entre deux occurrences de /e/ (voir formule 7.1).

Au début du chapitre, nous avons formulé l'hypothèse selon laquelle le déplacement de [e] à [i] est moindre que le changement de [i] à [e], vu que dans le premier cas, des contraintes articulatoires sont mises en place. De même, les cibles articulatoires, et, par là même, acoustiques de /i/, seront plus stables parce que les locuteurs peuvent plus facilement les retrouver à cause de leur caractère extrême (périphérique). Nous sommes alors amené à émettre les postulats suivants :

$$\Delta_{ii} \neq \Delta_{ee} \quad (7.1)$$

$$\Delta_{ie} > \Delta_{ei} \quad (7.2)$$

$$\Delta_{ii} < \Delta_{ei} \quad (7.3)$$

$$\Delta_{ii} < \Delta_{ie} \quad (7.4)$$

$$\Delta_{ee} < \Delta_{ie} \quad (7.5)$$

$$\Delta_{ee} < \Delta_{ei} \quad (7.6)$$

Les colonnes 2 et 3 du tableau indiquent que pour 12 locuteurs sur 19, l'hypothèse 7.1 n'est pas validée. Cette analyse, pourtant, ne nous permet pas de savoir la direction de la différence pour les 7 locuteurs. Nous allons voir plus tard comment interpréter ces résultats.

Les colonnes 4 et 5 montrent que la différence entre [ie] et [ei] est significative pour 13 sur 19 locuteurs. De même, nous ne savons pas si $\Delta_{ie} > \Delta_{ei}$ ou si $\Delta_{ie} < \Delta_{ei}$.

Les colonnes 6 et 7 mettent la formule 7.4 à l'épreuve. Le patron prédit est que le déplacement vers une autre catégorie sera plus important que le retour à la configuration originelle. Onze locuteurs adoptent cette stratégie. Il est évident alors que le mouvement [i]→[e] est restreint pour ces locuteurs. Si les deux catégories /i/ et /e/ sont équivalentes, la formule 7.6 doit aussi être valide. Pour 9 locuteurs, la

Tableau 7.4: Résultats des analyses multivariées (MANOVA) deux à deux pour les 19 locuteurs.

	ii / ee		ie / ei		ii / ie		ii / ei		ee / ie		ee / ei	
	<i>F</i>	<i>p</i>										
AR_F1	7.844	<.01	0.000	0.982	3.271	0.076	3.686	0.060	0.798	0.375	0.874	0.354
AR_M1	10.434	<.005	10.117	<.005	1.941	0.169	19.620	<.001	4.214	<.05	0.806	0.373
BRP_F1	32.829	<.001	10.438	<.005	2.611	0.112	1.799	0.185	23.359	<.001	61.299	<.001
BRP_F2	1.399	0.242	2.952	0.091	22.073	<.001	9.000	<.005	38.055	<.001	19.170	<.001
CZ_F1	1.026	0.315	41.457	<.001	5.481	<.05	11.978	<.005	1.430	0.237	20.740	<.001
FR_F1	18.211	<.001	10.831	<.005	0.018	0.893	5.786	<.05	47.084	<.001	63.259	<.001
FR_F18	32.521	<.001	80.645	<.001	78.512	<.001	1.829	0.181	25.864	<.001	31.194	<.001
FR_F19	1.447	0.234	14.114	<.001	6.255	<.05	27.867	<.001	17.543	<.001	42.864	<.001
FR_F2	1.708	0.196	4.187	<.05	17.657	<.001	5.221	<.05	9.647	<.005	1.138	0.291
FR_F3	1.235	0.271	129.448	<.001	89.819	<.001	1.482	0.228	99.245	<.001	0.998	<.001
FR_F6	8.447	<.01	30.371	<.001	25.048	<.001	0.001	0.977	4.642	<.05	10.539	<.001
FR_F7	3.197	0.079	36.959	<.001	38.576	<.001	0.183	0.671	23.468	<.001	2.092	0.153
FR_M1	1.514	0.223	1.515	0.223	1.216	0.275	4.334	<.05	0.002	0.966	1.854	0.179
FR_M2	7.899	<.01	7.659	<.01	24.748	<.001	3.786	0.057	7.270	<.01	0.194	0.661
FR_M3	10.952	<.005	0.033	0.855	0.609	0.438	0.665	0.418	9.666	<.005	5.321	<.05
FR_M4	5.746	<.05	0.092	0.762	3.588	0.063	1.948	0.168	0.380	0.540	0.636	0.428
IT_F1	35.712	<.001	5.747	<.05	66.198	<.001	45.035	<.001	6.524	<.05	0.097	0.757
PL_F1	5.399	<.05	0.962	0.331	0.203	0.654	1.560	0.217	11.174	<.005	13.609	<.001
RO_F1	17.218	<.001	17.766	<.001	53.429	<.001	15.442	<.001	9.641	<.005	0.633	0.430

Tableau 7.5: Résultats de l'analyse discriminante. Nous indiquons les 4 bandes critiques qui contribuent le plus à la fonction canonique.

	ii / ee				ie / ei			
AR_F1	5	3	12	16	5	14	12	13
AR_M1	5	15	14	16	17	5	19	16
BRP_F1	3	5	17	19	3	4	6	20
BRP_F2	13	17	20	14	14	13	5	11
CZ_F1	3	17	14	5	3	19	5	17
FR_F1	5	3	16	19	5	16	3	17
FR_F18	4	5	3	17	4	5	16	19
FR_F19	5	3	17	16	3	6	18	16
FR_F2	18	5	6	4	5	6	3	18
FR_F3	5	14	15	8	5	3	14	15
FR_F6	4	3	5	9	3	17	16	5
FR_F7	17	4	18	3	17	20	7	3
FR_M1	4	5	13		17	11	3	
FR_M2	5	15	3	7	4	3	15	5
FR_M3	4	5	13	17	14	16	5	6
FR_M4	3	4	15	19	15	17	3	13
IT_F1	8	6	15	9	8	13	14	6
PL_F1	3	18	15	16	16	5	6	13
RO_F1	12	7	14	6	12	3	18	17

différence entre [ee] et [ei] n'est pas significativement différente (colonnes 12 et 13). Il est donc difficile de soutenir le postulat de contrainte articulatoire.

Les colonnes 8–11 peuvent nous donner des informations sur la stabilité des cibles. Il s'agit de l'axiome de similarité de Tversky (voir p.54) Nous observons donc que la formule 7.3 n'est valide que pour 9 locuteurs. Ceci signifie que dans la séquence [ii], le déplacement est aussi important que [ei]. Nous remarquons, en revanche, que la formule 7.6 est valide pour 15 locuteurs. Nous pouvons conclure que le déplacement pour [ee] est moindre que pour [ii]. La cible de la catégorie /e/ est par conséquent plus stable.

Nous avons ensuite mené une analyse discriminante afin de déterminer les bandes critiques qui déclenchent la différence entre [ii] et [ee], et entre [ie] et [ei]. Une analyse séparée a été menée pour chacune des deux paires. Les paramètres de l'analyse ont été les suivants :

- prédicteur : les 18 bandes critiques
- variable de groupe : paires de voyelles
- a priori : tous groupes pareils
- tolérance : 0.001
- estimation : à pas et progressive ('stepwise', 'forward')
- probabilité (F à enlever ; F-to-remove) : 0.15
- force : 0

Le tableau 7.5 présente les résultats. Il faut noter que pour chaque analyse, 60 valeurs ont été utilisées (10 répétitions x 3 contextes consonantiques x 2 paires de voyelles). Dans une analyse discriminante, il est conseillé d'avoir au-moins 10 fois plus d'observations que de prédicteurs. Notre analyse n'est que préliminaire. Une étude

Tableau 7.6: Les vingt triphthongues de notre corpus.

$V_1V_2V_1$	iei	iyi	uou	uyu	εaε	ɔaɔ	eøe	iəi	uəu	aəa
$V_2V_1V_2$	eie	yiy	ouo	yuy	aεa	aɔa	øeø	əiə	əuə	əaə

plus poussée est nécessaire. Ici nous tentons d'illustrer la variance inter-sujet. Nous observons donc que même à l'intérieur du même groupe linguistique, une différence est retrouvée dans les basses fréquences (entre les bandes 3 et 6) correspondant au premier formant, ainsi que dans la distance spectrale dans les hautes fréquences (bandes 15–...).

7.2 Expérience 2 : Production de triphthongues

$V_1V_2V_1$ ¹

Cette tâche a eu comme but d'explorer la stabilité de la cible acoustique des voyelles françaises.

7.2.1 Participants

Parmi les dix-neuf locuteurs de l'expérience 1, huit ont participé à cette expérience. Certains n'ont pas mené la tâche jusqu'à son terme pour des raisons de temps.

7.2.2 Corpus

Nous avons inclus vingt triphthongues $V_1V_2V_1$, où V_1 et V_2 sont des voyelles françaises, occupant des positions adjacentes dans l'espace acoustique. Il s'agissait de dix paires de voyelles dans les combinaisons $V_1V_2V_1$ et $V_2V_1V_2$ (voir tableau 7.6).

7.2.3 Procédure

Les locuteurs ont prononcé des occurrences consécutives de la même triphthongue toutes les 2–2.5 secondes jusqu'à ce que quinze bonnes occurrences fussent adéquates pour nos analyses. Le nombre d'occurrences nécessaires pour une séquence donnée a varié de 15 à 45 répétitions, en fonction de la complexité des logatomes (des séquences telles que [uyu] et [eøe] sont plus difficiles à articuler). L'ordre des séances a varié en fonction des difficultés des locuteurs.

Les conditions externes ont été les mêmes que pour l'expérience 1. Nous avons appliqué la même méthode de calcul des bandes critiques.

1. L'idée d'enregistrer des triphthongues nous a été suggérée par René Carré pendant la conférence *Acoustics '08*.

Tableau 7.7: Valeurs p extraites d'une analyse *a posteriori* de Tukey.

	ɔ	ø	ɛ	a	e	i	ə	o	u	y
ɔ	1.000	<.001	0.235	<.001	1.000	1.000	0.322	<.001	<.001	<.005
ø		1.000	<.001	<.001	<.001	<.001	<.05	<.001	<.001	0.900
ɛ			1.000	<.05	<.01	0.118	<.001	<.001	<.001	<.001
a				1.000	<.001	<.001	<.001	<.001	<.001	<.001
e					1.000	0.965	0.431	<.001	<.001	<.001
i						1.000	<.01	<.001	<.001	<.001
ə							1.000	<.001	<.001	0.361
o								1.000	<.001	<.001
u									1.000	<.001
y										1.000

7.2.4 Résultats

Afin d'envisager la stabilité des cibles des différentes voyelles, nous avons réduit les 18 valeurs à une variable qui correspondaient à leur moyenne. Nous avons ensuite comparé les différentes triptongues pour chaque locuteur avec un test de Kruskal-Wallis. L'effet de Triptongue a été significatif pour tous les locuteurs, ce qui signifie que la stabilité des voyelles n'est pas fixe. Un test *a posteriori* de Tukey avec $\alpha = .01$ a été conduit sur tous les locuteurs mélangés (voir le DVD ci-joint pour les statistiques individuelles). Le tableau 7.7 présente les résultats. Nous observons que la variance est significativement différente pour la plupart des comparaisons 2x2.

La théorie quantique prédit que les voyelles les plus stables sont les voyelles quantiques, à savoir /i u a y/. Nous avons classé les dix voyelles en fonction de leur variance. La hiérarchie retrouvée est la suivante : /a ɛ i ɔ e ə y ø u o/. Nous remarquons que /a/ et /i/ illustrent en effet une variance peu considérable. /u/ et /y/ pourtant, ainsi que /o/ qui peut aussi être focal, sont les plus instables. Étant donné que les cinq voyelles les plus instables impliquent des mouvements des lèvres, nous postulons que leur variance est due aux mouvements labiaux et non pas aux déplacements de la langue.

7.3 Conclusions générales

Dans cette section, nous avons abordé deux aspects de la production : la stabilité des cibles acoustiques et les asymétries.

Nous avons mené deux expériences. Dans la première, nous avons enregistré des séquences [CiCi], [CeCe], [CiCe] et [CeCi]. Nous avons comparé la distance spectrale entre [CiCi] et entre [CeCe], puis la différence entre ces distances. Pour 12 locuteurs sur 19, cette différence a été significative. Une des deux voyelles manifeste moins de variance, à savoir plus de stabilité que l'autre. En regardant les données brutes, nous avons remarqué qu'il ne s'agit pas toujours de la même voyelle. Le postulat de périphérie ou de la théorie quantique qui suggère que les voyelles extrêmes ou focales sont plus stables n'est pas confirmé par nos données.

Le changement vocalique dans les séquences [CiCe] et [CeCi] est asymétrique. Dans la plupart des cas, la distance spectrale dans [CiCe] est plus importante. Nous

avons remarqué que pour 8 locuteurs, la différence de distance entre [CiCi] et [CiCe] n'est pas significative, ce qui infirme notre hypothèse selon laquelle dans la direction [CiCe], un déplacement plus large serait effectué, puisque les locuteurs ne seraient pas capables de retrouver facilement la cible de [e]. Nous avons postulé que dans [CeCi], le déplacement serait limité par des contraintes articulatoires. Ceci a été le cas pour 10 locuteurs seulement.

Dans la deuxième tâche, nous avons fait enregistrer vingt triptongues comprenant des voyelles françaises à huit locuteurs. Ici, nous avons envisagé l'hypothèse de stabilité. Nous nous attendions à ce que les voyelles périphériques ou les voyelles quantiques manifestent moins de variance. La hiérarchie que nous avons décelée était : /a ε i ɔ e ə y ø u o/. Nous avons à ce point suggéré qu'une grande partie de la différence dans la variance entre les dix voyelles est due aux gestes labiaux.

Discussion générale

Dans cette thèse, nous avons tenté de mettre à l'épreuve différentes théories sur l'origine des asymétries. Afin de mieux comprendre chacun de ces postulats, nous les avons d'abord interprétés à travers divers modèles de discriminabilité.

La périphéricité est le fait que certaines voyelles sont plus extrêmes au niveau articulatoire et, par là, acoustique (Polka & Bohn, 2003). Cette propriété leur donne un statut particulier. Une tendance est apparente dans la littérature : quand le changement vocalique se fait vers une voyelle plus périphérique, la discrimination est plus facile que dans l'ordre inverse. Dans le chapitre 5, nous avons observé que pour les catégories de stimuli correspondant à dix voyelles françaises, la discrimination est plus facile dans la direction de la périphérie (figure 5.4, page 104), confirmant ainsi le postulat originel. La même tendance a été retrouvée dans les données de Repp & Crowder (1990) pour les voyelles anglaises (figure 5.5, page 105). L'étude du chapitre 3, pourtant, a montré que les asymétries dépendaient de l'expérience linguistique. Chez les francophones et chez les italophones, la discrimination est plus facile dans la direction /e/ → /i/. Pour les hispanophones et les lusophones, l'ordre n'a pas joué de rôle. Ceci nous amène au postulat que la périphéricité n'est pas un biais universel mais est liée, soit aux mécanismes de décodage au niveau de la mémoire à court terme, soit aux représentations permanentes. Dans le premier cas, les auditeurs de certaines langues apprennent pendant leur exposition langagière à utiliser la périphérie comme point de référence. Par 'représentation mentale', nous entendons des informations emmagasinées dans la mémoire à long terme et qui concernent les limites de l'espace acoustique. Néanmoins, il est difficile de comprendre pourquoi une partie seulement des locuteurs utiliseraient de telles informations.

Une théorie similaire est celle de la focalisation. Schwartz & Escudier (1989) suggèrent que les traces auditives des voyelles focales sont plus stables au niveau de la mémoire à court terme. Schwartz *et al.* (2005) expliquent que les asymétries sont dues à la saillance perceptive des stimuli focaux qui servent de points de comparaison dans les tâches de discrimination. Nous pouvons interpréter cette stabilité de plusieurs façons. Il peut s'agir d'une meilleure activation neuronale accompagnée de bruit interne plus faible pour les stimuli focaux ou bien d'une dégradation plus lente des traces mémorielles. Il est aussi possible que les locuteurs puissent mieux décoder ces stimuli dans la mémoire à court terme. Les résultats de l'étude inter-langues montrent qu'il ne s'agit pas d'un facteur universel. Nous nous attendrions à ce que tous les locuteurs partagent cette stratégie en faveur des stimuli focaux. Schwartz *et al.* (2005, p.433) soulèvent le point que la focalisation peut être innée ou acquise. L'étude sur la variabilité (chapitre 6) apporte des données sur le caractère volatile de

la focalisation. Tandis que chez certains auditeurs, des effets solides sont retrouvés à l'extrémité /i/, ceci n'est pas le cas pour d'autres sujets. De même, ces effets d'ordre semblent dépendre du paradigme expérimental. Quant à la notion de stabilité, nos résultats ne révèlent pas d'interaction claire entre les effets d'ordre et l'intervalle Interstimulus.

Cowan & Morse (1986) et Repp & Crowder (1990) soutiennent que quand une voyelle est retenue dans la mémoire, elle se dégrade vers le centre de l'espace acoustique, c'est-à-dire qu'elle se 'neutralise'. Selon les auteurs, cette dégradation est à l'origine des asymétries mais ils ne précisent pas s'il s'agit de la trace mémorielle (image auditive) du stimulus ou bien du produit final du traitement de la mémoire à court terme. Nous supposons que le produit de traitement est plus résistant à la dégradation puisque les informations superflues ont été déjà filtrées et écartées. Il paraît plus logique que ce soit l'image auditive qui se détériore. Cette image, quoique partiellement abstraite et débarrassée d'indices fins, est emmagasinée sous forme analogique et est donc plus volatile. Les résultats du chapitre 5 montrent que ce postulat de la neutralisation n'est pas viable.

L'expérience linguistique semble jouer un rôle sur les asymétries. Comme nous l'expliquons plus haut, le chapitre 3 révèle des effets d'ordre pour les locuteurs de deux langues seulement. De même, nous avons trouvé des différences dans les effets de contexte et d'ordre dans notre étude sur la perception des configurations manuelles en Langue des Signes Française. Il n'est pourtant pas évident d'affirmer si les participants des divers groupes linguistiques décodent le signal de façon dissimilaire ou bien s'ils attribuent différents poids aux différents stimuli.

Dans le chapitre 7, nous avons mis à l'épreuve le rôle que la périphéricité et le caractère quantique pouvaient avoir sur la production des voyelles. Nous sommes parti de deux hypothèses : a. les voyelles périphériques et quantiques (focales) sont plus stables en production, et b. le déplacement des organes phonatoires est plus restreint pendant un changement vers une voyelle plus périphérique puisque nous arrivons vite à des plafonds et un déplacement ultérieur peut donner une autre catégorie de sons. Aucune des deux hypothèses n'a été validée par nos données. En ce qui concerne la stabilité (variance réduite), nous avons observé que les voyelles arrondies (antérieures et postérieures) présentent le plus de variance. Quant aux asymétries, la transition /i/→/e/ n'a pas été plus importante que la transition /e/→/i/.

En combinant les résultats de plusieurs expériences, nous remarquons que les asymétries ne semblent pas être directement liées à des représentations mentales permanentes. Nous n'avons pas d'indication en faveur de l'utilisation de prototypes. De même, les effets d'ordre ne sont pas corrélés aux effets de contexte, ce qui signifie qu'il ne s'agit pas d'un mécanisme d'étiquetage différentiel. Par conséquent, nous concluons que le phénomène en question est localisé au niveau de la mémoire à court terme, ou est lié à l'image auditive des stimuli. Nous ne croyons pas que la mémoire précatégorielle puisse jouer de rôle. Étant donné les différences entre les groupes linguistiques, nous supposons qu'un traitement, même élémentaire, du stimulus se met en place. L'image auditive est une version abstraite de la trace mémorielle qui est possiblement assujettie aux règles d'organisation perceptives qui sont localisées dans la mémoire à long terme.

Perspectives

Le phénomène d’asymétrie paraît être enraciné dans le décodage de l’information et dans la prise de décision. Dans cette thèse, nous avons tenté de démontrer qu’il s’agissait d’un mécanisme extrêmement complexe. La variabilité inter-sujet démontre que le déclencheur principal n’est pas un biais universel lié aux contraintes articulatoires, à la stabilité des traces mémorielles ou au traitement de bas niveau dans la mémoire à court terme. Dans le reste de cette section, nous allons proposer des méthodologies à adopter afin de mieux comprendre les effets d’ordre.

Nous avons suggéré dans le chapitre 2 que les confusions asymétriques étaient très fréquentes dans la reconnaissance des voyelles et des consonnes (voir tableaux 2.1 et 2.2). Les matrices de confusion sont une source riche d’informations dans laquelle des biais universaux peuvent émerger. Une étude systématique d’un grand nombre de matrices est en cours d’élaboration et une nouvelle méthode d’analyse est proposée. Il faut noter, pourtant, que le groupement de stimuli sous une étiquette (/i/, /p/, ...), à savoir l’abstraction de la nature physique des stimuli, peut souvent être trompeur. Imaginons que pour la paire des consonnes /p/ et /k/, /p/ est identifié correctement à 100% tandis que les auditeurs catégorisent les segments /k/ comme /p/ à 30%. Il nous est impossible de savoir si cette confusion asymétrique est due à la tendance des locuteurs à prononcer /k/ avec des gestes propres à /p/ (avec une légère labialisation par exemple) ou bien à la stratégie des auditeurs qui auraient tendance à donner la réponse /p/ plus souvent à cause d’un statut éventuellement privilégié de /p/.

Un problème inhérent aux études utilisant une tâche ‘même-différent’ est la nature binaire de ces données. Si la distance spectrale entre deux stimuli est à peine perceptible, les scores arrivent à un minimum (autour de 0%). Des analyses statistiques sont alors impossibles. Si cette distance spectrale est au dessus du seuil de discrimination, des effets de plafond sont immédiatement atteints et le même problème d’analyse émerge. Nous suggérons que le paradigme le plus approprié pour les tâches de discrimination est le jugement de similarité. En introduisant une échelle de 0 à 7 par exemple, on évite les asymptotes. En même temps, ce type de données se prête à des analyses statistiques plus solides. Il est possible, par exemple, de mener une analyse de régression en déclarant comme variable indépendante le jugement de similarité et comme variables dépendantes les différentes mesures estimant la différence spectrale entre les stimuli, qu’il s’agisse d’estimations des valeurs formantiques, de mesures des bandes critiques, de durée, etc.

Un troisième point que nous désirons soulever concerne les asymétries en production. Nous sommes conscient qu’une différence au niveau spectral peut provenir

de plusieurs gestes articulatoires. Notre étude acoustique ne nous donne que des indices indirects sur la production. L'inversion acoustique→articulatoire étant complexe, une analyse acoustique ne peut pas nous fournir d'informations sur les gestes s'exécutant de façon asymétrique ou sur leur plus grande stabilité. Nous avons récemment entrepris une étude physiologique des asymétries où nous utilisons des ultrasons et la caméra Qualisys pour étudier les mouvements de la langue et des lèvres respectivement. Nous croyons que cette méthode révélera les stratégies des locuteurs de façon beaucoup plus claire. De même, nous avons décidé d'élever le nombre de répétitions de chaque logatome à 200, ce qui nous permettra d'effectuer des analyses discriminantes, et aussi d'estimer la variance avec plus de fiabilité.

Le nombre de répétitions est la clé dans le travail sur les asymétries. Avec un paradigme 'même-différent', la seule analyse qui nous permet d'estimer l'effet d'ordre pour chaque paire est le test χ^2 . Pour que la relation entre l'ordre de présentation et le nombre de réponses 'différent' soit significative, une des deux conditions doit être satisfaite :

- si le nombre de répétitions pour chaque paire dans chaque ordre est restreint, l'effet d'ordre doit être relativement élevé, à savoir au-moins de 25%
- si le nombre de répétitions est élevé (≥ 50), la différence nécessaire entre les deux ordres est réduite mais nous risquons une surexposition des locuteurs aux stimuli. Il est possible que les asymétries se réduisent quand les stimuli deviennent très familiers.

Dans un futur article, nous comptons envisager la relation entre le degré d'asymétrie et le degré d'exposition aux stimuli.

Nous avons discuté pendant cette thèse de la stabilité des traces mémorielles des stimuli et d'une éventuelle différence entre les stimuli. Pour que ceci soit vérifié, une étude plus systématique de l'interaction entre les fausses alarmes et l'intervalle interstimulus est indispensable. Les asymétries peuvent être dues à un effet proactif ou rétroactif. Il faut donc déterminer si l'effet de récence est plus puissant que l'effet de primauté. Cowan & Morse (1986) rapportent que les effets d'ordre sont positivement corrélés à l'intervalle interstimulus, tandis que Repp *et al.* (1979) concluent que les asymétries ne sont pas affectées par l'interférence introduite dans l'Intervalle Interstimulus.

Dans le chapitre 2, nous avons présenté les différents modèles de discriminabilité qui soutiennent l'idée que le bruit interne (ou incertitude) dans l'activation neuronale est similaire pour tous les stimuli. Nous avons ensuite postulé qu'il est possible que pour certains stimuli, la trace mémorielle soit mieux enregistrée à cause de la structure interne de ces stimuli, et de la saillance de certaines fréquences. Cette question mérite plus de travail. Le bruit interne peut être estimé soit par des méthodes psychophysiques (voir la méthodologie de Kontsevich *et al.*, 2002), soit par la neurophysiologie.

Dans ce travail, nous n'avons pas exploité les données sur les temps de réaction. Il est, pourtant, obligatoire de faire une étude poussée de ce type de données afin d'exploiter les mécanismes sous-jacents de discrimination et de choisir le modèle théorique qui correspond le mieux à nos besoins. Chaque modèle fait des hypothèses différentes sur les patrons de temps de latence. Les données brutes ne sont pas suffisantes pour décomposer des phénomènes aussi complexes.

Le dernier point que nous voulons soulever concerne la relation entre les effets d'ordre et les effets de contexte. Nous avons soutenu dans les chapitres 4 et 6 qu'un tel lien n'était pas retrouvé. Néanmoins, notre méthode d'investigation est indirecte puisque la catégorisation des deux stimuli n'est pas explicite. Nous pouvons imaginer une tâche de discrimination dans laquelle le sujet doit, après le jugement de similarité, faire une catégorisation explicite, c'est-à-dire indiquer l'étiquette des deux stimuli.

Cette thèse a essayé de mettre en évidence le phénomène d'asymétrie et ses implications. Il s'agit d'un mécanisme qui semble agir sur la perception et la production de la parole, ainsi que sur la prise de décision. Étant donné que les effets d'ordre concernent une large gamme de jugements psychophysiques, il est indispensable d'adapter les protocoles expérimentaux et les méthodes d'investigation afin que ce type de tendances soit mieux pris en compte.

Bibliographie

- ARLINGER, S., JERLVALL, L., AHREN, L. & HOLMGREN, E. (1977). Thresholds for linear frequency ramps of a continuous pure tone. *Acta Oto-laryngologica*, 83:317–327.
- ATKINSON, R. & JUOLA, J. (1974). Search and decision processes in recognition memory. Dans KRANTZ, D., ATKINSON, R., LUCE, R. & SUPPES, P., (éditeurs) : *Contemporary developments in mathematical psychology*, volume 1, pp. 242–293. W.H. Freeman, San Francisco.
- ATKINSON, R. & SHIFFRIN, R. (1968). Human memory : A proposed system and its control processes. Dans BOWEN, G. & SPENCE, J., (éditeurs) : *The psychology of learning and motivation*, volume 2. Academic Press, New York.
- BAKER, S. A., IDSARDI, W. J., GOLINKOFF, R. M. & PETITTO, L.-A. (2005). The perception of handshapes in American Sign Language. *Memory & Cognition*, 33(5):887–904.
- BARSALOU, L. (1983). Ad hoc categories. *Memory & Cognition*, 11(3):211–227.
- BEALE, J. & KEIL, F. (1995). Categorical effects in the perception of faces. *Cognition*, 57(3):217–239.
- BELL, T. S., DIRKS, D. D. & CARTERETTE, E. C. (1989). Interactive factors in consonant confusion patterns. *Journal of the Acoustical Society of America*, 85(1): 339–346.
- BENKÍ, J. R. (2003). Analysis of English nonsense syllable recognition in noise. *Phonetica*, 60(2):129–157.
- BERLIN, B. & KAY, P. (1969). *Basic color terms : Their universality and evolution*. University of California Press, Berkeley.
- BEST, C. & FABER, A. (2000). Developmental increase in infants' discrimination of nonnative vowels that adults assimilate to a single native vowel. Dans *International Conference on Infant Studies*, pp. 221–231, Brighton, UK.
- BEST, C., SINGH, L., BOUCHARD, J., CONNELLY, G., COOK, A. & FABER, A. (1997). Developmental changes in infants' discrimination of non-native vowels that assimilate to two native categories. Dans *Meeting of the Society of Research in Child Development, Providence, RI, 3-6 April 1997*.

- BEST, K. (1995). A direct realist view of cross-language speech perception : New Directions in Research and Theory. Dans STRANGE, W., (éditeur) : *Speech perception and linguistic experience : Theoretical and methodological issues*, pp. 171–204. York Press, Baltimore.
- BHARUCHA, J. & KRUMHANSL, C. L. (1983). The representation of harmonic structure in music : Hierarchies of stability as a function of context. *Cognition*, 13(1):63–102.
- BHARUCHA, J. & PRYOR, J. (1986). Disrupting the isochrony underlying rhythm : An asymmetry in discrimination. *Perception & Psychophysics*, 40:137–141.
- BOERSMA, P. & WEENINK, D. (2001). PRAAT, a system for doing phonetics by computer. *Glott International*, 5(9/10):341–345.
- BOHN, O.-S. & POLKA, L. (2001). Target spectral, dynamic spectral, and duration cues in infant perception of german vowels. *Journal of the Acoustical Society of America*, 110:504–515.
- BORNSTEIN, M. & KORDA, N. (1984). Discrimination and matching within and between hues measured by reaction time : Some implications for categorical perception and levels of information processing. *Psychological Research*, 46:207–222.
- BOUTORA, L. (2007). Un inventaire des configurations manuelles pour une catégorisation des unités minimales en LSF. *Silexicales*, 5:17–30.
- BOUTORA, L. (2008). *Fondements historiques et implications théoriques d'une phonologie des langues des signes. Etude de la perception catégorielle des configurations manuelles en LSF et réflexion sur la transcription des langues des signes*. PhD thesis, Université Paris 8.
- BOUTORA, L. & KARYPIDIS, C. (2007). Are French Sign Language handshapes perceived categorically? Submitted.
- BOUTORA, L. & KARYPIDIS, C. (2009a). Establishing criteria for selecting stimuli and defining categorical perception. comparing methodologies from american and french sign languages. Submitted.
- BOUTORA, L. & KARYPIDIS, C. (2009b). Perception catégorielle et représentations mentales des configurations manuelles en Langues des Signes Française. To appear.
- BRAIDA, L., LIM, J., BERLINER, J., DURLACH, N., RABINOWITZ, W. & PURKS, S. (1984). Intensity perception. XIII. Perceptual anchor model of context-coding. *Journal of the Acoustical Society of America*, 76(3):722–731.
- BROADBENT, D. E. (1958). *Perception and Communication*. Pergamon, London.
- BURDICK, C. K. & MILLER, J. D. (1975). Speech perception by the chinchilla : discrimination of sustained |a| and |i|. *Journal of the Acoustical Society of America*, 58(2):415–427.

- CALLIOPE (1989). *La parole et son traitement automatique*. Masson, Paris.
- CARRÉ, R. (2004). Programme de catégorisation de voyelles isolées ou dans un contexte dynamique. Personal communication.
- CHAN, S., CHUANG, C.-K. & WANG, W.-Y. (1975). Cross-linguistic study of categorical perception for lexical tone. *Journal of the Acoustical Society of America*, 58:S119.
- CHIBA, T. & KAJIYAMA, M. (1941). *The vowel. Its nature and structure*. Tokyo-Kaiseikan Publishing Company, Tokyo.
- COWAN, N. & MORSE, P. (1986). The use of auditory and phonetic memory in vowel discrimination. *Journal of the Acoustical Society of America*, 79:500–507.
- CROWDER, R. & MORTON, J. (1969). Precategorical acoustic storage (PAS). *Perception & Psychophysics*, 5:365–373.
- CUTTING, J. (1978). There may be nothing peculiar to perceiving in a speech mode. Dans REQUIN, J., (éditeur) : *Attention and Performance*, volume VII, pp. 229–245. Lawrence Erlbaum Association, Hillsdale, New Jersey.
- CUTTING, J. E., ROSNER, B. S. & FOARD, C. F. (1976). Perceptual categories for musiclike sounds : Implications for theories of speech perception. *The Quarterly Journal of Experimental Psychology*, 28(3):361.
- DEJARDINS, R. & TRAINOR, L. (1998). Fundamental frequency influences vowel discrimination in 6-month-old infants. *Can. Acoust. Proc.*, 26:96–97.
- DOOLEY, G. & MOORE, B. (1988). Detection of linear frequency glides as a function of frequency and duration. *Journal of the Acoustical Society of America*, 84:2045–2057.
- DURLACH, N. I. & BRAIDA, L. D. (1969). Intensity Perception. I. Preliminary Theory of Intensity Resolution. *Journal of the Acoustical Society of America*, 46(2B):372–383.
- EHRENSTEIN, W. & EHRENSTEIN, A. (1999). Psychophysical methods. Dans WINDHORST, U. & JOHANSSON, H., (éditeurs) : *Modern Techniques in Neuroscience Research*, chapitre Psychophysical methods, pp. 1211–1241. Springer-Verlag, Berlin.
- EHRET, G. (1987). Categorical perception of sound signals : Facts and hypotheses from animal studies. Dans HARNAD, S., (éditeur) : *Categorical perception : The groundwork of cognition*, pp. 301–331. Cambridge University Press, New York.
- EMMOREY, K., MCCULLOUGH, S. & BRENTARI, D. (2003). Categorical perception in American Sign Language. *Language & Cognitive Processes*, 18:21–45.
- ESCUDERO, P., BOERSMA, P., RAUBER, A. S. & BION, R. A. H. (2009). A cross-dialect acoustic description of vowels : Brazilian and European Portuguese. *Journal of the Acoustical Society of America*, 126(3):1379–1393.

- ETCOFF, N. L. & MAGEE, J. J. (1992). Categorical perception of facial expressions. *Cognition*, 44(3):227–240.
- FECHNER, G. T. (1860). *Elemente der Psychophysik*. Breitkopf und Hèartel, Leipzig.
- FLANAGAN, J. L. (1957). Estimates of the Maximum Precision Necessary in Quantizing Certain “Dimensions” of Vowel Sounds. *Journal of the Acoustical Society of America*, 29(4):533–534.
- FLETCHER, H. (1940). Auditory patterns. *Rev Mod Physics*, 12:47–66.
- FOURCIN, A., AINSWORTH, W., FANT, G., FUJIMURA, O., FUJISAKI, H., HESS, W., HOLMES, J., ITAKURA, F., SCHROEDER, M. & STRUBE, H. (1977). Speech processing by man and machine. Dans BULLOCK, T., (éditeur) : *Recognition of Complex Acoustic Signals*, pp. 307–351, Berlin. Dahlem Workshops, Abakon Verlag.
- FOWLER, C. (1983). Realism and unrealism : A reply. *Journal of Phonetics*, 11:303–322.
- FRIEDA, E., WALLEY, A., FLEGE, J. & SLOANE, M. (1999). Adults’ perception of native and nonnative vowels : implications for the perceptual magnet effect. *Perception & Psychophysics*, 61(3):561–577.
- FRY, D., ABRAMSON, A., EIMAS, P. & LIBERMAN, A. (1962). The identification and discrimination of synthetic vowels. *Language and Speech*, 5:171–188.
- FUJISAKI, H. (1979). On the modes and mechanisms of speech perception - analysis and interpretation of categorical effects in discrimination. Dans LINDBLOM, B. & ÖHMAN, S., (éditeurs) : *Frontiers of Speech Communication Research*, pp. 177–189. Academic Press, London.
- FUJISAKI, H. & KAWASHIMA, T. (1969). On the modes and mechanisms of speech perception. *Annual Report of the Engineering Research Institute, Faculty of Engineering, University of Tokyo*, 28:67–73.
- FUJISAKI, H. & KAWASHIMA, T. (1970). Some experiments on speech perception and a model for the perceptual mechanism. *Annual Report of the Engineering Research Institute, Faculty of Engineering, University of Tokyo*, 29:207–214.
- FUJISAKI, H. & KAWASHIMA, T. (1971). A model of the mechanisms for speech perception quantitative analysis of categorical effects in discrimination. *Annual Report of the Engineering Research Institute, Faculty of Engineering, University of Tokyo*, 30:59–68.
- GARDNER, R. B. & WILSON, J. P. (1979). Evidence for direction-specific channels in the processing of frequency modulation. *Journal of the Acoustical Society of America*, 66(3):704–709.
- GREEN, D. M. & SWETS, J. A. (1966). *Signal Detection Theory and Psychophysics*. Wiley, New York.

- HABER, R. N. & HERSHENSON, M. (1980). *Psychology of Visual Perception*. Holt, Rinehart & Winston, New York, 27ème édition.
- HALLÉ, P. A., CHANG, Y.-C. & BEST, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32(3):395–421.
- HELLSTRÖM, Å. (1979). Time errors and differential sensation weighting. *Journal of Experimental Psychology : Human Perception and Performance*, 5:460–477.
- HELLSTRÖM, Å. (1985). The time-order error and its relatives : Mirrors of cognitive processes in comparing. *Psychological Bulletin*, 97:35–61.
- HELLSTRÖM, Å. (1989). What happens when we compare two successive stimuli? Dans LJUNGGREN, G. & DORNIC, S., (éditeurs) : *Psychophysics in action*, pp. 25–39. Springer-Verlag, Berlin.
- HELLSTRÖM, Å. (2003). Comparison is not just subtraction : Effects of time- and space-order on subjective stimulus difference. *Perception & Psychophysics*, 65: 1161–1177.
- HELLSTRÖM, Å. (2008). Temporal asymmetry and “magnet effect” in similarity and discrimination of prototypical and nonprototypical stimuli : consequences of differential sensation weighting. Dans *Fechner Day 2008. Proceedings of the 24th Annual Meeting of the International Society for Psychophysics*.
- HIENZ, R., ALESZCZYK, C. & MAY, B. (1996). Vowel discrimination in cats : acquisition, effects of stimulus level, and performance in noise. *Journal of the Acoustical Society of America*, 99:3656–3668.
- HIENZ, R., SACHS, M. & SINNOTT, J. (1981). Discrimination of steady-state vowel by blackbirds and pigeons. *Journal of the Acoustical Society of America*, 70:699–706.
- HILLENBRAND, J., GETTY, L. A., CLARK, M. J. & WHEELER, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97(5):3099–3111.
- HOWARD, D., ROSEN, S. & BROAD, V. (1992). Major/minor triad identification and discrimination by musically trained and untrained listeners. *Music Perception*, 10(2):205–220.
- IVERSON, P. & KUHL, P. (2000). Perceptual magnet and phoneme boundary effects in speech perception : Do they arise from a common mechanism? *Perception & Psychophysics*, 62(4):874–886.
- IVERSON, P. & KUHL, P. K. (1995). Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *Journal of the Acoustical Society of America*, 97:553–562.
- JOHNSON, K. (1997). Speech perception without speaker normalization : An exemplar model. Dans JOHNSON, K. & MULLENIX, J., (éditeurs) : *Talker variability in speech processing*, pp. 145–166. Academic Press, San Diego.

- KARYPIDIS, C. (2007a). Dégradation de la qualité des voyelles retenues en mémoire à court terme : la voyelle neutre et les points de référence. Dans *Colloque des Jeunes Chercheurs en Sciences Cognitives, Lyon, France*.
- KARYPIDIS, C. (2007b). Dégradation de la qualité des voyelles retenues en mémoire à court terme : l'hypothèse de neutralisation. Dans *Journées fondatrices du groupe Perception Sonore, Lyon, France*.
- KARYPIDIS, C. (2007c). Order effects and vowel decay in short-term memory : the neutralization hypothesis. Dans *Proceedings of the XVI International Congress of Phonetic Sciences*, pp. 657–660, Saarbrücken, Germany.
- KARYPIDIS, C. (2009). Les effets d'ordre de présentation et l'hypothèse de neutralisation. la voyelle neutre et les points de référence. To appear.
- KARYPIDIS, C., COLAZO-SIMON, A., COSTAGLIOLA, A. V., CECOTE STEIN, C. & GUGLIELMI, G. (2008). Order effects and peripherality : a cross-linguistic perceptual study using an [i]-[e] articulatory continuum. *Acoustics '08*, 123(5):3182.
- KARYPIDIS, C., COLAZO-SIMON, A., COSTAGLIOLA, A. V., STEIN, C. C. & GUGLIELMI, G. (2007). The presentation order effect and the role of linguistic experience : a cross-linguistic perceptual study on an [i]-[e] continuum. Dans *Proceedings of the 3o Convegno Nazionale AISV*, pp. 33–41, Trento, Italy.
- KARYPIDIS, C., COSTAGLIOLA, A. V. & COLAZO-SIMON, A. (2006a). Vowel-prototype assimilation : a cross-linguistic perceptual study of five-vowel systems. Dans *La Comunicazione Parlata, Naples, Italy*.
- KARYPIDIS, C., COSTAGLIOLA, A. V. & COLAZO-SIMON, A. (2009). Relative weight of perceptual cues : a cross-linguistic study of 5-vowel systems. Unpublished.
- KARYPIDIS, C., GUGLIELMI, G. & COSTAGLIOLA, A. V. (2006b). Asimmetria nella percezione vocalica di L1 : uno studio di sintesi articolatoria del continuum [i]-[e] (Asymmetry in vowel perception in L1 : evidence from articulatory synthesis of an [i]-[e] continuum). Dans SAVY, R. & CROCCO, C., (éditeurs) : *Proceedings of the 2o Convegno Nazionale AISV*, pp. 695–707.
- KARYPIDIS, C. & MONTAGU, J. (2009). Towards a phonetic and phonological definition of the feature [±round]. Submitted.
- KLATT, D. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, 67:737–793.
- KLUENDER, K., DIEHL, R. & KILLEEN, P. (1987). Japanese quail can learn phonetic categories. *Science*, 237(4819):1195–1197.
- KONTSEVICH, L. L., CHEN, C.-C. & TYLER, C. W. (2002). Separating the effects of response nonlinearity and internal noise psychophysically. *Vision Research*, 42(14):1771–1784.

- KROLL, J. & HERSHENSON, M. (1980). Two stages in visual matching. *Canadian Journal of Psychology*, 34(1):49–61. PMID : 7388695.
- KRUSCHKE, J. K. (1992). ALCOVE : an exemplar-based connectionist model of category learning. *Psychological Review*, 99:22–44.
- KUHL, P. (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. & PADDEN, D. (1982). Enhanced discriminability at the phonetic boundaries for the voicing feature in macaques. *Perception & Psychophysics*, 32:542–550.
- KUHL, P., WILLIAMS, K., LACERDA, F., STEVENS, K. & LINDBLOM, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255:606–608.
- KUHL, P. K. & MILLER, J. D. (1978). Speech perception by the chinchilla : Identification functions for synthetic vowel stimuli. *Journal of the Acoustical Society of America*, 63(3):905–917.
- LAHIRI, A. & REETZ, H. (2002). Underspecified recognition. Dans GUSSENHOVEN, C. & WERNER, N., (éditeurs) : *Laboratory Phonology VII*, pp. 637–675. Mouton, Berlin.
- LANE, H. (1965). The motor theory of speech perception. *Psychological Review*, 72:275–309.
- LIBERMAN, A. (1970). Some characteristics of perception in the speech mode. *Perception and its disorders*, 48:238–254.
- LIBERMAN, A., COOPER, F., HARRIS, K. & MACNEILAGE, P. (1963). A motor theory of speech perception. Dans *Speech Communication Seminar*, Stockholm. Royal Institute of Technology.
- LIBERMAN, A. & MATTINGLY, I. (1985). The motor theory of speech perception revisited. *Cognition*, 21:1–36.
- LIBERMAN, A. M., HARRIS, K. S., HOFFMAN, H. S. & GRIFFITH, H. S. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54:358–368.
- LIPSKI, S., LAHIRI, A. & EULITZ, C. (2007). Differential height specification in front vowels for German speakers and Turkish-German bilinguals : an electroencephalographic study. Dans *Proceedings of the XVI International Congress of Phonetic Sciences*, Saarbrücken, Germany.
- LIVELY, S. & PISONI, D. (1997). On prototype and phonetic categories : A critical assessment of the perceptual magnet effect in speech perception. *J. Exp. Psych. : Hum. Percept. Psychophys.*, 23:1665–1679.

- LOCKE, S. & KELLAR, L. (1973). Categorical perception in a non linguistic mode. *Cortex*, 9:355–369.
- LOCKHEAD, G. (1972). Processing dimensional stimuli : A note. *Psychological Review*, 79:410–419.
- LOTTO, A., KLUENDER, K. & HOLT, L. (1998). Depolarizing the perceptual magnet effect. *Journal of the Acoustical Society of America*, 103:3648–3655.
- MACMILLAN, N. (1977). The psychophysics of categorical perception. *Psychological Review*, 84:452–471.
- MACMILLAN, N., BRAIDA, L. & GOLDBERG, R. (1987). Central and peripheral processes in the perception of speech and nonspeech sounds. Dans SCHOUTEN, M., (éditeur) : *The Psychophysics of Speech Perception*, pp. 28–45. Martinus Nijhoff, The Hague.
- MACMILLAN, N. A., GOLDBERG, R. F. & BRAIDA, L. D. (1988). Resolution for speech sounds : Basic sensitivity and context memory on vowel and consonant continua. *Journal of the Acoustical Society of America*, 84(4):1262–1280. Journal Article City U New York, Brooklyn Coll, US. English.
- MADDIESON, I. (1991). Testing the universality of phonological generalizations with a phonetically specified segment database : results and limitations. *Phonetica*, 48:193–206.
- MAEDA, S. (1990). Compensatory articulation during speech ; evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model. Dans MARCHAL, A. & HARDCASTLE, W., (éditeurs) : *Speech Production and Speech Modelling, NATO ASI Series*. Kluwer Academic Publishers.
- MANTAKAS, M., SCHWARTZ, J. L. & ESCUDIER, P. (1986). Modèle de prédiction du ‘deuxième formant effectif’ F₂ — application à l’étude de la labialité des voyelles avant du français. Dans *Proceedings of 15th Journées d’Étude sur la parole*, pp. 157–161.
- MASSARO, D. (1970). Preperceptual auditory images. *Journal of Experimental Psychology*, 85(3):411–417.
- MASSARO, D. (1972). Preperceptual images, processing time, and perceptual units in auditory perception. *Psychological Review*, 79(2):124–145.
- MASSARO, D. W. (1989). Testing between the TRACE model and the fuzzy logical model of speech perception. *Cognitive Psychology*, 21(3):398–421.
- MATHUR, G. & BEST, C. T. (2007). Categorical perception study of an ASL hand-shape contrast in dynamic nonce signs. Unpublished.
- MAY, B., MOODY, D. B. & STEBBINS, W. C. (1989). Categorical perception of conspecific communication sounds by japanese macaques, *macaca fuscata*. *Journal of the Acoustical Society of America*, 85(2):837–847.

- MILLER, G. (1956). The magical number seven, plus or minus two : some limits on our capacity for processing information. *Psychological Review*, 63:81–97.
- MILLER, G. & NICELY, P. (1955). Analysis of perceptual confusions among some english consonants. *Journal of the Acoustical Society of America*, 27:338–353.
- MORTON, J. (1969). The interaction of information in word recognition. *Psychological Review*, 76:165–178.
- MORTON, J. (1970). A functional model for memory. Dans NORMAN, D., (éditeur) : *Models for human memory*. Academic Press, New York.
- NICKERSON, R. (1972). Binary-classification reaction time : A review of some studies of human information-processing capabilities. *Psychonomic Monograph Supplements*, 4:275–317.
- NICKERSON, R. (1973). The use of binary-classification tasks in the study of human information processing : A tutorial survey. Dans KORNBLUM, S., (éditeur) : *Attention and Performance*, volume 4, pp. 449–476. Academic Press, New York.
- NICKERSON, R. (1978). On the time it takes to tell things apart. Dans REQUIN, J., (éditeur) : *Attention and Performance*, volume 7, pp. 77–88. Erlbaum, Hillsdale, N.J.
- NORMAN, D. A. (1970). *Models of Human Memory*. Academic Press, New York.
- ODEN, G. & MASSARO, D. (1978). Integration of featural information in speech perception. *Psychological Review*, 85:172–191.
- Op de BEECK, H., WAGEMANS, J. & VOGELS, R. (2003). Asymmetries in stimulus comparisons by monkey and man. *Current Biology*, 13:1803–1808.
- OZGEN, E. & DAVIES, I. (2002). Acquisition of categorical color perception : A perceptual learning approach to the linguistic relativity hypothesis. *Journal of Experimental Psychology : General*, 131:477–493.
- PETERSON, G. & BARNEY, H. (1952). Control methods used in the study of the vowels. *Journal of the Acoustical Society of America*, 24:175–184.
- PISONI, D. (1971). *On the nature of the categorical perception of speech sounds*. Unpublished doctoral dissertation, University of Michigan.
- PISONI, D. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics*, 13:253–260.
- PISONI, D. B. (1976). Some effects of discrimination training on the identification and discrimination of rapid spectral changes. *Research on Speech Perception Progress Report, Indiana University*, 3:121–141.
- PISONI, D. B. & TASH, J. (1975). Auditory property detectors and processing place features in stop consonants. *Perception & Psychophysics*, 16(6):401–408.

- PLAUCHÉ, M., DELOGU, C. & OHALA, J. (1997). Asymmetries in consonant confusion. Dans *Proceeding of the Eurospeech '97*, volume 4, pp. 2187–2190, 5th European Conference on Speech Communication and Technology.
- POLK, T. A., BEHENSKY, C., GONZALEZ, R. & SMITH, E. (2002). Rating the similarity of simple perceptual stimuli : asymmetries induced by manipulating exposure frequency. *Cognition*, 82:B75–B88.
- POLKA, L. & BOHN, O.-S. (1996). A cross-language comparison of vowel perception in english-learning and german-learning infants. *Journal of the Acoustical Society of America*, 100:577–592.
- POLKA, L. & BOHN, O.-S. (2003). Asymmetries in vowel perception. *Speech Communication*, 41(1):221–231.
- POLKA, L. & WERKER, J. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology : Human Perception and Performance*, 20:421–435.
- REPP, B., HEALY, A. F. & CROWDER, R. G. (1979). Categories and context in the perception of isolated steady-state vowels. *Journal of Experimental Psychology : Human Perception and Performance*, 5:129–145.
- REPP, B. H. & CROWDER, R. G. (1990). Stimulus order effects in vowel discrimination. *Journal of the Acoustical Society of America*, 88(5):2080–2090.
- ROBERSON, D. & DAVIDOFF, J. (2000). The "categorical perception" of colors and facial expressions : The effect of verbal interference. *Memory & Cognition*, 28:977–986.
- ROSCH, E. (1975). Cognitive reference points. *Cognitive Psychology*, 7(4):532–547.
- ROSCH, E. (1978). Principles of categorization. Dans ROSCH, E. & LLOYD, B., (éditeurs) : *Cognition and categorization*, pp. 27–48. Lawrence Erlbaum, Hillsdale, NJ.
- ROSCH, E. H. (1973). Natural categories. *Cognitive Psychology*, 4(3):328–350.
- SAMUEL, A. (1982). Phonetic prototypes. *Perception & Psychophysics*, 31(4):307–314.
- SCHELLENBERG, E. (2002). Asymmetries in the discrimination of musical intervals : Going out-of-tune is more noticeable than going in-tune. *Music Perception*, 19:223–248.
- SCHELLENBERG, E. G. (2001). Asymmetries in the discrimination of musical intervals : Going Out-of-Tune is more noticeable than going In-Tune. *Music Perception*, 19(2):223–248.
- SCHELLENBERG, E. G. & TREHUB, S. E. (1999). Culture-General and Culture-Specific factors in the discrimination of melodies,. *Journal of Experimental Child Psychology*, 74(2):107–127.

- SCHOUTEN, B., GERRITS, E. & van HESSEN, A. (2003). The end of categorical perception as we know it. *Speech Communication*, 41(1):71–80.
- SCHOUTEN, M. (1980). The case against a speech mode of perception. *Acta Psychologica*, 44:71–98.
- SCHOUTEN, M. (1986). Three-way identification of sweep tones. *Perception & Psychophysics*, 40:359–361.
- SCHOUTEN, M. & van HESSEN, A. (1992). Modeling phoneme perception. I : Categorical perception. *Journal of the Acoustical Society of America*, 92(4):1841–1855.
- SCHOUTEN, M. E. (1985). Identification and discrimination of sweep tones. *Perception & Psychophysics*, 37:369–376.
- SCHROEDER, M., ATAL, B. & HALL, J. (1979). Optimizing digital speech coders by exploiting masking properties of the human ear. *Journal of the Acoustical Society of America*, 66:1647–1652.
- SCHWARTZ, J.-L., ABRY, C., BOË, L.-J., MÉNARD, L. & VALLÉE, N. (2005). Asymmetries in vowel perception, in the context of the Dispersion-Focalisation Theory. *Speech Communication*, 45(4):425–434.
- SCHWARTZ, J. L., BOË, L. J., VALLÉE, N. & ABRY, C. (1997). The dispersion-focalization theory of vowel systems. *Journal of Phonetics*, 25:255–286.
- SCHWARTZ, J. L. & ESCUDIER, P. (1989). A strong evidence for the existence of a large scale integrated spectral representation in vowel perception. *Speech Communication*, 8:235–259.
- SHANNON, C. & WEAVER, W. (1949). *The mathematical theory of communication*. University of Illinois Press, Urbana.
- SMITH, E. & MEDIN, D. (1981). *Categories and concepts*. Harvard University Press, Cambridge, MA.
- SPERLING, G. (1963). A model for visual memory tasks. *Human Factors*, 5:19–31.
- SPERLING, G. (1967). Successive approximation to a model for short-term memory. *Acta Psychologica*, 27:285–292.
- STEVENS, K. (1989). On the quantal nature of speech. *Journal of Phonetics*, 17:3–45.
- STEVENS, K. N. (1972). The quantal nature of speech : evidence from articulatory-acoustic data. Dans DAVID, E. & DENES, P., (éditeurs) : *Human Communication : a unified view*, pp. 51–66. Mac Graw-Hill, New York.
- STUDDERT-KENNEDY, M., LIBERMAN, A. M., HARRIS, K. S. & COOPER, F. S. (1970). Motor theory of speech perception : A reply to Lane’s critical review. *Psychological Review*, 77:234–249.

- SUSSMAN, J. E. & LAUCKNER-MORANO, V. J. (1995). Further tests of the “perceptual magnet effect” in the perception of [i] : Identification and change/no-change discrimination. *Journal of the Acoustical Society of America*, 97(1):539–552.
- SWOBODA, P., KASS, J., MORSE, P. & LEAVITT, L. (1978). Memory factors in infant vowel discrimination of normal and at-risk infants. *Child Development*, 49:332–339.
- SWOBODA, P., MORSE, P. & LEAVITT, L. (1976). Continuous vowel discrimination in normal and at risk infants. *Child Development*, 47:459–465.
- TANNER, W. & SWETS, J. (1954). A decision-making theory of visual detection. *Psychological Review*, 61(6):401–409.
- THYER, N., HICKSON, L. & DODD, B. (2000). The perceptual magnet effect in Australian English vowels. *Perception & Psychophysics*, 62:1–20.
- TJOMOV, V. (1971). A model to describe the results of psychoacoustical experiments on steady-state stimuli. Dans GERSHUNI, G., (éditeur) : *Analiz Rechevykh Signalov Chelovekom*, pp. 36–49. Nauka, Leningrad.
- TRAINOR, L. (1997). The effect of frequency ratio on infants and adults discrimination of simultaneous intervals. *Journal of Experimental Psychology : Human Perception and Performance*, 23:1427–1438.
- TRAUNMÜLLER, H. (1990). Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustical Society of America*, 88(1):97–100.
- TSUMURA, T., SONE, T. & NIMURA, T. (1973). Auditory detection of frequency transition. *Journal of the Acoustical Society of America*, 53(1):17–25.
- TVERSKY, A. (1977). Features of similarity. *Psychological Review*, 84:327–352.
- UITENBROEK, D. (1997). Binomial, sisa. <http://home.clara.net/sisa/binomial.htm>, Accessed January 1, 2006.
- VALLÉE, N. (1994). *Systèmes vocaliques : de la typologie aux prédictions*. Thèse de doctorat, Université Stendhal.
- WERTHEIMER, M. (1938 [1912]). Numbers and numerical concepts in primitive peoples. Dans WILLIS, D., (éditeur) : *A Source Book of Gestalt Psychology*, pp. 265–273. Routledge and Kegan Paul, London.
- WOZNIAK, R. (1999). *Classics in Psychology, 1855-1914 : Historical Essays*. Thoemmes Press, Bristol, UK.
- ZADEH, L. (1965). Fuzzy sets. *Information and Control*, 8(3):338–353.
- ZWICKER, E. & TERHARDT, E. (1980). Analytical expressions for critical band rate and critical bandwidth as a function of frequency. *Journal of the Acoustical Society of America*, 68:1523–1525.

Table des matières

Table des figures	11
Liste des tableaux	15
Introduction	18
I Partie théorique	23
1 Discriminabilité et modules mémoriels	25
1.1 Fechner et la psychophysique	25
1.2 Les modules de la mémoire	27
1.3 Catégorisation et modèles de discriminabilité	28
1.3.1 Perception catégorielle	29
1.3.2 Le rôle des prototypes	31
1.3.3 Mode de perception unique pour la parole?	32
1.3.4 Le modèle du processus duel	33
1.3.5 La théorie trace-contexte	34
1.3.6 Le modèle à deux étapes	35
1.4 Théorie de la détection du signal	36
1.5 Conclusion	40
2 Asymétries et hypothèses	41
2.1 Asymétries et stimuli non-linguistiques	41
2.1.1 Formes géométriques	41
2.1.2 Couleurs	41
2.1.3 Tons modulés	42
2.1.4 Accords musicaux	42
2.2 Asymétries et stimuli linguistiques	42
2.2.1 Matrices de confusion et asymétries en reconnaissance	42
2.2.2 L'hypothèse de la neutralisation	44
2.2.3 Périphéricité et focalisation	48
2.2.4 L'effet des aimants perceptifs	51
2.2.5 Sous-spécification et <i>mismatch negativity</i>	53
2.3 Modèles génériques	54
2.3.1 Tversky (1977) et les asymétries de similarité	54

2.3.2	L'erreur de l'ordre temporel	56
2.4	Recadrage des théories et hypothèses formulées	57
2.4.1	Perception catégorielle	57
2.4.2	Modèle unifié pour parole et non-parole	58
2.4.3	Le modèle du processus duel	58
2.4.4	La théorie trace-contexte	59
2.4.5	Le modèle à deux étapes de Kroll & Hershenson (1980) . . .	59
2.4.6	Récapitulatif des hypothèses	59
 II Partie expérimentale		 61
3	Asymétries et expérience linguistique	63
3.1	Expérience 1 : Catégorisation	65
3.1.1	Participants	65
3.1.2	Stimuli	65
3.1.3	Procédure	68
3.1.4	Résultats	69
3.2	Expérience 2 : Discrimination	70
3.2.1	Participants	70
3.2.2	Stimuli	70
3.2.3	Procédure	70
3.2.4	Résultats	71
3.3	Expérience 3 : Catégorisation AXB	78
3.3.1	Participants	78
3.3.2	Stimuli	79
3.3.3	Procédure	79
3.3.4	Résultats	80
3.4	Conclusions générales	84
 4	 Asymétries et configurations manuelles	 85
4.1	Expérience 1 : Catégorisation	86
4.1.1	Participants	86
4.1.2	Stimuli	86
4.1.3	Procédure	87
4.1.4	Résultats	88
4.2	Expérience 2 : Discrimination	90
4.2.1	Participants	90
4.2.2	Stimuli	91
4.2.3	Procédure	91
4.2.4	Résultats	91
4.3	Conclusions générales	94
 5	 L'hypothèse de la neutralisation	 95
5.1	Expérience 1 : Discrimination	95
5.1.1	Sujets	95
5.1.2	Stimuli	96

5.1.3	Procédure	98
5.1.4	Résultats	99
5.1.5	Discussion	101
5.2	Conclusions générales	105
6	La variabilité inter-sujet	107
6.1	Sujets	108
6.2	Stimuli	108
6.3	Procédure	109
6.3.1	Identification X (iX)	109
6.3.2	Identification avec présentation des deux extrémités (iAXi et iAXe)	109
6.3.3	Identification avec appariement des stimuli (iAX1r et iAX2r)	109
6.3.4	Discrimination avec présentation des deux extrémités (dAXi et dAXe)	110
6.3.5	Discrimination avec appariement des stimuli (dAX1r et dAX2r)	110
6.3.6	Discrimination avec appariement des stimuli filtrés (dAX2r_fil)	110
6.3.7	Discrimination et Intervalle Interstimulus (dAX2r_250, dAX2r_500 et dAX2r_1000)	110
6.3.8	Discrimination et protocole 'fixed' (dAX2f)	110
6.3.9	Discrimination de tons purs (f1i, f1ii, f2i, f2ii, f3i, f3ii)	110
6.4	Résultats	111
6.4.1	Effets de contexte – ordre et les extrémités /i/ et /e/	111
6.4.2	Effets de contexte et distance acoustique (iAX1r et iAX2r)	111
6.4.3	Effets d'ordre et distance acoustique (dAX1r et dAX2r)	116
6.4.4	Effets d'ordre et apport des formants supérieurs (dAX2r et dAX2r_fil)	117
6.4.5	Effets d'ordre et différence entre paradigmes (dAX2r et dAX2f)	123
6.4.6	Effets d'ordre et l'Intervalle Interstimulus	124
6.4.7	Effets d'ordre et tons purs	127
6.5	Conclusions générales	130
7	Asymétries et cibles acoustiques	145
7.1	Expérience 1 : Production de logatomes CVCV	146
7.1.1	Participants	146
7.1.2	Corpus	146
7.1.3	Procédure	147
7.1.4	Résultats	147
7.2	Expérience 2 : Production de triptongues $V_1V_2V_1$	153
7.2.1	Participants	153
7.2.2	Corpus	153
7.2.3	Procédure	153
7.2.4	Résultats	154
7.3	Conclusions générales	154
	Discussion générale	156

Perspectives	159
Bibliographie	163
Table des matières	175

Asymétries en perception et traitement de bas
niveau : traces auditives, mémoire à court terme
et représentations mentales

Asymmetries in perception and low-level
processing : auditory traces, short-term memory
an mental representations

L'effet d'ordre de la présentation des stimuli a un impact sur leur discriminabilité. Dans cette thèse, nous étudions le phénomène d'asymétrie et comment celui-ci réagit avec divers facteurs cognitifs localisés dans différentes étapes du traitement. Nous nous focalisons sur le rôle de la périphéricité acoustique, de la dégradation auditive et des représentations mentales permanentes. Des données sur la perception visuelle complètent la partie principale de notre travail. Nous envisageons aussi la variabilité inter- et intra-sujet dans la perception et dans la production des voyelles. Nos données révèlent que les effets d'ordre sont un mécanisme complexe qui ne peut pas être réduit à un seul facteur déclencheur.

Mots-clé : effet d'ordre ; asymétrie ; discriminabilité ; voyelles ; périphéricité ; neutralisation ; traces auditives ; mémoire à court- et à long-terme ; représentations mentales

It has been known for 150 years that the order of presentation of stimuli has an impact on their discriminability. In this thesis, we have examined the phenomenon of asymmetries in vowel perception and how it interacts with various cognitive factors located at different steps of the stimulus decoding. We have focussed on the role of acoustic peripherality, of auditory decay and of permanent mental representations. Data from visual perception complements the main body of our work. Inter-subject variability in vowel perception and production is also dealt with. Our data shows that order effects are a complex mechanism which cannot be reduced to a single trigger factor.

Keywords : order effect ; asymmetry ; discriminability ; vowels ; peripherality ; neutralisation ; auditory traces ; short- and long-term memory ; mental representations