

Discrimination de sons vocaliques isolés français, de durée variable, par des non-natifs

Véronique Delplancq *, Bernard Harmegnies **,

* Instituto Politécnico de Viseu, Escola Superior de Educação

** Université de Mons-Hainaut, Institut de Linguistique

1. Introduction

Les exemples de recherches centrées sur le caractère catégoriel de la perception sont très nombreux depuis les années 60. Classiquement, les travaux montrent que les adultes non-natifs témoignent de difficultés sensibles à maîtriser des distinctions qui n'ont pas un caractère contrastif dans leur langue d'origine. Cette performance est cependant variable en fonction d'effets systématiques d'expériences. Pour Kuhl (1992), la catégorisation des sons de la parole est strictement liée à la L1. Selon le modèle de l'assimilation perceptuelle (Best, 1994), le contraste de la L2 est assimilé à une catégorie de la L1. Même dans le cas d'une distinction, l'auditeur peut opérer une catégorisation différente de ce que ferait l'auditeur natif, si le phonème paraît acoustiquement différent dans les deux langues considérées. Flege *et al.* (1999) pensent que l'interaction entre les systèmes de la L1 et de la L2 ont une incidence certaine sur les effets liés à l'âge, interaction variable en fonction de l'exposition à la L2. La production des phonèmes de la L2 apparaît donc très fortement dépendante de la perception.

La question de savoir comment et à quel point la L1 interagit avec l'apprentissage de la L2 a été maintes fois posée et discutée pour plusieurs langues. Llisterri et Poch (1986) étudient les interférences du castillan (L2) et du catalan (L1) dans l'acquisition du timbre vocalique du français (L3). Ils constatent que la fausse interprétation des sons de la L1 est conditionnée par l'interférence de la L1 et d'autres interférences par la proximité de la L2. Les résultats permettent de dresser une liste du système de fautes des bilingues, base de propositions de correction phonétique. L'article de Mc Allister (2001) discute de l'effet de l'expérience dans l'emploi d'une L2 en tant que facteur favorisant l'acquisition de la L2. Dans une étude sur l'acquisition du contraste en suédois, le chercheur observe une grande variabilité intersujet dans l'habileté à produire et à percevoir les phonèmes de la L2. Dans le projet «Parole», l'équipe de l'INRIA (Bonneau *et al.*, 2002) se donne pour objectif d'approfondir les connaissances sur les indices essentiels à l'identification des sons de la parole ainsi qu'aux mécanismes de perception. Le renforcement d'indices forts, c'est-à-dire très discriminants d'un point de

vue phonétique et bien marqués acoustiquement, permet aux apprenants de mieux percevoir les sons de la L2 et d'ainsi de mieux les assimiler. Eu égard à l'accroissement de la mobilité des personnes et de l'usage des télécommunications, Hazan (2002) s'interroge sur les outils à développer pour améliorer l'apprentissage de la langue orale. Elle rappelle les différences entre l'apprentissage précoce et l'apprentissage tardif et conclut que les problèmes de perception et de production de sons d'une LE (langue étrangère) ne se résolvent pas spontanément même après une longue immersion. La perception s'améliore nettement au cours d'un programme d'entraînement qui favorise l'acquisition des contrastes entre les sons de la L2 difficiles à acquérir. Baqué *et al.* (2003) s'intéressent à la perception et au traitement de la matière phonique d'une langue romane abordée sans la médiation de l'écrit. Un apprenant peut se réapproprier de façon erronée le système de la LE à cause du fait qu'une même matière phonique peut correspondre à deux catégories phonologiques différentes dans les deux langues. Kim (2004) tente d'établir le système d'interférence entre le coréen et le français. Une des conclusions est que, malgré la bonne compréhension d'un énoncé produit par des étudiants coréens, le francophone repère les particularités vocaliques de durée et de timbre. Une diminution forte de durée vocalique entraîne la non compréhension de la production. L'auteur préconise un travail d'apprentissage sur la durée des syllabes accentuées et inaccentuées. Kalliorinne (2004), se basant sur les écrits de Best (1994), constate que, dans des expériences d'identification des voyelles postérieures françaises par des apprenants finnophones, l'auditeur rencontre d'énormes difficultés à distinguer des phonèmes distincts en L2 mais appartenant à une seule catégorie de la L1. L'analyse des effets de l'expérience linguistique de locuteurs du portugais européen et brésilien et de Japonophones sur leur perception de diverses LE est menée par Parlato-Oliveira en 2005. Kamiyama travaille en 2006 la production et la perception de voyelles orales françaises par les japonophones. Certains auteurs se posent des questions plus spécifiques comme par exemple celle de l'effet facilitateur de l'expérience précoce en langue sur l'acquisition de phonèmes non-natifs à l'âge adulte (Iverson *et al.* par exemple, 2003).

Cependant, le rôle des traits segmentaux est souvent souligné, occultant ainsi le rôle des facteurs temporels. Les paradigmes expérimentaux courants dans les recherches conduisant à des essais de modélisation des stratégies perceptives sont l'identification (le plus usité), la discrimination, l'évaluation des similarités des langues et la recherche de l'intrus. Dans les recherches les plus classiques, les paradigmes utilisés conduisent à convier un grand nombre d'auditeurs à identifier une suite de stimuli sonores généralement de durée brève. Une phase d'apprentissage est parfois proposée dans le but d'habituer le sujet au matériel acoustique et au plan expérimental. Les scores d'identification dépendent de la durée de l'extrait et de la familiarité des sujets avec les langues testées.

Les langues romanes ont déjà fait l'objet de démarches visant à mettre en évidence des stratégies perceptives. Vasilescu (2000, 2002) conclut que la familiarité qu'a le sujet avec la langue test et la pertinence de traits linguistiques forts jouent un rôle primordial. Il pose l'hypothèse que la complexité des systèmes vocaliques de langues romanes est un élément significatif pour la poursuite des recherches. La situation en portugais

européen est compliquée par le caractère évanescent de certaines réalisations vocaliques sous l'effet de l'accent. Delgado-Martins (1988) établit les valeurs formantiques des voyelles orales du portugais européen. Le système vocalique oral portugais s'avère de structure complexe de par sa division en trois sous-systèmes créés par la valeur phonologique de l'accent et de par la présence de réalisations difficilement observables, conséquence du phénomène de réduction vocalique. A partir de 2003, Vaz *et al.* observent quelles sont les voyelles du portugais européen problématiques pour les hispanophones.

Depuis plusieurs années, nos recherches ont comme base commune l'évaluation de l'évolution des structures vocaliques à l'oral (Delplancq, Harmegnies, 1999; Delplancq, Harmegnies, 2002; Delplancq, Harmegnies, 2005; Harmegnies, Delplancq, 2005). Partant ici du constat que le portugais est caractérisé par un phénomène très marqué de réduction vocalique sous l'effet de l'accentuation (Delgado-Martins *et al.*, 1996; Delplancq, Harmegnies, 1997), nous nous interrogeons sur le développement, par l'auditeur lusophone, de capacités perceptuelles spécifiques, propres à compenser la durée vocalique très réduite. La présente étude a pour objectif principal de contraster les possibilités perceptuelles d'auditeurs lusophones, en fonction de leur maîtrise langagière, conviés à participer à une tâche d'identification de stimuli vocaliques de la langue française émis par un francophone, avec des durées variables. La question est de savoir dans quelle mesure l'expertise de la langue favorise ou non la reconnaissance de timbres vocaliques sur base d'un segment de signal très court.

2. Expérimentation

2.1 Échantillon d'auditeurs

Quarante et un auditeurs se sont prêtés à cette expérimentation. Tous sont lusophones de Viseu ou de sa région (Beira Alta) et étudiants de l'École Supérieure d'Éducation de Viseu (Portugal). Les recueils de données ont été menés durant le 2^{ème} semestre de l'année académique 2004/2005. A cette époque, 13 étudiants fréquentaient la 4^{ème} année de la licence en vue de devenir professeurs de 2^{ème} cycle en portugais/français. Les 28 autres auditeurs étaient inscrits en 2^{ème} année de la licence en Communication Sociale (2^{ème} semestre de l'année académique 2005/2006).

Sous la responsabilité du professeur de FLE, chaque étudiant a rempli une fiche d'identification personnelle relative à son contact avec la langue française. Il en ressort que 3 des sujets sont nés en France et y ont résidé entre 18 et 20 ans avant de rentrer au Portugal; 2 autres sujets ont vécu plusieurs années (5 et 6 ans) dans un pays francophone (France et Suisse). Les autres sujets ont effectué toutes leurs études au Portugal et n'ont eu de contact avec le français que dans le cadre de leur cursus et, rarement, lors de séjours touristiques brefs en France et/ou dans le cadre familial et/ou amical.

En marge de cette enquête, une évaluation du niveau de maîtrise de la langue a également été menée. Les diverses compétences en français (compréhension et

expression) de chaque sujet ont ainsi été classées selon les niveaux du *Cadre européen commun de référence pour les langues* (2001) par un panel de 3 experts.

2.2 Stimuli

Les voyelles à l'étude sont au nombre de 6, à savoir [i, e, ε, u, o, ɔ]. Elles font partie du système accentuel de l'inventaire des voyelles du portugais européen (Delgado-Martins, 1988; Mateus & Andrade, 2000; Delplancq, Harmegnies, 1997).

Ces voyelles ont été choisies parce que:

- elles existent aussi bien dans l'inventaire phonologique du français que dans celui du portugais;
- elles présentent, dans le sous-système accentuel, des oppositions similaires dans la série des antérieures (i, e, ε) et dans celle des postérieures (u, o, ɔ);
- elles ont des degrés de présence variables dans les différents sous-systèmes.

Les stimuli ont été traités selon les étapes suivantes:

- enregistrement digital d'un échantillon de chacune des 6 voyelles produites en isolation;
- normalisation d'amplitude par le logiciel Audacity;
- création des fichiers tests.

Pour cette expérience, nous avons retenu les durées suivantes: 20 ms, 40 ms, 60 ms, c'est-à-dire que nous avons eu recours à des sons tests de durée comparable aux plus brèves durées observables en parole spontanée (consulter par exemple Gendrot *et al.*). Néanmoins, pour habituer le sujet à la tâche, et dans un but de calibration, nous avons également créé des sons tests d'une durée de 150 ms.

2.3 Recueil de données

Il résulte de ces décisions que 24 sons tests (6 voyelles et 4 durées) doivent être présentés aux sujets. Nous les avons disposés dans une série dont

- les 6 premiers éléments sont les sons à 150 ms, en ordre aléatoire;
- les 18 autres sons sont, ensuite, présentés aléatoirement.

Par ailleurs, pour étudier la stabilité des comportements des sujets, nous leur proposons la même tâche à six reprises différentes. Le sujet n'est pas averti du fait qu'il doit effectuer la même tâche à plusieurs reprises. Son travail est divisé en trois phases, avec une possibilité de repos entre les phases. Chaque étudiant est installé devant un ordinateur et est autonome dans la procédure d'audition et d'identification des sons.

Plus de 50000 données ont ainsi été encodées en excel.

3. Résultats

3.1 Scores d'identification

Dans un premier temps, nous avons procédé à une étude des scores de reconnaissance correcte des stimuli par les locuteurs.

DISCRIMINATION DE SONS VOCALIQUES ISOLÉS FRANÇAIS, DE DURÉE VARIABLE

		N_sum						Total
		[i]	[e]	[ɛ]	[u]	[o]	[ɔ]	
N_rps	[i]	879	116	7	7	2	2	1013
	[e]	49	568	187	24	16	14	858
	[ɛ]	26	276	759	18	12	11	1102
	[u]	17	9	9	221	129	25	410
	[o]	10	10	9	604	602	341	1576
	[ɔ]	3	5	13	110	223	591	945
Total		984	984	984	984	984	984	5904

Tableau 1: Nombre de réponses données (colonne) en fonction du stimulus (ligne), tous les locuteurs confondus

[i] est la voyelle la mieux reconnue (sur 984 sons proposés pour tous les sujets, 879 ont été correctement reconnus) tandis que le [u] est la moins bien identifiée: moins du quart des réponses sont correctes (sur 984 [u] produits, l'identification correcte est réalisée dans 221 des cas).

La figure 1 montre ainsi les pourcentages de succès en fonction des degrés d'aperture et d'antériorité du stimulus. Elle montre que le succès de l'identification est d'autant meilleur que le stimulus est antérieur; le degré d'aperture n'apparaît pas ici avoir un rôle critique.

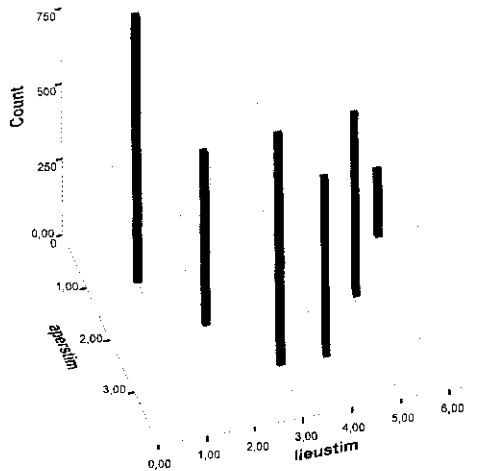


Figure 1: nombres d'identifications correctes des voyelles en fonction du lieu d'articulation (de 1: antérieur à 6: postérieur) et de l'aperture (de 1: fermé à 3: mi-ouvert)

Tous locuteurs confondus, la confirmation est ainsi donnée que [i] est toujours bien identifiée tandis que le [u] est la voyelle la moins bien reconnue.

L'examen des matrices de confusion correspondant à chacune des durées de stimulus confirme globalement les constats établis au départ de la matrice globale (tableau 2). On note que le score de reconnaissance le plus faible pour [e] et [o] s'observe pour les durées de 20 ms, le [o] se reconnaissant de mieux en mieux lorsque la durée augmente. Cette tendance est, par contre, complètement inversée pour [u] qui est d'autant mieux identifié qu'il est court.

dur		[i]	[e]	[ɛ]	[u]	[o]	[ɔ]	Total	
20	N_rps	[i]	193	30	1	4	1	0	229
		[e]	31	103	51	16	5	10	216
		[ɛ]	11	99	179	5	7	4	305
		[u]	7	6	4	88	8	18	131
		[o]	4	5	3	102	80	65	259
	[ɔ]	0	3	8	31	145	149	336	
	Total	246	246	246	246	246	246	1476	
40	N_rps	[i]	224	28	5	0	0	0	257
		[e]	12	157	54	6	7	1	237
		[ɛ]	6	55	184	4	1	4	254
		[u]	3	3	1	57	66	3	133
		[o]	0	3	1	151	147	110	412
	[ɔ]	1	0	1	28	25	128	183	
	Total	246	246	246	246	246	246	1476	
60	N_rps	[i]	230	32	1	0	1	1	265
		[e]	5	158	36	2	2	2	205
		[ɛ]	4	55	201	6	2	2	270
		[u]	4	0	3	45	34	3	89
		[o]	1	0	3	168	180	71	423
	[ɔ]	2	1	2	25	27	167	224	
	Total	246	246	246	246	246	246	1476	
150	N_rps	[i]	232	26	0	3	0	1	262
		[e]	1	150	46	0	2	1	200
		[ɛ]	5	67	195	3	2	1	273
		[u]	3	0	1	31	21	1	57
		[o]	5	2	2	183	195	95	482
	[ɔ]	0	1	2	26	26	147	202	
	Total	246	246	246	246	246	246	1476	

Tableau 2: Synthèse du nombre de réponses données (colonne) pour les stimuli proposés (ligne) compte tenu de leurs durées (dur), tous locuteurs confondus

[u] est bien la voyelle la moins bien identifiée, et l'est de moins en moins au fur et à mesure que la durée du stimulus augmente. Cette voyelle est majoritairement confondue avec [o]. Le taux de confusion avec ce dernier son vocalique augmente avec la durée.

3.2 Distance stimulus-réponse

3.2.1 Méthode

L'analyse des scores de reconnaissance limite cependant les observations au constat binaire de succès ou de l'échec de la tâche d'identification, compte non tenu de l'ampleur de l'éventuelle erreur d'identification. Afin de pallier cette limitation, nous avons mis au point une procédure d'objectivation basée sur le calcul d'une distance non paramétrique entre stimulus et réponse dans l'espace bidimensionnel discret formé par les axes d'aperture et de lieu, sur chacun desquels les écarts entre, respectivement, types de lieu consécutifs et types d'aperture consécutifs sont définis unitaires. Compte tenu du fait que l'expérience se centre sur six voyelles qui relèvent de 3 degrés d'aperture et de 6 lieux d'articulation, on obtient l'espace discret représenté à la figure 2. La distance entre stimulus et réponse peut y être définie comme:

$$\text{Distance} = [(\text{différence d'aperture})^2 + (\text{différence de lieu})^2]^{1/2}$$

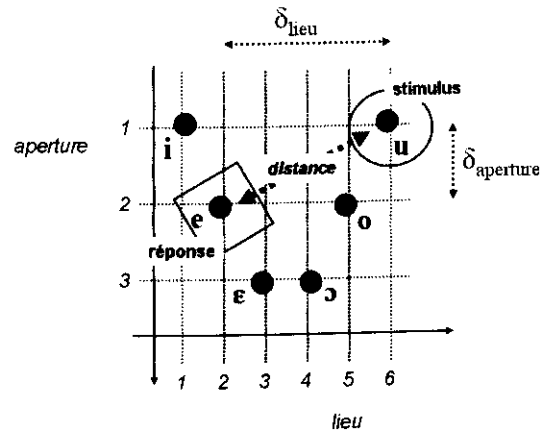


Figure 2: représentation graphique de la distance euclidienne entre les coordonnées du stimulus à l'écoute et celles de la réponse donnée

L'utilisation de la distance permet d'apprécier plus finement l'efficacité des sujets dans la tâche d'identification. Les statistiques obtenues sont donc ici non seulement influencées par le nombre d'erreurs (comme c'était le cas dans l'analyse précédente) mais aussi par leur ampleur: la distance moyenne est d'autant plus faible que les sujets identifient correctement les voyelles qui leur sont présentées.

Un autre intérêt du recours à cette distance tient dans le fait que des procédures inférentielles peuvent lui être appliquées, et ainsi permettre de nouvelles lectures des comportements des sujets.

3.2.1 Effets du lieu d'articulation et de l'aperture sur la distance

Ainsi, nous avons soumis à une analyse de variance à une dimension la distance considérée comme variable dépendante. La procédure a d'abord été appliquée avec le lieu comme variable indépendante, et ensuite avec l'aperture comme variable indépendante.

On constate, ainsi que l'avait suggéré l'étude des scores de reconnaissance, que les performances sont d'autant meilleures que les stimuli sont antérieurs ($F=211, \alpha < .001$). Comme le montre la figure 3, les différences sont particulièrement sensibles aux extrémités de l'axe de lieu d'articulation.

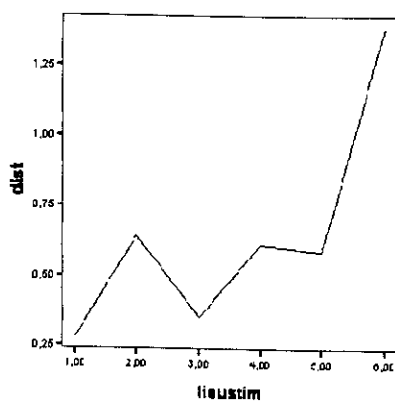


Figure 3: distance euclidienne («dist») entre les coordonnées du stimulus à l'écoute et celles de la réponse donnée en fonction du lieu d'articulation du stimulus («lieustim», de 1: antérieur à 6: postérieur)

DISCRIMINATION DE SONS VOCALIQUES ISOLÉS FRANÇAIS, DE DURÉE VARIABLE

Appliquée aux mêmes données, mais cette fois avec l'aperture comme variable indépendante, l'analyse de variance révèle un effet, certes ténu en ampleur, mais néanmoins significatif ($F=75,428$, $\alpha < .001$).

Cet effet, qui n'avait pu être mis en évidence par la simple analyse des scores de reconnaissance, prend la forme d'une légère diminution de la distance à mesure que s'accroît l'aperture, ainsi que le montre la figure 4.

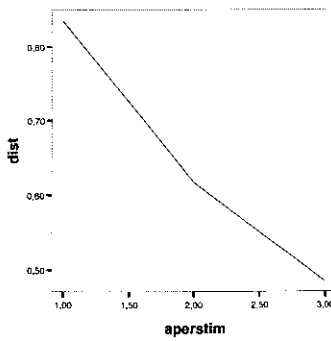


Figure 4: distance euclidienne entre les coordonnées du stimulus à l'écoute et celles de la réponse donnée en fonction de l'aperture (de 1: fermé à 3: mi-ouvert) du stimulus

3.2.2 Effets de l'expertise langagière et de la durée du stimulus

Nous avons procédé à une analyse de variance à deux dimensions croisées visant à apprécier les effets de la *durée du stimulus* et celui de l'*expertise langagière*. Pour contrôler l'effet de la durée, les valeurs ont simplement été réparties en 4 classes mutuellement exclusives correspondant aux durées des stimuli (20 ms, 40 ms, 60ms et 150 ms). Pour le contrôle de la variable d'expertise, deux groupes ont été constitués. Dans le groupe "0", les sujets (au nombre de 36) n'ont eu de contacts avec le français qu'à l'occasion de leur cursus académique, alors que dans le groupe "1", les sujets (au nombre de 5) ont passé une période importante en immersion en pays francophone. Comme on pouvait s'y attendre, la variable durée exerce un effet significatif sur la distance entre le stimulus réel et le stimulus identifié ($F = 24,465$, $\alpha = .000$). On constate, par ailleurs, un effet faiblement significatif de la variable d'expertise langagière ($F = 3,418$, $\alpha = .065$) mais pas de l'interaction expertise/durée ($F = 1,619$, $\alpha = .183$).

Comme le montre la figure 5, la distance tend à décroître au fur et à mesure que croît la durée; en d'autres termes, l'identification est d'autant meilleure que la durée est importante. Par ailleurs, on constate une tendance ténue du groupe d'experts à présenter

des valeurs de distance légèrement supérieures à celles constatées dans l'autre groupe (valable pour toutes les durées sauf 60 ms). Curieusement, les sujets considérés comme experts de la langue française rencontrent donc plus de problèmes à identifier correctement les stimuli.

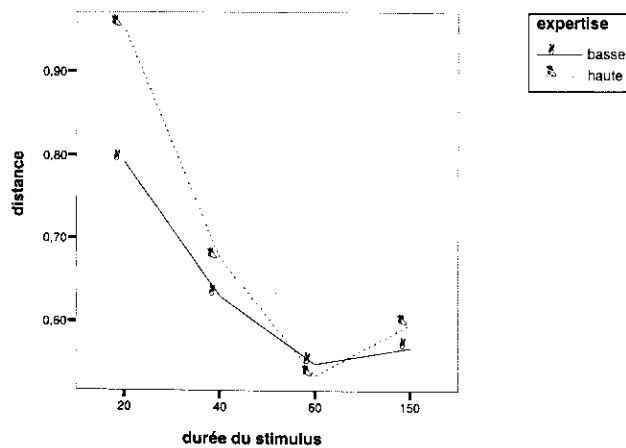


Figure 5: moyenne des distances obtenues (en ordonnée), toutes voyelles confondues, en fonction des durées des stimuli (en abscisse) et de l'expertise langagière du sujet sur base de son histoire linguistique

Etant donné le caractère quelque peu étonnant du constat d'une supériorité dans l'identification des sujets à plus faible expertise du français, et vu le caractère faiblement significatif de l'effet révélé par l'analyse de variances, nous avons décidé de réitérer l'analyse en nous basant cette fois non plus sur les informations objectives concernant l'histoire linguistique des sujets, mais bien sur les évaluations subjectives effectuées par les experts. Parmi les différentes informations disponibles, nous avons retenu le classement des sujets par les experts en les macro-catégories A, B et C du *Cadre européen commun de référence pour les langues*, sur base du critère de *compréhension à l'oral*.

On obtient ainsi une répartition en groupes d'effectifs respectivement égaux à 7 (A), 28 (B) et 6 (C): la partition est ici non seulement plus fine, mais aussi mieux équilibrée (même si la catégorie médiane accueille plus de la moitié des observations).

DISCRIMINATION DE SONS VOCALIQUES ISOLÉS FRANÇAIS, DE DURÉE VARIABLE

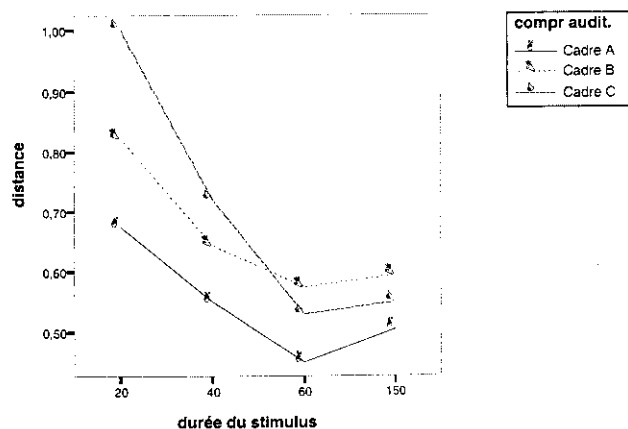


Figure 6: moyenne des distances obtenues (en ordonnée), toutes voyelles confondues, en fonction des durées des stimuli (en abscisse) et de l'évaluation de la compétence du sujet en compréhension orale

Comme le montre la figure 6, l'effet de la durée se confirme. En marge de celui-ci apparaît de plus une action très nette du niveau de compétence en compréhension orale sur la distance stimulus-réponse; l'effet est significatif ($F = 8,142, \alpha < .001$), mais on n'observe pas d'interaction significative entre les deux variables indépendantes ($F = 1,501, \alpha = .173$). L'analyse confirme donc la supériorité des sujets à faible compétence dans la tâche d'identification.

4. Conclusions

Ces premiers résultats obtenus grâce au calcul d'une distance dans le plan lieu/aperture pour un échantillon de sujets significatifs (41 étudiants lusophones) contribuent à remettre en question les relations existant entre l'identification correcte de sons vocaliques, la variation en durées de ces stimuli et la maîtrise de la langue française.

Les tendances observées lors d'une expérience préliminaire (Harmegnies, Delplancq, 2005) sont ici confirmées: les auditeurs classés comme étant des experts de la langue française selon le Cadre de Référence du Conseil de l'Europe ne sont pas ceux qui accomplissent nécessairement le mieux les tâches d'identification, surtout lorsque

les durées des stimuli vocaliques s'éloignent des durées généralement compatibles avec des sons de la parole.

Cette 1^{ère} expérimentation soulève plusieurs interrogations qui motivent la poursuite de la recherche. Le débat est ouvert quant aux multiples interprétations que ces résultats suscitent: faut-il évoquer un manque d'attention dans certaines phases de l'expérimentation chez les sujets aux compétences proches de celles des natifs ou s'agit-il plutôt d'un problème de non reconnaissance de sons beaucoup trop brefs que pour être perçus comme sons de la parole ? Des analyses supplémentaires des données ainsi recueillies ainsi qu'un traitement de l'information basé sur un échantillon d'auditeurs plus diversifié sont nécessaires pour cerner la problématique de l'influence de l'expertise de la langue sur la reconnaissance de timbres vocaliques sur base d'un segment de signal très court.

Les perspectives sont multiples, notamment au niveau pédagogique, la recherche contribuant aux réflexions sur les sources d'interférences à l'audition de sons de la parole chez les bilingues et à la proposition d'exercices ciblés de correction phonétique.

5. Références

- Baqué, L., Le Besnerais M., Masperi, M. (2003) Entraînement à la compréhension orale des langues romanes: quelques repères pour faciliter la prise en compte de la matière phonique. *LIDIL*, 28, Grenoble, pp. 137-152;
- Best, C. T. (1994) The emergence of native-language phonological influences in infants: a perceptual assimilation model. In *The development of speech perception: the transition from speech sounds to spoken words*. Cambridge: MIT press, pp. 167-224;
- Bonneau, A. & al. (2002) *Analyse de la parole. Projet parole, Rapport d'activités*. Inria; *Cadre européen commun de référence pour les langues* (1988), Paris, Conseil de l'Europe / Les éditions Didier;
- Delgado Martins, M.R. (1988) *Ouvrir Falar; Introdução à Fonética do Português*. Lisboa: Caminho, Coleção Universitária, Linguística;
- Delgado-Martins, M. R., Harmegnies, B., Poch-Olivé, D. (1996) Mudança fonética em curso no português europeu. In Delgado-Martins M. R., *Fonética do Português, trinta anos de investigação*, Lisboa: Caminho, pp. 283-293;
- Delgado-Martins, M. R., Mira Mateus, M. H., Harmegnies, B., Poch-Olivé, D. (1996) Fenómenos de reestruturação silábica em português de Lisboa. In Delgado-Martins M. R., *Fonética do Português, trinta anos de investigação*, Lisboa: Caminho, pp. 301-318;
- Delplancq, V., Harmegnies, B. (1997) Les phonétiques et les phonologies du vocalisme portugais. *Paroles* 1, pp. 41-58;
- Delplancq, V., Harmegnies, B. (1999) Regards sur la formation des professeurs de FLE à la pratique de l'oral. In *Cahiers du Centre Interdisciplinaire des Sciences du Langage* 14, pp. 35-52;
- Delplancq, V., Harmegnies, B. (2002) Évolution des structures de l'oral en formation de formateurs de FLE. In *Actes des Journées d'étude sur la parole*, Nancy, pp. 181-184;
- Delplancq, V., Harmegnies, B. (2005) Évolution des rapports de durée vocalique dans l'apprentissage de la langue seconde; restructurations cognitives dans l'apprentissage du français par les lusophones. In *Actes du 1er Colloque International de Didactique Cognitive (Didgog)*, Université de Toulouse-Le-Mirail (France), pp 120-125;

- Flege, J.E. & al. (1994), Factors affecting strength of perceived foreign accent in a second language. In *JASA* 97, pp. 3125-3134;
- Gendrot, C. & al. (2005), Impact of duration on F1/F2 formant values of oral vowels: an automatic analysis of large broadcast news corpora in French and German. In *Actes Eurospeech*. Lisboa, pp. 2453-2456;
- Harmegnies, B., Delplanq, V. (2005), Expérience d'identification de voyelles en isolation par des Lusophones. In *Actes de la XXIème Rencontre de l'Association Portugaise de Linguistique*. Porto, pp. 435-443;
- Hazan, V. (2002), Apprentissages des langues. In *Actes des Journées d'étude sur la parole*. Nancy, pp. 225-229;
- Iverson, P. & al. (2003), A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*. 87, pp. 47-57;
- Kalliorinne, V. (2004) *Perception des voyelles postérieures françaises par les apprenants finnophones*. Université de Turku;
- Kamiyama, T. (2006), La production et la perception des voyelles orales françaises par les apprenants japonophones. In *Actes des Journées d'étude sur la parole*. Dinard, pp. 333-336;
- Kim, H.Z. (2004) *Perception du français prononcé par des étudiants coréens*. Université de Dankook;
- Kuhl, P. (1992) Speech prototypes: studies on the nature, function, ontogeny and phylogeny of the centers of speech categories. In: Tohkura, Y. et al. (eds) *Speech perception, production and linguistic structure*. Tokyo: Ohmsha. pp. 239-264;
- Listerri, J., Poch, D. (1986) Influence de la L1 (catalan) et de la L2 (castillan) sur l'apprentissage du système phonologique d'une troisième langue (français). In *Actes de las IX Jornadas Pedagógicas sobre la Enseñanza del Francés en España*. Barcelona, pp. 153-167;
- Mateus, M. H., Andrade, E. (2000) *The phonology of Portuguese*. Oxford: Oxford University Press;
- Mc Allister, R. (2001) Experience as a factor in l2 phonological acquisition. In *Working papers*. Lund University, pp. 116-119;
- Parlato-Oliveira, E. (2005) *Une étude inter-langue d'un phénomène d'illusion dans la communication verbale; le cas de l'épenthèse vocalique*. Thèse de doctorat, Ecole des Hautes Etudes en Sciences Sociales, Pontificia Universidade Católica de São Paulo;
- Vasilescu, I. & al. (2000) Détermination expérimentale d'indices linguistiques pour la discrimination des langues romanes. In *Actes des Journées d'étude sur la parole*. Aussois, pp. 337-340;
- Vasilescu, I. (2002) Stratégies perceptives en identification des langues. In *Actes des Journées d'étude sur la parole*. pp. 97-100;
- Vaz, A.M. & al. (2003) Quanto mais pior: considerações acerca da aquisição das vogais orais do português europeu por hispanófonos». In *Cadernos de PLE*. 2, Universidade de Aveiro, pp. 155-170.