



HAL
open science

Temps de réaction et identification perceptuelle des langues: étude préparatoire en vue de l'optimisation d'un protocole expérimental en IRMf

Melissa Barkat-Defradas, Jorge Guterrez-Celaya, Samia Belaid

► To cite this version:

Melissa Barkat-Defradas, Jorge Guterrez-Celaya, Samia Belaid. Temps de réaction et identification perceptuelle des langues: étude préparatoire en vue de l'optimisation d'un protocole expérimental en IRMf. 28^e Journées d'Étude sur la Parole (JEP), May 2010, Mons, Belgique. pp.197. halshs-00778675

HAL Id: halshs-00778675

<https://shs.hal.science/halshs-00778675v1>

Submitted on 21 Jan 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.

Temps de réaction et identification perceptuelle des langues : étude préparatoire en vue de l'optimisation d'un protocole expérimental en IRMf

Melissa BARKAT-DEFRADAS, Jorge GUTIERREZ-CELAYA, Samia BELAID

¹ Laboratoire Praxiling UMR 5237 CNRS & Université de Montpellier

melissa.barkat@univ-montp3.fr ; jorge.gutierrez@univ-montp3.fr ; samia.belaid@etu.univ-montp3.fr

ABSTRACT

The aim of this study is to investigate to which extent Arabic listeners can assess some reliable perceptual characteristics from samples of speech of different durations so as to identify the Arabic dialect in which they are produced. Our goal is to determine the minimum duration threshold that is required to achieve correctly a perceptual identification task. Results – analyzed in terms of reaction time (RT), Correct identification rate and errors analysis – are discussed in the field of experimental research and will be used to draw up the experimental design of an ongoing fMRI study dealing with the perceptual treatment of linguistic distance applied to the Arabic dialectal continuum.

Keywords: Perceptual identification, Arabic dialects, Reaction time, Methodology, Experimental design, fMRI.

1. INTRODUCTION

Avant toute expérience, il faut impérativement concevoir un protocole expérimental. Cette étape, primordiale, permet de s'assurer que les obligations légales sont remplies, que le matériel est disponible, que l'on connaît les procédures techniques à utiliser pour confirmer ou infirmer l'hypothèse à tester et que l'on a déterminé les différentes étapes expérimentales pour obtenir des résultats scientifiquement valides. Le protocole expérimental doit également inclure les risques potentiels et envisager d'éventuels moyens de prévention pour, entre autres, répondre aux exigences éthiques auxquelles la recherche est soumise. Cet aspect est d'autant plus important qu'avec le développement de la psychologie expérimentale et des sciences cognitives, on assiste à une profusion de recherches faisant appel à des sujets volontaires naïfs. Ces expérimentations font parfois appel à des protocoles lourds, voire à des techniques d'investigation coûteuses. C'est par exemple le cas des études qui utilisent l'IRMf, technique d'imagerie médicale permettant de visualiser précisément et en temps réel l'activité du cerveau.

Depuis une dizaine d'années, l'IRMf est devenue une technique couramment utilisée pour l'étude du traitement de la parole et du langage [1]. Cette technique – qui repose sur l'observation des

modifications de concentration d'oxygène dans le sang – présente de nombreux avantages par rapport à d'autres techniques plus anciennes. En effet, outre sa bonne résolution spatiale et temporelle, cette technique s'avère totalement inoffensive et les résultats qu'elle permet d'observer ont souvent une portée considérable. Toutefois, l'IRM comporte certains inconvénients liés notamment au dispositif expérimental peu banal qu'elle requiert et aux contraintes qu'elle fait peser sur le sujet qui – bien que volontaire – se trouve le temps de l'expérimentation en bien mauvaise posture du fait du confinement dans le tunnel étroit de l'imageur et de la nécessité absolue d'éviter tout mouvement de la tête pouvant conduire à des activations parasites (Figure 1).



Figure 1 : Imageur Bruker Medspec 30/80 Avance - Aimant à haut champ (3 Tesla) du centre IRMf de l'Hôpital de la Timone à Marseille (photographie empruntée à J-L. Anton).

Aussi, le choix du nombre de séquences d'enregistrements et du nombre d'essais constitue un facteur expérimental non négligeable. En effet, on présuppose qu'au delà de quelques minutes la fatigue et les mouvements du sujet risquent fort de détériorer la qualité des données. Toutefois, la lecture attentive des articles scientifiques relatifs à l'étude du langage par imagerie médicale révèle que cet aspect n'est pas toujours clairement présenté dans la partie méthodologique. C'est pourquoi une réflexion portant sur le temps d'expérimentation nous a semblé intéressante. L'expérience que nous présentons ici a pour objectif d'étudier le temps nécessaire à l'identification perceptuelle de différents parlars arabes

par des locuteurs naïfs. Les résultats observés nous permettront de concevoir au mieux le design expérimental d'une expérience en IRMf actuellement en cours de réalisation¹.

1.1. Identification perceptuelle des langues et des dialectes

La capacité de l'être humain à identifier et à discriminer perceptuellement des langues variées a donné lieu à de nombreux travaux (pour un état des lieux, voir [2]). Des expériences conduites en parole naturelle se sont ainsi intéressées à mesurer les performances d'identification de différentes langues indépendamment des conditions d'écoute et/ou des particularités linguistiques des langues proposées. Lorch et Maera [3] ont par exemple testé la capacité d'auditeurs britanniques à identifier 6 langues inconnues après 20 secondes d'écoute. Muthusamy et al. [4] se sont intéressés à la capacité de sujets monolingues anglophones à identifier 10 langues à partir de stimuli de durées variables (1-6 secondes). Le score moyen d'identification correcte varie en fonction de la durée des stimuli et selon que l'on inclut ou non les performances relatives à l'identification de la langue maternelle (LM). Cette étude démontre également l'effet de l'apprentissage sur l'amélioration des performances des sujets. Les résultats observés dans les deux études précédemment citées portent sur l'identification de langues plus ou moins éloignées. C'est pourquoi, au delà des scores d'identification correcte, l'analyse des erreurs s'avère particulièrement informative pour notre propos. En effet, il est intéressant de noter que les échecs à la tâche ont toujours concerné les langues les plus proches linguistiquement (i.e. mandarin/vietnamien, langues toutes deux tonales et coréen/japonais, langues attestant une large part de vocabulaire commun). Suite à ces travaux, nous avons conçu une série d'expériences dont l'objectif était de tester l'effet des variables 'proximité' vs. 'distance' linguistique sur les performances d'identification. Le continuum dialectal arabe permet de tester de façon rigoureuse ce paramètre puisqu'il existe, sur ce domaine, une corrélation forte entre proximité linguistique et proximité géographique. Nous avons ainsi conçu un paradigme expérimental dont la tâche était d'identifier une série de langues/dialectes² plus ou moins proches

¹ Recherche réalisée en collaboration avec le laboratoire Parole et Langage (UMR 6057 CNRS & Université de Provence) et le centre IRMf de l'Hôpital de la Timone de Marseille, avec le soutien du CNRS.

² En linguistique, il est difficile de distinguer clairement 'langues' et 'dialectes'. On dit généralement que lorsqu'il y a *intercompréhension* entre deux parlers on a affaire à deux 'dialectes' de la même langue. Au contraire si cette intercompréhension n'existe pas, il s'agit alors de deux 'langues' différentes. Cette définition n'est pas pleinement satisfaisante, en particulier lorsque l'on rencontre des

du parler maternel (Barkat et al., 1999 [5] et Barkat-Defradas, 2001 [6]). Une partie de nos résultats est présentée ci-dessous.

Outre des scores d'identification significativement plus élevés pour l'identification de la langue maternelle, nos résultats ont également montré que les sujets obtiennent des scores supérieurs pour l'identification des parlers les plus proches de leur dialecte d'origine. La notion de distance et/ou de proximité linguistique peut donc être considérée comme un critère d'étude pertinent dans la mesure où elle influe fortement sur les performances des sujets.

1.2. Corrélât cérébral lié au traitement perceptuel de la variabilité linguistique

Faisant suite à la caractérisation des parlers arabes aux niveaux segmental (Barkat-Defradas, 2000 [7]), supra-segmental (Hamdi & al., 2005 [8]) et à leur identification perceptuelle à partir de parole naturelle (Barkat-Defradas 2001 [6]) ou synthétisée (Barkat et al., 1999 [5]), notre étude en IRMf a pour objectif d'étudier le corrélât cérébral lié au traitement perceptuel de la variabilité dialectale en fonction de la distance linguistique existant entre le dialecte maternel et les autres parlers³. Pour répondre à cette question, nous entendons concevoir un paradigme expérimental où des sujets arabophones natifs – stimulés auditivement à l'intérieur du scanner – auront à identifier la nature dialectale de différents échantillons de parole. En regard des résultats obtenus en perception, 3 conditions sont envisagées : (i) dialecte maternel, (ii) dialecte proche du dialecte maternel et (iii) dialecte éloigné du dialecte maternel. Pour chaque condition, 5 locuteurs différents seront enregistrés, chaque locuteur produira 6 échantillons de parole différents dans son dialecte maternel. Ceci conduit à un total de 90 échantillons de parole à traiter. On comprend dès lors que la détermination du temps minimal nécessaire pour réaliser la tâche d'identification est un facteur non négligeable lors de

situations de '*chaînes dialectales*', où il peut y avoir intercompréhension entre une zone linguistique A et une autre zone B adjacente, de même entre une zone B et une zone C sans pour autant avoir une intercompréhension entre les locuteurs de la zone A et de la zone C, trop distinctes l'une de l'autre tant du point de vue géographique que linguistique.

³ Le domaine arabophone s'avère particulièrement adapté à l'investigation de cette question. En effet, la plupart des travaux sur les traitements des langues chez le sujet bilingue comparent l'activité cérébrale générée par des langues parfois très proches (par exemple, le catalan et l'espagnol), parfois beaucoup plus éloignées (comme par exemple, le chinois et l'anglais). Il nous semble cependant très important de mieux contrôler ce paramètre pour examiner s'il existe une relation entre la proximité linguistique des langues traitées par le cerveau d'un individu et l'activation des zones cérébrales concernées par le traitement associé à L1 ou L2.

la conception du paradigme expérimental, notamment pour ce qui concerne la réduction du temps d'expérimentation.

2. TEMPS DE REACTION ET IDENTIFICATION DIALECTALE

L'expérience réalisée ici a pour objectif de déterminer le temps minimal nécessaire pour réaliser, avec le moins d'erreurs possibles, la tâche d'identification dialectale qui sera proposée aux sujets lors de l'expérimentation en IRMF. Les résultats observés devant permettre *in fine* l'optimisation de notre design expérimental. En effet, selon que les échantillons de parole présentent une durée moyenne de 1 à 6 secondes, la durée totale d'une session passe de 2,25 à 13,5 minutes en prenant en compte les intervalles inter stimulations (ISI) de durée aléatoirement variable (2-4 secondes) lesquels, couplés à une présentation aléatoire des stimuli, permettent d'améliorer l'amplitude de la réponse cérébrale du sujet en réduisant l'effet d'habituation à la tâche (Table 1).

Table 1 : Evolution du temps d'expérimentation en fonction de la durée des stimuli soumis à identification.

Durée du stimulus (sec)	Nombre d'essais	Durée des stimuli (sec)	Durée ISI (sec)	Durée totale d'une session (min)
1	90	90	45	2.25
2	90	180	90	4.5
3	90	270	135	6.75
4	90	360	180	9
5	90	450	225	11.25
6	90	540	270	13.5

2.1. Matériel, sujets, méthode

Dans le cadre de cette étude préparatoire, nous avons retenu les 3 mêmes conditions linguistiques que celles envisagées pour le protocole IRMF, soit : (i) dialecte maternel des sujets testés (i.e. arabe marocain), (ii) dialecte proche du dialecte maternel (i.e. arabe tunisien) et (iii) dialecte éloigné du dialecte maternel (i.e. arabe syrien). Afin d'éviter que les sujets testés ne s'appuient sur les caractéristiques vocales des sujets enregistrés pour constituer le matériel de stimulation auditive, 3 locuteurs par dialectes ont été enregistrés (i.e. 9 voix différentes) alors qu'ils se livraient à une tâche de narration dans leur dialecte maternel respectif. Chaque enregistrement présentait une durée totale de 10 minutes environ. Les enregistrements numériques (22KHz, 16bits, mono) ont été réalisés en chambre anéchoïque à l'aide d'un micro externe directionnel (microphone dynamique professionnel HQ) placé à une distance d'environ 15 cm de la bouche du locuteur. De ces enregistrements, nous avons extrait des échantillons de parole correspondant à des énoncés complets et dont la durée variait de 1 à 4 secondes (3 échantillons / durée). Le matériel de stimulation est ainsi constitué de 36 échantillons différents. Nous

avons utilisé ces échantillons pour concevoir une expérience perceptuelle d'identification dialectale avec mesure du temps de réaction (TR). L'interface expérimentale a été réalisée sous E-Prime. Les 36 échantillons ont été présentés par trois fois au cours de 3 sessions distinctes (i.e. 108 stimuli). Afin de familiariser les sujets au matériel et à la tâche, deux sessions d'entraînement, conduites à partir de voix (3) et d'échantillons sonores différents (9), ont préalablement été proposées aux sujets. Treize sujets volontaires arabophones natifs, locuteurs d'arabe marocain ont participé à l'étude. La consigne visuellement fournie aux sujets via l'écran d'accueil de l'ordinateur utilisé pour la passation de l'expérience était la suivante : « *vous allez entendre différents extraits de parole en arabe dialectal. Ces extraits sont produits soit en arabe marocain, soit en arabe tunisien, soit en arabe syrien. Vous devez devinez le plus rapidement possible de quel dialecte il s'agit. Si, selon vous, il s'agit d'arabe marocain, taper sur la touche étiquetée 'M', s'il s'agit d'arabe tunisien, taper sur la touche étiquetée 'T', s'il s'agit d'arabe syrien, taper sur la touche étiquetée 'S'.* ». Trois des quatre touches de flèches directionnelles du clavier ont été garnies d'une étiquette autocollante portant les lettres précédemment citées. En référence à la géographie du monde arabe, nous avons attribué à la touche ← la valeur 'M' (le Maroc étant situé à l'Ouest du domaine), à la touche → la valeur 'S' (la Syrie étant située à l'Est du domaine) et enfin, à la flèche ↑ la valeur 'T' (la Tunisie étant située entre le Maroc et la Syrie). Cette motivation symbolique a été clairement explicitée aux participants. L'écoute des différents stimuli – présentés en ordre aléatoire pour chaque sujet et à l'intérieur de chaque session – a été faite via un casque Sennheiser HD580 à un niveau d'écoute confortable (60dB).

2.2. Résultats

Les résultats montrent d'une part que les scores d'identification sont élevés pour l'ensemble des parlars étudiés avec, pour l'arabe marocain (correspondant à la langue maternelle des sujets) 93,16% d'identification correcte, pour l'arabe tunisien, (correspondant au parler proche du dialecte maternel) 85,68% d'identification correcte et, pour l'arabe syrien (dialecte le plus éloigné – et donc linguistiquement le plus différent – du dialecte maternel) 88,03% d'identification correcte. D'autre part, l'analyse des erreurs en regard de la notion de distance vs. proximité linguistique montre que l'identification du dialecte le plus éloigné du dialecte maternel est une tâche moins complexe à réaliser que l'identification du parler le plus proche de la langue maternelle. En effet, les confusions (voire l'absence de réponse) concernent le plus souvent l'identification du tunisien (Table 4).

La difficulté à discriminer le dialecte maternel du parler qui lui est le plus proche transparait également sur le temps de traitement. Le temps de réaction moyen est significativement plus long pour l'identification du

tunisien que pour l'identification du syrien ($p < .05$) (Figure 2).

Table 4. Matrice de confusion (en nombre d'essais / 1404 essais) et en % d'identification.

Stimuli	MA	TU	SY	Pas de réponse
MA	436 (93,1%)	21 (44,8%)	10 (2,1%)	1 (0,21%)
TU	34 (7,2%)	401 (85,6%)	30 (6,1%)	3 (0,64%)
SY	7 (1,4%)	47 (88%)	412 (88%)	2 (0,42%)

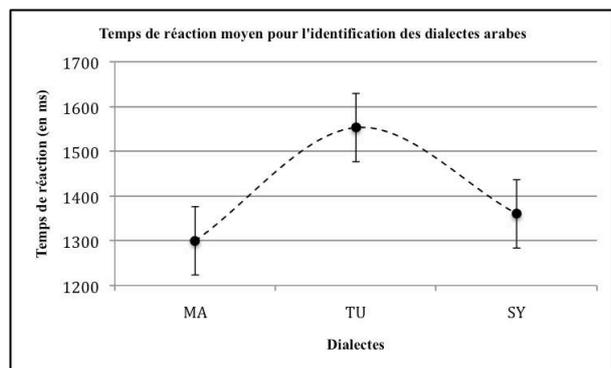


Figure 2. Temps de réaction moyen (en ms) pour l'identification en fonction du dialecte.

Les résultats montrent également un effet de la session sur les performances des sujets : les scores augmentent entre chaque session, révélant ainsi que l'apprentissage se poursuit au cours du temps ($p < .01$) (Figure 3).

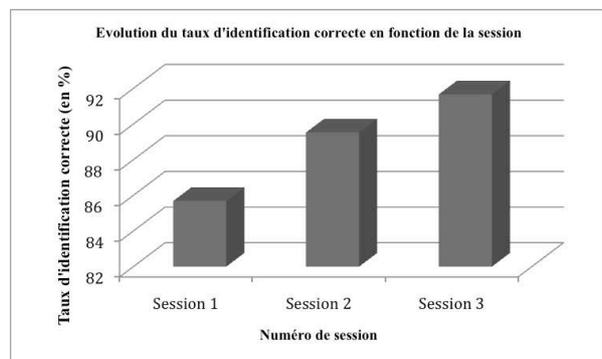


Figure 3. Evolution du taux d'identification (en %) en fonction de la session.

En revanche, contrairement à ce qui est rapporté dans la littérature, les scores d'identification ne sont pas corrélés à la durée des stimuli (coeff. 0.05). Ce résultat inattendu est très probablement lié au fait que nous avons explicitement demandé aux sujets de réaliser la tâche d'identification le plus rapidement possible (cf. consigne). Enfin, afin de répondre à l'objectif méthodologique de cette étude, nous avons mis en regard les scores d'identification correcte et la durée moyenne normalisée du temps de réaction (Figure 4). Ces derniers résultats suggèrent l'existence d'un seuil de stimulation minimal à partir duquel il est possible d'observer des résultats d'identification probants ($\approx 1,75$ seconde).

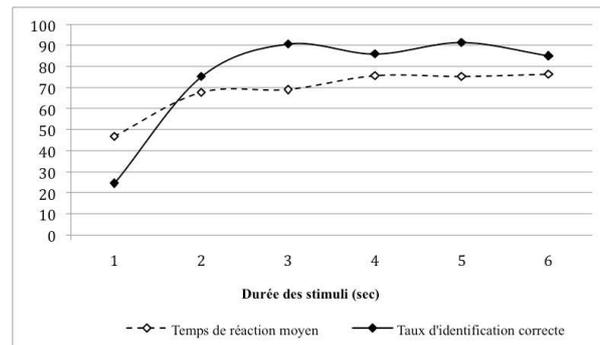


Figure 4. Temps de réaction moyen et taux d'identification correcte en fonction de la durée des stimuli.

3. CONCLUSIONS

Les résultats de cette étude préparatoire nous ont conduits à envisager différents types d'amélioration pour notre protocole en IRMf. Concrètement, nous envisageons d'une part d'intégrer en amont une tâche comportementale d'identification perceptuelle. D'autre part, nous ferons le choix d'intégrer un parler proche du dialecte maternel qui soit suffisamment différent du dialecte d'origine pour ne pas porter à confusion, ce qui pourrait conduire à des chevauchements d'activations entre langue maternel et dialecte proche. Enfin, la détermination d'un point d'équilibre correspondant au temps minimal de stimulation nécessaire à la bonne réalisation de la tâche nous permet d'envisager l'augmentation du nombre de séquences et d'essais sans pour autant augmenter inutilement le temps d'expérimentation en IRMf.

4. REFERENCES

- [1] A., Marchal, Ch., Cavé, *L'imagerie médicale pour l'étude de la parole*, Hermès Lavoisier, 2009.
- [2] M., Barkat-Defradas, I., Vasilescu, I. F., Pellegrino, F., Stratégies perceptuelles et identification automatique des langues: application au continuum dialectal arabe, *Revue Parole*, pp. 1-44, 2003.
- [3] M., Lorch & P., Maera, "How People Listen to Languages They Don't Know", *Language Sciences*. Volume I, Number 4, pp. 343-353, 1989.
- [4] Y., Muthusamy, N., Jain, R., Cole, "Perceptual benchmarks for Automatic Language Identification", in *Proceedings ICASSP*, 1994.
- [5] M., Barkat, J.J., Ohala, F., Pellegrino, "Prosody as a Distinctive Feature for the Recognition of Arabic Dialects." in *Proceedings of Interspeech*, Budapest, 1999.
- [6] M., Barkat, Vers l'identification automatique des parlers arabes, *Revue Internationale 'Langues et Linguistique'*, vol. 7, pp. 47-75, 2001.
- [7] M. Barkat, Détermination d'indices acoustiques robustes pour l'identification automatique des parlers arabes, Editions du Septentrion, Lille, 2000.
- [8] R. Hamdi, R., M., Barkat-Defradas, E., Ferragne, F., Pellegrino, (2004), "Speech Timing and Rhythmic structure in Arabic dialects", in proceeding of 8th ICSLP, pp. 331-334, 2004.