



Rapport de stage de Master 2 Recherche
Spécialité : Signal, image, parole, télécoms

Evaluation perceptive
de l'apport de la vision de la langue
à la compréhension de la parole

par

Yuliya Tarabalka

réalisé au laboratoire GIPSA

Maîtres de stage : Pierre Badin et Gérard Bailly,
équipe MPACIF (Machines Parlantes, Agents
Communicants et Interaction Face à Face)

Juin 2007

Table des matières

Remerciements.....	5
Résumé.....	9
Abstract	10
Introduction	11
Contexte.....	11
Problématique.....	11
Chapitre 1 – Analyse de la littérature	14
1.1 Apport de la vision des composantes typiquement visibles du visage à la lecture faciale	14
1.2 Vision de la langue.....	18
1.2.1 Visualisation de langue par vidéofluoroscopie.....	18
1.2.2 Visualisation de langue en synthèse 3D	19
1.3 Conscience phonologique et articulatoire	20
Chapitre 2 – Elaboration du test perceptif.....	22
2.1 Modèle de langue	22
2.2 Acquisition des données articulatoires : Enregistrements EMA	23
2.3 Détermination des paramètres de contrôle des articulateurs.....	24
2.3.1 Association entre bobines EMA et points du maillage 3D du modèle de langue ...	24
2.3.2 Inversion	25
2.3.3 Détermination des paramètres de contrôle pour les stimuli du test.....	28
2.4 Elaboration du test de perception	29
2.4.1 Corpus.....	29
2.4.2 Structure du test.....	30
2.4.3 Préparation de stimuli.....	31
2.4.4 Logiciel de passation du test.....	31
Chapitre 3 – Analyse des résultats du test de perception	33
3.1 Déroulement du test de perception	33
3.2 Analyse des résultats	33
3.2.1 Résultats globaux.....	33
3.2.2 Résultats détaillés	37
3.2.3 Pourcentage d’Information Relative Transmise.....	43
Conclusions et perspectives	46
Conclusions	46
Applications.....	46
Perspectives	46
Références bibliographiques.....	47
Annexe A : Résultats de l’inversion.....	49
Annexe B : Traitement du corpus EMA enregistré à l’ICP le 18 avril 2003.....	55

<i>Annexe C : Exemples du texte affiché durant le test de perception</i>	<i>56</i>
<i>Annexe D : Exemples de fenêtres de présentation du test de perception</i>	<i>58</i>
<i>Annexe E : Pourcentage d'identification correcte de consonne en différents contextes vocaliques pour les différentes séries du test.....</i>	<i>59</i>

Remerciements

Je souhaite remercier tous ceux qui ont contribué directement ou indirectement à la réalisation de ce projet.

Tout d'abord, je remercie mes parents qui m'ont donné la vie, qui est tellement merveilleuse et passionnante, et qui m'ont aidée à grandir pour que je puisse étudier le monde.

J'exprime ma gratitude à Pierre Badin et à Gérard Bailly qui m'ont accueillie à l'ICP (actuellement : département Parole & Cognition de GIPSA-lab) et qui m'ont permis de travailler sur un sujet intéressant et enrichissant.

J'adresse un grand merci à Christophe Savariaux pour l'enregistrement des données EMA et à Frédéric Elisei pour sa contribution à la création du programme de test de perception et pour ses nombreux conseils. Je remercie aussi Jean-Luc Schwartz et Amélie Rochet-Capellan pour leur aide à l'analyse statistique des résultats de notre expérience.

Je remercie tous mes enseignants pour leurs cours et leurs nombreuses explications, ainsi que tout le personnel de l'ICP qui m'a permis de travailler dans de bonnes conditions.

Merci à tous ceux qui m'ont aidée à venir en France et à passer le temps dans une ambiance qui m'a donné de l'inspiration pour faire mes recherches.



La science consiste à passer
d'un étonnement à un autre.

Aristote

Résumé

De nombreux travaux ont établi que la vision du visage du locuteur facilite la compréhension de la parole par les humains et montré que les articulateurs typiquement visibles (lèvres, mâchoire, visage, partie antérieure de la langue, dents) contribuent significativement à un meilleur taux de reconnaissance. Nous avons utilisé une tête parlante virtuelle audiovisuelle qui permet d'afficher l'ensemble des articulateurs de la parole, y compris la langue – dans un mode de réalité augmentée – pour déterminer si la vision directe de la forme complète de la langue pouvait être perçue et utilisée par les sujets.

Pour mener une série de tests de perception audiovisuelle, nous avons construit un ensemble de stimuli audiovisuels voyelle – consonne – voyelle en pilotant la tête parlante à partir de paramètres obtenus par inversion d'enregistrements articulatoires d'un sujet. Nous avons mesuré les scores de reconnaissance de la consonne dans les séquences VCV présentées suivant quatre modalités : (1) audio seul, (2) audiovisuel avec intérieur du visage *sans* langue, (3) audiovisuel avec intérieur du visage *avec* langue et (4) audiovisuel avec visage complet de profil. Quatre niveaux de bruit blanc étaient aussi ajoutés au signal audio.

Il apparaît que : (1) toutes les modalités de présentation vidéo améliorent la compréhension de parole par rapport à l'audio seul ; (2) la modalité avec la vision de la langue n'est pas significativement mieux perçue que la modalité avec le visage complet. Ceci suggère que la capacité naturelle des sujets à « lire sur les lèvres et sur le visage » est au moins égale à leur capacité naturelle à « lire la langue ». Il arrive, dans la condition sans audio, que le score avec une langue visible soit supérieur au score avec le visage complet. Ceci suggère que lorsque l'information audio n'est plus suffisante pour aider la lecture labiale, la lecture « linguale » peut prendre le relais si nécessaire. Un effet d'apprentissage implicite de la lecture linguale semble avoir eu lieu. Des tests complémentaires sont nécessaires pour montrer qu'un apprentissage rapide peut être fait.

Mots-clés : Têtes parlantes, mesures articulatoires, inversion articulatoire, réalité augmentée, perception audiovisuelle, modélisation et animation de la langue

Abstract

Numerous researches have established that the vision of the speaker's face eases the speech comprehension. They have shown that the typically visible articulators (lips, jaw, face, back of the tongue, teeth) contribute significantly to better recognition rates. The audiovisual virtual talking head was used to display the speech organs, tongue included, and to determine if the direct vision of the complete tongue shape can be perceived and used by the subjects.

To conduct the audiovisual perception tests, a set of audiovisual vowel – consonant – vowel stimuli was constructed by driving the talking head. The control parameters were obtained by inverting articulatory recordings of the subject. We have measured the consonant recognition scores of the VCV sequences, presented in four modalities: (1) audio only, (2) audiovisual with the interior of the head but *without* tongue, (3) audiovisual with the interior of the head *with* tongue, (4) audiovisual with full-face in profile. Four levels of noise were added to the audio signal.

It appears that: (1) all modalities of video presentation improve the speech comprehension when compared with audio only, (2) the modality with visible tongue was not significantly more perceptible than the modality with the full-face. This suggests that the natural capacity of subjects to “read on the lips” is at least equal to their natural capacity to “read on the tongue”. With no audio, the score with the visible tongue is sometimes superior to the score with the full-face. This suggests that when the audio information isn't sufficient to complete the lipreading, the “lingual” reading can add intelligibility.

An effect of the implicit training of the “lingual” reading seems to appear. Supplementary tests are necessary to verify that a fast training can be achieved.

Key words: Talking heads, articulatory measurements, articulatory inversion, augmented reality, audiovisual perception, tongue modelling and animation

Introduction

Contexte

Depuis longtemps, l'ICP (actuellement : département Parole & Cognition de GIPSA-lab) développe des têtes parlantes audiovisuelles virtuelles capables d'afficher les mouvements articulatoires du visage et des lèvres (Le Goff et Benoît, 1996 ; Revéret et al., 2000; cf. exemple de la modélisation de tête parlante : figure 1, 2). A des fins d'évaluation de ces modèles, l'apport de la vision de tout ou partie du visage (lèvres, mâchoire, visage complet, synthétique ou naturel) pour l'intelligibilité de la parole a été largement étudié (Mohamadi, 1993 ; Benoît et Le Goff, 1998).

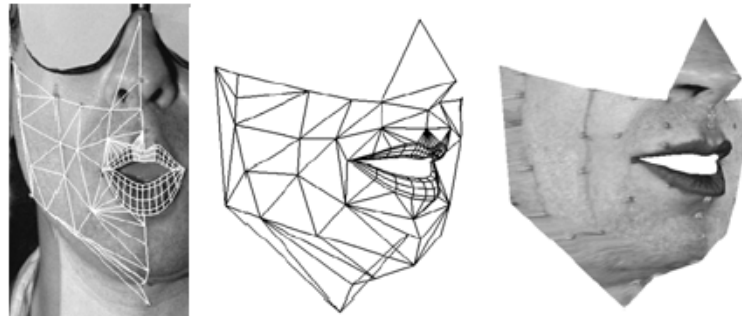


Fig. 1 – Exemple de modélisation de tête parlante

Tout récemment, un modèle de langue a été construit à l'ICP à partir d'IRM tridimensionnelles (Badin et al 2002, Badin & Serrurier, 2006).

Il est donc possible désormais d'utiliser ce modèle (cf. exemple du modèle de visage et de langue: figure 2) pour l'évaluation perceptive de l'apport de la vision de l'organe peu ou pas visible que constitue la langue pour la compréhension de la parole dans le bruit.

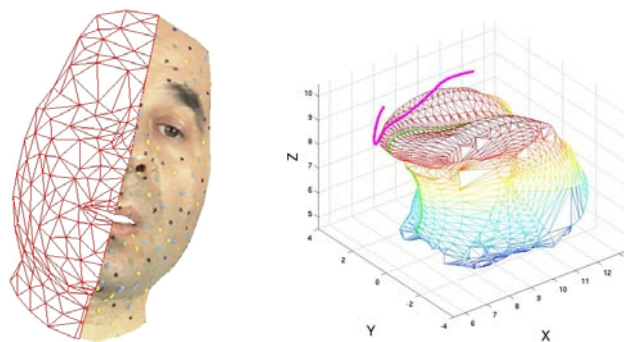


Fig. 2 – Affichage du modèle de visage et de langue

Problématique

Un aspect important du mécanisme de perception humaine est sa capacité à intégrer l'entrée de différents systèmes sensoriels (par exemple, vision, audition, toucher, goût). Les processus de perception de l'environnement par l'être humain sont essentiellement multimodaux, mais le cerveau synthétise l'information de différentes sources pour produire une sensation cohérente.

La parole est un signal audiovisuel généré des mouvements articulatoires et destiné à la communication orale entre humains (Abry et al., 1994). Dodd (1987) constate que “les enfants ne semblent pas traiter séparément la perception des modalités auditive et visuelle”, et conclut que la perception de la parole est un processus bimodal dès le début. Reisberg et al. (1987) postulent que la perception de la parole est un processus « amodal », i.e. l'information des

sources auditive et visuelle converge dans tous les niveaux de l'analyse. Le signal de parole est redondant, aussi bien au niveau du son qu'au niveau de la complémentarité entre le système acoustique et le système visuel. Ceci rend ce signal robuste, ce qui est nécessaire pour un signal de communication. Son degré de redondance est adaptable en fonction des conditions environnementales de bruit et de la quantité d'information contenue dans le message à transmettre (Badin et al., 2002).

La vision du visage du locuteur constitue un apport important pour la perception de la parole par les humains. La visualisation des composantes du visage en cohérence avec l'information auditive facilite la compréhension de parole. En particulier, la vision du visage est essentielle pour la perception de la parole par les malentendants (« lecture labiale »), mais elle est utilisée aussi par les bien-entendants, pour l'amélioration de l'intelligibilité, quand la parole audio est dégradée (Sumbly et Pollack, 1954 ; Erber, 1969 ; MacLeod et Summerfield, 1987 ; Benoît et al., 1994). En outre, la vision des mouvements du visage du locuteur peut améliorer la compréhension de la parole même quand le signal audio est clair et intact (Arnold et Hill, 2001 ; Reisberg et al., 1987). Il est connu que la contribution visuelle est importante pour la compréhension et l'apprentissage d'une langue étrangère (Reisberg et al., 1987 ; Kellerman, 1990).

En revanche, l'incohérence entre l'information auditive et visuelle peut compliquer la compréhension de la parole. L'effet McGurk montre la puissance de l'influence de la parole vidéo sur la perception : par exemple, un signal /ba/ auditif fusionné avec un signal /ga/ visuel est très souvent perçu comme un signal audiovisuel /da/ (McGurk et MacDonald, 1976).

Les visages synthétiques augmentent aussi l'intelligibilité de la parole, quand les mouvements visibles du visage (générés à la base par un modèle 2D ou 3D) sont cohérents et synchronisés avec le signal audio de parole, qu'il soit naturel ou synthétisé (Brooke et Summerfield, 1983 ; Mohamadi, 1993 ; Benoît et al., 1996 ; Benoit et Le Goff, 1998).

Il est intéressant d'évaluer quelles composantes du visage dans la lecture faciale (« speechreading »)¹ améliorent de façon plus ou moins significative la compréhension de la parole. Plusieurs séries de travaux ont été réalisées pour évaluer l'apport des différentes composantes du visage naturel et synthétique telles que les lèvres, les dents, la langue (la partie buccale) et les yeux, le front, les joues, la mâchoire, le nez (la partie non-buccale) (e.g., Summerfield, 1979 ; Ijsseldijk, 1992 ; Tye-Murray et al., 1993 ; Benoît et al., 1996 ; Rosenblum et al., 1996 ; Thomas et Jordan, 2004). Tous les résultats indiquent que la vision de la partie buccale du visage fournit toute ou presque toute l'information visuelle de parole que le visage parlant entier peut apporter. Cependant, il est plus difficile de lire les lèvres seules que le visage entier. En conclusion, la langue et les dents apportent un certain gain d'intelligibilité de parole.

La plupart des investigations ont évalué l'apport de la vision des composantes typiquement visibles du visage. En effet, il est beaucoup plus difficile de réaliser la visualisation des articulateurs typiquement non visibles en cohérence avec la parole audio.

Une publication de Tye-Murray et al. (1993) décrit les résultats des expériences effectuées pour déterminer si l'information articulaire visible fournie par vidéofluoroscopie (donnant une vision de profil des articulateurs typiquement cachés tels que la langue, les dents, la mâchoire et le velum) augmente le taux de reconnaissance de parole. Dans les conditions de cette expérience, les résultats montrent que la vision des articulateurs typiquement non visibles par vidéofluoroscopie n'améliore pas la compréhension de parole.

Notons que la vidéofluoroscopie de profil donne la projection de toute la tête sur un plan sagittal. Tous les organes apparaissent ainsi superposés les uns aux autres, comme la mâchoire sur la langue. Ceci ne facilite pas la vision des articulateurs, et pourrait expliquer l'absence

¹ Il faut distinguer les termes « lecture labiale » (« lipreading ») et « lecture faciale » (« speechreading »). « Lipreading involves the extraction of meaning from movements of the lips, jaw, and tongue. Speechreading is a broader term that involves lipreading together with the interpretation of body language, facial expressions, and linguistic and situational cues » (Arnold, 1997).

d'amélioration apportée par ce signal visuel. La présentation des mouvements des articulateurs à l'aide de modèles permettrait d'éviter ce problème de superposition.

Massaro et Light (2003, 2004) ont évalué l'efficacité de la tête parlante Baldi (qui peut illustrer les articulateurs typiquement invisibles en rendant sa peau transparente) comme outil d'apprentissage pour les malentendants et bien-entendants qui apprennent une langue étrangère. Ils ont conclu que la perception et la production de parole ont été améliorées après l'apprentissage avec Baldi. Mais ils n'ont pas publié la contribution spécifiquement due à la visualisation de mouvements des organes typiquement invisibles.

Le présent projet vise donc à évaluer l'apport perceptif, pour l'intelligibilité de la parole dans le bruit, de l'affichage de la langue en réalité augmentée², à l'aide du modèle 3D de langue construit à l'ICP (cette vision est évidemment plus nette que les images de vidéofluoroscopie).

S'il s'avère que la vision de la langue améliore la compréhension de la parole dans le bruit, nous pourrions envisager de développer des applications pour la correction phonétique (correction des défauts de langage pour les enfants, apprentissage d'articulations en langue étrangère, réhabilitation de la parole des malentendants) et pour la communication en conditions de bruit.

² Par un système de réalité augmentée on entend un système qui rend possible la superposition de l'image d'un modèle virtuel 3D ou 2D sur une image de la réalité

Chapitre 1 – Analyse de la littérature

1.1 Apport de la vision des composantes typiquement visibles du visage à la lecture faciale

Sumby et Pollack (1954) ont examiné la contribution des composantes visuelles à la compréhension de parole en présence de bruit et conclu que plus le rapport Signal / Bruit diminue plus la contribution visuelle devient importante (la contribution visuelle varie de 5 à 23 dB, selon le rapport Signal / Bruit et la taille de vocabulaire employé). Ils proposent d'utiliser l'observation visuelle du locuteur pour une amélioration de perception de parole dans les conditions bruitées. Le protocole de l'expérience et l'index de la contribution visuelle au rehaussement de l'information acoustique proposés par Sumby et Pollack deviennent classiques pour l'estimation de la contribution visuelle à la compréhension de parole.

Les recherches de O'Neill (1954), Neely (1956), Erber (1969), Ewertsen et Birk Nielsen (1971) indiquent aussi que les bien-entendants obtiennent des scores d'intelligibilité de parole (avec le faible rapport Signal / Bruit) plus élevés s'ils perçoivent l'information audiovisuelle qu'avec la seule information auditive.

Dès la fin des années 1970, Summerfield, McGrath et MacLeod ont commencé à étudier l'apport des composants individuels du visage à la perception de messages audio dégradés. Summerfield (1979) donne une première évaluation de l'apport de la vision des lèvres à l'intelligibilité de la parole. Il a analysé cinq conditions : sans vidéo, vidéo complète (vidéo monochrome du visage naturel), lèvres isolées (du même visage), 4 points lumineux (placés dans les centres des lèvres supérieure et inférieure et dans les coins de la bouche du locuteur), cercle (représentant l'enveloppe de l'amplitude du signal audio par la variation de son diamètre). Pour chaque condition, le signal audio des phrases test (le corpus contenait 6 assortiments de 30 phrases test), dégradé par un bruit du fond sonore constitué de texte lu au niveau de RSB = -12 dB, était présenté aux bien-entendants, qui devaient restituer ces phrases. Le tableau 1 représente les résultats, classés selon le gain de performance.

Tableau 1 – Pourcentages moyens et écart-types de mots transcrits correctement pour chaque condition (Summerfield, 1979)

	Condition				
	Cercle	Sans vidéo	4 points lumineux	Lèvres isolées	Vidéo complète
Correct, %	20.8	22.7	30.7	54.0	65.3
Ecart-type	10.0	8.59	16.2	14.5	19.7

Conformément au critère statistique adopté, seules les vidéos en conditions Vidéo complète et Lèvres isolées améliorent l'intelligibilité par rapport à l'audio seul. Une explication de la différence de 11.3 % entre ces 2 conditions est suggérée par les commentaires des sujets qui trouvent que c'est l'absence de dents et de mouvement de la langue, plutôt que l'absence du reste du visage, qui rend la compréhension de la parole plus difficile. La représentation des lèvres par 4 points lumineux n'apporte que 8 % de plus de performance par rapport à l'audio seul. Cependant, des expériences ultérieures ont montré qu'un placement plus judicieux d'un plus grand nombre des points lumineux donne plus d'information sur le signal de parole. La vidéo de l'enveloppe de l'amplitude du signal audio n'apporte rien à la compréhension (Summerfield a proposé 2 raisons : le manque de validité anatomique et le caractère inhabituel de la présentation).

Summerfield et al. (1989) ont évalué la contribution des dents à la distinction de voyelles. Pour ces expériences ils ont utilisé un corpus de 11 syllabes: « beeb », « bib », « beb », « bab », « barb », « berb », « bub », « bob », « boub », « borb », « boob » et sept conditions de présentation visuelle. Le tableau 2 montre la précision d'identification de voyelles pour chaque condition.

Tableau 2 – Pourcentages moyens et écart-types de voyelles identifiées correctement pour chaque condition (Summerfield et al., 1989)

	Naturel			Synthétique			
	Visage entier	Lèvres et dents seules	Lèvres seules	Visage entier	Visage entier sans les dents	Lèvres et dents seules	Lèvres seules
Correct, %	77.8	56.1	50.3	56.7	51.1	58.6	52.9
Ecart-type	12.4	10.3	12.0	15.5	10.3	12.6	12.1

Ainsi, les dents apportent en moyenne un gain d'intelligibilité de 6 %, mais beaucoup plus pour certains voyelles (sans les dents il y avait la confusion entre « beeb » et « berb », « beb » et « bub », etc.). On peut supposer que la différence de 21.7 % entre les présentations du Visage entier et Lèvres et dents seules indique une importance de la vision de la langue (sa partie visible dans la bouche). La performance avec le visage entier synthétique (dans ces expériences le visage synthétique est une animation de dessins) est bien plus faible que la performance avec le visage entier naturel. Summerfield et al. supposent que ceci est dû à l'absence de la langue dans le visage de synthèse et à l'imprécision de la forme des lèvres synthétisées.

Ijsseldijk (1992) a étudié l'apport du visage entier et des lèvres seules à la lecture faciale de la parole par les enfants sourds. Il a testé la compréhension des mots et phrases (corpus de trois listes, équilibrés en termes de complexité, chacune contenant 4 mots multisyllabiques, 2 phrases de 3 mots et 4 phrases de 5 à 7 mots) à partir de la vidéo du Visage Entier-Face (0°), du Visage à 60° et des Lèvres Seules. Les scores du tableau 3 indiquent le nombre moyen des syllabes correctement restituées.

Tableau 3 – Les scores de lecture faciale pour chaque condition (Ijsseldijk, 1992)

Nombre des syllabes total	Présentation	Nombre moyen des syllabes imitées	Ecart-type
54	Visage Entier-Face	25.42	8.24
61	Visage à 60°	24.39	9.91
53	Lèvres Seules	22.93	9.08

Ijsseldijk a conclu que les scores de la condition Lèvres Seules sont légèrement inférieurs aux scores obtenus pour les deux autres conditions, mais pas de façon significative. Le fait que pour chaque condition l'autre liste de syllabes soit employée rend les résultats moins précis. Les sujets de cette expérience étant sourds, la comparaison des résultats obtenus par Ijsseldijk avec les expériences décrites précédemment est difficile.

Benoît et Le Goff (1998) mentionnent que Mohamadi (1993) a évalué la performance de lecture des lèvres synthétiques pour un modèle de lèvres « key-frame » (les images représentatives de 24 visèmes ont été stockées en mémoire; la zone des lèvres était présentée en rose, les dents en blanc, et le menton était représenté par une parabole; cf. exemples des images des visèmes: figure 3) et un modèle paramétrique.



Fig. 3 – Exemples des images des visèmes du modèle des lèvres « key-frame » (Mohamadi, 1993)

Mohamadi a utilisé le même protocole de test que Sumbly et Pollack (1954), en présentant soit l'audio seul, soit l'audio en synchronie avec deux types des lèvres synthétisées. Le corpus contenait 18 phrases du type « C'est pas VCVCVz ? », où V = {/a/, /i/, /y/} et C = {/b/, /v/, /z/},

/□/, /r/, /l/}. La figure 4 représente les scores d'intelligibilité de parole, qui sont remarquablement similaires dans les deux cas de présentation des lèvres différentes. La cadence de présentation d'animation étant différente dans les deux cas (25 fps pour le modèle « key-frame » et 10 fps pour le modèle paramétrique), l'approche paramétrique semble plus prometteuse.

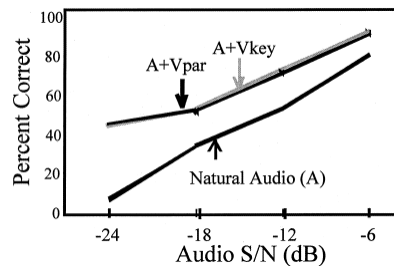


Fig. 4 – Score d'intelligibilité en fonction du bruit pour les différentes conditions (Mohamadi, 1993)

Le Goff et al. (1996) ont évalué la performance des modèles paramétriques 3D des lèvres et du visage entier en utilisant le même paradigme que dans l'expérience précédente. La figure 5 compare les scores d'intelligibilité obtenus pour les deux modèles paramétriques, le visage naturel et l'audio seul. Notons que la langue n'était pas visible dans le modèle des lèvres et visible, mais pas animée, dans le modèle du visage.

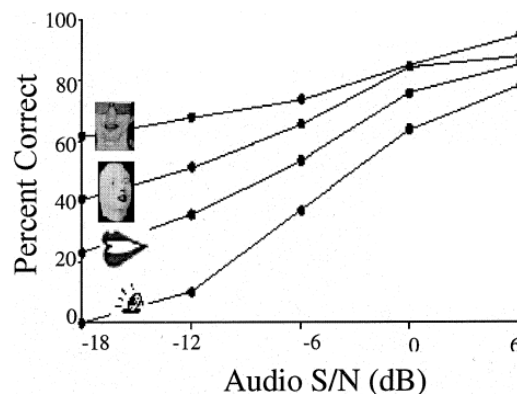


Fig. 5 – Scores d'intelligibilité en fonction du bruit pour les différentes conditions (de bas en haut) : audio seul ; audio + modèle des lèvres ; audio + modèle du visage ; audio + visage naturel (Le Goff et al., 1996)

Dans le cas d'un signal audio sévèrement dégradé, ce modèle de lèvres (piloté par 5 paramètres de contrôle) restitue environ 1/3 de l'intelligibilité observée pour le visage naturel, tandis que si le RSB est plus élevé, le modèle de lèvres restitue environ les 2/3 de l'information visuelle. Le modèle du visage, piloté ainsi par 6 paramètres de contrôle, apporte environ les 2/3 de l'intelligibilité observée pour le visage naturel (dans le cas audio dégradé sévèrement).

En développant leur synthétiseur de parole visuel, Le Goff et Benoît (1996) ont ajouté au modèle du visage le modèle de langue basique (contrôlé par un seul paramètre) qui a été intégré par Cohen et Massaro (1993). Ils ont évalué l'apport de la vision du nouveau modèle à la compréhension de la parole (figure 6). Les auteurs constatent l'amélioration de l'intelligibilité apportée par ce modèle par rapport au modèle précédent (il n'est pas facile de déduire cet apport en comparant les figures 5 et 6).

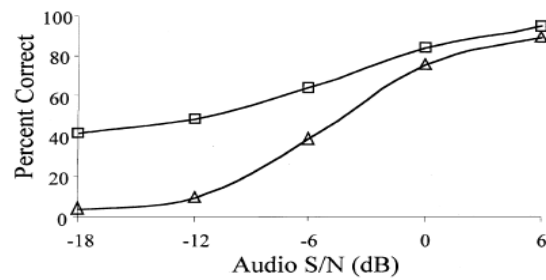


Fig. 6 – Scores d'intelligibilité en fonction du bruit pour les conditions de l'audio seule (en bas) et de l'audio + modèle du visage (en haut), (Benoit et Le Goff, 1998)

Rosenblum et al. (1996) ont employé la technique des points lumineux (39 points au total) pour isoler les aspects cinématiques et géométriques du visage articulatoire et pour étudier ainsi le rôle des images obtenus par cette technique dans la lecture faciale. Les phrases (60 phrases, de six à sept syllabes, chacune avec 3 mots-clé) étaient présentées avec du bruit dans 5 conditions différentes (Audio seul, Audio + 3 configurations différentes des points lumineux, Audio + le visage illuminé entièrement) à des sujets bien-entendants. Le tableau 4 représente les performances de la compréhension de parole obtenues pour chaque condition et mesurées en SRTN (« speech-reception threshold in noise », le RSB minimal, en dB, pour lequel le sujet perçoit les trois mots-clés de la phrase; les phrases étaient présentées en utilisant la méthode des limites ascendantes – « ascending method of limits », MacLeod et Summerfield, 1987).

Tableau 4 – Les scores SRTN moyens et écart-types pour chaque condition (Rosenblum et al., 1996)

Condition	Audio seul	Points des lèvres	Points des lèvres, dents, langue	Tous les points	Illumination complète
SRTN moyen, dB	-4.7	-8.4	-9.6	-10.00	-15.5
Ecart-type, dB	4.0	5.5	6.7	7.2	9.1

Rosenblum et al. (1996) trouvent que toutes les comparaisons deux à deux sont significatives, sauf *Points des lèvres, dents, langue* vs. *Tous les points*, et ils concluent que l'ajout de points sur le front, les joues, la mâchoire et le nez n'améliore pas substantiellement la performance si on observe les points des lèvres, des dents et de la langue. Ils admettent toutefois qu'un autre placement de points pourrait donner un résultat plus performant, i.e. plus proche du score de l'*Illumination complète*. Les auteurs constatent aussi que l'ajout d'information sur la langue en mouvement aide la compréhension de parole. Dans cette recherche un seul point lumineux sur l'apex de la langue a caractérisé les mouvements de langue (la forme de langue n'était pas présentée).

Les études récentes montrent cependant une importance des composantes de la partie non buccale (les yeux, le front, la mâchoire, le nez, le contour) pour la perception de parole (Vatikiotis-Bateson et al., 1998; Thomas et Jordan, 2004). Les mouvements des composantes de la partie non buccale sont en partie corrélés fortement avec les mouvements des articulateurs buccaux (Munhall et Vatikiotis-Bateson, 1998).

En conclusion, toutes les études montrent que la vidéo du visage naturel entier, dans différents angles d'observation, donne la meilleure performance dans la lecture faciale. Probablement, la vision de la texture des dents, de la langue, des lèvres et de la peau du visage est-elle importante pour le rehaussement de l'intelligibilité de la parole dans le bruit. Ceci montre qu'il est important d'utiliser des modèles de visage les plus naturels possibles.

La série de ces études indique aussi un certain apport de la vision de la langue en mouvement à la compréhension de la parole, même avec les modèles de la langue pilotés par un seul paramètre ou en ne visualisant qu'un seul point de l'apex de langue. Nous pouvons faire

l'hypothèse qu'une visualisation de la langue basée sur un modèle 3D de langue plus complet pourrait apporter une plus grande contribution à l'intelligibilité de la parole.

1.2 Vision de la langue

Beaucoup de chercheurs et de cliniciens pensent que la difficulté de la lecture faciale de la parole (en moyenne, les malentendants n'identifient que 40% de mots en contexte de phrase; Tyler et Lowder, 1992) est due à l'absence de visibilité de plusieurs des articulateurs qui produisent la parole. Par exemple, on ne peut pas observer l'activité du corps et du dos de la langue, et on ne peut voir l'apex que dans certains cas où le locuteur est situé en face du lecteur. On peut supposer que l'ajout d'information articulatoire visible améliorera l'intelligibilité de la parole.

1.2.1 Visualisation de langue par vidéofluoroscopie

Tye-Murray et al. (1993) ont évalué l'apport de la vision par vidéofluoroscopie des articulateurs typiquement cachés à la reconnaissance de la parole. La vidéofluoroscopie est une procédure radiographique qui permet de visualiser en 2D (Tye-Murray et al., 1993 ont utilisé la vidéofluoroscopie de profil, qui donne des images de la projection de la tête sur un plan sagittal) les mouvements des articulateurs (à 60 images / seconde). Ceci permet d'observer les mouvements de la langue, des dents, de la mâchoire et du velum en temps réel, en synchronie avec le signal audio.

Cinq expériences ont été menées, sur un corpus de phrases à deux ou trois mots-clé, avec des sujets bien-entendants et malentendants, en utilisant les films vidéo de profil de la tête du locuteur (à 60 images / seconde) et la vidéofluoroscopie. Les sujets ont observé les films de vidéo et vidéofluoroscopie 2 fois (2 présentations) en ordre alternant (ex., vidéo, vidéofluoroscopie, vidéo, vidéofluoroscopie) et devaient répéter la phrase après chaque film. Les scores étaient les pourcentages de mots-clé correctement restitués.

Les résultats obtenus montrent que la visualisation des articulateurs supralaryngés par vidéofluoroscopie typiquement non visibles n'améliore pas la compréhension de la parole. La vidéo donne toujours un meilleur score de reconnaissance que la vidéofluoroscopie. Les scores obtenus par les sujets bien-entendants pour la vidéo varient entre 0% et 17% de mots-clé corrects (5.5 % en moyenne pour la première présentation), et pour la vidéofluoroscopie entre 0% et 11% de mots-clé corrects (3.1 % en moyenne pour la première présentation). Les scores pour les sujets malentendants variaient pour la vidéo entre 0% et 88 % de mots-clé corrects (41.7 % en moyenne pour la première présentation), et pour la vidéofluoroscopie entre 0% et 89% de mots-clé corrects (27.4 % en moyenne pour la première présentation). L'entraînement des sujets n'améliore pas la compréhension de la parole dans ces expériences (les gens expérimentés en lecture des images cinéradiographiques ne montrent pas de meilleure compréhension de la parole que les novices).

Cependant on peut faire quelques considérations sur ces résultats :

Les images de vidéofluoroscopie en 2D n'étaient pas très claires : elles apparaissaient comme des images cinéradiographiques que les auditeurs n'étaient pas habitués à lire. Même si les sujets ont été soumis à un entraînement à la lecture des vidéos cinéradiographiques, pour la compréhension de la parole il fallait lire les images de vidéofluoroscopie rapidement : cette procédure était difficile. Un entraînement plus extensif pourrait donner d'autres résultats.

Les sujets n'ont observé que le plan de profil, alors que les auditeurs sont habituellement situés en face du locuteur. Mais cette considération n'est pas très importante. Jordan, Sergeant, Martin, Thomas et Thow (Jordan et al., 1997) ont en effet montré que la reconnaissance de la parole visuelle et audio-visuelle n'était pas affectée par le changement de l'angle horizontal de l'observation (pour les angles de 0 degrés, 45 degrés, 90 degrés). D'un autre côté, si on visualise la langue, il est évident que le plan de profil apporte la plus grande quantité d'information sur la forme de langue.

Une des expériences de Tye-Murray et al. (1993) a étudié l'impact de la vision de l'activité de langue par vidéofluoroscopie sur l'intelligibilité de la parole. Pour cela, un deuxième locuteur a été introduit, avec des plombages en métal sur les dents qui cachaient sa langue en vidéofluoroscopie. Le deuxième locuteur prononçait une liste de phrases différente de celle prononcée par le premier locuteur. Les sujets malentendants ont observé la première présentation du film vidéo et vidéofluoroscopie du premier locuteur (avec la langue visible – locuteur V), et ensuite la première présentation de deux films du deuxième locuteur (avec la langue non visible – locuteur NV). Ensuite ils ont observé la seconde présentation de deux films pour deux locuteurs. Les scores obtenus pour le locuteur V variaient pour la vidéo – entre 4 % et 98 % de mots-clé corrects (46.2 % en moyenne pour la première présentation), pour la vidéofluoroscopie – entre 0% et 96 % de mots-clé corrects (23.3 % en moyenne pour la première présentation). Pour le locuteur NV les scores variaient pour la vidéo – entre 4 % et 98 % de mots-clé corrects (55.4 % en moyenne pour la première présentation), pour la vidéofluoroscopie – entre 5% et 86 % de mots-clé corrects (36.6 % en moyenne pour la première présentation). Tye-Murray et al. ont conclu qu'il n'y a pas d'avantage à rendre visible l'activité de langue par vidéofluoroscopie.

En comparant des films avec langue visible et avec langue non visible on doit prendre en compte que les locuteurs et les phrases des deux cas étaient différents. La différence entre le score moyen avec présentation vidéo et le score moyen avec vidéofluoroscopie est plus grande pour le locuteur V (22,9 %) que pour le locuteur NV (18.8 %). De plus, le plombage en métal sur les dents ne cache pas parfaitement la langue, ses parties apparaissent quand leur projection sur un plan sagittal n'est pas superposée avec la projection des dents. Ainsi, la conclusion de Tye-Murray et al. n'est pas très crédible.

Une avantage de l'utilisation du modèle 3D pour la visualisation de la langue et des autres parties du visage est que l'on peut comparer les résultats de l'apport de la vision de l'organe non visible en effectuant les expériences sans et avec vision de l'organe sur le même modèle, c'est-à-dire avec un unique locuteur. De plus, on peut rendre les parties de la tête modélisée transparentes pour observer nettement les mouvements des organes typiquement non visibles, effectuer la vision des organes en réalité augmentée et visualiser facilement les coupes particulières de la tête synthétisée. Dans ce cadre, on peut penser que la visualisation synthétisée des organes non visibles pourrait améliorer l'intelligibilité de la parole dans le bruit.

1.2.2 Visualisation de langue en synthèse 3D

Les équipes de Massaro (Perceptual Science Laboratory, University of California – Santa Cruz), d'Engwall (Centre for Speech Technology/TMH, Stockholm) et de Gérard Bailly (GIPSA-lab, Grenoble) développent des têtes parlantes animées 3D qui peuvent illustrer les mouvements des organes typiquement non visibles.

Le projet de développement de système d'entraînement de prononciation ARTUR (Bälter et al., 2005) élabore un tuteur de correction d'articulation, qui utilise les animations 3D du visage et les parties intérieures de la bouche (langue, palais, mâchoire, etc.) pour montrer le placement des articulateurs et ainsi pour donner la différence entre la déviation de l'utilisateur et la prononciation correcte. Trois enfants âgés de 9-14 ans, bien-entendants, mais avec des désordres de parole, ayant une bonne expérience de l'apprentissage de parole, ont été retenus pour évaluer le système ARTUR. Tous les enfants ont donné une note positive à ce système et ils ont considéré que la meilleure partie était la correction des mouvements de langue plus en arrière ou plus en avant. Mais l'évaluation de l'apport de la vision des parties intérieures de la bouche à l'intelligibilité de la parole avec ce système n'est pas publiée.

L'équipe de Massaro développe une tête parlante animée Baldi en utilisant l'approche géométrique, c'est-à-dire qu'elle restitue les déformations faciales en termes géométriques, sans prendre en compte les mécanismes physiologiques de production de la parole (Massaro, 2004). Baldi peut parler lentement, peut illustrer les articulateurs typiquement invisibles (langue, dents et palais) en rendant sa peau transparente, et peut visualiser les caractéristiques articulatoires

supplémentaires comme la vibration du cou pour montrer le voisement et l'expulsion de l'air pour signaler les bruits de friction.

Massaro et Light (2004) ont évalué la perception et production de la parole par les enfants malentendants avant et après l'apprentissage avec Baldi. Sept enfants malentendants âgés de 8 à 13 ans ont été entraînés pendant 21 semaines (6 heures par semaines) avec les 8 catégories de segments (4 distinctions voisés vs. non voisés, 3 distinctions des clusters de consonne (ex. : clusters de consonne initiaux pour /s/ - *smile*, *slit*, *stare*), 1 distinction fricative vs. affricative). Les 4 conditions de présentation différentes de Baldi qui ont été utilisées pour montrer les différents perspectives de l'activité sont : de l'arrière, latéral, sagittal, de devant. La figure 7 représente deux des quatre conditions de présentation.

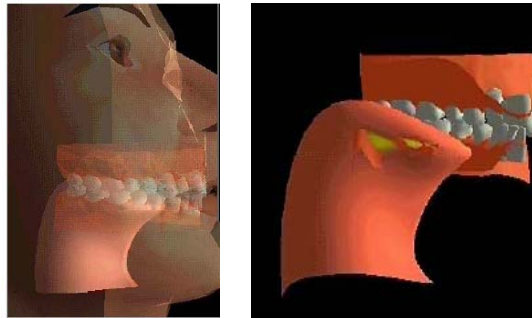


Fig. 7 – Deux des quatre conditions de présentation de Baldi : latéral (à gauche) et sagittal (à droite), (Massaro et Light, 2004)

La figure 8 illustre le pourcentage d'identification correcte pour le test avant et après l'apprentissage.

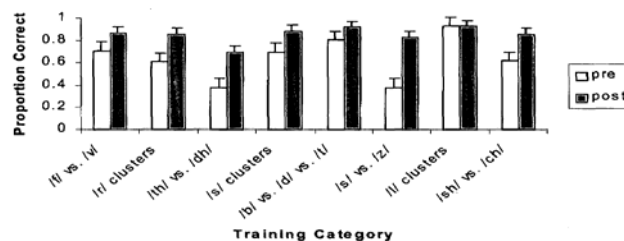


Fig. 8 – Pourcentage d'identification correcte pour le test avant et après l'apprentissage (Massaro et Light, 2004)

La perception et la production de la parole ont été améliorées pour tous les enfants.

Massaro et Light (2003) ont aussi évalué l'efficacité de Baldi pour l'apprentissage des contrastes phonétiques de la langue étrangère, en utilisant l'apprentissage avec la visualisation des mouvements articulatoires de cavité orale typiquement invisibles et avec seulement la visualisation du visage du tuteur. Les onze japonais qui parlent anglais en deuxième langue ont été entraînés par Baldi à identifier et produire les /r/ et /l/ de l'anglais américain. Les auteurs ont conclu que l'identification et la production de la parole ont été améliorées en utilisant chaque méthode d'apprentissage, bien que la visualisation des articulateurs internes ne donne pas d'avantage supplémentaire.

En conclusion, les études effectuées ne donnent pas de conclusion nette sur l'apport de la vision des articulateurs typiquement non visibles, en particulier de la langue, à la compréhension de la parole, et il est intéressant d'évaluer cet apport.

1.3 Conscience phonologique et articulatoire

Pour évaluer l'apport potentiel de la vision de la langue à la compréhension de la parole il est intéressant d'analyser les aspects de la conscience articulatoire, c'est-à-dire la capacité des humains à associer les phonèmes avec la position correspondante des articulateurs. Cette notion

est liée à la notion de conscience phonologique, qui inclut la capacité de distinguer à partir du signal de parole audio les parties de cette parole telles que les syllabes et les phonèmes.

La théorie motrice de la perception de la parole décrit l'importance des gestes articulatoires pour la perception (Heilman et al., 1996). Cette théorie avance que les gestes articulatoires ne constituent pas seulement les structures élémentaires de la production de la parole, mais qu'ils sont aussi cruciaux pour la perception de la parole. Ces gestes articulatoires fournissent la base pour les catégories phonémiques. A chaque phonème d'un mot correspondent certains mouvements articulatoires, et pour la perception de certaines séquences de parole il faut retrouver la séquence spécifique des gestes intentionnels. Ainsi, la théorie motrice de la perception de la parole explique comment on développe la conscience phonologique.

Montgomery (1981) a étudié l'importance de la conscience articulatoire pour la perception de la parole en comparant la performance d'enfants normaux et dyslexiques (la dyslexie est une déficience qui touche la conscience phonologique), du même âge « de lecture » (environ 8 ans), dans une tâche d'association de phonèmes aux images des positions des articulateurs pour ces phonèmes. Il a montré que les enfants dyslexiques ont des performances beaucoup plus faibles que les enfants normaux, qui ont bien réussi à résoudre ce problème.

Le test de Montgomery pour la mesure de la conscience articulatoire a été repris par Griffiths et Frith (2002) avec des sujets adultes (17 normaux d'âge moyen 21.36 ans et 17 dyslexiques d'âge moyen 21.67 ans). On a montré aux sujets les 9 images des coupes médiosagittales³ en couleurs du visage humain (2D, 3 x 5 pouces), chaque image représentant la position particulière de la langue, des dents et des lèvres pendant l'articulation du phonème particulier. Après l'audition du phonème (ils ont été présentés en ordre /a/, /t/, /th/, /p/, /l/, /d/, /f/, /e/, /z/, /s/), le sujet l'a répété plusieurs fois jusqu'à la bonne prononciation de ce phonème. Ensuite le sujet a associé une des 9 images à l'articulation de ce phonème particulier. Les scores moyens (écart-types) obtenus dans ce test sont les suivants (max = 10) : pour les adultes normaux 7.24 (1.52); pour les dyslexiques 5.06 (1.98). La figure 9 représente la distribution des scores pour les dyslexiques et normaux et la distribution des scores pour chaque phonème.

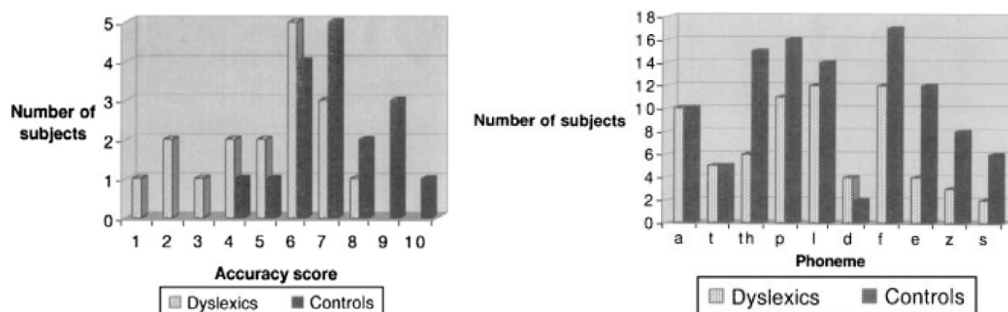


Fig. 9 – Distribution des scores du test de la conscience articulatoire (à gauche) et nombre des articulations identifiées correctement pour chaque phonème (à droite), (Griffiths et Frith, 2002)

Les adultes dyslexiques ont des scores faibles, tandis que les adultes normaux ont bien identifié l'image correspondante dans 72.4 % des cas en moyenne. On peut penser que ce score pourrait être amélioré par l'utilisation d'images plus claires et plus précises, et par l'utilisation des animations 3D.

Les résultats obtenus par Griffiths et Frith nous montrent cependant que la vision des articulateurs (en particulier, de la langue) peut aider à la perception de la parole.

Pour vérifier l'importance de toutes ces considérations, nous allons visualiser le modèle 3D de la langue en réalité augmentée à la base du modèle 3D de langue et évaluer cet apport à la compréhension de la parole.

³ En anatomie, le plan médiosagittal est le plan qui sépare la moitié gauche de la moitié droite du corps

Chapitre 2 – Elaboration du test perceptif

L'objectif de ce chapitre est l'élaboration du test perceptif pour évaluer l'apport de la vision de la langue à la compréhension de la parole.

Tout d'abord, le modèle articulatoire que nous avons utilisé pour la visualisation de la langue en réalité augmentée est décrit. Pour piloter ce modèle d'une manière réaliste, nous avons réalisé l'acquisition des données articulatoires par la technique d'articulographie électromagnétique. A partir de ces mesures nous avons déterminé les paramètres de contrôle des articulateurs (pour les 64 logatomes VCV – voyelle – consonne – voyelle, choisis comme les stimuli du test de perception), en inversant les modèles de langue et du visage.

Ensuite, le protocole et le programme du test de perception sont élaborés. L'expérience consiste à tester la compréhension de la parole dans quatre conditions de présentations des stimuli (audio seul, audiovisuel avec visualisation de l'intérieur du visage sans langue et avec langue, audiovisuel avec visualisation du visage complet) en condition de bruit de différents niveaux. L'analyse des scores d'intelligibilité pour ces différentes conditions permettra de conclure sur l'apport de la vision de la langue à l'intelligibilité de la parole dans le bruit.

2.1 Modèle de langue

Pour évaluer l'apport de la vision de la langue en réalité augmentée à la compréhension de la parole, nous allons visualiser les différents articulateurs en mouvement. Pour cela, nous utilisons les modèles articulatoires statistiques 3D de langue, de mouvements de la mâchoire et des lèvres. Ceci permet de contrôler facilement une présentation réaliste des articulateurs.

Les modèles articulatoires 3D de langue, de mouvements de la mâchoire et des lèvres utilisés dans notre recherche sont des modèles statistiques linéaires orientés sur un sujet. L'approche de la modélisation est décrite en détails dans Badin et al. (2002) et Badin et Serrurier (2006).

Des images 3D de la langue ont été obtenues par IRM sur le sujet qui a prononcé le corpus de 46 articulations des allophones du français. Des images supplémentaires du palais dur, mâchoire et os hyoïde ont été prises par tomographie (CT). La surface 3D de la langue a été représentée pour chacune de 46 articulations du corpus par le maillage (à partir du maillage générique ajusté à l'ensemble de contours de la langue tracés à partir d'images IRM). L'ensemble des coordonnées 3D des sommets de maillage de la langue a constitué les variables et l'analyse en composantes linéaires a été appliquée sur l'ensemble des articulations. Six composantes indépendantes ont été trouvées qui expliquent 87 % de la variance totale des données de la langue. Les paramètres correspondants qui contrôlent le modèle articulatoire linéaire de la langue (jaw height JH, tongue body TB, tongue dorsum TD, tongue tip vertical TTV, tongue tip horizontal TTH, hyoïde HY) sont associés à la position de la mâchoire et de l'os hyoïde, et aux actions des muscles de la langue (genioglossus, hyoglossus et styloglossus). On peut observer l'effet de ces paramètres sur la figure 10 qui représente l'intersection médiosagittale des nomogrammes 3D avec les variations linéaires des paramètres de contrôle.

Badin et Serrurier (2006) ont aussi montré que la surface complète 3D de la langue est prédictible, dans le cas d'une activité de parole, à partir de son contour médiosagittal avec une augmentation de l'erreur RMS totale de reconstruction 3D ne dépassant pas 13.6 %.

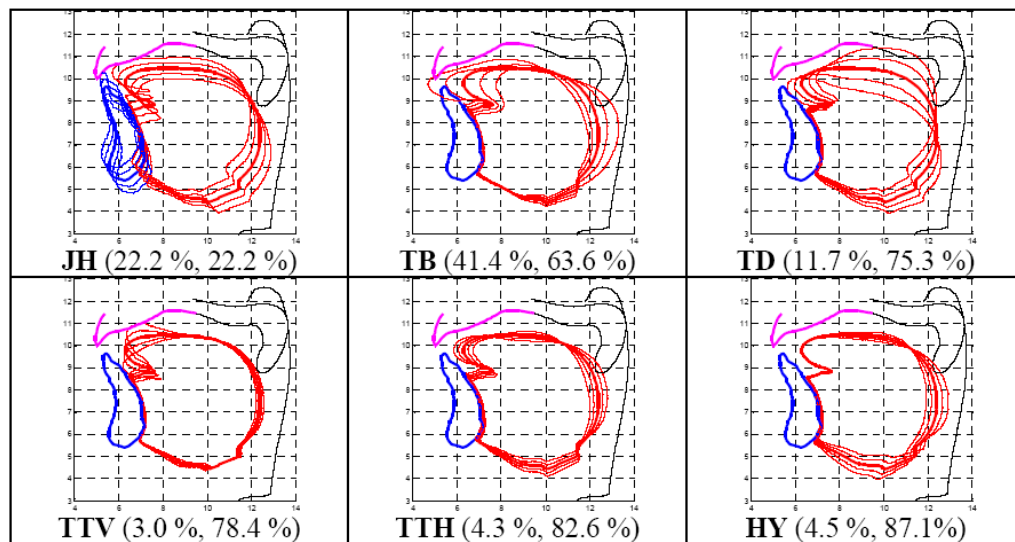


Fig. 10 – Section médiosagittale des nomogrammes 3D avec les variations linéaires des paramètres entre -2 et +2 (entre parenthèses : explication de variance, explication de variance cumulée)

2.2 Acquisition des données articulatoires : Enregistrements EMA

Pour piloter les modèles articulatoires de langue et des lèvres d'une manière réaliste, nous allons déterminer les valeurs des paramètres de contrôle, en utilisant des données dynamiques de mouvements des articulatoires mesurées sur le même sujet.

Des données dynamiques ont été enregistrées sur le sujet âgé de 51 ans par l'articulographe électromagnétique Carstens AG 200 le 5 avril 2007. Cette technique permet d'analyser les mouvements articulatoires à partir de la position de plusieurs points de chair du conduit vocal dans le plan médiosagittal au cours du temps (Perkell et al., 1992 ; Hoole, 1996). Huit bobines réceptrices ont été attachées aux articulatoires du sujet (cf. préparation du sujet pour l'enregistrement EMA : figure 11):

- Une bobine sur le nez et une bobine sur les incisives supérieures pour servir de référence.
- Une bobine sur les incisives inférieures « jaw » pour enregistrer la position de la mâchoire.
- Trois bobines sur la langue (une à 1.2 cm de l'apex « tip », une au milieu à 4.2 cm de l'apex « mid » et une en arrière à 7.3 cm de l'apex « bck ») pour enregistrer la position de ces points de la langue au cours du temps.
- Deux bobines sur les lèvres supérieure et inférieure pour enregistrer la position de ces points des lèvres.

D'autres « artifices » apparaissent sur le visage du locuteur maquillé, mais ils ne servent pas dans le cadre de ce stage.



Fig. 11 –Préparation du sujet pour l'enregistrement EMA

Le **corpus** enregistré est composé des 14 voyelles du français et de séquences VCV (voyelle – consonne – voyelle), où $C = \{/p/, /t/, /k/, /f/, /s/, /□/, /b/, /d/, /g/, /v/, /z/, /□/, /m/, /n/, /□/, /l/, /w/, /□/, /j/\}$ (les 16 consonnes + 3 semi-consonnes du français), $V = \{/a/, /□/, /e/, /i/, /y/, /u/, /o/, /ø/, /□/, /œ/, /ã/, /ĩ/, /ẽ/, /ĩ/\}$ (l'ensemble des 14 voyelles du français), de 190 mots monosyllabiques (extraits du test de rime de Peckels et Rossi, 1973), de 68 courtes phrases phonétiquement équilibrées et de phrases plus longues (la durée totale du corpus est d'environ une heure). Les enregistrements supplémentaires ont été conduits pour le calage des données.

La tête est fixe pendant l'enregistrement, mais il faut caler les données dans un repère physiologique : ce repère est fixé aux structures osseuses de la tête ; les coordonnées du bout des incisives supérieures sont (5 ; 0 ; 10) et le plan X est parallèle au plan occlusal. Nous avons appliqué une transformation rigide (rotation + translation) aux coordonnées des bobines. Les données des enregistrements supplémentaires (déplacement de bobine le long du palais) ont montré que les mesures de la bobine « tip » semblaient décalées. Nous avons déplacé les coordonnées de cette bobine de 1.5 mm en arrière (axe horizontal) et de 1.5 mm vers le bas (axe vertical).

Notons qu'avant de pouvoir effectuer l'enregistrement EMA avec le corpus et les positions des bobines souhaitées, nous disposions d'un autre enregistrement EMA effectué sur le même sujet, le 18 avril 2003, avec seulement cinq bobines (incisives supérieures et inférieures, « tip », « mid » et « bck » sur la langue). Une partie de la mise au point de l'inversion a été réalisée sur cet ancien corpus.

2.3 Détermination des paramètres de contrôle des articulateurs

L'étape suivante consiste à déterminer, par inversion, les paramètres de contrôle du modèle articulatoire de la langue à partir des mesures articulatoires.

Le but de l'inversion est de trouver à partir des données articulatoires EMA (coordonnées des trois bobines sur la langue et de la bobine sur les incisives inférieures) les paramètres pour piloter le modèle de langue 3D (JH, TB, TD, TTV, TTH, HY).

2.3.1 Association entre bobines EMA et points du maillage 3D du modèle de langue

Pour établir la relation entre le modèle 3D de langue et les coordonnées des bobines EMA, on cherche les indices de 3 sommets de maillage générique de la langue auxquels on peut considérer que les 3 bobines EMA ont été attachées (points pseudo-EMA).

Nous avons cherché ces indices à partir de 22 articulations⁴ que nous avons choisies pour assurer la bonne position du point pseudo-EMA « tip ».

La procédure de détermination de l'indice de chacun des trois points pseudo-EMA est la suivante. Pour chacune des 22 articulations, on apparie les articulations du corpus IRM avec celles du corpus EMA. Les sommets de maillage de la langue sont classés (en ordre ascendant) selon leur distance à la position de la bobine EMA. Le rang total d'un sommet est défini comme le rang maximal de ce sommet sur toutes les listes des sommets classés des 22 articulations. Le sommet avec le plus petit rang total est choisi comme le point pseudo-EMA.

Nous avons trouvé les indices des points pseudo-EMA: 1615, 354, 151 pour «tip», «mid» et «bck» respectivement.

La figure 12 représente les nuages des points pseudo-EMA superposés avec les nuages des traces de bobines EMA « tip », « mid » et « bck » pour les 44 articulations du corpus IRM qui a servi à établir le modèle articulatoire de la langue.

⁴ $\{/a/, /□/, /e/, /i/, /y/, /ø/, /ĩ/, /ẽ/, /ĩ/, /ta/, /sa/, /□a/, /ma/, /la/, /pi/, /ti/, /□i/, /mi/, /ni/, /tu/, /su/, /□u/, /lu/\}$

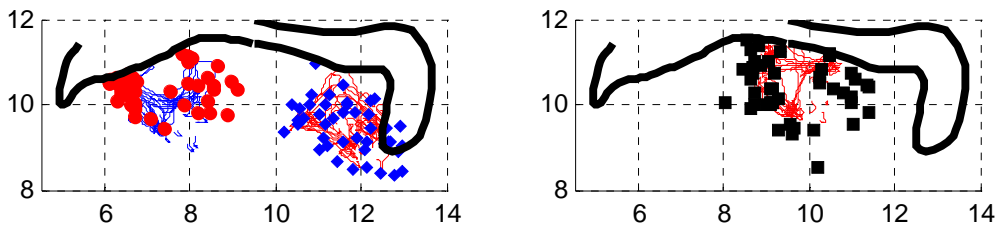


Fig. 12 – Nuages des points pseudo-EMA superposés avec les nuages des traces de bobines EMA pour les 44 articulations du corpus IRM (« tip » et « bck » – à gauche, « mid » – à droite)

La figure 13 représente les distances entre les points pseudo-EMA et les bobines EMA correspondantes pour chacune des 44 articulations.

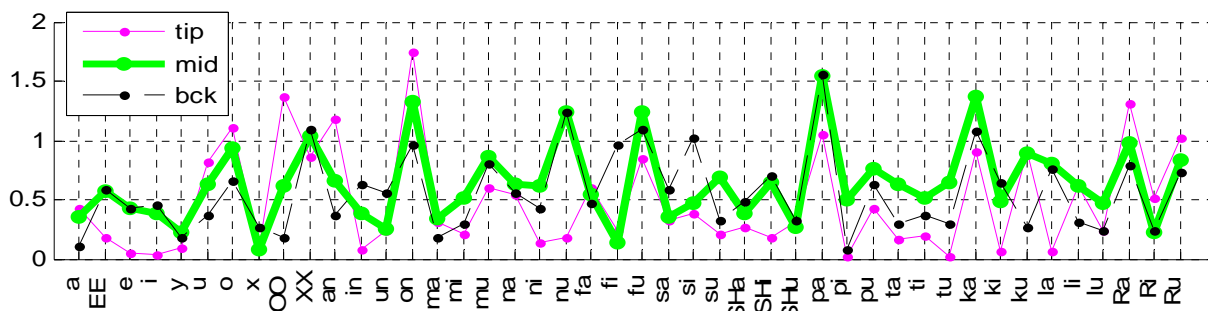


Fig. 13 – Distances entre les points pseudo EMA et les bobines EMA pour 44 articulations (cm)

Pour la suite, si on a des erreurs $err[1..N]$ pour l'ensemble des données de taille N , on calcule l'erreur RMS (« root mean squared ») sur l'ensemble de ces données par :

$$RMS = \sqrt{\frac{1}{N} \sum_{ind=1}^N err_{ind}^2}, \quad (1)$$

Les erreurs RMS sur l'ensemble des 44 articulations pour les bobines « tip », « mid » et « bck » sont égales à 0.64, 0.72, 0.65 cm respectivement.

2.3.2 Inversion

Ayant défini ces 3 indices, nous cherchons les paramètres de contrôle du modèle de la langue 3D (JH, TB, TD, TTV, TTH, HY) en utilisant les mesures EMA.

La reconstitution de la forme de la langue à partir des positions des bobines EMA en utilisant le modèle articulatoire, a été effectuée auparavant par Badin et al. (1997) et par Engwall et Beskow (2003). Badin et al. (1997) ont restitué le contour 2D médiosagittal de la langue, tandis que Engwall et Beskow (2003) ont obtenu la forme 3D de la langue. Dans les deux cas, tout d'abord le paramètre JH a été estimé (en mesurant l'ouverture de la mâchoire), et les 4 autres paramètres ont été déterminés par optimisation en minimisant les distances entre les 3 bobines EMA collées sur la langue et le contour médiosagittal de la langue restitué par le modèle articulatoire.

La procédure de détermination des paramètres du contrôle dans notre cas a été la suivante :

Nous supposons le paramètre $HY = 0$ et les autres 5 paramètres ont été cherchés (la matrice des paramètres $PRED_{1 \times 5} = [JH, TB, TD, TTV, TTH]$). Comme le modèle de langue 3D (Badin et al., 2002) est linéaire et construit par ACP guidée, on dispose de la matrice des coefficients de dimensions $6 \times 1640 \times 3$ (le maillage 3D est constitué de 1640 points), et de la matrice des moyennes (la forme neutre de langue) de dimensions 1640×3 . Si $HY = 0$, la contribution de ce paramètre dans le calcul des coordonnées des points à partir des paramètres du contrôle est nulle, et on n'utilise que la matrice de coefficients de dimensions $5 \times 1640 \times 3$.

Remarquons aussi qu'à partir de données EMA on n'a que les coordonnées X et Z des bobines. De toutes ces considérations et de notre supposition que les 3 bobines EMA sont attachées aux 3 points de maillage avec les indices trouvés, on peut retenir des matrices des coefficients et moyennes les sous-matrices de coefficients de dimensions $5 \times 3 \times 2$ ($COEF_{5 \times 3 \times 2}$) et des moyennes de dimensions 3×2 ($MOY_{3 \times 2}$), qui correspondent aux mesures des coordonnées X et Z des 3 bobines EMA ($MES_{3 \times 2}$) suivant la formule :

$$PRED_{1 \times 5} \cdot COEF_{5 \times 3 \times 2} + MOY_{3 \times 2} = MES_{3 \times 2}, \quad (2)$$

où les paramètres (prédicteurs) $PRED_{1 \times 5}$ sont ceux que l'on recherche.

Tout d'abord, le paramètre JH a été estimé. Comme le modèle de mouvements de la mâchoire est piloté par un seul paramètre JH, les mouvements de mâchoire sont limités. Ainsi, les mouvements de l'apex de mâchoire (le point le plus haut des incisives dans le plan médiosagittal) sont limités par la ligne représentée sur la figure 14 (en étoiles), et si en réalité la bobine EMA « jaw » est située à côté de cette ligne (ex.: losange dans la figure 14), l'erreur de reconstruction est inévitable.

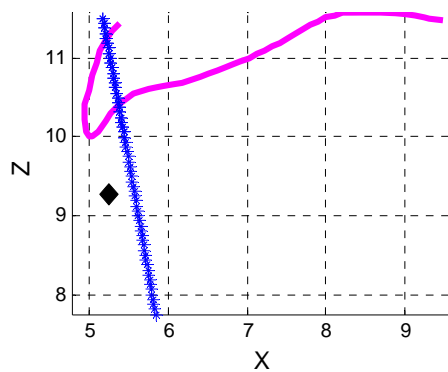


Fig. 14 – Trajectoire des mouvements de l'apex de mâchoire pour les valeurs de JH qui varient de -6 à 6 (en étoiles), exemple de placement de la bobine EMA « jaw » (losange)

Nous avons déterminé JH en utilisant la fonction d'optimisation *fgoalattain* de Matlab (l'optimisation par un algorithme fondé sur la programmation quadratique séquentielle), pour que l'erreur de reconstruction de mâchoire sur l'axe Z soit minimale.

Quand le paramètre JH est estimé (JH_{est}), sa contribution aux coordonnées des points pseudo-EMA est calculée (en utilisant les coefficients du modèle de langue $COEF_{1 \times 3 \times 2}^{JH}$ qui correspondent au paramètre JH) et cette contribution est soustraite des mesures de bobines «tip», «mid» et «bck», en supprimant en même temps les moyennes. On obtient les résidus $RES_{3 \times 2}$, tels que :

$$PRED_{1 \times 4} \cdot COEF_{4 \times 3 \times 2} = RES_{3 \times 2}, \quad (3)$$

où $PRED_{1 \times 4} = [TB, TD, TTV, TTH]$ sont les paramètres inconnus et $COEF_{4 \times 3 \times 2}$ la sous matrice des coefficients qui correspondent à ces paramètres et aux coordonnées X et Z de 3 points pseudo-EMA.

Les paramètres $PRED_EST_{1 \times 4}$ sont déterminés en réorganisant les matrices (en supprimant une dimension) et en inversant la matrice des coefficients $COEF_{4 \times 6}$:

$$PRED_EST_{1 \times 4} = RES_{1 \times 6} \cdot COEF_{6 \times 4}^{-1}, \quad (4)$$

Ainsi, les paramètres du contrôle sont trouvés pour 44 articulations. A partir de ces paramètres, la langue et les mouvements de la mâchoire (rotation autour du point LT – « incisives inférieures » et translation par rapport au point de référence (5; 0; 10)) sont modélisés pour chaque articulation (cf. exemple de modélisation de la langue et de la mâchoire pour la configuration /□/ : figure 15; coupes médiosagittales de la langue et de la mâchoire

modélisées, superposées au contour obtenu en coupant dans le plan médiosagittal le maillage 3D obtenu à partir de données IRM et avec la position des bobines : annexe A).

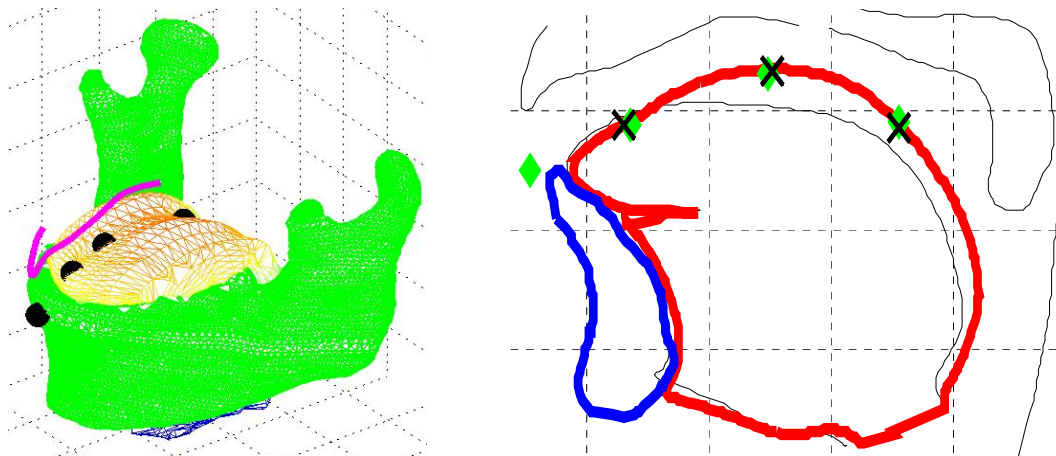


Fig. 15 – Exemple de modélisation de langue et mâchoire pour la configuration /□/: reconstruction 3D (à gauche ; sphères noires représentent les positions des bobines EMA) et coupe médiosagittale (à droite ; superposée au contour obtenu en coupant dans le plan médiosagittal le maillage 3D obtenu à partir de données IRM (ligne continue mince), la position des bobines (losanges) et les points pseudo EMA (croix))

On observe que la mâchoire modélisée est bien placée, que les 3 bobines, attachées sur la langue, collent bien avec les points pseudo-EMA et que la forme de la langue modélisée est correcte et coïncide bien avec le contour obtenu à partir de données IRM dans presque tous les cas (sauf deux : configurations /la/ et /lu/, où la langue n’arrive pas à toucher le palais, et configurations /□a/ et /□i/, où la langue traverse un peu la mâchoire).

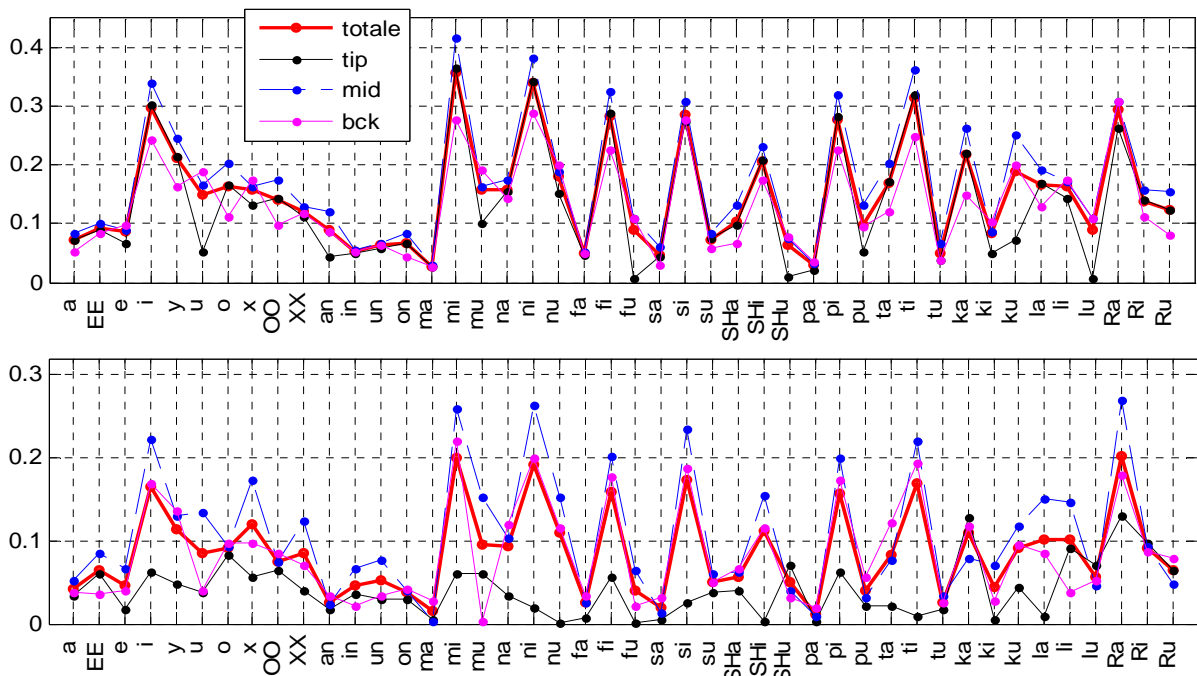


Fig. 16 – Graphes des erreurs d’inversion pour chaque articulation : en haut – erreurs point à point, en bas – erreurs point à contour (cm)

La figure 16 représente les graphes des erreurs d’inversion pour 44 articulations, calculées comme les distances entre les bobines EMA, attachées sur la langue, et les points pseudo-EMA correspondants (erreurs point à point) et les distances entre les bobines EMA, attachées sur la langue, et les contours des coupes médiosagittales de la langue modélisée (erreurs point à

contour). Les erreurs totales pour chaque configuration sont calculées comme les erreurs RMS sur l'ensemble des 3 bobines. Les erreurs RMS totales point à point et point à contour sur 44 configurations sont égales à 0.17 cm et 0.10 cm respectivement, ces erreurs sont admissibles (le tableau 5 représente les erreurs de reconstruction de langue). L'erreur RMS de reconstitution de la position de mâchoire sur 44 articulations sur l'axe X est égale à 0.40 cm.

Tableau 5 – Erreurs de reconstruction de langue à partir de paramètres trouvés par inversion de matrice des coefficients (sur 44 articulations, cm). « pp » : distance point à point ; « pc » : distance point à contour.

	RMS _{tot} pp	RMS _{tot} pc	Max pp	Max pc	RMS _{tip} pp	RMS _{mid} pp	RMS _{bck} pp	RMS _{tip} pc	RMS _{mid} pc	RMS _{bck} pc
Inversion	0.17	0.10	0.42	0.27	0.17	0.20	0.16	0.05	0.13	0.10

Ayant trouvé les paramètres articulatoires [JH, TB, TD, TTV, TTH, HY] par la pseudo inversion de matrice des coefficients, on peut les modifier pour que les erreurs d'inversion point à point ou point à contour soient minimales. Nous avons testé les différentes variantes d'optimisation sur le corpus EMA enregistré à l'ICP le 18 avril 2003 (en optimisant l'erreur point à point pour la (les) bobine et l'erreur point à contour pour l'autre (les autres) bobine, cf. annexe B), l'optimisation multi objectifs a été effectuée par la fonction de Matlab *fminimax*. Pour le corpus de 44 articulations du 5 avril 2007 l'optimisation des erreurs point à point pour la bobine « tip » et point à contour pour les bobines « mid » et « bck » a été appliquée. La forme de langue restituée et les paramètres du contrôle ne sont pas très différents de ceux qui ont été trouvés par pseudo inversion de matrice. Les erreurs RMS totales point à point et point à contour sur 44 configurations (après cette optimisation) sont égales à 0.30 cm et 0.02 cm respectivement.

Nous avons montré qu'en choisissant bien les points pseudo-EMA, les paramètres articulatoires peuvent être trouvés directement en inversant la matrice des coefficients du modèle, ce qui est une solution plus rapide et plus robuste (l'optimisation peut donner un résultat qui n'est pas meilleur). Ceci est intéressant pour la restitution de la forme de la langue en temps réel. C'est cette méthode que nous avons finalement appliquée au corpus enregistré en avril 2007.

2.3.3 Détermination des paramètres de contrôle pour les stimuli du test

Les séquences des paramètres pour le modèle de langue pour 64 logatomes VCV⁵ (où $C = \{/b/, /d/, /g/, /v/, /z/, /□/, /□/, /l/\}$, $V = \{/a/, /i/, /y/, /u/, /□/, /e/, /o/, /œ/\}$) en dynamique (à 50 trames / seconde) ont été calculées par l'inversion de la matrice des coefficients. La langue restituée à partir de ces paramètres a traversé le palais dans le cas des logatomes /ada/, /idi/, /ydy/, /udu/, /□d□/, /ede/, /ili/, /ulu/, /aga/, /i□i/, /y□y/. Ce problème a été mentionné dans la littérature (Cohen et al., 1998). Pour le résoudre, nous proposons la procédure suivante. Les paramètres de contrôle correspondant aux trames où la langue traversait le palais ont été corrigés en utilisant la fonction de Matlab *fminimax*, l'objet de l'optimisation étant la minimisation des erreurs point à point pour les 3 bobines avec la contrainte que le contour médiosagittal de la langue restituée par le modèle articulatoire ne traverse pas le contour médiosagittal du palais. La figure 17 montre les coupes médiosagittales de la langue pour la configuration /d/ en contexte /a/, restituée à partir des paramètres trouvés par l'inversion de la matrice des coefficients (à gauche) et à partir des paramètres corrigés par optimisation (à droite).

⁵ Ces stimuli sont utilisés dans le test perceptif

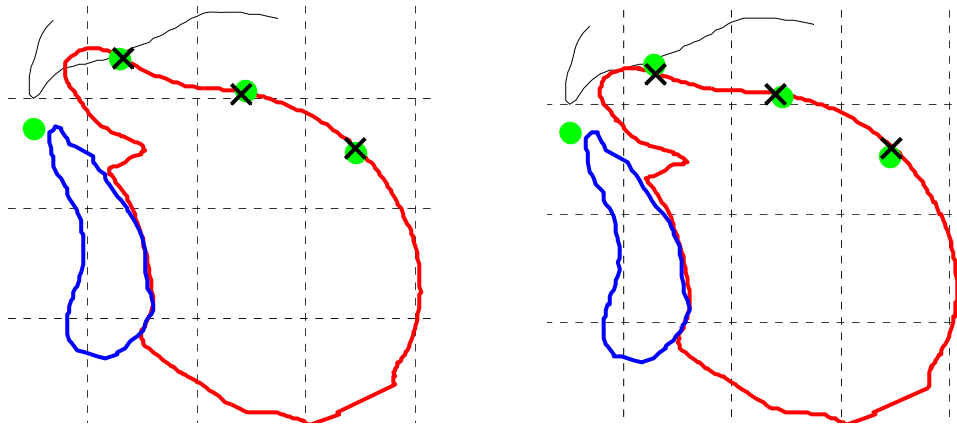


Fig. 17 – Coupes médiosagittales de la langue et mâchoire modélisés (configuration /d/ en contexte /a/): à gauche – le problème de pénétration du palais n’est pas résolu, à droite – le problème de pénétration du palais est résolu

Les séquences de 6 paramètres pour le modèle de lèvres ont été calculées par l’inversion de la matrice des coefficients de ce modèle (le paramètre JH est aligné en cohérence avec son analogue du modèle de la langue ; l’inversion est réalisée par P. Badin).

2.4 Elaboration du test de perception

L’objet de cette partie du travail est d’élaborer le protocole et de développer le programme du test de perception.

2.4.1 Corpus

Nous avons choisi de tester la compréhension de 8 consonnes $\{/b/, /d/, /g/, /v/, /z/, /□/, /□/, /l/\}$ – soit toutes les consonnes voisées sauf les consonnes nasales $\{/m/, /n/\}$. Ce choix a été fait, en prenant en compte notre objectif d’évaluer l’apport de la vision de la langue à la compréhension de parole. Les consonnes voisées et non voisées ne se différencient pas au niveau de la perception visuelle. Les consonnes nasales $\{/m/, /n/\}$ diffèrent visuellement des consonnes $\{/b/, /d/\}$ par la position du velum qui est fixe dans notre présentation des organes internes.

Le lieu d’articulation définit l’endroit où se produit une consonne. Ce trait permet de regrouper les consonnes en 3 grandes classes : labiales et labiodentales ($/p/, /b/, /f/, /v/, /m/$), dentales/ alvéolaires/ post-alvéolaires ($/t/, /d/, /s/, /z/, /n/, /l/, /□/, /□/$), palatales et palato-vélaires ($/k/, /g/, /□/$).

Les consonnes labiales et labiodentales sont articulées soit au niveau des lèvres, soit au niveau des lèvres et des dents. Les consonnes dentales et post-alvéolaires impliquent un contact entre la langue et les alvéoles du palais dur. Les consonnes palatales et palato-vélaires peuvent être produites par le contact de la langue avec le palais dur ou par le contact de la langue avec le velum.

Nous allons étudier si la vision de la langue aide à reconnaître le lieu d’articulation de la consonne (dans notre cas labiales et labiodentales $\{/b/, /v/\}$, dentales/ alvéolaires/ post-alvéolaires $\{/d/, /z/, /l/, /□/\}$, palatales et palato-vélaires $\{/g/, /□/\}$).

Le corpus du test est composé de 64 stimuli de type VCV (voyelle – consonne – voyelle), où $C = \{/b/, /d/, /g/, /v/, /z/, /□/, /□/, /l/\}$, $V = \{/a/, /i/, /y/, /u/, /□/, /e/, /o/, /œ/\}$.

Les voyelles $\{/a/, /i/, /y/, /u/\}$ correspondent aux positions extrêmes du triangle vocalique. La perception des consonnes en contexte de ces voyelles est testée selon différentes modalités de présentation et avec des niveaux de bruit différents.

Les stimuli avec les voyelles $\{/□/, /e/, /o/\}$ sont utilisés à la fin du test, pour tester la capacité de généralisation, ce qui est expliqué plus tard.

Les stimuli avec $V = \{/œ/\}$ sont choisis pour la séance de familiarisation avec la procédure et les conditions du test.

2.4.2 Structure du test

Le **principe général du test** est le suivant : les stimuli audiovisuels de parole (de type VCV) sont présentés successivement au sujet. L'objectif du test est de reconnaître la consonne. Après chaque présentation le sujet doit indiquer la consonne reconnue (en choisissant parmi les 8 consonnes $\{/b/, /d/, /g/, /v/, /z/, /□/, /□/, /l/\}$). Le sujet ne peut pas réécouter le stimulus. S'il hésite entre plusieurs réponses, il doit cliquer sur une des cases possibles au hasard.

Nous avons décidé de tester **4 modalités de présentation** des stimuli :

1. Audio seul ;
2. Audiovisuel avec visualisation de l'intérieur du visage sans langue (dans la suite « Intérieur sans langue ») ;
3. Audiovisuel avec visualisation de l'intérieur du visage avec langue (dans la suite « Intérieur avec langue ») ;
4. Audiovisuel avec visualisation du visage complet (synthétique) de l'extérieur (dans la suite « Visage complet »).

Pour la présentation audiovisuelle avec visualisation de l'intérieur du visage sans langue et avec langue, la tête parlante est coupée par le plan médiosagittal. La présentation de l'intérieur du visage est composée de la visualisation du contour du visage, du palais dur et du voile du palais, de la paroi arrière du pharynx, ainsi que de la mâchoire et des lèvres qui sont en mouvement. De plus, dans la modalité « Intérieur avec langue » on visualise la langue en mouvement (voir figure 18, annexe D).

L'autre question posée est le choix des angles d'observation de la tête (angle horizontal et vertical). Nous avons décidé de présenter les stimuli de profil (subjectivement, cette vue donne le plus d'information).

Dans chaque modalité de présentation les stimuli sont présentés avec **quatre niveaux du rapport signal/bruit** : - infini (i.e. vidéo seul), -9 dB, +3 dB, +infini (i.e. signal audio sans bruit). On ne présente pas le cas où il n'y a ni information vidéo ni information audio. Les 2 niveaux intermédiaires de bruit -9 dB et +3 dB ont été choisis pour que le pourcentage des consonnes reconnues dans la modalité « Audio seul » soit égal à environ 33 % et 66 % respectivement (un test préliminaire avec 2 sujets et des niveaux de bruit variant de -18 dB jusqu'à 15 dB par pas de 3 dB a permis d'établir ces niveaux).

Pour familiariser les participants avec la procédure de test, **la séance débute** par une démonstration de différents modes de présentation et une courte séance d'entraînement, qui consiste à écouter et regarder 5 stimuli (VCV avec la voyelle /œ/) dans le mode de présentation « Intérieur avec langue » avec le son avec RSB = - 9 dB, et à choisir la consonne reconnue.

Le test est composé des séries successives.

Chaque série est caractérisée par sa modalité et son niveau de bruit (i.e. les différentes modalités de présentation et niveaux de bruit ne sont pas mélangées). Au début de chaque série de test, 2 stimuli VCV avec la voyelle /œ/ sont présentés pour que le sujet s'habitue aux nouvelles conditions (on ne les prend donc pas en compte pour l'analyse des résultats). Ensuite, les 32 stimuli VCV (où $C = \{/b/, /d/, /g/, /v/, /z/, /□/, /□/, /l/\}$, $V = \{/a/, /i/, /y/, /u/\}$) sont présentés, en ordre aléatoire et différent pour chaque sujet.

L'ordre des séries du test est:

- 3 séries avec la condition « Audio seul » et 3 niveaux de bruit différents ;
- 4 séries avec la condition « Intérieur sans langue » et 4 niveaux de bruit différents ;
- 4 séries avec la condition « Intérieur avec langue » et 4 niveaux de bruit différents ;
- 4 séries avec la condition « Visage complet » et 4 niveaux de bruit différents.

Cet ordre a été choisi en faisant l'hypothèse que « Intérieur sans langue » apporte plus d'information que « Audio seul », et que « Intérieur avec langue » en apporte plus que « Intérieur

sans langue », et donc en ajoutant de l'information chaque fois ; la modalité « Visage complet » a été placée à la fin sans aucune hypothèse précise *a priori*.

En considérant que les effets d'apprentissage de l'association audio et vidéo se font préférentiellement lorsque l'audio est plus clair, nous organisons deux ordres différents des tests pour deux séries de sujets :

1. Pour chaque modalité les séries avec du bruit descendant (bruit maximal vers bruit minimal).

2. Pour chaque modalité les séries avec du bruit ascendant (bruit minimal vers bruit maximal).

Les différences de résultats permettront une évaluation de l'effet d'apprentissage implicite au test lorsque le niveau de bruit est nul.

La dernière série du test présente les 24 nouveaux stimuli VCV (où $C = \{/b/, /d/, /g/, /v/, /z/, /□/, /□/, /l/\}$, $V = \{/□/, /e/, /o/\}$) dans la modalité « Intérieur avec langue » et avec un $RSB = -9$ dB (plus 2 stimuli de familiarisation au début). L'objectif de cette partie du test et de tester la généralisation, i.e. analyser si les sujets ont appris à lire l'information sur la langue (et n'ont pas seulement retenu les stimuli concrets du test).

2.4.3 Préparation de stimuli

La détermination des séquences des paramètres pour le modèle de langue et des lèvres pour 64 logatomes VCV (où $C = \{/b/, /d/, /g/, /v/, /z/, /□/, /□/, /l/\}$, $V = \{/a/, /i/, /y/, /u/, /□/, /e/, /o/, /œ/\}$) a été décrite dans la partie « Détermination des paramètres de contrôle des articulateurs ».

Les stimuli audio sont préparés en 2 étapes :

– addition de l'audio original et du bruit blanc en proportion pour atteindre rapport signal / bruit (à partir de formule 5).

$$RSB(dB) = 20 \log_{10} \left(\frac{A_{signal}}{A_{noise}} \right), \quad (5)$$

où A_{signal} et A_{noise} sont les amplitudes du signal original et du bruit respectivement.

– normalisation des stimuli, pour que la puissance moyenne de chaque stimulus sur sa durée soit la même pour tous les stimuli.

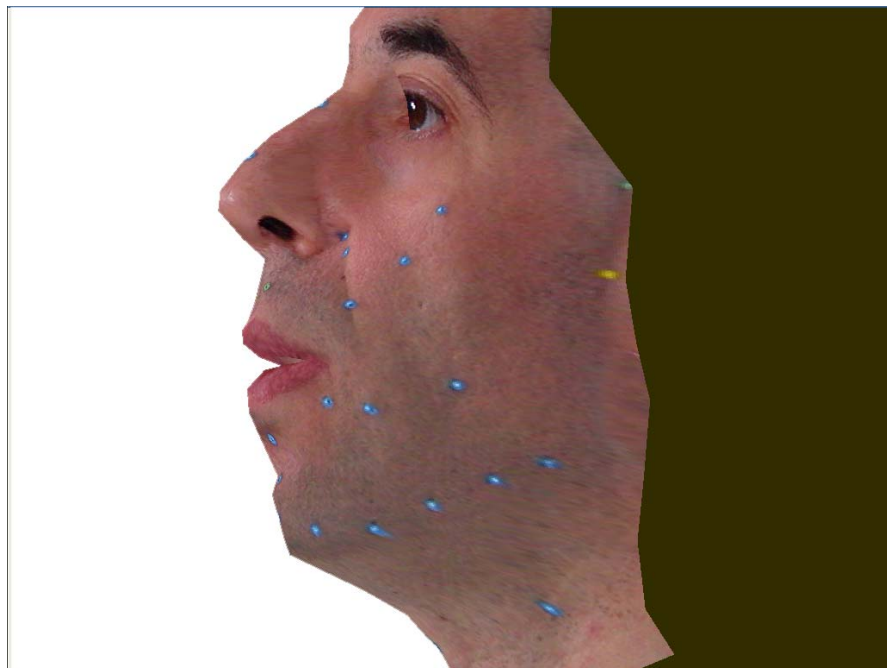
2.4.4 Logiciel de passation du test

La bibliothèque « Minim_lib » a été créée par Frédéric Elisei à l'ICP (actuellement : département Parole & Cognition de GIPSA-lab) pour visualiser les modèles de « têtes parlantes » du laboratoire (visage et lèvres, langue...) avec des modèles d'apparence plus ou moins réalistes. Elle permet aussi de les animer en temps-réel, en lui fournissant un fichier de séquence des paramètres et le fichier audio.

Autour de cette bibliothèque, Frédéric Elisei a aussi développé Minim_expe, un programme pour réaliser des tests de perception sous Linux. Le déroulement du programme est géré par un script qui décrit l'enchaînement des actions : texte à afficher, séquences à visualiser, questions proposées au sujet. Les scripts générés sont différents pour chaque sujet (afin que les stimuli des séries du test soient dans un ordre différent pour les différents sujets).



a)



b)

Fig. 18 – Exemples de fenêtres de présentation du test de perception : a – dans la modalité « Intérieur avec langue », b – dans la modalité « Visage complet »

La figure 18 représente les fenêtres du programme avec les présentations « Intérieur avec langue » et « Visage complet » (cf. les exemples de texte affiché sur l'écran au cours de programme et les exemples des autres fenêtres du test : annexes C et D).

Chapitre 3 – Analyse des résultats du test de perception

3.1 Déroutement du test de perception

Nous avons sélectionné 27 sujets français pour passer le test perceptif (12 femmes et 15 hommes, âgés de 19 à 59 ans), bien-entendants et bien-voyants. Parmi ces sujets, 23 sujets n'avaient pas d'expérience préliminaire de l'analyse des organes de parole, et 4 n'étaient pas naïfs. Par la suite nous avons analysé les résultats des sujets naïfs.

Parmi ces 23 sujets naïfs, 12 des participants ont passé le test « avec du bruit descendant » (groupe I, 7 femmes et 5 hommes, d'âge moyen = 27.25 ans) et 11 participants « avec du bruit ascendant » (groupe II, 4 femmes et 7 hommes, d'âge moyen = 26.91 ans)

La durée d'une séance de test complète variait de 30 à 50 minutes.

Les commentaires informels des sujets sont résumés ci-après. La présentation vidéo des stimuli a été estimée de très bonne qualité. Il n'était pas facile de regarder en même temps les mouvements des lèvres et de la langue. Les stimuli étaient courts et n'avaient pas de sens, ce qui rendait la compréhension plus difficile. Quand il y avait le son, même avec beaucoup de bruit, les sujets avaient l'impression que la langue ne servait pas beaucoup, mais dans la condition de vidéo seule la langue aidait à reconnaître la consonne. Certains sujets ont indiqué qu'il était difficile de focaliser son regard et de suivre simultanément les mouvements de la langue et des lèvres : un compromis qui semblait envisageable consistait à fixer la région des incisives, afin d'avoir la langue d'un côté du champ visuel et les lèvres de l'autre côté. La dernière série du test a paru plus facile que la série avec les mêmes conditions pendant le test.

3.2 Analyse des résultats

3.2.1 Résultats globaux

Nous analysons d'abord l'intelligibilité globale de parole pour les différentes conditions de présentation des stimuli. La figure 19 et les tableaux 6-7 représentent les scores d'intelligibilité (pourcentage de consonnes identifiées correctement) pour les différentes séries du test pour les deux groupes de sujets.

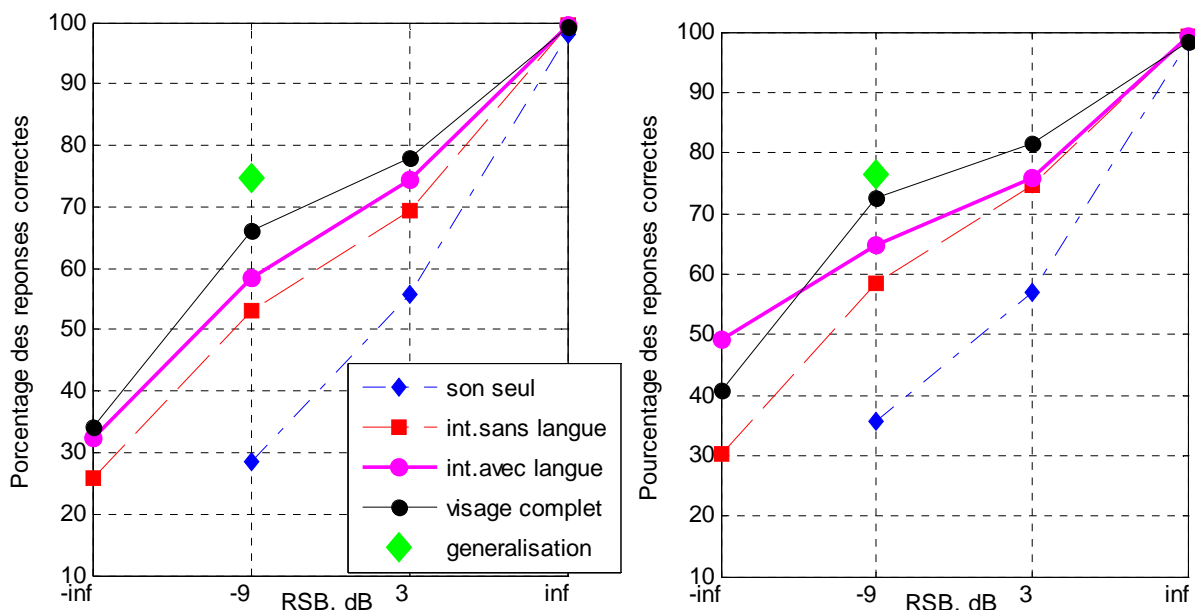


Fig. 19 – Scores d'intelligibilité en fonction du bruit (à gauche, groupe I ; à droite, groupe II) pour les différentes conditions (de bas en haut) : audio seul; intérieur sans langue; intérieur avec langue; visage complet ; série de généralisation

Tableau 6 – Moyennes (écart-types) des scores d'intelligibilité du groupe I pour les différentes séries du test (%)

RSB/ condition	Audio seul	Intérieur sans langue	Intérieur avec langue	Visage complet	Généralisation
- infini		25.78 (6.12)	32.29 (7.46)	34.11 (8.26)	
- 9 dB	28.65 (7.28)	53.13 (6.67)	58.33 (6.71)	66.14 (7.40)	74.65 (9.96)
+ 3 dB	55.73 (8.41)	69.27 (2.93)	74.48 (6.64)	77.86 (6.85)	
+ infini	98.18 (3.63)	99.48 (1.22)	99.48 (1.22)	99.22 (1.94)	

Tableau 7 – Moyennes (écart-types) des scores d'intelligibilité du groupe II pour les différentes séries du test (%)

RSB/ condition	Audio seul	Intérieur sans langue	Intérieur avec langue	Visage complet	Généralisation
- infini		30.40 (5.25)	49.15 (13.41)	40.63 (12.42)	
- 9 dB	35.80 (9.72)	58.52 (4.44)	64.77 (12.03)	72.44 (5.90)	76.52 (5.03)
+ 3 dB	56.82 (9.36)	74.72 (5.13)	75.85 (6.26)	81.53 (6.62)	
+ infini	98.30 (2.56)	99.15 (1.46)	99.15 (1.46)	98.30 (2.15)	

On observe sur la figure 19 que les scores du groupe II sont plus élevés que ceux du groupe I. Toutes les modalités de présentation vidéo améliorent la compréhension de la parole par rapport à l'audio seul. Les courbes des différentes présentations vidéo sont proches l'une de l'autre, les scores de présentation « Intérieur avec langue » sont supérieurs à ceux obtenus pour la présentation « Intérieur sans langue », et la courbe des scores de présentation « Visage complet » est au-dessus des deux autres courbes presque pour tous les niveaux de bruit.

Les écart-types des scores (cf. tableaux 6-7) sont assez importants, comme le montre la figure 20 qui représente les scores d'intelligibilité avec l'étendue des scores pour l'ensemble des sujets.

Les scores pour la modalité « Audio seul » sont significativement différents des scores pour les autres modalités (la différence entre les scores « Audio seul » et « Intérieur sans langue » est significative [groupe I : $F(1, 10)=77.75$; $p<0.0001$; groupe II : $F(1, 10)=99.51$; $p<0.0001$]). Comme nous nous intéressons à la comparaison de scores obtenus pour les différentes modalités de présentation avec vidéo, et comme il n'y avait que les 3 séries du test avec la condition « Audio seul », l'analyse de variance (ANOVA) a été effectuée, en ne prenant en compte que les 3 modalités de présentation avec vidéo. Pour cela, la fonction de Matlab *anovan* est utilisée, pour tester l'effet des facteurs multiples ; la procédure de comparaison multiple est effectuée par la fonction *multcompare*. Les facteurs intra sujet de l'ANOVA ont été la modalité de présentation (3 modalités) et le niveau S/B (4 niveaux) ; le facteur inter sujet a été le groupe (l'analyse est effectuée avec 11 sujets dans chaque groupe).

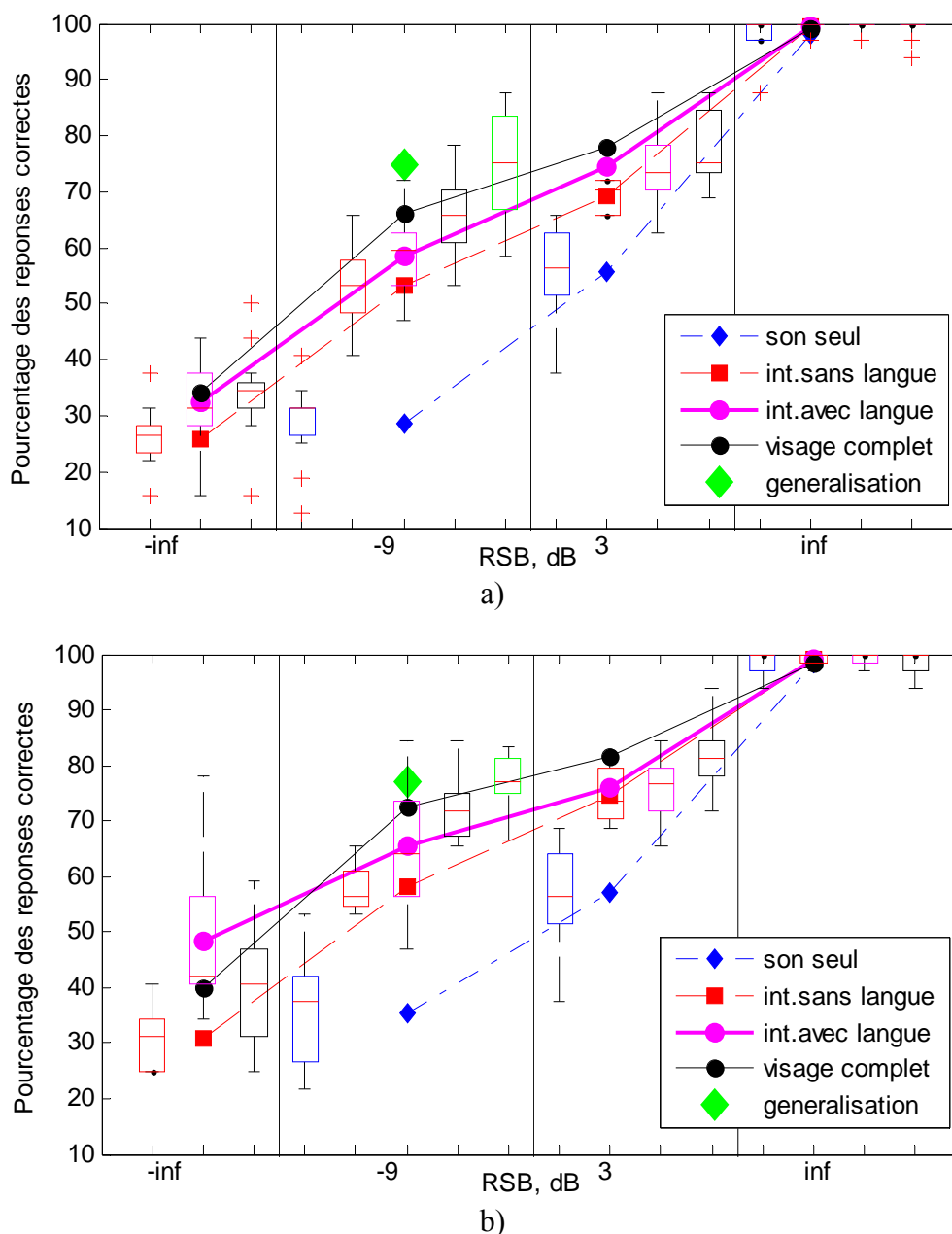


Fig. 20 – Scores d’intelligibilité en fonction du bruit (a, groupe I ; b, groupe II) pour les différentes conditions (de bas en haut, de gauche à droite dans chaque secteur) : audio seul; intérieur sans langue; intérieur avec langue; visage complet ; série de généralisation. Les barres dans chaque secteur représentent l’étendue des scores pour l’ensemble des sujets, les valeurs de la médiane et des deux quartiles pour ces scores

Les scores des 2 groupes sont significativement différents [$F(1,10)=35.59$; $p<0.0001$] ; on a donc intérêt à analyser l’influence des facteurs pour chaque groupe.

L’analyse a montré que les différences de scores des modalités de présentation sont significatives [groupe I : $F(2, 10)=27.1$; $p<0.05$; groupe II : $F(2, 10)=21.03$; $p<0.0001$], et les scores de toutes les paires des modalités sont significativement différentes. Le niveau S/B est aussi significatif [groupe I : $F(2, 10)=1198.6$; $p<0.0001$; groupe II : $F(2, 10)=567.54$; $p<0.0001$]. L’interaction significative entre ces 2 facteurs a été observée [groupe I : $F(6, 60)=4.13$; $p<0.002$; groupe II : $F(6, 60)=8.74$; $p<0.0001$].

Les différences entre les scores de la présentation « Intérieur avec langue » et chacune des deux autres présentations pour chaque niveau de bruit ne sont pas significatives, sauf dans deux cas : entre présentations « Intérieur avec langue » et « Visage complet » pour le niveau du

bruit -9 dB pour le groupe I [$F(1, 10)=12.6$; $p<0.05$], et entre présentations « Intérieur avec langue » et « Intérieur sans langue » pour le niveau de bruit -infini pour le groupe II [$F(1, 10)=9.28$; $p<0.05$].

Le score de la présentation « Visage complet » est significativement plus grand que celui de la présentation « Intérieur sans langue » pour chaque niveau de bruit (avec $p<0.05$ pour chaque paire), sauf pour le groupe II pour le niveau de bruit +3 dB ($F(1, 10)=7.29$; $p>0.05$). Ceci peut être dû au fait que la texture du visage apporte une information supplémentaire liée à la nature redondante des mouvements de la mâchoire et des lèvres avec ceux des joues. Mais comme le sujet observait les stimuli en modalité « Visage complet » après les présentations « Intérieur sans langue » et « Intérieur avec langue », il est possible que le sujet ait appris les stimuli et qu'il ait mieux réussi grâce à l'observation du visage complet.

Ainsi, même si pour le groupe II dans la condition de vidéo seule, le meilleur score est obtenu en modalité « Intérieur avec langue » (ce qui est cohérent avec les commentaires des sujets selon lesquels la langue aide beaucoup quand il n'y a pas de son), on ne peut pas considérer ce résultat comme significativement différent de celui obtenu pour la présentation « Visage complet ». La figure 21 illustre les histogrammes des scores pour ces 2 séries. On observe qu'un sujet (naïf) a très bien réussi par rapport aux autres en modalité « Intérieur avec langue », et l'histogramme pour cette modalité est décalé à droite par rapport à l'histogramme pour la modalité « Visage complet ».

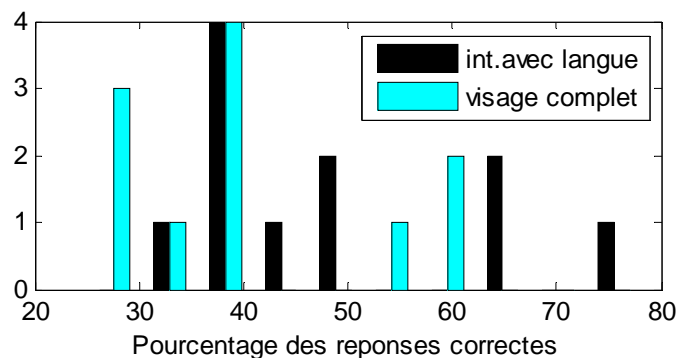


Fig. 21 – Histogrammes des scores d'intelligibilité en modalité « Intérieur avec langue » et « Visage complet » dans la condition de vidéo seule pour le groupe II

Les scores pour la série de généralisation (qui a été introduite pour tester si les sujets apprennent à lire l'information sur la langue, ou apprennent les stimuli concrets du test) sont les plus élevés parmi les scores pour le niveau du bruit -9 dB et sont proches pour les 2 groupes [$F(1, 10)=0.61$; $p>0.44$]. La différence entre les scores de cette série et de la modalité « Visage complet » avec niveau du bruit -9 dB n'est pas significative [groupe I : $F(1, 10)=4.74$; $p>0.05$; groupe II : $F(1, 10)=3.04$; $p>0.05$]. La différence entre les scores de la série de généralisation et de la modalité « Intérieur avec langue » avec niveau du bruit -9 dB est significative [groupe I : $F(1, 10)=23.68$; $p<0.001$; groupe II : $F(1, 10)=8.92$; $p>0.01$].

Comme l'ordre des modes de présentation dans le test a été constant pour tous les sujets, on ne peut pas conclure, si les scores montent en passant d'un mode de présentation à un autre parce que ce mode de présentation permet de transmettre plus d'information au sujet, ou à cause de l'effet de l'autre facteur. Les autres raisons possibles sont : (1) les sujets apprennent les stimuli du test ; (2) les sujets s'habituent au visage synthétique (par exemple, les lèvres sont présentées dans tous les modes de présentation).

Le fait que les scores pour la série de généralisation soient significativement plus élevés que ceux qui correspondent aux mêmes conditions audiovisuelles (modalité « Intérieur avec langue » avec le niveau du bruit -9 dB) semble confirmer l'hypothèse selon laquelle pendant la session du test les sujets apprennent implicitement à lire sur la langue. Il faut cependant considérer cette interprétation avec prudence. En effet, Benoît et al. (1994b) ont montré que le

contexte vocalique influence l'intelligibilité des consonnes voisines : l'amélioration du score pourrait être due au fait que dans la série de généralisation des stimuli avec d'autres voyelles sont utilisés.

3.2.2 Résultats détaillés

Il existe des sons qui sont plus facilement confondus entre eux que d'autres. MacLeod et Summerfield, 1987, ont construit l'arbre de confusion audio pour les consonnes de l'anglais (figure 22). Il y a des consonnes qui sont plus ou moins proches au niveau de la perception visuelle. L'arbre de confusion visuel a été construit par Walden et al. (1977) pour les consonnes de l'anglais (figure 22).

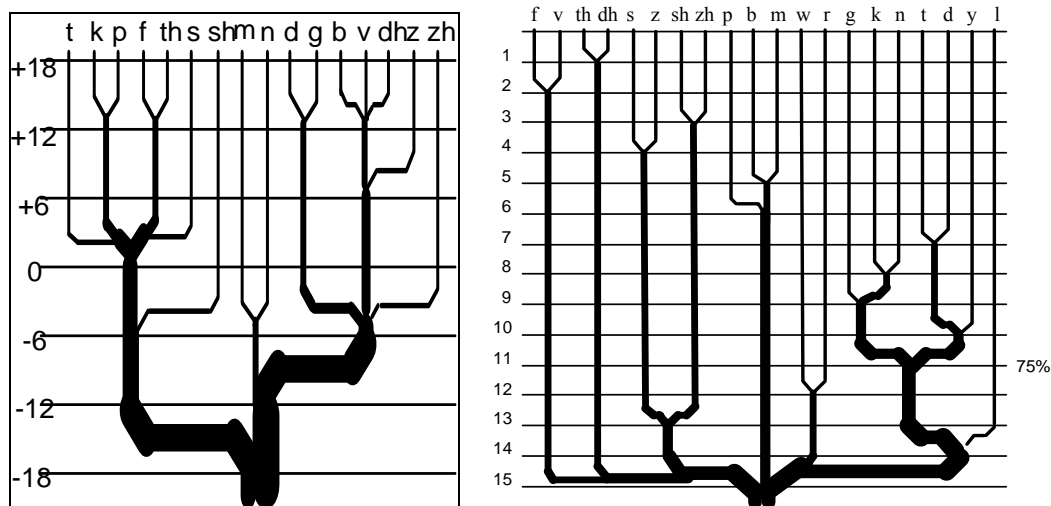


Fig. 22 – Arbre de confusion audio (MacLeod et Summerfield, 1987, à gauche) et arbre de confusion visuel (Walden et al, 1977, à droite)

Analysons les confusions des consonnes dans notre expérience. Les figures 23-24 représentent les matrices de confusion pour les 2 groupes pour chaque série du test. Chaque ligne de matrice explique la distribution des réponses pour la consonne à l'entrée (en pourcentage).

On observe que les sujets confondaient le plus souvent les paires suivantes : /b/ et /v/; /z/ et /□/; /d/ et /g/. Dans la condition « Audio seul », outre ces paires, les participants confondaient aussi les consonnes /z/, /□/ et /v/.

Ces résultats sont cohérents avec les arbres de confusion présentés sur la figure 22. Les consonnes /d/ et /g/, /z/ et /□/ sont proches dans les arbres de confusion audio et visuelle. Les confusions au niveau audio seul entre /b/ et /v/ sont très largement réduites pour toutes les modalités vidéo (ces consonnes sont situées près l'une de l'autre dans l'arbre de confusion audio, mais loin l'une de l'autre dans l'arbre de confusion visuelle).

Analysons maintenant si la visualisation de la langue a diminué la confusion entre les 3 groupes des consonnes, classés par lieu d'articulation. On observe dans les figures 23-24 qu'en modalité « Intérieur avec langue » les participants continuent à confondre les consonnes /d/ et /g/, pour lesquelles la position de la langue est significativement différente (/d/ est la consonne dentale et /g/ est palato-vélaire).

	- infini	-9 dB	+3 dB	+ infini																																																																																																																																																																																																																																																																																																																																				
Son seul		<table border="1"> <tr><td>B</td><td>17</td><td>15</td><td>13</td><td>6</td><td>-</td><td>42</td><td>8</td><td>-</td></tr> <tr><td>V</td><td>13</td><td>25</td><td>23</td><td>8</td><td>-</td><td>19</td><td>10</td><td>2</td></tr> <tr><td>J</td><td>10</td><td>25</td><td>19</td><td>15</td><td>4</td><td>4</td><td>4</td><td>19</td></tr> <tr><td>Z</td><td>6</td><td>33</td><td>23</td><td>15</td><td>-</td><td>8</td><td>6</td><td>8</td></tr> <tr><td>L</td><td>2</td><td>17</td><td>15</td><td>-</td><td>56</td><td>4</td><td>2</td><td>4</td></tr> <tr><td>D</td><td>13</td><td>10</td><td>10</td><td>10</td><td>-</td><td>42</td><td>15</td><td>-</td></tr> <tr><td>G</td><td>15</td><td>15</td><td>8</td><td>10</td><td>2</td><td>38</td><td>13</td><td>-</td></tr> <tr><td>R</td><td>13</td><td>10</td><td>19</td><td>2</td><td>-</td><td>8</td><td>4</td><td>44</td></tr> <tr><td></td><td>B</td><td>V</td><td>J</td><td>Z</td><td>L</td><td>D</td><td>G</td><td>R</td></tr> </table>	B	17	15	13	6	-	42	8	-	V	13	25	23	8	-	19	10	2	J	10	25	19	15	4	4	4	19	Z	6	33	23	15	-	8	6	8	L	2	17	15	-	56	4	2	4	D	13	10	10	10	-	42	15	-	G	15	15	8	10	2	38	13	-	R	13	10	19	2	-	8	4	44		B	V	J	Z	L	D	G	R	<table border="1"> <tr><td>B</td><td>46</td><td>19</td><td>2</td><td>2</td><td>-</td><td>17</td><td>15</td><td>-</td></tr> <tr><td>V</td><td>4</td><td>60</td><td>17</td><td>4</td><td>-</td><td>-</td><td>-</td><td>15</td></tr> <tr><td>J</td><td>-</td><td>10</td><td>46</td><td>13</td><td>-</td><td>2</td><td>2</td><td>27</td></tr> <tr><td>Z</td><td>-</td><td>25</td><td>21</td><td>29</td><td>2</td><td>-</td><td>2</td><td>21</td></tr> <tr><td>L</td><td>4</td><td>4</td><td>2</td><td>2</td><td>83</td><td>2</td><td>-</td><td>2</td></tr> <tr><td>D</td><td>15</td><td>6</td><td>2</td><td>4</td><td>-</td><td>52</td><td>21</td><td>-</td></tr> <tr><td>G</td><td>17</td><td>19</td><td>4</td><td>-</td><td>2</td><td>23</td><td>31</td><td>4</td></tr> <tr><td>R</td><td>-</td><td>-</td><td>-</td><td>-</td><td>2</td><td>-</td><td>98</td><td>-</td></tr> <tr><td></td><td>B</td><td>V</td><td>J</td><td>Z</td><td>L</td><td>D</td><td>G</td><td>R</td></tr> </table>	B	46	19	2	2	-	17	15	-	V	4	60	17	4	-	-	-	15	J	-	10	46	13	-	2	2	27	Z	-	25	21	29	2	-	2	21	L	4	4	2	2	83	2	-	2	D	15	6	2	4	-	52	21	-	G	17	19	4	-	2	23	31	4	R	-	-	-	-	2	-	98	-		B	V	J	Z	L	D	G	R	<table border="1"> <tr><td>B</td><td>98</td><td>2</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>V</td><td>-</td><td>98</td><td>2</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>J</td><td>-</td><td>-</td><td>98</td><td>-</td><td>-</td><td>-</td><td>-</td><td>2</td></tr> <tr><td>Z</td><td>-</td><td>-</td><td>-</td><td>100</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>L</td><td>-</td><td>-</td><td>2</td><td>-</td><td>98</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>D</td><td>-</td><td>-</td><td>-</td><td>4</td><td>-</td><td>94</td><td>2</td><td>-</td></tr> <tr><td>G</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>100</td></tr> <tr><td>R</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>100</td></tr> <tr><td></td><td>B</td><td>V</td><td>J</td><td>Z</td><td>L</td><td>D</td><td>G</td><td>R</td></tr> </table>	B	98	2	-	-	-	-	-	-	V	-	98	2	-	-	-	-	-	J	-	-	98	-	-	-	-	2	Z	-	-	-	100	-	-	-	-	L	-	-	2	-	98	-	-	-	D	-	-	-	4	-	94	2	-	G	-	-	-	-	-	-	-	100	R	-	-	-	-	-	-	-	100		B	V	J	Z	L	D	G	R																																																																																	
B	17	15	13	6	-	42	8	-																																																																																																																																																																																																																																																																																																																																
V	13	25	23	8	-	19	10	2																																																																																																																																																																																																																																																																																																																																
J	10	25	19	15	4	4	4	19																																																																																																																																																																																																																																																																																																																																
Z	6	33	23	15	-	8	6	8																																																																																																																																																																																																																																																																																																																																
L	2	17	15	-	56	4	2	4																																																																																																																																																																																																																																																																																																																																
D	13	10	10	10	-	42	15	-																																																																																																																																																																																																																																																																																																																																
G	15	15	8	10	2	38	13	-																																																																																																																																																																																																																																																																																																																																
R	13	10	19	2	-	8	4	44																																																																																																																																																																																																																																																																																																																																
	B	V	J	Z	L	D	G	R																																																																																																																																																																																																																																																																																																																																
B	46	19	2	2	-	17	15	-																																																																																																																																																																																																																																																																																																																																
V	4	60	17	4	-	-	-	15																																																																																																																																																																																																																																																																																																																																
J	-	10	46	13	-	2	2	27																																																																																																																																																																																																																																																																																																																																
Z	-	25	21	29	2	-	2	21																																																																																																																																																																																																																																																																																																																																
L	4	4	2	2	83	2	-	2																																																																																																																																																																																																																																																																																																																																
D	15	6	2	4	-	52	21	-																																																																																																																																																																																																																																																																																																																																
G	17	19	4	-	2	23	31	4																																																																																																																																																																																																																																																																																																																																
R	-	-	-	-	2	-	98	-																																																																																																																																																																																																																																																																																																																																
	B	V	J	Z	L	D	G	R																																																																																																																																																																																																																																																																																																																																
B	98	2	-	-	-	-	-	-																																																																																																																																																																																																																																																																																																																																
V	-	98	2	-	-	-	-	-																																																																																																																																																																																																																																																																																																																																
J	-	-	98	-	-	-	-	2																																																																																																																																																																																																																																																																																																																																
Z	-	-	-	100	-	-	-	-																																																																																																																																																																																																																																																																																																																																
L	-	-	2	-	98	-	-	-																																																																																																																																																																																																																																																																																																																																
D	-	-	-	4	-	94	2	-																																																																																																																																																																																																																																																																																																																																
G	-	-	-	-	-	-	-	100																																																																																																																																																																																																																																																																																																																																
R	-	-	-	-	-	-	-	100																																																																																																																																																																																																																																																																																																																																
	B	V	J	Z	L	D	G	R																																																																																																																																																																																																																																																																																																																																
Intérieur sans langue	<table border="1"> <tr><td>B</td><td>48</td><td>10</td><td>4</td><td>2</td><td>8</td><td>8</td><td>4</td><td>15</td></tr> <tr><td>V</td><td>2</td><td>71</td><td>6</td><td>2</td><td>6</td><td>4</td><td>6</td><td>2</td></tr> <tr><td>J</td><td>10</td><td>15</td><td>6</td><td>13</td><td>15</td><td>6</td><td>17</td><td>19</td></tr> <tr><td>Z</td><td>8</td><td>6</td><td>13</td><td>8</td><td>10</td><td>10</td><td>27</td><td>17</td></tr> <tr><td>L</td><td>13</td><td>10</td><td>8</td><td>4</td><td>19</td><td>13</td><td>15</td><td>19</td></tr> <tr><td>D</td><td>6</td><td>4</td><td>10</td><td>6</td><td>29</td><td>15</td><td>19</td><td>10</td></tr> <tr><td>G</td><td>10</td><td>6</td><td>6</td><td>6</td><td>29</td><td>6</td><td>13</td><td>23</td></tr> <tr><td>R</td><td>8</td><td>4</td><td>6</td><td>10</td><td>23</td><td>10</td><td>10</td><td>27</td></tr> <tr><td></td><td>B</td><td>V</td><td>J</td><td>Z</td><td>L</td><td>D</td><td>G</td><td>R</td></tr> </table>	B	48	10	4	2	8	8	4	15	V	2	71	6	2	6	4	6	2	J	10	15	6	13	15	6	17	19	Z	8	6	13	8	10	10	27	17	L	13	10	8	4	19	13	15	19	D	6	4	10	6	29	15	19	10	G	10	6	6	6	29	6	13	23	R	8	4	6	10	23	10	10	27		B	V	J	Z	L	D	G	R	<table border="1"> <tr><td>B</td><td>67</td><td>23</td><td>-</td><td>-</td><td>-</td><td>6</td><td>-</td><td>4</td></tr> <tr><td>V</td><td>4</td><td>85</td><td>6</td><td>-</td><td>-</td><td>-</td><td>-</td><td>4</td></tr> <tr><td>J</td><td>-</td><td>13</td><td>17</td><td>15</td><td>6</td><td>2</td><td>8</td><td>40</td></tr> <tr><td>Z</td><td>-</td><td>13</td><td>29</td><td>15</td><td>6</td><td>-</td><td>17</td><td>21</td></tr> <tr><td>L</td><td>-</td><td>6</td><td>2</td><td>-</td><td>75</td><td>-</td><td>8</td><td>8</td></tr> <tr><td>D</td><td>2</td><td>4</td><td>8</td><td>-</td><td>-</td><td>23</td><td>52</td><td>10</td></tr> <tr><td>G</td><td>2</td><td>6</td><td>13</td><td>4</td><td>-</td><td>21</td><td>50</td><td>4</td></tr> <tr><td>R</td><td>-</td><td>2</td><td>-</td><td>-</td><td>-</td><td>2</td><td>2</td><td>94</td></tr> <tr><td></td><td>B</td><td>V</td><td>J</td><td>Z</td><td>L</td><td>D</td><td>G</td><td>R</td></tr> </table>	B	67	23	-	-	-	6	-	4	V	4	85	6	-	-	-	-	4	J	-	13	17	15	6	2	8	40	Z	-	13	29	15	6	-	17	21	L	-	6	2	-	75	-	8	8	D	2	4	8	-	-	23	52	10	G	2	6	13	4	-	21	50	4	R	-	2	-	-	-	2	2	94		B	V	J	Z	L	D	G	R	<table border="1"> <tr><td>B</td><td>71</td><td>23</td><td>-</td><td>-</td><td>-</td><td>2</td><td>4</td><td>-</td></tr> <tr><td>V</td><td>-</td><td>96</td><td>2</td><td>-</td><td>-</td><td>-</td><td>-</td><td>2</td></tr> <tr><td>J</td><td>-</td><td>10</td><td>60</td><td>-</td><td>2</td><td>-</td><td>-</td><td>27</td></tr> <tr><td>Z</td><td>-</td><td>21</td><td>25</td><td>17</td><td>2</td><td>-</td><td>6</td><td>29</td></tr> <tr><td>L</td><td>-</td><td>-</td><td>-</td><td>-</td><td>100</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>D</td><td>-</td><td>2</td><td>-</td><td>-</td><td>4</td><td>42</td><td>50</td><td>2</td></tr> <tr><td>G</td><td>2</td><td>15</td><td>4</td><td>-</td><td>-</td><td>10</td><td>69</td><td>-</td></tr> <tr><td>R</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>100</td></tr> <tr><td></td><td>B</td><td>V</td><td>J</td><td>Z</td><td>L</td><td>D</td><td>G</td><td>R</td></tr> </table>	B	71	23	-	-	-	2	4	-	V	-	96	2	-	-	-	-	2	J	-	10	60	-	2	-	-	27	Z	-	21	25	17	2	-	6	29	L	-	-	-	-	100	-	-	-	D	-	2	-	-	4	42	50	2	G	2	15	4	-	-	10	69	-	R	-	-	-	-	-	-	-	100		B	V	J	Z	L	D	G	R	<table border="1"> <tr><td>B</td><td>100</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>V</td><td>-</td><td>100</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>J</td><td>-</td><td>-</td><td>100</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>Z</td><td>-</td><td>-</td><td>-</td><td>100</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>L</td><td>-</td><td>-</td><td>-</td><td>-</td><td>100</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>D</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>96</td><td>4</td><td>-</td></tr> <tr><td>G</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>100</td></tr> <tr><td>R</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>100</td></tr> <tr><td></td><td>B</td><td>V</td><td>J</td><td>Z</td><td>L</td><td>D</td><td>G</td><td>R</td></tr> </table>	B	100	-	-	-	-	-	-	-	V	-	100	-	-	-	-	-	-	J	-	-	100	-	-	-	-	-	Z	-	-	-	100	-	-	-	-	L	-	-	-	-	100	-	-	-	D	-	-	-	-	-	96	4	-	G	-	-	-	-	-	-	-	100	R	-	-	-	-	-	-	-	100		B	V	J	Z	L	D	G	R
B	48	10	4	2	8	8	4	15																																																																																																																																																																																																																																																																																																																																
V	2	71	6	2	6	4	6	2																																																																																																																																																																																																																																																																																																																																
J	10	15	6	13	15	6	17	19																																																																																																																																																																																																																																																																																																																																
Z	8	6	13	8	10	10	27	17																																																																																																																																																																																																																																																																																																																																
L	13	10	8	4	19	13	15	19																																																																																																																																																																																																																																																																																																																																
D	6	4	10	6	29	15	19	10																																																																																																																																																																																																																																																																																																																																
G	10	6	6	6	29	6	13	23																																																																																																																																																																																																																																																																																																																																
R	8	4	6	10	23	10	10	27																																																																																																																																																																																																																																																																																																																																
	B	V	J	Z	L	D	G	R																																																																																																																																																																																																																																																																																																																																
B	67	23	-	-	-	6	-	4																																																																																																																																																																																																																																																																																																																																
V	4	85	6	-	-	-	-	4																																																																																																																																																																																																																																																																																																																																
J	-	13	17	15	6	2	8	40																																																																																																																																																																																																																																																																																																																																
Z	-	13	29	15	6	-	17	21																																																																																																																																																																																																																																																																																																																																
L	-	6	2	-	75	-	8	8																																																																																																																																																																																																																																																																																																																																
D	2	4	8	-	-	23	52	10																																																																																																																																																																																																																																																																																																																																
G	2	6	13	4	-	21	50	4																																																																																																																																																																																																																																																																																																																																
R	-	2	-	-	-	2	2	94																																																																																																																																																																																																																																																																																																																																
	B	V	J	Z	L	D	G	R																																																																																																																																																																																																																																																																																																																																
B	71	23	-	-	-	2	4	-																																																																																																																																																																																																																																																																																																																																
V	-	96	2	-	-	-	-	2																																																																																																																																																																																																																																																																																																																																
J	-	10	60	-	2	-	-	27																																																																																																																																																																																																																																																																																																																																
Z	-	21	25	17	2	-	6	29																																																																																																																																																																																																																																																																																																																																
L	-	-	-	-	100	-	-	-																																																																																																																																																																																																																																																																																																																																
D	-	2	-	-	4	42	50	2																																																																																																																																																																																																																																																																																																																																
G	2	15	4	-	-	10	69	-																																																																																																																																																																																																																																																																																																																																
R	-	-	-	-	-	-	-	100																																																																																																																																																																																																																																																																																																																																
	B	V	J	Z	L	D	G	R																																																																																																																																																																																																																																																																																																																																
B	100	-	-	-	-	-	-	-																																																																																																																																																																																																																																																																																																																																
V	-	100	-	-	-	-	-	-																																																																																																																																																																																																																																																																																																																																
J	-	-	100	-	-	-	-	-																																																																																																																																																																																																																																																																																																																																
Z	-	-	-	100	-	-	-	-																																																																																																																																																																																																																																																																																																																																
L	-	-	-	-	100	-	-	-																																																																																																																																																																																																																																																																																																																																
D	-	-	-	-	-	96	4	-																																																																																																																																																																																																																																																																																																																																
G	-	-	-	-	-	-	-	100																																																																																																																																																																																																																																																																																																																																
R	-	-	-	-	-	-	-	100																																																																																																																																																																																																																																																																																																																																
	B	V	J	Z	L	D	G	R																																																																																																																																																																																																																																																																																																																																
Intérieur avec langue	<table border="1"> <tr><td>B</td><td>48</td><td>13</td><td>2</td><td>4</td><td>8</td><td>6</td><td>4</td><td>15</td></tr> <tr><td>V</td><td>6</td><td>63</td><td>4</td><td>2</td><td>8</td><td>4</td><td>6</td><td>6</td></tr> <tr><td>J</td><td>2</td><td>6</td><td>15</td><td>2</td><td>33</td><td>21</td><td>8</td><td>13</td></tr> <tr><td>Z</td><td>2</td><td>4</td><td>15</td><td>13</td><td>25</td><td>23</td><td>10</td><td>8</td></tr> <tr><td>L</td><td>2</td><td>-</td><td>6</td><td>-</td><td>29</td><td>27</td><td>17</td><td>19</td></tr> <tr><td>D</td><td>-</td><td>-</td><td>6</td><td>10</td><td>31</td><td>29</td><td>8</td><td>15</td></tr> <tr><td>G</td><td>2</td><td>4</td><td>6</td><td>2</td><td>15</td><td>13</td><td>27</td><td>31</td></tr> <tr><td>R</td><td>-</td><td>8</td><td>6</td><td>2</td><td>13</td><td>2</td><td>33</td><td>35</td></tr> <tr><td></td><td>B</td><td>V</td><td>J</td><td>Z</td><td>L</td><td>D</td><td>G</td><td>R</td></tr> </table>	B	48	13	2	4	8	6	4	15	V	6	63	4	2	8	4	6	6	J	2	6	15	2	33	21	8	13	Z	2	4	15	13	25	23	10	8	L	2	-	6	-	29	27	17	19	D	-	-	6	10	31	29	8	15	G	2	4	6	2	15	13	27	31	R	-	8	6	2	13	2	33	35		B	V	J	Z	L	D	G	R	<table border="1"> <tr><td>B</td><td>69</td><td>19</td><td>-</td><td>-</td><td>-</td><td>4</td><td>4</td><td>4</td></tr> <tr><td>V</td><td>2</td><td>88</td><td>2</td><td>-</td><td>-</td><td>-</td><td>4</td><td>4</td></tr> <tr><td>J</td><td>-</td><td>10</td><td>21</td><td>8</td><td>10</td><td>6</td><td>8</td><td>35</td></tr> <tr><td>Z</td><td>2</td><td>15</td><td>31</td><td>13</td><td>4</td><td>4</td><td>13</td><td>19</td></tr> <tr><td>L</td><td>-</td><td>2</td><td>4</td><td>-</td><td>81</td><td>6</td><td>4</td><td>2</td></tr> <tr><td>D</td><td>4</td><td>2</td><td>4</td><td>4</td><td>2</td><td>44</td><td>33</td><td>6</td></tr> <tr><td>G</td><td>4</td><td>4</td><td>4</td><td>10</td><td>-</td><td>15</td><td>58</td><td>4</td></tr> <tr><td>R</td><td>-</td><td>2</td><td>-</td><td>2</td><td>-</td><td>-</td><td>2</td><td>94</td></tr> <tr><td></td><td>B</td><td>V</td><td>J</td><td>Z</td><td>L</td><td>D</td><td>G</td><td>R</td></tr> </table>	B	69	19	-	-	-	4	4	4	V	2	88	2	-	-	-	4	4	J	-	10	21	8	10	6	8	35	Z	2	15	31	13	4	4	13	19	L	-	2	4	-	81	6	4	2	D	4	2	4	4	2	44	33	6	G	4	4	4	10	-	15	58	4	R	-	2	-	2	-	-	2	94		B	V	J	Z	L	D	G	R	<table border="1"> <tr><td>B</td><td>75</td><td>17</td><td>-</td><td>-</td><td>-</td><td>4</td><td>4</td><td>-</td></tr> <tr><td>V</td><td>-</td><td>90</td><td>4</td><td>-</td><td>-</td><td>2</td><td>-</td><td>4</td></tr> <tr><td>J</td><td>-</td><td>8</td><td>67</td><td>2</td><td>4</td><td>-</td><td>-</td><td>19</td></tr> <tr><td>Z</td><td>-</td><td>10</td><td>29</td><td>31</td><td>6</td><td>2</td><td>2</td><td>19</td></tr> <tr><td>L</td><td>-</td><td>-</td><td>2</td><td>-</td><td>96</td><td>-</td><td>2</td><td>-</td></tr> <tr><td>D</td><td>4</td><td>2</td><td>-</td><td>4</td><td>2</td><td>54</td><td>33</td><td>-</td></tr> <tr><td>G</td><td>-</td><td>10</td><td>-</td><td>-</td><td>2</td><td>4</td><td>83</td><td>-</td></tr> <tr><td>R</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>100</td></tr> <tr><td></td><td>B</td><td>V</td><td>J</td><td>Z</td><td>L</td><td>D</td><td>G</td><td>R</td></tr> </table>	B	75	17	-	-	-	4	4	-	V	-	90	4	-	-	2	-	4	J	-	8	67	2	4	-	-	19	Z	-	10	29	31	6	2	2	19	L	-	-	2	-	96	-	2	-	D	4	2	-	4	2	54	33	-	G	-	10	-	-	2	4	83	-	R	-	-	-	-	-	-	-	100		B	V	J	Z	L	D	G	R	<table border="1"> <tr><td>B</td><td>100</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>V</td><td>-</td><td>100</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>J</td><td>-</td><td>-</td><td>98</td><td>-</td><td>-</td><td>-</td><td>-</td><td>2</td></tr> <tr><td>Z</td><td>-</td><td>-</td><td>-</td><td>100</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>L</td><td>-</td><td>-</td><td>-</td><td>-</td><td>100</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>D</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>98</td><td>2</td><td>-</td></tr> <tr><td>G</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>100</td></tr> <tr><td>R</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>100</td></tr> <tr><td></td><td>B</td><td>V</td><td>J</td><td>Z</td><td>L</td><td>D</td><td>G</td><td>R</td></tr> </table>	B	100	-	-	-	-	-	-	-	V	-	100	-	-	-	-	-	-	J	-	-	98	-	-	-	-	2	Z	-	-	-	100	-	-	-	-	L	-	-	-	-	100	-	-	-	D	-	-	-	-	-	98	2	-	G	-	-	-	-	-	-	-	100	R	-	-	-	-	-	-	-	100		B	V	J	Z	L	D	G	R
B	48	13	2	4	8	6	4	15																																																																																																																																																																																																																																																																																																																																
V	6	63	4	2	8	4	6	6																																																																																																																																																																																																																																																																																																																																
J	2	6	15	2	33	21	8	13																																																																																																																																																																																																																																																																																																																																
Z	2	4	15	13	25	23	10	8																																																																																																																																																																																																																																																																																																																																
L	2	-	6	-	29	27	17	19																																																																																																																																																																																																																																																																																																																																
D	-	-	6	10	31	29	8	15																																																																																																																																																																																																																																																																																																																																
G	2	4	6	2	15	13	27	31																																																																																																																																																																																																																																																																																																																																
R	-	8	6	2	13	2	33	35																																																																																																																																																																																																																																																																																																																																
	B	V	J	Z	L	D	G	R																																																																																																																																																																																																																																																																																																																																
B	69	19	-	-	-	4	4	4																																																																																																																																																																																																																																																																																																																																
V	2	88	2	-	-	-	4	4																																																																																																																																																																																																																																																																																																																																
J	-	10	21	8	10	6	8	35																																																																																																																																																																																																																																																																																																																																
Z	2	15	31	13	4	4	13	19																																																																																																																																																																																																																																																																																																																																
L	-	2	4	-	81	6	4	2																																																																																																																																																																																																																																																																																																																																
D	4	2	4	4	2	44	33	6																																																																																																																																																																																																																																																																																																																																
G	4	4	4	10	-	15	58	4																																																																																																																																																																																																																																																																																																																																
R	-	2	-	2	-	-	2	94																																																																																																																																																																																																																																																																																																																																
	B	V	J	Z	L	D	G	R																																																																																																																																																																																																																																																																																																																																
B	75	17	-	-	-	4	4	-																																																																																																																																																																																																																																																																																																																																
V	-	90	4	-	-	2	-	4																																																																																																																																																																																																																																																																																																																																
J	-	8	67	2	4	-	-	19																																																																																																																																																																																																																																																																																																																																
Z	-	10	29	31	6	2	2	19																																																																																																																																																																																																																																																																																																																																
L	-	-	2	-	96	-	2	-																																																																																																																																																																																																																																																																																																																																
D	4	2	-	4	2	54	33	-																																																																																																																																																																																																																																																																																																																																
G	-	10	-	-	2	4	83	-																																																																																																																																																																																																																																																																																																																																
R	-	-	-	-	-	-	-	100																																																																																																																																																																																																																																																																																																																																
	B	V	J	Z	L	D	G	R																																																																																																																																																																																																																																																																																																																																
B	100	-	-	-	-	-	-	-																																																																																																																																																																																																																																																																																																																																
V	-	100	-	-	-	-	-	-																																																																																																																																																																																																																																																																																																																																
J	-	-	98	-	-	-	-	2																																																																																																																																																																																																																																																																																																																																
Z	-	-	-	100	-	-	-	-																																																																																																																																																																																																																																																																																																																																
L	-	-	-	-	100	-	-	-																																																																																																																																																																																																																																																																																																																																
D	-	-	-	-	-	98	2	-																																																																																																																																																																																																																																																																																																																																
G	-	-	-	-	-	-	-	100																																																																																																																																																																																																																																																																																																																																
R	-	-	-	-	-	-	-	100																																																																																																																																																																																																																																																																																																																																
	B	V	J	Z	L	D	G	R																																																																																																																																																																																																																																																																																																																																
Visage complet	<table border="1"> <tr><td>B</td><td>56</td><td>17</td><td>4</td><td>-</td><td>8</td><td>-</td><td>8</td><td>6</td></tr> <tr><td>V</td><td>-</td><td>88</td><td>4</td><td>4</td><td>2</td><td>-</td><td>-</td><td>2</td></tr> <tr><td>J</td><td>4</td><td>17</td><td>19</td><td>4</td><td>15</td><td>10</td><td>23</td><td>8</td></tr> <tr><td>Z</td><td>2</td><td>6</td><td>15</td><td>23</td><td>15</td><td>19</td><td>19</td><td>2</td></tr> <tr><td>L</td><td>-</td><td>4</td><td>10</td><td>4</td><td>35</td><td>13</td><td>10</td><td>23</td></tr> <tr><td>D</td><td>6</td><td>6</td><td>6</td><td>15</td><td>19</td><td>19</td><td>15</td><td>15</td></tr> <tr><td>G</td><td>2</td><td>17</td><td>15</td><td>4</td><td>17</td><td>15</td><td>19</td><td>13</td></tr> <tr><td>R</td><td>2</td><td>10</td><td>13</td><td>8</td><td>31</td><td>2</td><td>19</td><td>15</td></tr> <tr><td></td><td>B</td><td>V</td><td>J</td><td>Z</td><td>L</td><td>D</td><td>G</td><td>R</td></tr> </table>	B	56	17	4	-	8	-	8	6	V	-	88	4	4	2	-	-	2	J	4	17	19	4	15	10	23	8	Z	2	6	15	23	15	19	19	2	L	-	4	10	4	35	13	10	23	D	6	6	6	15	19	19	15	15	G	2	17	15	4	17	15	19	13	R	2	10	13	8	31	2	19	15		B	V	J	Z	L	D	G	R	<table border="1"> <tr><td>B</td><td>79</td><td>15</td><td>-</td><td>-</td><td>-</td><td>6</td><td>-</td><td>-</td></tr> <tr><td>V</td><td>-</td><td>94</td><td>-</td><td>-</td><td>-</td><td>4</td><td>2</td><td>-</td></tr> <tr><td>J</td><td>-</td><td>15</td><td>38</td><td>13</td><td>-</td><td>2</td><td>10</td><td>23</td></tr> <tr><td>Z</td><td>2</td><td>10</td><td>27</td><td>31</td><td>2</td><td>10</td><td>6</td><td>10</td></tr> <tr><td>L</td><td>-</td><td>2</td><td>2</td><td>-</td><td>90</td><td>-</td><td>-</td><td>6</td></tr> <tr><td>D</td><td>-</td><td>4</td><td>2</td><td>8</td><td>-</td><td>44</td><td>33</td><td>8</td></tr> <tr><td>G</td><td>-</td><td>2</td><td>8</td><td>4</td><td>2</td><td>15</td><td>58</td><td>10</td></tr> <tr><td>R</td><td>-</td><td>4</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>96</td></tr> <tr><td></td><td>B</td><td>V</td><td>J</td><td>Z</td><td>L</td><td>D</td><td>G</td><td>R</td></tr> </table>	B	79	15	-	-	-	6	-	-	V	-	94	-	-	-	4	2	-	J	-	15	38	13	-	2	10	23	Z	2	10	27	31	2	10	6	10	L	-	2	2	-	90	-	-	6	D	-	4	2	8	-	44	33	8	G	-	2	8	4	2	15	58	10	R	-	4	-	-	-	-	-	96		B	V	J	Z	L	D	G	R	<table border="1"> <tr><td>B</td><td>85</td><td>15</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>V</td><td>-</td><td>98</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>2</td></tr> <tr><td>J</td><td>-</td><td>13</td><td>71</td><td>-</td><td>8</td><td>-</td><td>-</td><td>8</td></tr> <tr><td>Z</td><td>-</td><td>21</td><td>21</td><td>35</td><td>4</td><td>-</td><td>4</td><td>15</td></tr> <tr><td>L</td><td>-</td><td>-</td><td>8</td><td>-</td><td>92</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>D</td><td>-</td><td>-</td><td>-</td><td>2</td><td>2</td><td>67</td><td>29</td><td>-</td></tr> <tr><td>G</td><td>2</td><td>10</td><td>2</td><td>-</td><td>-</td><td>10</td><td>75</td><td>-</td></tr> <tr><td>R</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>100</td></tr> <tr><td></td><td>B</td><td>V</td><td>J</td><td>Z</td><td>L</td><td>D</td><td>G</td><td>R</td></tr> </table>	B	85	15	-	-	-	-	-	-	V	-	98	-	-	-	-	-	2	J	-	13	71	-	8	-	-	8	Z	-	21	21	35	4	-	4	15	L	-	-	8	-	92	-	-	-	D	-	-	-	2	2	67	29	-	G	2	10	2	-	-	10	75	-	R	-	-	-	-	-	-	-	100		B	V	J	Z	L	D	G	R	<table border="1"> <tr><td>B</td><td>100</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>V</td><td>-</td><td>100</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>J</td><td>-</td><td>-</td><td>98</td><td>-</td><td>-</td><td>-</td><td>-</td><td>2</td></tr> <tr><td>Z</td><td>-</td><td>-</td><td>-</td><td>100</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>L</td><td>-</td><td>-</td><td>-</td><td>-</td><td>100</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>D</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>96</td><td>4</td><td>-</td></tr> <tr><td>G</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>100</td></tr> <tr><td>R</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>100</td></tr> <tr><td></td><td>B</td><td>V</td><td>J</td><td>Z</td><td>L</td><td>D</td><td>G</td><td>R</td></tr> </table>	B	100	-	-	-	-	-	-	-	V	-	100	-	-	-	-	-	-	J	-	-	98	-	-	-	-	2	Z	-	-	-	100	-	-	-	-	L	-	-	-	-	100	-	-	-	D	-	-	-	-	-	96	4	-	G	-	-	-	-	-	-	-	100	R	-	-	-	-	-	-	-	100		B	V	J	Z	L	D	G	R
B	56	17	4	-	8	-	8	6																																																																																																																																																																																																																																																																																																																																
V	-	88	4	4	2	-	-	2																																																																																																																																																																																																																																																																																																																																
J	4	17	19	4	15	10	23	8																																																																																																																																																																																																																																																																																																																																
Z	2	6	15	23	15	19	19	2																																																																																																																																																																																																																																																																																																																																
L	-	4	10	4	35	13	10	23																																																																																																																																																																																																																																																																																																																																
D	6	6	6	15	19	19	15	15																																																																																																																																																																																																																																																																																																																																
G	2	17	15	4	17	15	19	13																																																																																																																																																																																																																																																																																																																																
R	2	10	13	8	31	2	19	15																																																																																																																																																																																																																																																																																																																																
	B	V	J	Z	L	D	G	R																																																																																																																																																																																																																																																																																																																																
B	79	15	-	-	-	6	-	-																																																																																																																																																																																																																																																																																																																																
V	-	94	-	-	-	4	2	-																																																																																																																																																																																																																																																																																																																																
J	-	15	38	13	-	2	10	23																																																																																																																																																																																																																																																																																																																																
Z	2	10	27	31	2	10	6	10																																																																																																																																																																																																																																																																																																																																
L	-	2	2	-	90	-	-	6																																																																																																																																																																																																																																																																																																																																
D	-	4	2	8	-	44	33	8																																																																																																																																																																																																																																																																																																																																
G	-	2	8	4	2	15	58	10																																																																																																																																																																																																																																																																																																																																
R	-	4	-	-	-	-	-	96																																																																																																																																																																																																																																																																																																																																
	B	V	J	Z	L	D	G	R																																																																																																																																																																																																																																																																																																																																
B	85	15	-	-	-	-	-	-																																																																																																																																																																																																																																																																																																																																
V	-	98	-	-	-	-	-	2																																																																																																																																																																																																																																																																																																																																
J	-	13	71	-	8	-	-	8																																																																																																																																																																																																																																																																																																																																
Z	-	21	21	35	4	-	4	15																																																																																																																																																																																																																																																																																																																																
L	-	-	8	-	92	-	-	-																																																																																																																																																																																																																																																																																																																																
D	-	-	-	2	2	67	29	-																																																																																																																																																																																																																																																																																																																																
G	2	10	2	-	-	10	75	-																																																																																																																																																																																																																																																																																																																																
R	-	-	-	-	-	-	-	100																																																																																																																																																																																																																																																																																																																																
	B	V	J	Z	L	D	G	R																																																																																																																																																																																																																																																																																																																																
B	100	-	-	-	-	-	-	-																																																																																																																																																																																																																																																																																																																																
V	-	100	-	-	-	-	-	-																																																																																																																																																																																																																																																																																																																																
J	-	-	98	-	-	-	-	2																																																																																																																																																																																																																																																																																																																																
Z	-	-	-	100	-	-	-	-																																																																																																																																																																																																																																																																																																																																
L	-	-	-	-	100	-	-	-																																																																																																																																																																																																																																																																																																																																
D	-	-	-	-	-	96	4	-																																																																																																																																																																																																																																																																																																																																
G	-	-	-	-	-	-	-	100																																																																																																																																																																																																																																																																																																																																
R	-	-	-	-	-	-	-	100																																																																																																																																																																																																																																																																																																																																
	B	V	J	Z	L	D	G	R																																																																																																																																																																																																																																																																																																																																
Généralisation		<table border="1"> <tr><td>B</td><td>78</td><td>8</td><td>-</td><td>-</td><td>-</td><td>8</td><td>6</td><td>-</td></tr> <tr><td>V</td><td>-</td><td>92</td><td>-</td><td>3</td><td>-</td><td>-</td><td>3</td><td>3</td></tr> <tr><td>J</td><td>-</td><td>6</td><td>56</td><td>11</td><td>3</td><td>6</td><td>17</td><td>3</td></tr> <tr><td>Z</td><td>-</td><td>-</td><td>33</td><td>42</td><td>11</td><td>3</td><td>8</td><td>3</td></tr> <tr><td>L</td><td>-</td><td>-</td><td>3</td><td>-</td><td>97</td><td>-</td><td>-</td><td>-</td></tr> <tr><td>D</td><td>3</td><td>-</td><td>-</td><td>3</td><td>-</td><td>56</td><td>39</td><td>-</td></tr> <tr><td>G</td><td>6</td><td>-</td><td>-</td><td>-</td><td>-</td><td>14</td><td>78</td><td>3</td></tr> <tr><td>R</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>100</td></tr> <tr><td></td><td>B</td><td>V</td><td>J</td><td>Z</td><td>L</td><td>D</td><td>G</td><td>R</td></tr> </table>	B	78	8	-	-	-	8	6	-	V	-	92	-	3	-	-	3	3	J	-	6	56	11	3	6	17	3	Z	-	-	33	42	11	3	8	3	L	-	-	3	-	97	-	-	-	D	3	-	-	3	-	56	39	-	G	6	-	-	-	-	14	78	3	R	-	-	-	-	-	-	-	100		B	V	J	Z	L	D	G	R																																																																																																																																																																																																																																																					
B	78	8	-	-	-	8	6	-																																																																																																																																																																																																																																																																																																																																
V	-	92	-	3	-	-	3	3																																																																																																																																																																																																																																																																																																																																
J	-	6	56	11	3	6	17	3																																																																																																																																																																																																																																																																																																																																
Z	-	-	33	42	11	3	8	3																																																																																																																																																																																																																																																																																																																																
L	-	-	3	-	97	-	-	-																																																																																																																																																																																																																																																																																																																																
D	3	-	-	3	-	56	39	-																																																																																																																																																																																																																																																																																																																																
G	6	-	-	-	-	14	78	3																																																																																																																																																																																																																																																																																																																																
R	-	-	-	-	-	-	-	100																																																																																																																																																																																																																																																																																																																																
	B	V	J	Z	L	D	G	R																																																																																																																																																																																																																																																																																																																																

Fig. 23 – Matrices de confusion des consonnes pour le **groupe I** pour chaque série du test (B - /b/, V - /v/, J - /j/, Z - /z/, L - /l/, D - /d/, G - /g/, R - /r/ ; stimuli = ligne, réponses = colonne)

	- infini	-9 dB	+3 dB	+ infini
Son seul	B 14 2 11 16 9 25 18 5 V 11 20 20 16 5 9 14 5 J - 9 27 23 9 2 7 23 Z 2 11 36 23 2 2 9 14 L - 7 9 2 77 2 2 - D 5 2 20 9 - 34 25 5 G - 5 39 11 5 25 14 2 R 7 5 2 5 - 5 - 77	B 34 7 7 2 2 25 23 - V - 39 16 16 5 5 7 14 J - 16 59 5 14 - 5 2 Z - 18 32 27 7 7 5 5 L - 2 5 - 91 - 2 - D 7 5 5 7 - 45 32 - G 7 7 2 2 7 16 59 - R - - - - - - - 100	B 98 - - - - - 2 - V - 100 - - - - - - J - - 100 - - - - - Z - - - 100 - - - - L - - - - 100 - - - D - - - - 7 - 89 5 - G - - - - - - - 100 - R - - - - - - - - 100	
	B 64 18 2 - 5 7 2 2 V 2 89 - - - 7 2 - J 2 30 7 2 18 5 20 16 Z - 16 18 11 9 9 18 18 L 7 18 9 5 27 5 2 27 D 2 9 16 7 16 11 23 16 G 5 20 18 5 16 9 9 18 R - 11 16 2 27 11 7 25	B 70 23 - - - 2 2 2 V 5 93 - - - 2 - - J - 14 23 16 2 2 14 30 Z - 16 23 25 7 2 9 18 L - - - - 95 - 2 2 D - 2 7 9 - 27 45 9 G - 11 16 2 2 20 43 5 R - - 5 - - 5 - 91	B 68 14 2 - 2 11 2 - V - 98 - 2 - - - - J - 7 61 9 9 - - 14 Z - 7 25 34 5 - 11 18 L - - - - 100 - - - D - - 2 - - 61 36 - G - 5 2 2 2 9 77 2 R - - 2 - - - - 98	B 100 - - - - - - V - 100 - - - - - - J - - 100 - - - - - Z - - - 100 - - - - L - - - - 100 - - - D - - - - 2 - 93 5 - G - - - - - - 100 - R - - - - - - - 100
	B 55 11 5 2 9 2 7 9 V 9 86 2 - 2 - - - J 2 7 16 - 45 14 14 2 Z 5 7 11 25 18 25 7 2 L - 2 7 - 66 14 5 7 D - 5 5 5 9 66 5 7 G - 5 - 2 16 5 32 41 R - 5 7 2 16 5 18 48	B 61 20 5 2 - 5 7 - V 5 82 5 2 2 2 2 - J - - 34 9 25 5 5 23 Z - 2 23 30 9 18 7 11 L - 2 - 2 91 2 2 - D - 2 7 9 - 68 11 2 G - 7 7 5 2 18 57 5 R - 2 - - - 2 - 95	B 61 14 2 - - 11 11 - V - 84 7 - 5 - - 5 J - 2 73 7 11 - - 7 Z - - 34 36 11 11 2 5 L - - - - 100 - - - D - - 2 2 - 77 18 - G 2 2 - 2 5 9 77 2 R - - 2 - - - - 98	B 100 - - - - - - V - 100 - - - - - - J - - 100 - - - - - Z - - - 100 - - - - L - - - - 100 - - - D - - - - 2 - 93 5 - G - - - - - - 100 - R - - - - - - - 100
	B 78 8 - - - 8 6 - V - 92 - 3 - - 3 3 J - 6 56 11 3 6 17 3 Z - - 33 42 11 3 8 3 L - - 3 - 97 - - - D 3 - - 3 - 56 39 - G 6 - - - - 14 78 3 R - - - - - - - 100	B 77 20 - - - - 2 - V - 98 - 2 - - - - J - 9 52 9 2 - 5 23 Z - 18 25 41 - 7 7 2 L - 2 - - 95 2 - - D - - 7 7 - 66 20 - G 2 7 14 2 - 23 50 2 R - - - - - - - 100	B 84 11 - - 2 2 - - V - 100 - - - - - - J - 11 73 7 7 - - 2 Z - 7 30 43 9 5 7 - L - - - - 100 - - - D - - 2 5 - 77 16 - G 2 7 2 2 - 7 80 - R - 2 2 - - - - 95	B 98 - - - - 2 - - V - 100 - - - - - - J - - 98 - 2 - - - Z - - - 100 - - - - L - - - - 100 - - - D - - - 5 - 91 5 - G - - - - - - 100 - R - - - - - - - 100
	B 88 3 - - - 3 6 - V 9 79 6 3 - - 3 - J - - 64 9 12 3 12 - Z - - 33 33 18 9 6 - L - - 3 - 97 - - - D - - - 3 - 73 24 - G - - 6 - - 9 82 3 R - 3 - - - - - 97			

Fig. 24 – Matrices de confusion des consonnes pour le **groupe II** pour chaque série du test (B - /b/, V - /v/, J - /□/, Z - /z/, L - /l/, D - /d/, G - /g/, R - /□/ ; stimuli = ligne, réponses = colonne)

La figure 25 visualise le pourcentage d'identification correcte du lieu d'articulation pour les différentes séries du test. On observe que ces scores sont plus élevés que les scores d'identification des consonnes (ceci est logique, car on a 8 consonnes à choisir, et on n'a que 3 lieu d'articulation à choisir ; donc la probabilité de répondre au hasard correctement est plus grande). On voit aussi que le placement des courbes l'une par rapport à l'autre est très semblable à celui des scores d'identification de consonne. Les scores pour les séries avec la présentation « Intérieur avec langue » restent inférieurs aux scores du « Visage complet », sauf dans le cas de la vidéo seule pour le groupe II.

L'ANOVA a été effectuée sur l'ensemble de ces scores (les facteurs intra sujet étant la modalité de présentation (3 modalités) et le niveau S/B (4 niveaux), le facteur inter sujet étant le groupe). Les conclusions de signifiante des facteurs sont les mêmes que pour l'ensemble des scores globaux. Mais ici encore, le score de la série avec la présentation « Intérieur avec langue » en condition de vidéo seule pour la groupe II est significativement plus grand que celui de la série avec la présentation « Visage complet » [$F(1, 10)=4.48$; $p<0.01$].

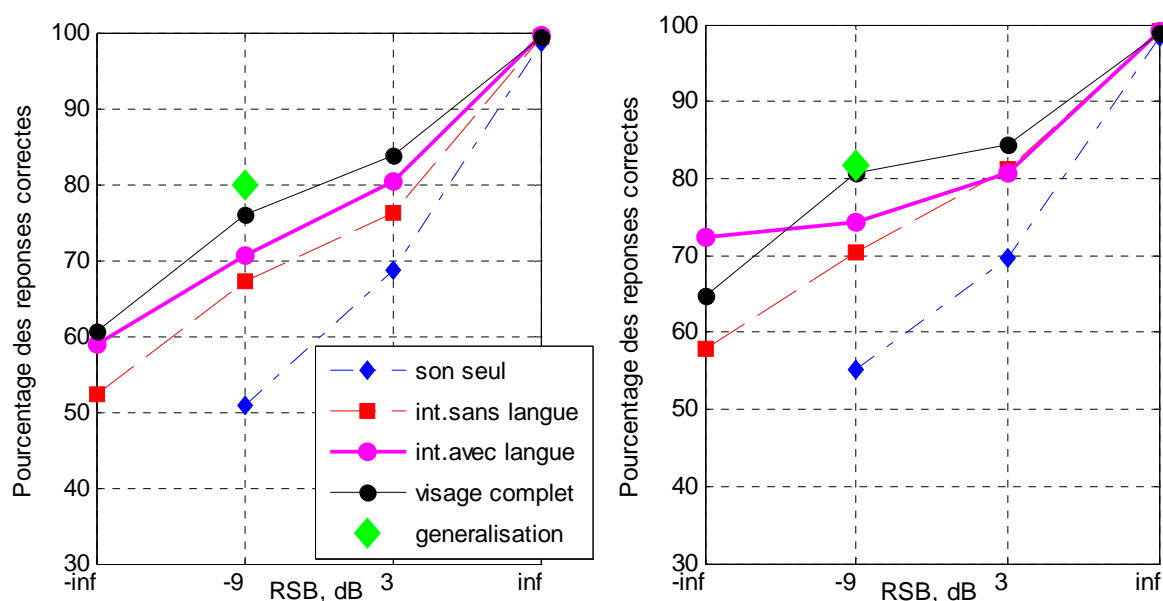


Fig. 25 – Pourcentage d'identification correcte de lieu d'articulation en fonction du bruit (à gauche, groupe I ; à droite, groupe II) pour les différentes conditions (de bas en haut) : audio seul; intérieur sans langue; intérieur avec langue; visage complet ; série de généralisation

Les figures 26-27 représentent les matrices de confusion de lieu d'articulation pour chaque série du test pour les 2 groupes. On observe qu'en modalité de présentation « Intérieur avec langue » les sujets ne confondent presque pas les consonnes labiales/ labiodentales avec les consonnes de 2 autres classes, mais confondent assez souvent les consonnes dentales/ alvéolaires/ post-alvéolaires avec palatales / palato-vélaires.

On tire des matrices de confusion les mêmes conclusions qu'avec les graphes : en condition vidéo seule la présentation « Intérieur avec langue » permet de mieux différencier le lieu d'articulation (les résultats de groupe II étant meilleurs que ceux du groupe I) ; en série de généralisation les sujets confondent moins le lieu d'articulation que dans les séries précédentes avec le même niveau du bruit.

	- infini	-9 dB	+3 dB	+ infini																																																																
Son seul		<table border="1"> <tr><td>BV</td><td>34</td><td>55</td><td>10</td></tr> <tr><td>JZLD</td><td>29</td><td>56</td><td>15</td></tr> <tr><td>GR</td><td>26</td><td>44</td><td>30</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	34	55	10	JZLD	29	56	15	GR	26	44	30		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>65</td><td>21</td><td>15</td></tr> <tr><td>JZLD</td><td>16</td><td>65</td><td>19</td></tr> <tr><td>GR</td><td>18</td><td>16</td><td>67</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	65	21	15	JZLD	16	65	19	GR	18	16	67		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>99</td><td>1</td><td>-</td></tr> <tr><td>JZLD</td><td>-</td><td>99</td><td>1</td></tr> <tr><td>GR</td><td>-</td><td>-</td><td>100</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	99	1	-	JZLD	-	99	1	GR	-	-	100		BV	JZLD	GR																
BV	34	55	10																																																																	
JZLD	29	56	15																																																																	
GR	26	44	30																																																																	
	BV	JZLD	GR																																																																	
BV	65	21	15																																																																	
JZLD	16	65	19																																																																	
GR	18	16	67																																																																	
	BV	JZLD	GR																																																																	
BV	99	1	-																																																																	
JZLD	-	99	1																																																																	
GR	-	-	100																																																																	
	BV	JZLD	GR																																																																	
Intérieur sans langue	<table border="1"> <tr><td>BV</td><td>66</td><td>21</td><td>14</td></tr> <tr><td>JZLD</td><td>18</td><td>46</td><td>35</td></tr> <tr><td>GR</td><td>15</td><td>49</td><td>36</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	66	21	14	JZLD	18	46	35	GR	15	49	36		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>90</td><td>6</td><td>4</td></tr> <tr><td>JZLD</td><td>9</td><td>49</td><td>41</td></tr> <tr><td>GR</td><td>5</td><td>20</td><td>75</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	90	6	4	JZLD	9	49	41	GR	5	20	75		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>95</td><td>2</td><td>3</td></tr> <tr><td>JZLD</td><td>8</td><td>63</td><td>29</td></tr> <tr><td>GR</td><td>8</td><td>7</td><td>84</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	95	2	3	JZLD	8	63	29	GR	8	7	84		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>100</td><td>-</td><td>-</td></tr> <tr><td>JZLD</td><td>-</td><td>99</td><td>1</td></tr> <tr><td>GR</td><td>-</td><td>-</td><td>100</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	100	-	-	JZLD	-	99	1	GR	-	-	100		BV	JZLD	GR
BV	66	21	14																																																																	
JZLD	18	46	35																																																																	
GR	15	49	36																																																																	
	BV	JZLD	GR																																																																	
BV	90	6	4																																																																	
JZLD	9	49	41																																																																	
GR	5	20	75																																																																	
	BV	JZLD	GR																																																																	
BV	95	2	3																																																																	
JZLD	8	63	29																																																																	
GR	8	7	84																																																																	
	BV	JZLD	GR																																																																	
BV	100	-	-																																																																	
JZLD	-	99	1																																																																	
GR	-	-	100																																																																	
	BV	JZLD	GR																																																																	
Intérieur avec langue	<table border="1"> <tr><td>BV</td><td>65</td><td>20</td><td>16</td></tr> <tr><td>JZLD</td><td>4</td><td>71</td><td>24</td></tr> <tr><td>GR</td><td>7</td><td>29</td><td>64</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	65	20	16	JZLD	4	71	24	GR	7	29	64		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>89</td><td>3</td><td>8</td></tr> <tr><td>JZLD</td><td>9</td><td>61</td><td>30</td></tr> <tr><td>GR</td><td>5</td><td>16</td><td>79</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	89	3	8	JZLD	9	61	30	GR	5	16	79		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>91</td><td>5</td><td>4</td></tr> <tr><td>JZLD</td><td>6</td><td>75</td><td>19</td></tr> <tr><td>GR</td><td>5</td><td>3</td><td>92</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	91	5	4	JZLD	6	75	19	GR	5	3	92		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>100</td><td>-</td><td>-</td></tr> <tr><td>JZLD</td><td>-</td><td>99</td><td>1</td></tr> <tr><td>GR</td><td>-</td><td>-</td><td>100</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	100	-	-	JZLD	-	99	1	GR	-	-	100		BV	JZLD	GR
BV	65	20	16																																																																	
JZLD	4	71	24																																																																	
GR	7	29	64																																																																	
	BV	JZLD	GR																																																																	
BV	89	3	8																																																																	
JZLD	9	61	30																																																																	
GR	5	16	79																																																																	
	BV	JZLD	GR																																																																	
BV	91	5	4																																																																	
JZLD	6	75	19																																																																	
GR	5	3	92																																																																	
	BV	JZLD	GR																																																																	
BV	100	-	-																																																																	
JZLD	-	99	1																																																																	
GR	-	-	100																																																																	
	BV	JZLD	GR																																																																	
Visage complet	<table border="1"> <tr><td>BV</td><td>80</td><td>11</td><td>8</td></tr> <tr><td>JZLD</td><td>11</td><td>60</td><td>29</td></tr> <tr><td>GR</td><td>16</td><td>52</td><td>32</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	80	11	8	JZLD	11	60	29	GR	16	52	32		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>94</td><td>5</td><td>1</td></tr> <tr><td>JZLD</td><td>8</td><td>67</td><td>24</td></tr> <tr><td>GR</td><td>3</td><td>15</td><td>82</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	94	5	1	JZLD	8	67	24	GR	3	15	82		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>99</td><td>-</td><td>1</td></tr> <tr><td>JZLD</td><td>8</td><td>78</td><td>14</td></tr> <tr><td>GR</td><td>6</td><td>6</td><td>88</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	99	-	1	JZLD	8	78	14	GR	6	6	88		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>100</td><td>-</td><td>-</td></tr> <tr><td>JZLD</td><td>-</td><td>98</td><td>2</td></tr> <tr><td>GR</td><td>-</td><td>-</td><td>100</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	100	-	-	JZLD	-	98	2	GR	-	-	100		BV	JZLD	GR
BV	80	11	8																																																																	
JZLD	11	60	29																																																																	
GR	16	52	32																																																																	
	BV	JZLD	GR																																																																	
BV	94	5	1																																																																	
JZLD	8	67	24																																																																	
GR	3	15	82																																																																	
	BV	JZLD	GR																																																																	
BV	99	-	1																																																																	
JZLD	8	78	14																																																																	
GR	6	6	88																																																																	
	BV	JZLD	GR																																																																	
BV	100	-	-																																																																	
JZLD	-	98	2																																																																	
GR	-	-	100																																																																	
	BV	JZLD	GR																																																																	
Généralisation		<table border="1"> <tr><td>BV</td><td>89</td><td>6</td><td>6</td></tr> <tr><td>JZLD</td><td>2</td><td>81</td><td>17</td></tr> <tr><td>GR</td><td>3</td><td>7</td><td>90</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	89	6	6	JZLD	2	81	17	GR	3	7	90		BV	JZLD	GR																																																		
BV	89	6	6																																																																	
JZLD	2	81	17																																																																	
GR	3	7	90																																																																	
	BV	JZLD	GR																																																																	

Fig. 26 – Matrices de confusion de lieu d’articulation du **groupe I** pour chaque série du test (B - /b/, V - /v/, J - /□/, Z - /z/, L - /l/, D - /d/, G - /g/, R - /□/ ; stimuli = ligne, réponses = colonne)

	- infini	-9 dB	+3 dB	+ infini																																																																
Son seul		<table border="1"> <tr><td>BV</td><td>24</td><td>56</td><td>20</td></tr> <tr><td>JZLD</td><td>9</td><td>70</td><td>21</td></tr> <tr><td>GR</td><td>8</td><td>45</td><td>47</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	24	56	20	JZLD	9	70	21	GR	8	45	47		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>40</td><td>39</td><td>22</td></tr> <tr><td>JZLD</td><td>12</td><td>76</td><td>13</td></tr> <tr><td>GR</td><td>7</td><td>14</td><td>80</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	40	39	22	JZLD	12	76	13	GR	7	14	80		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>99</td><td>-</td><td>1</td></tr> <tr><td>JZLD</td><td>-</td><td>99</td><td>1</td></tr> <tr><td>GR</td><td>-</td><td>-</td><td>100</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	99	-	1	JZLD	-	99	1	GR	-	-	100		BV	JZLD	GR																
BV	24	56	20																																																																	
JZLD	9	70	21																																																																	
GR	8	45	47																																																																	
	BV	JZLD	GR																																																																	
BV	40	39	22																																																																	
JZLD	12	76	13																																																																	
GR	7	14	80																																																																	
	BV	JZLD	GR																																																																	
BV	99	-	1																																																																	
JZLD	-	99	1																																																																	
GR	-	-	100																																																																	
	BV	JZLD	GR																																																																	
Intérieur sans langue	<table border="1"> <tr><td>BV</td><td>86</td><td>10</td><td>3</td></tr> <tr><td>JZLD</td><td>21</td><td>44</td><td>35</td></tr> <tr><td>GR</td><td>18</td><td>52</td><td>30</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	86	10	3	JZLD	21	44	35	GR	18	52	30		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>95</td><td>2</td><td>2</td></tr> <tr><td>JZLD</td><td>8</td><td>60</td><td>32</td></tr> <tr><td>GR</td><td>6</td><td>25</td><td>69</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	95	2	2	JZLD	8	60	32	GR	6	25	69		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>90</td><td>9</td><td>1</td></tr> <tr><td>JZLD</td><td>3</td><td>77</td><td>20</td></tr> <tr><td>GR</td><td>2</td><td>9</td><td>89</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	90	9	1	JZLD	3	77	20	GR	2	9	89		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>100</td><td>-</td><td>-</td></tr> <tr><td>JZLD</td><td>-</td><td>99</td><td>1</td></tr> <tr><td>GR</td><td>-</td><td>-</td><td>100</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	100	-	-	JZLD	-	99	1	GR	-	-	100		BV	JZLD	GR
BV	86	10	3																																																																	
JZLD	21	44	35																																																																	
GR	18	52	30																																																																	
	BV	JZLD	GR																																																																	
BV	95	2	2																																																																	
JZLD	8	60	32																																																																	
GR	6	25	69																																																																	
	BV	JZLD	GR																																																																	
BV	90	9	1																																																																	
JZLD	3	77	20																																																																	
GR	2	9	89																																																																	
	BV	JZLD	GR																																																																	
BV	100	-	-																																																																	
JZLD	-	99	1																																																																	
GR	-	-	100																																																																	
	BV	JZLD	GR																																																																	
Intérieur avec langue	<table border="1"> <tr><td>BV</td><td>81</td><td>11</td><td>8</td></tr> <tr><td>JZLD</td><td>7</td><td>81</td><td>12</td></tr> <tr><td>GR</td><td>5</td><td>26</td><td>69</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	81	11	8	JZLD	7	81	12	GR	5	26	69		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>84</td><td>11</td><td>5</td></tr> <tr><td>JZLD</td><td>2</td><td>83</td><td>15</td></tr> <tr><td>GR</td><td>5</td><td>17</td><td>78</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	84	11	5	JZLD	2	83	15	GR	5	17	78		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>80</td><td>13</td><td>8</td></tr> <tr><td>JZLD</td><td>1</td><td>91</td><td>8</td></tr> <tr><td>GR</td><td>2</td><td>9</td><td>89</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	80	13	8	JZLD	1	91	8	GR	2	9	89		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>100</td><td>-</td><td>-</td></tr> <tr><td>JZLD</td><td>-</td><td>99</td><td>1</td></tr> <tr><td>GR</td><td>-</td><td>-</td><td>100</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	100	-	-	JZLD	-	99	1	GR	-	-	100		BV	JZLD	GR
BV	81	11	8																																																																	
JZLD	7	81	12																																																																	
GR	5	26	69																																																																	
	BV	JZLD	GR																																																																	
BV	84	11	5																																																																	
JZLD	2	83	15																																																																	
GR	5	17	78																																																																	
	BV	JZLD	GR																																																																	
BV	80	13	8																																																																	
JZLD	1	91	8																																																																	
GR	2	9	89																																																																	
	BV	JZLD	GR																																																																	
BV	100	-	-																																																																	
JZLD	-	99	1																																																																	
GR	-	-	100																																																																	
	BV	JZLD	GR																																																																	
Visage complet	<table border="1"> <tr><td>BV</td><td>94</td><td>5</td><td>1</td></tr> <tr><td>JZLD</td><td>15</td><td>60</td><td>25</td></tr> <tr><td>GR</td><td>19</td><td>45</td><td>35</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	94	5	1	JZLD	15	60	25	GR	19	45	35		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>98</td><td>1</td><td>1</td></tr> <tr><td>JZLD</td><td>7</td><td>78</td><td>14</td></tr> <tr><td>GR</td><td>5</td><td>19</td><td>76</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	98	1	1	JZLD	7	78	14	GR	5	19	76		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>98</td><td>2</td><td>-</td></tr> <tr><td>JZLD</td><td>5</td><td>89</td><td>6</td></tr> <tr><td>GR</td><td>6</td><td>7</td><td>88</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	98	2	-	JZLD	5	89	6	GR	6	7	88		BV	JZLD	GR	<table border="1"> <tr><td>BV</td><td>99</td><td>1</td><td>-</td></tr> <tr><td>JZLD</td><td>-</td><td>99</td><td>1</td></tr> <tr><td>GR</td><td>-</td><td>-</td><td>100</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	99	1	-	JZLD	-	99	1	GR	-	-	100		BV	JZLD	GR
BV	94	5	1																																																																	
JZLD	15	60	25																																																																	
GR	19	45	35																																																																	
	BV	JZLD	GR																																																																	
BV	98	1	1																																																																	
JZLD	7	78	14																																																																	
GR	5	19	76																																																																	
	BV	JZLD	GR																																																																	
BV	98	2	-																																																																	
JZLD	5	89	6																																																																	
GR	6	7	88																																																																	
	BV	JZLD	GR																																																																	
BV	99	1	-																																																																	
JZLD	-	99	1																																																																	
GR	-	-	100																																																																	
	BV	JZLD	GR																																																																	
Généralisation		<table border="1"> <tr><td>BV</td><td>89</td><td>6</td><td>5</td></tr> <tr><td>JZLD</td><td>-</td><td>89</td><td>11</td></tr> <tr><td>GR</td><td>2</td><td>8</td><td>91</td></tr> <tr><td></td><td>BV</td><td>JZLD</td><td>GR</td></tr> </table>	BV	89	6	5	JZLD	-	89	11	GR	2	8	91		BV	JZLD	GR																																																		
BV	89	6	5																																																																	
JZLD	-	89	11																																																																	
GR	2	8	91																																																																	
	BV	JZLD	GR																																																																	

Fig. 27 – Matrices de confusion de lieu d’articulation du **groupe II** pour chaque série du test (B - /b/, V - /v/, J - /□/, Z - /z/, L - /l/, D - /d/, G - /g/, R - /□/ ; stimuli = ligne, réponses = colonne)

3.2.3 Pourcentage d'Information Relative Transmise

A partir des matrices de confusion, nous calculons le pourcentage d'information transmise. Contrairement au pourcentage de réponses correctes, le calcul de l'information transmise est fondé sur la totalité des informations fournies par la matrice de confusion. Ce calcul est fondé sur la notion d'entropie (issue de la théorie de l'information développée par Shannon, 1949). L'entropie H est définie comme la quantité moyenne d'information liée à un événement élémentaire (cf. formule 6).

$$H(A) = -\sum_{i=1}^n p_i \log(p_i), \quad (6)$$

où p_i est la probabilité que l'événement élémentaire A_i qui constitue A se réalise, et que $\sum_{i=1}^n p_i = 1$.

L'information transmise T correspond à la quantité moyenne d'information partagée par le couple stimulus A / réponse B . Elle est définie par la relation suivante (formule 7) :

$$T(A, B) = H(A) + H(B) - H(AB) = \sum_{i,j} p_{ij} \log\left(\frac{p_{ij}}{p_i \cdot p_j}\right), \quad (7)$$

où p_{ij} est la probabilité d'une réponse j lors de la présentation d'un stimulus i .

Pour une mesure du pourcentage d'information relative transmise (PIRT) par une série donnée de l'expérience, il suffit de normaliser la mesure de l'information transmise et de la ramener à un pourcentage (formule 8) :

$$T = 100 \cdot \frac{T(A, B)}{H(A)}. \quad (8)$$

Si la réponse est mal corrélée au stimulus présenté, alors le PIRT sera faible, et à l'opposé, si la réponse peut être prédite avec une grande précision à partir du stimulus présenté, le PIRT sera élevé (proche de 100 %).

Les figures 28-29 illustrent les graphes de PIRT de consonne et de lieu d'articulation pour les deux groupes. Les positions des courbes les unes par rapport aux autres sont similaires à celles des courbes du pourcentage des réponses correctes. Ces résultats prouvent que la confusion du lieu d'articulation en condition de bruit fort a été assez élevée. On voit aussi que le groupe II, qui a écouté les stimuli avec du bruit ascendant, a mieux réussi à reconnaître le lieu d'articulation en modalité « Intérieur avec langue » quand le bruit est très fort, que le groupe I.

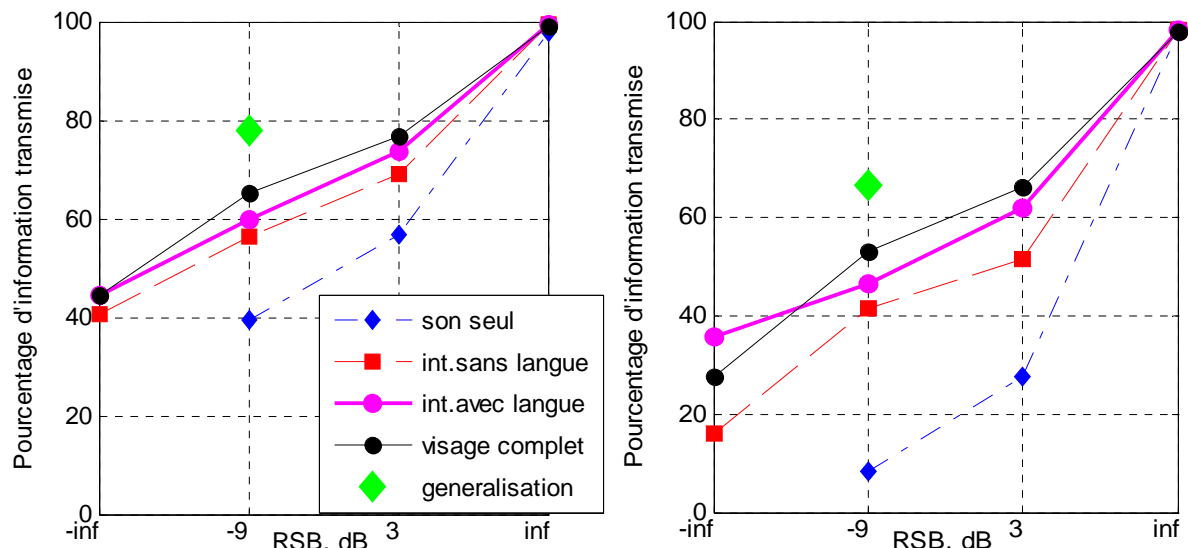


Fig. 28 – Pourcentage d’information relative transmise sur la consonne (à gauche) et le lieu d’articulation (à droite) pour le **groupe I** selon les différentes conditions (de bas en haut) : audio seul; intérieur sans langue; intérieur avec langue; visage complet ; série de généralisation

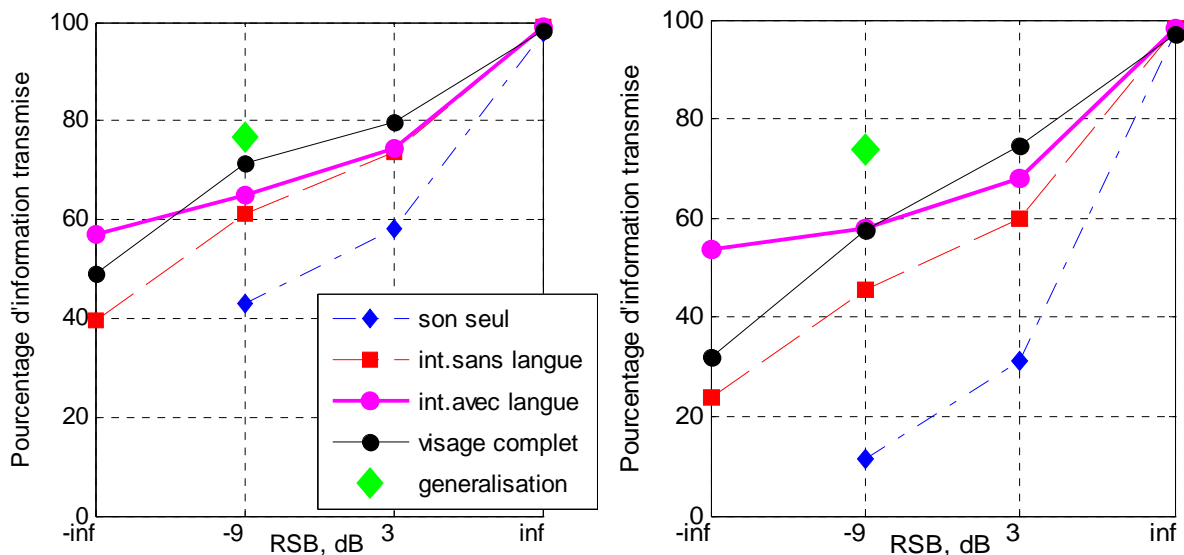


Fig. 29 – Pourcentage d’information relative transmise sur la consonne (à gauche) et le lieu d’articulation (à droite) pour le **groupe II** pour selon les différentes conditions (de bas en haut) : audio seul; intérieur sans langue; intérieur avec langue; visage complet ; série de généralisation

Nous avons également analysé le pourcentage d’identification correcte de consonne dans différents contextes vocaliques pour les différentes séries du test (cf. les graphes des scores : annexe E). Nous avons obtenu la même tendance d’effet du contexte vocalique sur l’intelligibilité des consonnes que Benoît et al. (1994b). Pour les différents modes de présentation et de niveaux de bruit, les consonnes ont été plus intelligibles en contexte /a/ qu’en contexte /i/, et plus intelligibles en contexte /i/ qu’en contexte /y/. Les scores d’identification de consonnes en contextes /y/ et /u/ sont similaires, sauf en condition RSB = +3 dB. L’intelligibilité des consonnes en mode de présentation audiovisuel est plus faible en contexte /u/ qu’en contexte /y/. On observe aussi que pour le groupe II en condition de vidéo seule avec la visualisation de la langue, les scores d’identification correcte des consonnes en contextes /i/ et /u/ sont bien plus élevés que les scores en modalité « Visage complet » [en contexte /i/ : $F(1, 10)=8.89$; $p<0.01$; en contexte /u/ : $F(1, 10)=8.76$; $p<0.01$].

L'analyse des scores individuels des sujets un par un fait apparaître des taux de réussite différents au cours des différentes séries du test (cf. écart-types de tableau 6-7). Certains sujets du groupe II ont obtenu des scores pour les séries avec la présentation « Intérieur avec langue » bien plus élevés que les scores obtenus pour les autres séries.

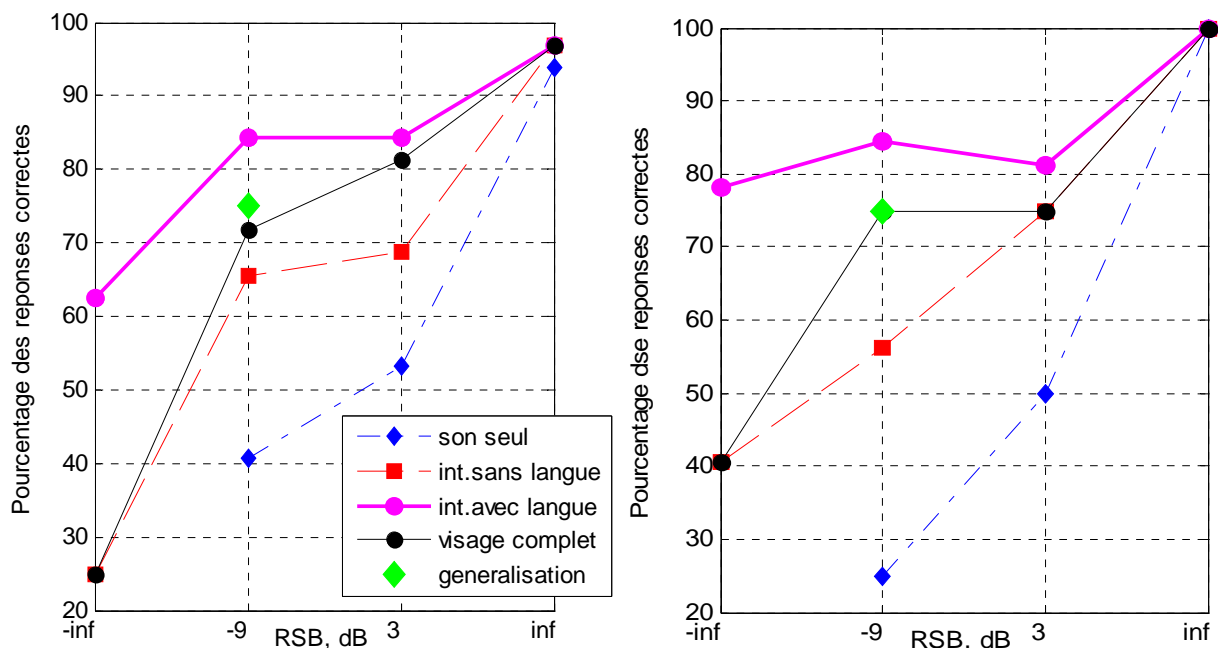


Fig. 30 – Scores d'intelligibilité de 2 sujets particuliers (du groupe II) en fonction du bruit pour les différentes conditions (de bas en haut) : audio seul; intérieur sans langue; intérieur avec langue; visage complet ; série de généralisation

La figure 30 montre les scores d'intelligibilité de 2 sujets particuliers (du groupe II) pour les différentes séries du test. On voit que ces sujets ont bien perçu l'information apportée par la visualisation de langue. Ainsi, leurs scores pour les séries avec visualisation de langue pour les différents niveaux du bruit sont bien meilleurs que les scores pour les séries avec les autres présentations, et meilleurs que les moyennes de scores des différentes présentations sur l'ensemble des sujets de groupe II. Cela pourrait être dû au fait que ces sujets possèderaient un meilleur niveau de conscience articulatoire que les autres, ou qu'ils apprennent plus vite que les autres. On observe aussi que le score pour la série de généralisation de ces sujets est inférieur au score de la série avec les mêmes conditions de bruit et de présentation. Cela peut être dû à la fatigue en fin du test, ou au fait, que les sujets ont appris de manière intrinsèque les stimuli du test en passant du bruit minimal au bruit maximal, et qu'ils n'ont pas vraiment appris à lire l'information sur la langue.

Finalement, l'analyse montre qu'en condition de bruit très fort la visualisation de la langue peut être utilisée comme information supplémentaire de la parole. L'apprentissage de la lecture des mouvements de la langue pourrait améliorer la compréhension de la parole en conditions audiovisuelles avec visualisation de la langue et audio dégradé.

Conclusions et perspectives

Conclusions

Dans le cadre de ce stage, un test de perception a été effectué pour évaluer l'apport de la vision d'une langue représentée par un modèle 3D à la compréhension de la parole. L'analyse des résultats de ce test ne permet pas de trancher de manière définitive la question de l'apport de la visualisation de la langue en réalité augmentée à la compréhension de la parole en condition de bruit. On peut cependant tirer de cette expérience les conclusions suivantes.

La synthèse des stimuli audiovisuels est de bonne qualité. La langue et les lèvres synthétiques sont réalistes et transmettent bien l'information sur le phonème.

Toutes les modalités de présentation vidéo améliorent la compréhension de parole par rapport à l'audio seul.

Si le niveau du bruit n'est pas très fort, grâce à l'effet de complémentarité audio-visuel, les humains récupèrent l'information de parole à partir du son dégradé et du visage (de l'extérieur ; les lèvres transmettent la partie importante de cette information). Cette perception de parole est naturelle car apprise dès le plus jeune âge. Ainsi, la vision de langue n'est pas utile quand le bruit est faible.

En revanche, le cas où le score « Intérieur avec langue » est supérieur au score « Visage complet » lorsque le signal audio est absent suggère que lorsque l'information audio n'est plus suffisante pour aider la lecture labiale, la lecture « linguale » peut prendre le relais si nécessaire.

Un effet d'apprentissage implicite de la lecture linguale semble avoir eu lieu. Des tests complémentaires sont nécessaires pour montrer qu'un apprentissage rapide peut être fait.

Applications

Le système de synthèse des stimuli audio-visuels, élaboré dans le cadre de ce projet, peut être utilisé pour le développement d'applications pour la correction phonétique, i.e. la correction des défauts de langage pour les enfants, l'apprentissage d'articulations en langue étrangère, la réhabilitation de la parole des malentendants.

Les animations de la langue peuvent être utilisées dans le domaine de l'enseignement des sciences cognitives, pour la démonstration des mouvements de la langue.

Perspectives

Tout d'abord, il serait intéressant d'élaborer d'autres tests de perception, afin de compléter les résultats présents. En particulier il faudrait :

– développer la séance d'apprentissage de la lecture sur la langue, et évaluer l'apport de la vision de la langue à la compréhension de la parole avant et après l'apprentissage ;

– ajouter une modalité de présentation audio-visuelle avec vision de la langue mais sans vision des lèvres ;

– enregistrer les points de fixation des yeux des sujets sur l'écran pendant que ces derniers passent le test ; cela permettrait de comprendre les stratégies de suivi des articulateurs, et de proposer éventuellement de meilleures méthodes (par exemple focaliser le regard sur la région des incisives, afin d'avoir la langue d'un côté du champ visuel et les lèvres de l'autre côté, comme suggéré par certains sujets) ;

– passer les différentes séries du test en ordre aléatoire pour chaque sujet, afin de pouvoir déterminer plus clairement si les sujets ont appris à lire sur la langue, ou s'ils ont appris les stimuli du test.

Toutes ces modifications du test permettraient de mieux comprendre l'effet de la vision de la langue sur la compréhension de la parole.

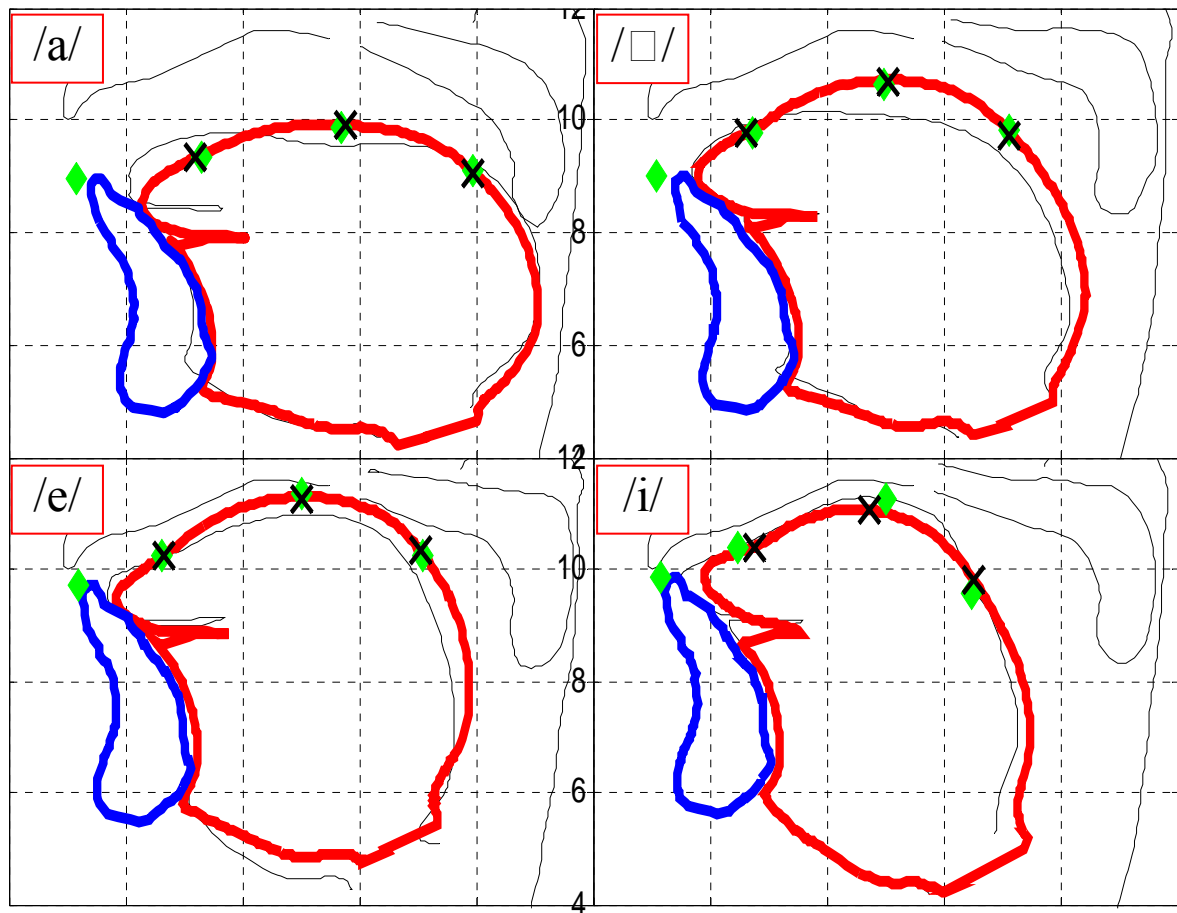
Références bibliographiques

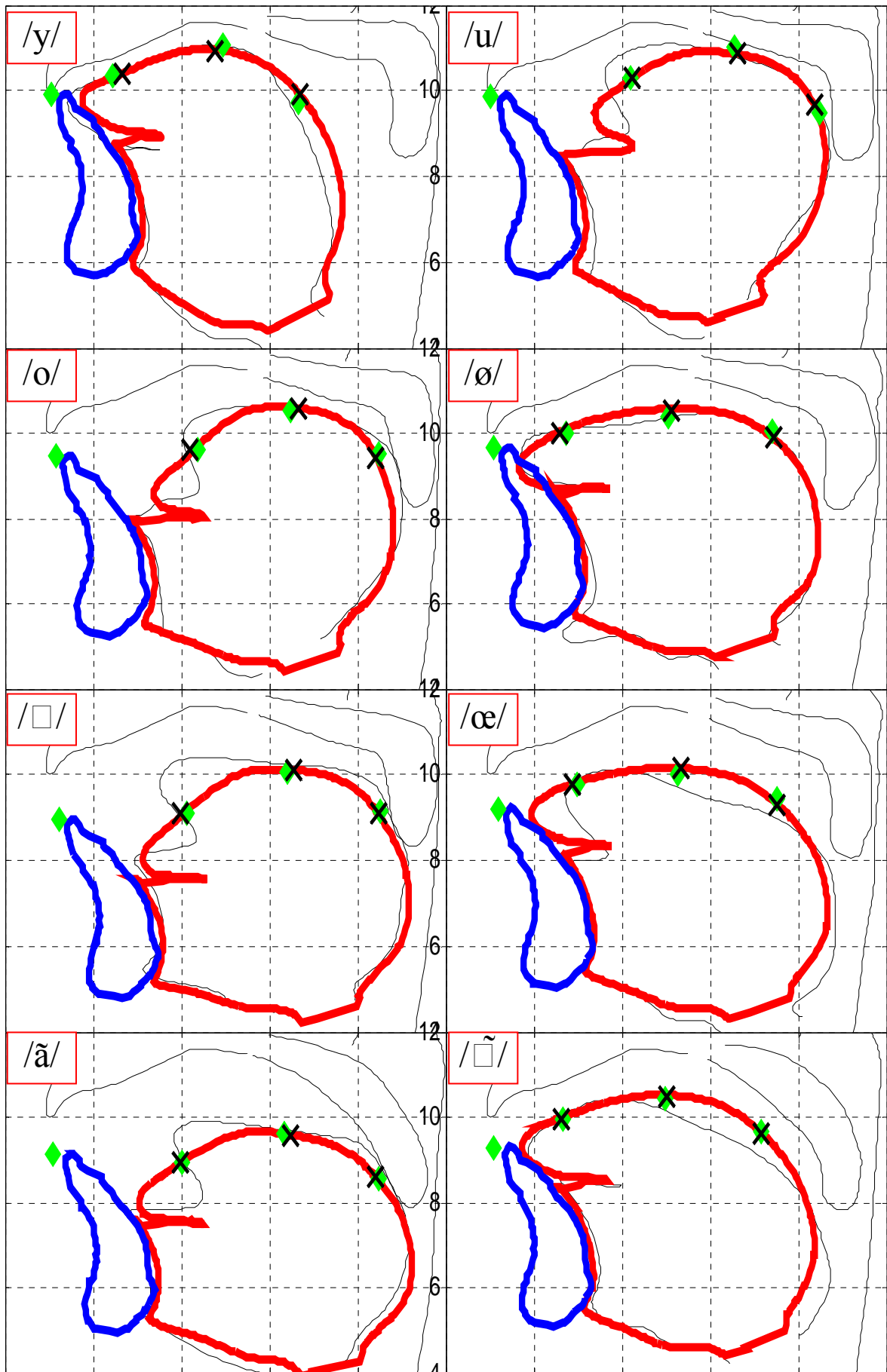
- [Abry et al., 1994] Abry, C., Badin P., Scully, C. (1994). Sound-to-gesture inversion in speech: The speech Maps approach, in *Advanced Speech applications*, edited by K. Vargnese, S. Pflieger and J. P. Lefèvre (Springer, Berlin), 182-196.
- [Arnold, 1997] Arnold, P. (1997). The Structure and Optimization of Speechreading. *Journal of Deaf Studies and Deaf Education*, 2:4, 199-211.
- [Arnold et Hill, 2001] Arnold, P., Hill, F. (2001). Bisensory augmentation: A speechreading advantage when speech is clearly audible and intact. *British Journal of Psychology*, 92, 339-355.
- [Badin et al, 1997] Badin, P., Baricchi, E., Vilain, A. (1997). Determining tongue articulation: from discrete fleshpoints to continuous shadow. In *Proceedings of the 5th EuroSpeech Conference*, vol. 1, 47-50.
- [Badin et al., 2002] Badin, P., Bailly, G., Revéret, L., Baciú, M., Segebarth C., et Savariaux C. (2002). Three-dimensional articulatory modeling of tongue, lips and face, based on MRI and video images. *Journal of Phonetics*, 30, 533-553.
- [Badin et Serrurier, 2006] Badin, P., Serrurier, A. (2006). Three-dimensional linear modeling of tongue: Articulatory data and models. In *Proceedings of the 7th International Seminar on Speech Production, ISSP7* (H.C. Yehia, D. Demolin & R. Laboissière, Eds.). Ubatuba, SP, Brazil. UFMG, Belo Horizonte, Brazil. 395-402.
- [Bälter et al., 2005] Bälter, O., Engwall, O., Öster, A.-M., Kjellström, H. (2005). Wizard-of-Oz Test of ARTUR - a Computer-Based Speech Training System with Articulation Correction. In *proceedings of ASSETS 2005*.
- [Benoît et al., 1994a] Benoît, C., Mohamadi, T., Kandel, S. (1994). Audio-visual intelligibility of French speech in noise. *Journal of Speech and Hearing Research*, 37, 1195-1203.
- [Benoît et al., 1994b] Benoît, C., Mohamadi, T., Kandel, S. (1994). Effects of phonetic context on audio-visual intelligibility of French. *Journal of Speech and Hearing Research*, 37, 1195-1203.
- [Benoît et al., 1996] Benoît, C., Guiard-Marigny, T., Le Goff, B., Adjoudani, A. (1996). Which components of the face do humans and machines best speechread?: In Stork, D., Hennecke, M. (Eds.), *Speechreading by Humans and Machines*, NATO-ASI Series 150 Springer, Berlin, 315-328.
- [Benoit et Le Goff, 1998] Benoît, C., Le Goff, B. (1998). Audio-visual speech synthesis from French text: Eight years of models, designs and evaluation at the ICP. *Speech Communication*, 26, 117-129.
- [Brooke et Summerfield, 1983] Brooke, N.M., Summerfield, A.Q., 1983. Analysis, synthesis and perception of visible articulatory movements. *Journal of Phonetics*, 11, 63-76.
- [Cohen et al., 1998] Cohen, M. M., Beskow, J., Massaro, D.W. (1998). Recent Developments in Facial Animation: An Inside View. In *Proceedings of the International Conference on Auditory-Visual Speech Processing - AVSP'98*, Terrigal, Australia, pp. 201-206.
- [Cohen et Massaro, 1993] Cohen, M. M., Massaro, D. W. (1993). Modeling coarticulation in synthetic visual speech. Models and techniques. In: Thalmann, N. M., Thalmann, D. (eds.), *Computer Animation*. Springer, Tokyo, 139-156.
- [Dodd, 1987] Dodd, B., (1987). The acquisition of lip-reading skills by normally-hearing children: In Dodd et Campbell, 1987.
- [Dodd et Campbell, 1987] Dodd, B., Campbell, R. (eds.) (1987). *Hearing by Eye: the Psychology of Lip-reading*. London: Lawrence Erlbaum Associates Ltd.
- [Engwall et Beskow, 2003] Engwall, O., Beskow, J. (2003). Resynthesis of 3D tongue movements from facial data. In *Proceedings of Eurospeech 2003*, 2261-2264.
- [Erber, 1969] Erber, N. P. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *Journal of speech and hearing research*, 12, 423-425.
- [Ewertsen et Birk Nielsen, 1971] Ewertsen, H. W., Birk Nielsen, H. (1971). A comparative analysis of the audiovisual, auditive, and visual perception of speech. *Acta oto- laryngologica.*, 72, 201-205.
- [Griffiths et Frith, 2002] Griffiths, S., Frith, U. (2002). Evidence for an articulatory awareness deficit in adult dyslexics. *Dyslexia*, 8, 14-21.
- [Heilman et al., 1996] Heilman, K. M., Voeller, K., Alexander, A. W. Developmental dyslexia: a motor-articulatory feedback hypothesis. *Ann Neurol*, 1996, 407-412.
- [Hoole, 1996] Hoole, P. (1996). Issues in the acquisition, processing, reduction and parameterization of articulographic data, *FIPKM* 34, 158-173.
- [Ijsseldijk, 1992] Ijsseldijk, F. J. (1992). Speechreading performance under different conditions of video image, repetition and speech rate. *Journal of Speech and Hearing Research*, 35, 466-471.
- [Jordan et al., 1997] Jordan, T. R., Sergeant, P. C., Martin, C., Thomas, S. M., Thow, E. (1997). Effects of horizontal viewing angle on visual and audiovisual speech perception. In *Computational cybernetics and simulation: 1997 IEEE International Conference on Systems, Man, and Cybernetics*, 2, 626-1631. Piscataway, NJ: Institute of Electrical and Electronics Engineers.
- [Kellerman, 1990] Kellerman, S. (1990). Lip service: the contribution of the visual modality to speech perception and its relevance to the teaching and testing of foreign language listening comprehension , *Applied Linguistics*, 11, 272-280.

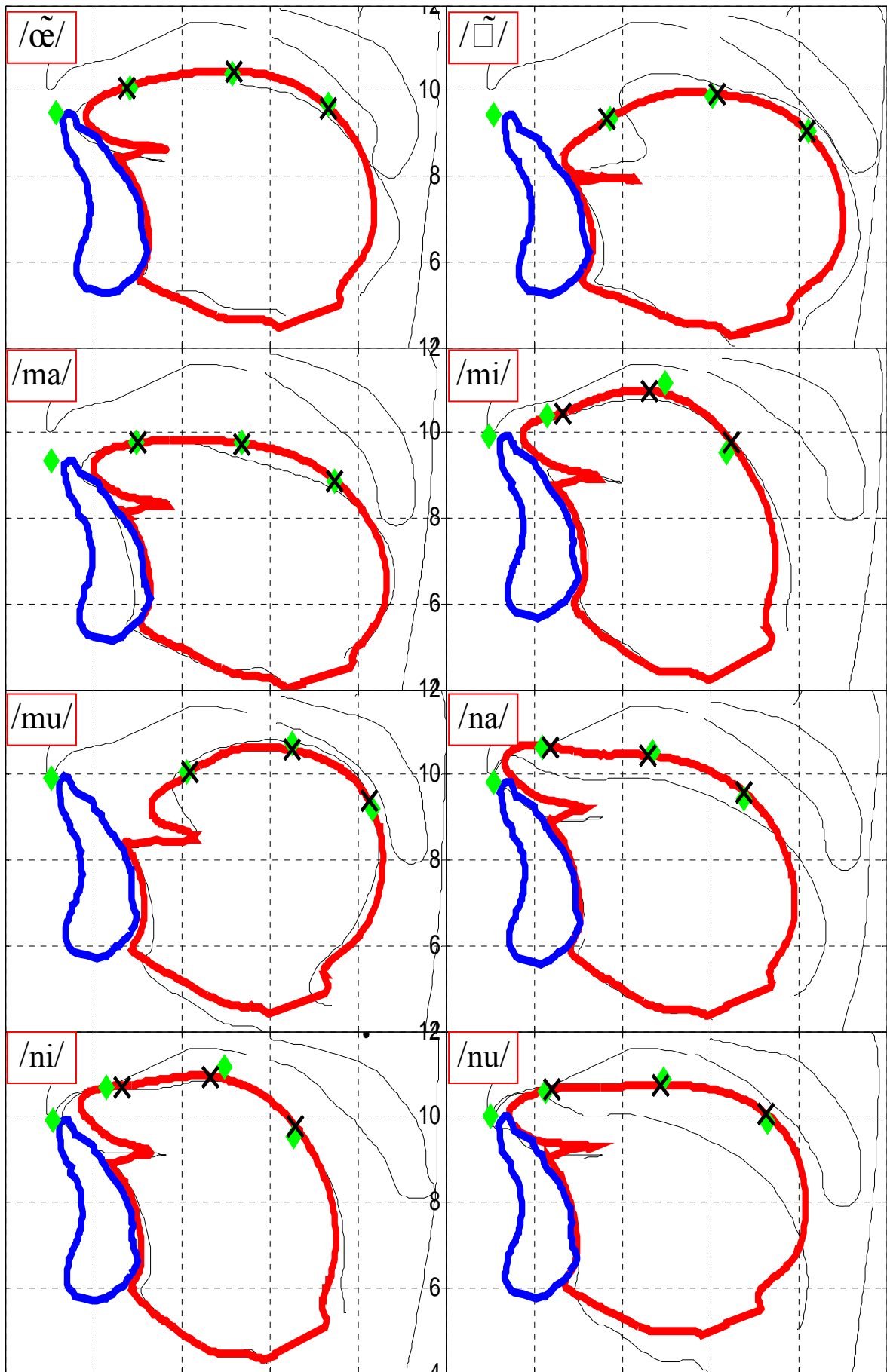
- [Le Goff et al., 1996] Le Goff, B., Guiard-Marigny, T., Benoît, C. (1996). Analysis-synthesis and intelligibility of a talking face. In : Van Santen, J.P.H., Sproat, R. W., Olive, J. P., Hirschberg, J. (eds.), *Progress in Speech Synthesis*. Springer. New York, 235-246.
- [Le Goff et Benoît, 1996] Le Goff, B., Benoît, C. (1996). A text-to-audiovisual-speech synthesizer for French. In: *Proceedings of the 4th International Conference on Spoken Language Processing*, Philadelphia, PA, USA, Vol. 4, 2163-2166.
- [MacLeod et Summerfield, 1987] MacLeod, A., Summerfield, Q. (1987). Quantifying the contribution of vision to speech perception in noise. *British journal of audiology*, 1987, 12, 131-141.
- [Massaro, 2004] Massaro, D. W. (2004). From multisensory integration to talking heads and language learning. In: *Handbook of Multisensory Processes*, Calvert, G., Spence, C., Stein, B. E., Eds., pp. 153–176, MIT Press, Cambridge, Mass, USA.
- [Massaro et Light, 2003] Massaro, D. W., Light, J. (2003). Read My Tongue Movements: Bimodal Learning To Perceive And Produce Non-Native Speech /t/ and /l/. In *Proceedings of Eurospeech (Interspeech)*, 8th European Conference on Speech Communication and Technology, Geneva, Switzerland.
- [Massaro et Light, 2004] Massaro, D. W., Light, J. (2004). Using visible speech for training perception and production of speech for hard of hearing individuals. *Journal of Speech, Language, and Hearing Research*, vol. 47, no. 2, 304–320.
- [McGurk et MacDonald, 1976] McGurk, H., MacDonald, J. (1976). Hearing lips and seeing voices, *Nature*, 264, 746-748.
- [Mohamadi, 1993] Mohamadi, T. (1993). Synthèse à partir du texte de visages parlants: Réalisation d'un prototype et mesures d'intelligibilité bimodale. Mémoire de thèse, INP, Grenoble, 174 pp.
- [Montgomery, 1981] Montgomery, D. (1981). Do dyslexics have difficulty accessing articulatory information? *Psychological Research*, 43, 235-243.
- [Munhall et Vatikiotis-Bateson, 1998] Munhall, K. G., Vatikiotis-Bateson, E. (1998). The moving face during speech communication. In Cambell, R., Dodd, B., Burnham, D. *Hearing by eye II: Advances in the psychology of speechreading and auditory-visual speech* (eds.). London: Psychology Press, 123-139.
- [Neely, 1956] Neely, K. K. (1956). Effect of visual factors on the intelligibility of speech. *Journal of Acoustic Society of America*, 28, 1275-1277.
- [O'Neill, 1954] O'Neill, J. J. (1954). Contributions of the visual components of oral symbols to speech comprehension. *Journal of Speech Hearing Disorders*, 19, 429-439.
- [Peckels et Rossi, 1973] Gibert, G., G. Bailly and F. Elisei (2006). Evaluation d'un système de synthèse 3D de Langue française Parlée Complétée. *Journées d'Etudes sur la Parole*, Rennes - France: accepted.
- [Perkell et al., 1992] Perkell, J., Cohen, M., Svirsky, M., Matthies, M., Garabieta, I., Jackson, M. (1992). Electro-magnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements. *Journal of the Acoustical Society of America*, 92, 3078-3096.
- [Reisberg et al., 1987] Reisberg, D., McLean, J., Goldfield, A. (1987). Easy to hear but hard to understand: A lipreading advantage with inact auditory stimuli: In *Dodd et Campbell*, 1987.
- [Revéret et al., 2000] Revéret, L., G. Bailly and P. Badin (2000). MOTHER: a new generation of talking heads providing a flexible articulatory control for video-realistic speech animation. *International Conference on Speech and Language Processing*, Beijing - China: 755-758.
- [Rosenblum et al., 1996] Rosenblum, L. D., Johnson, J. A., Saldaña, H. M. (1996). Pointlight facial displays enhance comprehension of speech in noise. *Journal of Speech and Hearing Research*, 39, 1159-1170.
- [Sumbly et Pollack, 1954] Sumbly, W. H., Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, 26(2), 212-215.
- [Summerfield, 1979] Summerfield, A.Q. (1979). Use of visual information for phonetic perception. *Phonetica*, 36, 314-331.
- [Summerfield et al., 1989] Summerfield, Q., MacLeod, A., McGrath, M., Brooke, M. (1989). Lips, teeth, and the benefits of lipreading. In Young, A. W., Ellis, H. D. *Handbook of Research on Face Processing* (eds.). Elsevier Science Publishers B. V., 223-233.
- [Thomas et Jordan, 2004] Thomas, S. M., Jordan, T. R. (2004). Contributions of Oral and Extraoral Facial Movement to Visual and Audiovisual Speech Perception. *Journal of Experimental Psychology: Human Perception and Performance*, 30, No. 5, 873-888.
- [Tye-Murray et al., 1993] Tye-Murray, N., Kirk, K. I., Schum, L. (1993). Making typically obscured articulatory activity available to speechreaders by means of videofluoroscopy. *NCVS Status and Progress Report*, 4, 41-63.
- [Tyler et Lowder, 1992] Tyler, R. S., Lowder, M. W. (1992). Audiological management and performance of adult cochlear-implant patients. *Ear, Nose and Throat Journal*, 71 (3), 117-128.
- [Vatikiotis-Bateson et al., 1998] Vatikiotis-Bateson, E., Eigsti, I. M., Yano, S., Munhall, K. G. (1998). Eye movements of perceivers during audiovisual speech perception. *Perception & Psychophysics*, 60, 926-940.
- [Walden et al., 1977] Walden, B., Prosek, R., Montgomery, A., Scherr, C., Jones, C. (1977). Effects of training on the visual recognition of consonants. In *Journal of Speech and Hearing Research*, 135-145.

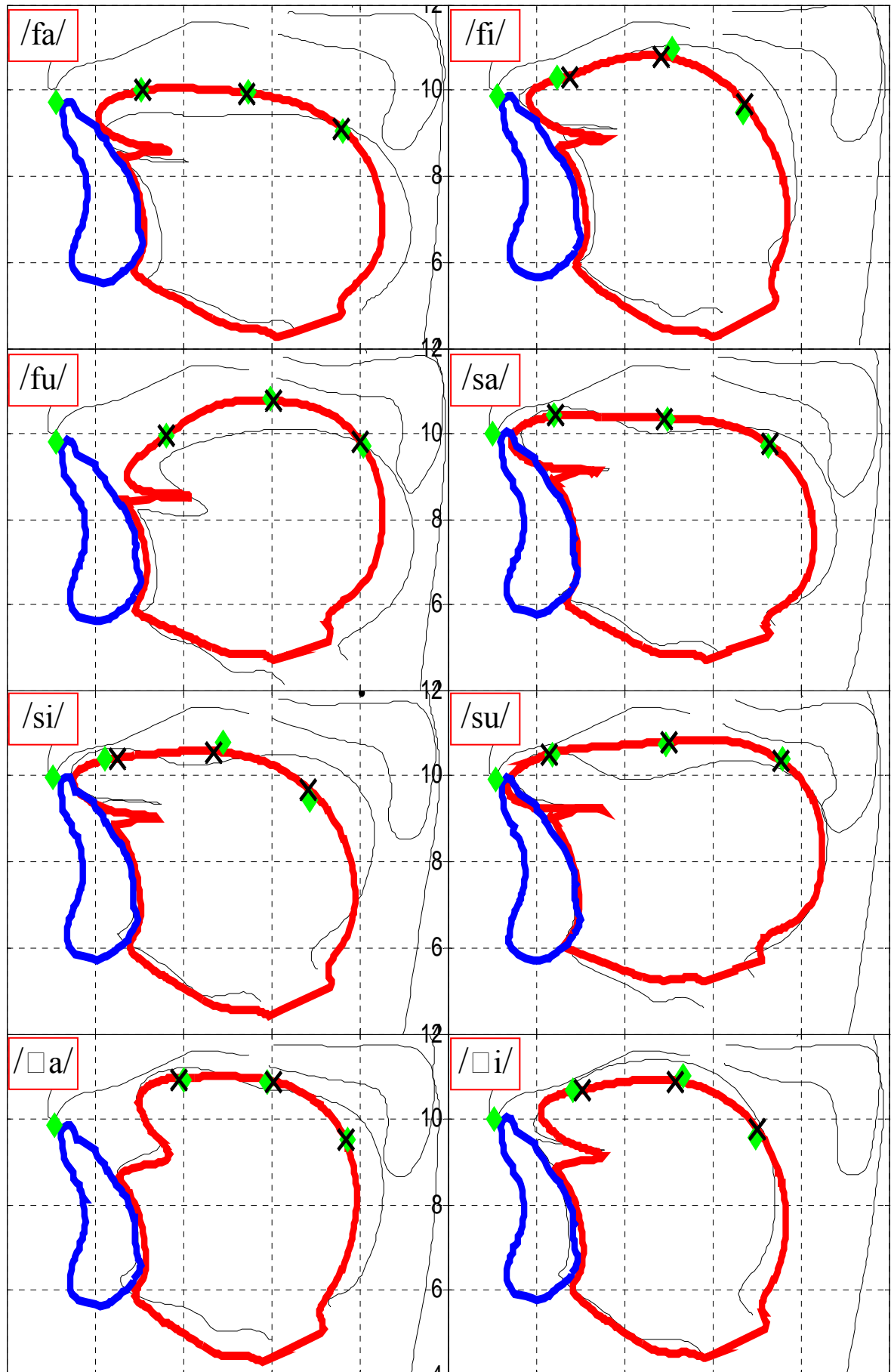
Annexe A : Résultats de l'inversion

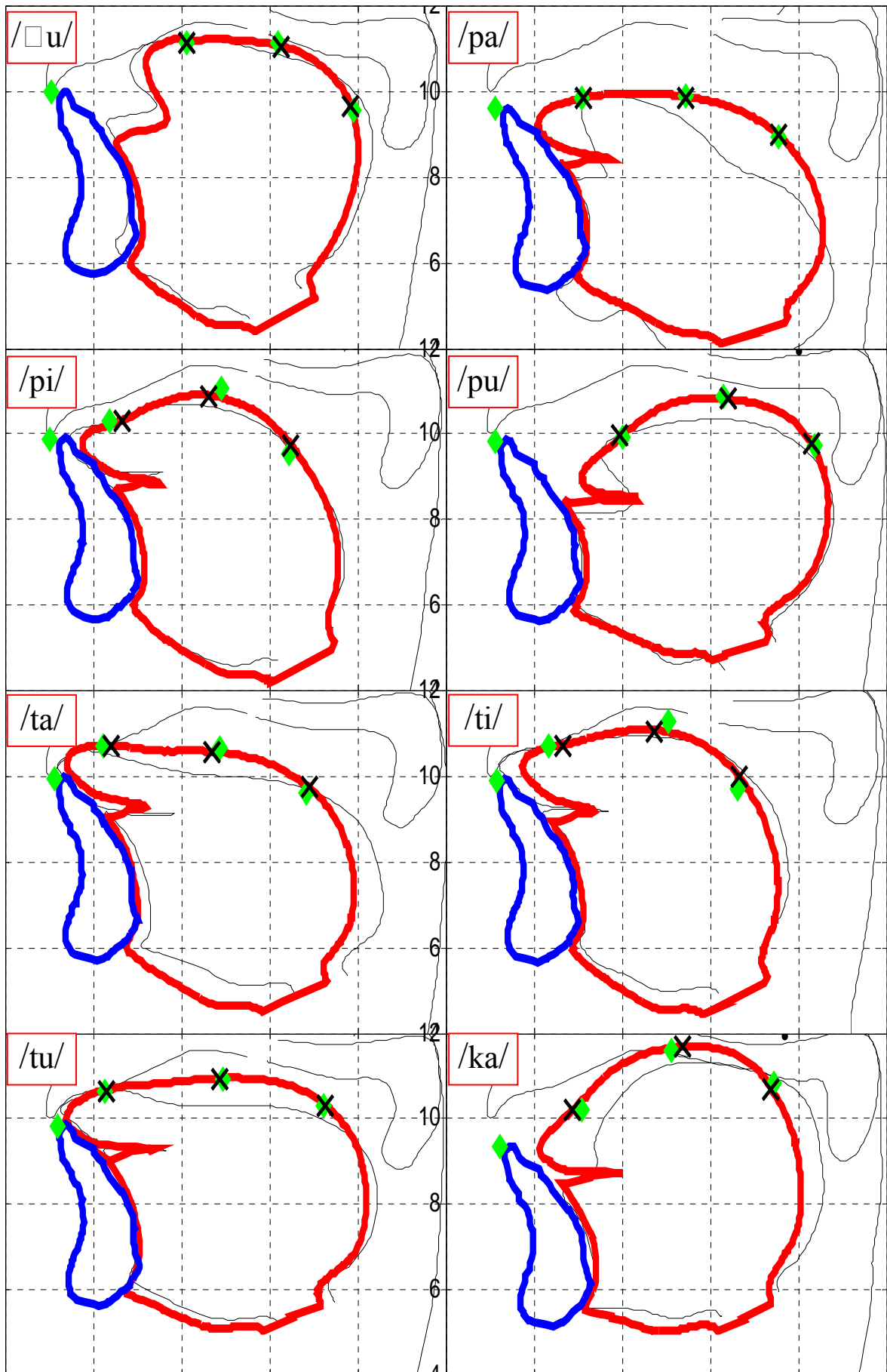
Toutes les figures suivantes représentent les coupes médiosagittales de la langue et de la mâchoire modélisées en utilisant la méthode d'inversion directe de la matrice des coefficients du modèle (ligne épaisse), superposées au contour obtenu en coupant dans le plan médiosagittal le maillage 3D obtenu à partir de données IRM (ligne continue mince), la position des bobines (losanges) et les points pseudo EMA (croix) pour les 44 articulations (traitement du corpus du 5 avril 2007).

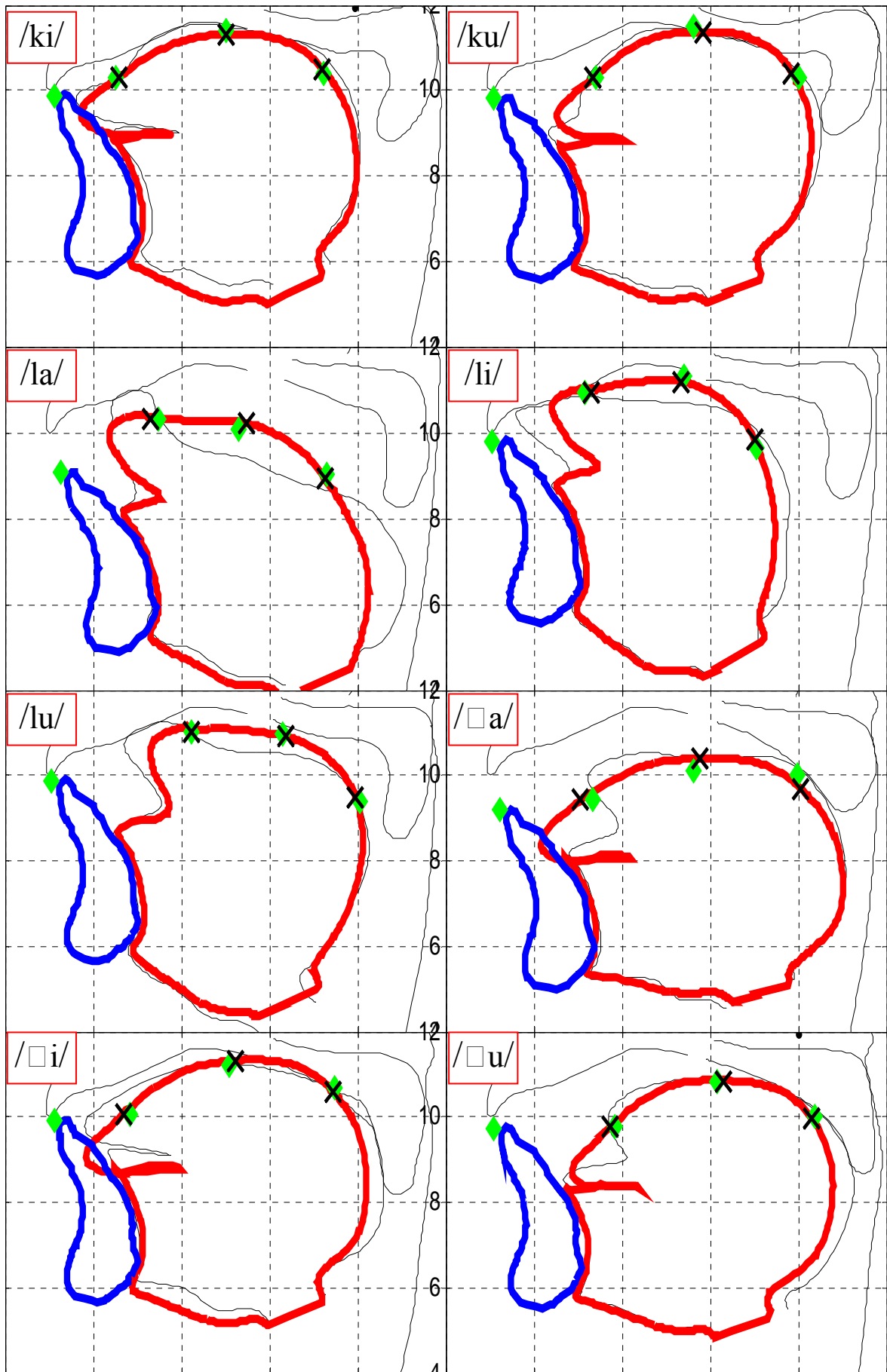












Annexe B : Traitement du corpus EMA enregistré à l'ICP le 18 avril 2003

Différentes méthodes d'inversion ont été testées sur le corpus EMA enregistré à l'ICP le 18 avril 2003 (26 configurations {/a/, /□/, /e/, /i/, /y/, /u/, /o/, /œ/, /fa/, /fi/, /fu/, /sa/, /si/, /su/, /□a/, /□i/, /□u/, /pa/, /pi/, /pu/, /ta/, /ti/, /tu/, /ka/, /ki/, /ku/}), y compris la méthode d'inversion de la matrice des coefficients du modèle et les 5 variantes d'optimisation (par la fonction de Matlab *fminimax*), qui différaient par le type d'erreur à optimiser pour chaque bobine (cf. tableau B.1). Les paramètres à estimer ont été contraints dans [-6 ; 6].

Tableau B.1 – Caractéristiques des 5 variantes d'optimisation

	Type d'erreur à optimiser		
	tip	mid	bck
1	Point à point	Point à point	Point à point
2	Point à point	Point à point	Point à contour
3	Point à point	Point à contour	Point à point
4	Point à point	Point à contour	Point à contour
5	Point à contour	Point à contour	Point à contour

Comme le système de coordonnées a été fixé par rapport aux structures osseuses du sujet, le palais dur ne changeait pas de position et la coordonnée X de n'importe quel point de langue ne pouvait pas être inférieure à 5 cm et supérieure à 13.5 cm, et la coordonnée Z de n'importe quel point de langue ne pouvait pas être supérieure à 11.8 cm. Nous avons utilisé ces contraintes supplémentaires en optimisation.

En analysant les coupes médiosagittales de la langue restituée dans les différents cas et les erreurs de reconstruction de la langue à partir de paramètres trouvés par optimisation et par inversion de matrice des coefficients, on a observé que la langue n'était pas bien restituée pour certaines articulations (/i/, /y/, /pi/, /□i/, /ti/) et que la forme de la langue était parfois irréaliste.

Nous supposons que ceci est dû la non cohérence des données (les données EMA n'étaient pas bien calées dans le système des coordonnées). Ainsi, on n'est pas arrivé à savoir quelle est la meilleure méthode pour l'inversion de modèle.

Annexe C : Exemples du texte affiché durant le test de perception

ECRAN 1

Evaluation d'une tête parlante audiovisuelle
Vous allez voir et entendre une tête parlante.
Plusieurs présentations de cette tête seront utilisées.
Les tests consistent à écouter et regarder ce qu'elle prononce,
et à indiquer la consonne que vous avez reconnue.

Vous devrez répondre **CORRECTEMENT**, mais en essayant d'être **RAPIDE**.
Vous n'aurez pas la possibilité de réécouter le mot prononcé.
Note: un bruit plus ou moins fort sera ajouté à chaque mot.

ECRAN 2

Déroulement de l'expérience
Au cours des différentes phases du test,
quatre types de présentation seront utilisés :
- son tout seul
- son avec intérieur du visage sans langue
- son avec intérieur du visage et langue
- son avec visage complet

ECRAN 3

Illustration visuelle :
Cliquez pour découvrir l'apparence de la tête parlante
en mode "Visage complet"...

ECRAN 3

Vue interne :
Les différents articulateurs visibles EN COUPE seront :
- palais dur et voile du palais en haut,
- paroi arrière du pharynx à droite,
- et mâchoire en mouvement

ECRAN 4

Vue avec la langue en mouvement :

ECRAN 5

Principe du test:
Les mots que vous allez entendre n'ont pas de signification.
Ils sont simplement formés d'une suite
Voyelle + CONSONNE + Voyelle.
L'objectif du test est de reconnaître la CONSONNE.
Après chaque présentation, vous devrez cliquer sur la case
correspondant à la CONSONNE reconnue.
Si vous hésitez entre plusieurs réponses,
choisissez en une au hasard.

ECRAN 6

Les huit consonnes testées sont:

- B (comme dans Bon, roBe)
- D (comme dans Dans, aiDe)
- G (comme dans Gare, baGue)
- J (comme dans Je, Gilet, Geôle)
- L (comme dans Lent, soL)
- R (comme dans Rue, veniR)
- V (comme dans Vous, rêVe)
- Z (comme dans Zéro, maiSon, roSe)

ECRAN 7

Entraînement :

La courte séance d'entraînement qui suit vous permet de vous familiariser avec la procédure de test.

Cliquez sur suite lorsque vous êtes prêt.

ECRAN 8

Votre entraînement est terminé.

Si vous avez des questions, posez-les maintenant à Yuliya.
Sinon, vous allez pouvoir enchaîner sur la PREMIERE partie.

ECRAN 9

Test son seul

Dans cette série, SEUL LE SON sera joué avec UN BRUIT ASSEZ GENANT.

Vous aurez la possibilité de faire une pause à la fin de cette série de 34 mots.

Cliquez sur suite lorsque vous êtes prêt.

Annexe D : Exemples de fenêtres de présentation du test de perception

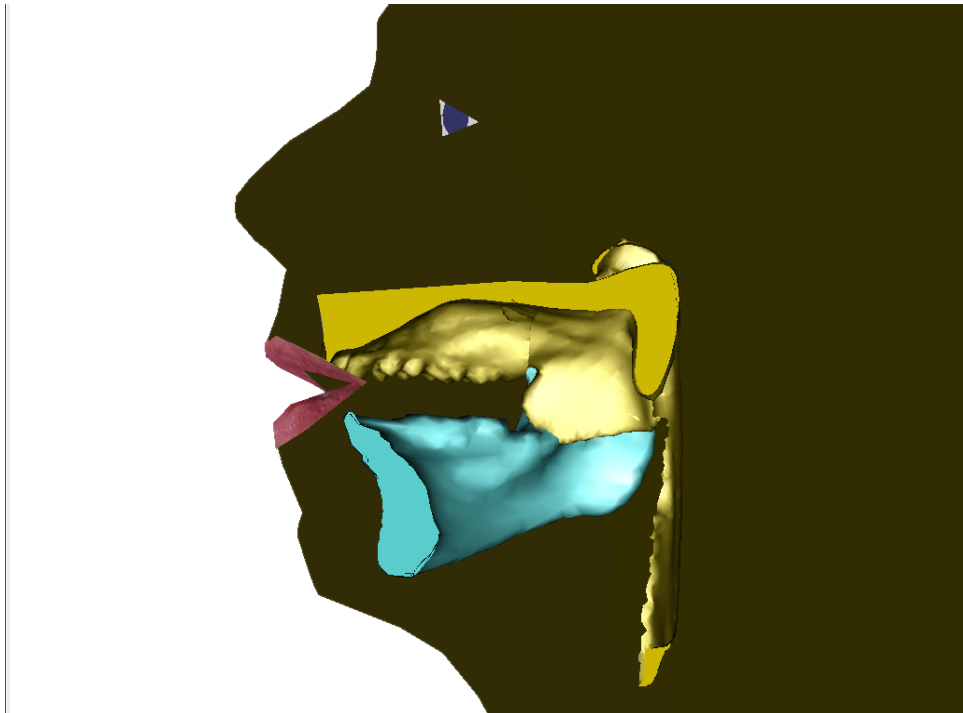


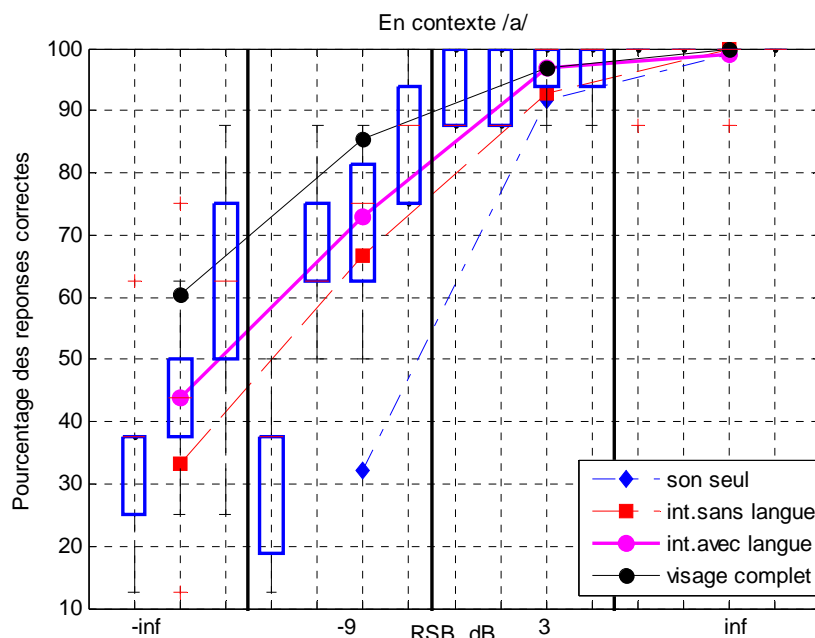
Fig. D.1 – Présentation dans la modalité « Intérieur sans langue »

Qu'avez-vous perçu ?

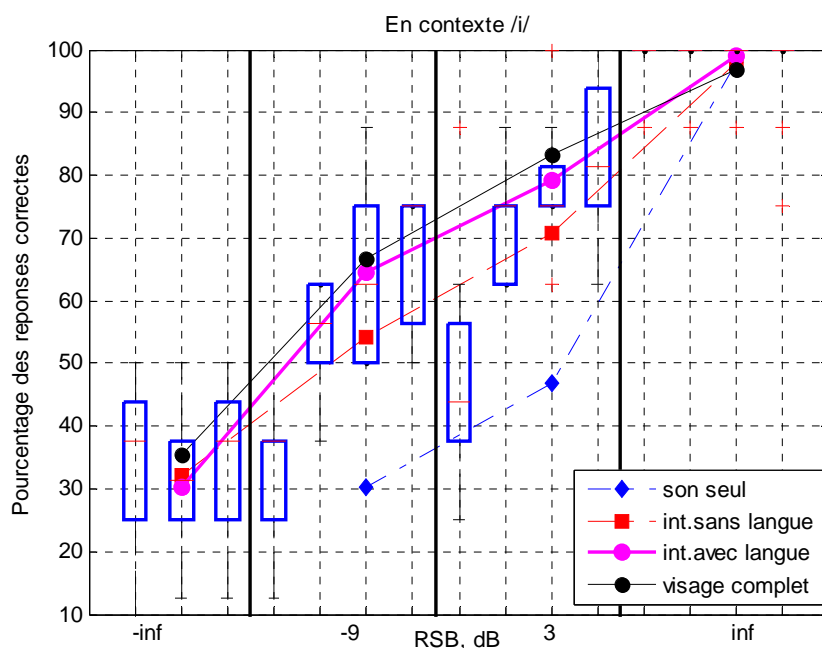
<input type="checkbox"/>	B (comme dans Bon, roBe)
<input type="checkbox"/>	D (comme dans Dans, aiDe)
<input type="checkbox"/>	G (comme dans Gare, baGue)
<input type="checkbox"/>	J (comme dans Je, Gilet, Geôle)
<input type="checkbox"/>	L (comme dans Lent, soL)
<input type="checkbox"/>	R (comme dans Rue, veniR)
<input checked="" type="checkbox"/>	V (comme dans Vous, rêVe)
<input type="checkbox"/>	Z (comme dans Zéro, maiSon, roSe)

Fig. D.2 – Fenêtre du programme pour choisir la consonne reconnue

Annexe E : Pourcentage d'identification correcte de consonne en différents contextes vocaliques pour les différentes séries du test

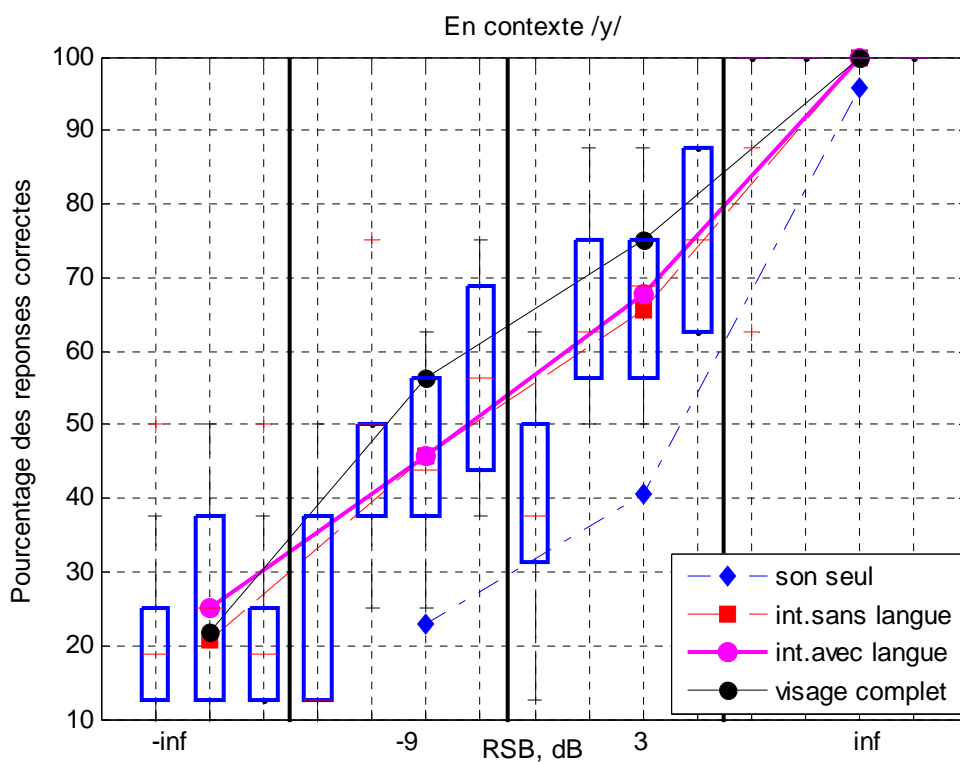


a)

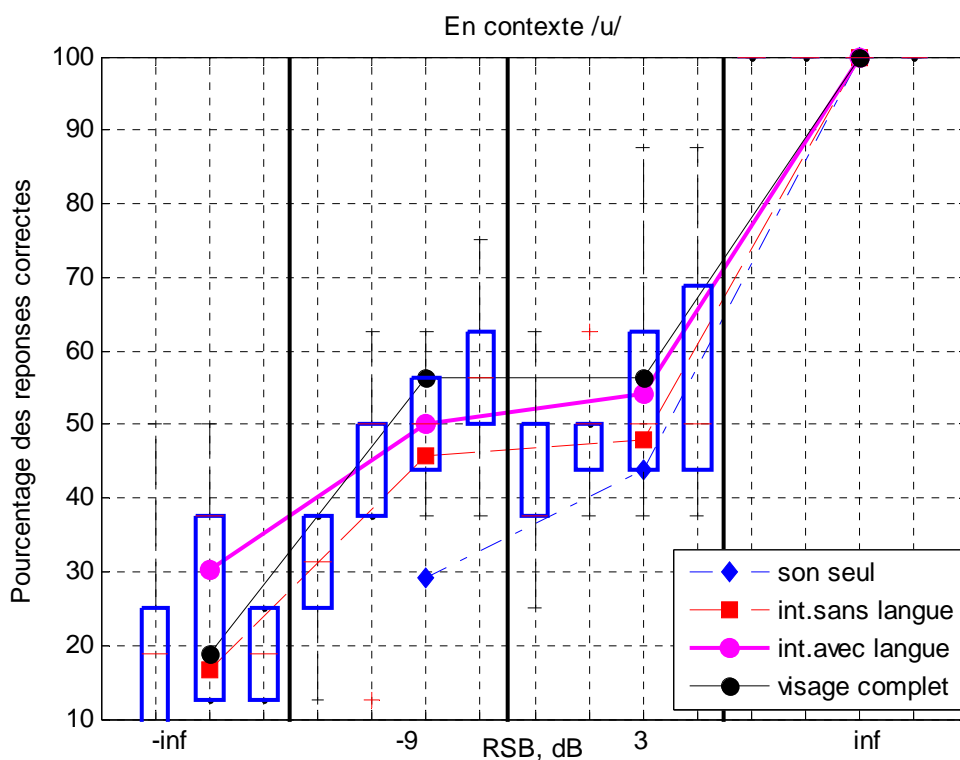


b)

Fig. E.1 – Scores d’intelligibilité du groupe I en fonction du bruit (a – en contexte /a/; b – en contexte /i/) pour les différentes conditions (de bas en haut, de gauche à droite dans chaque secteur) : audio seul; intérieur sans langue; intérieur avec langue; visage complet ; série de généralisation ; les barres dans chaque secteur représentent l’étendue des scores pour l’ensemble des sujets, les valeurs de médiane et des 2 quartiles pour ces scores

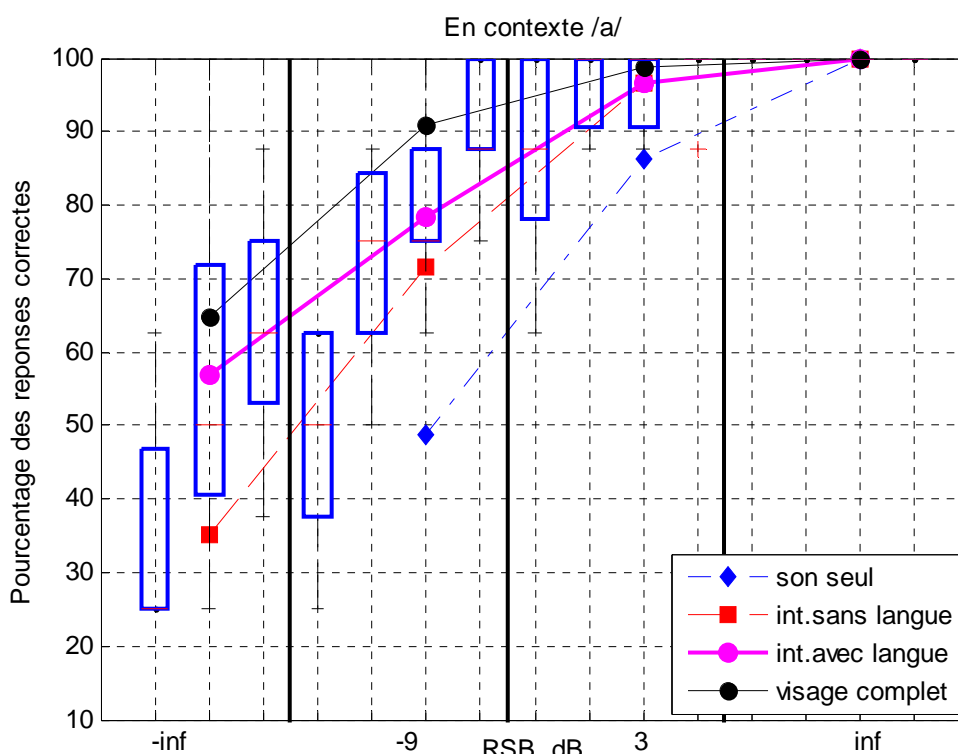


a)

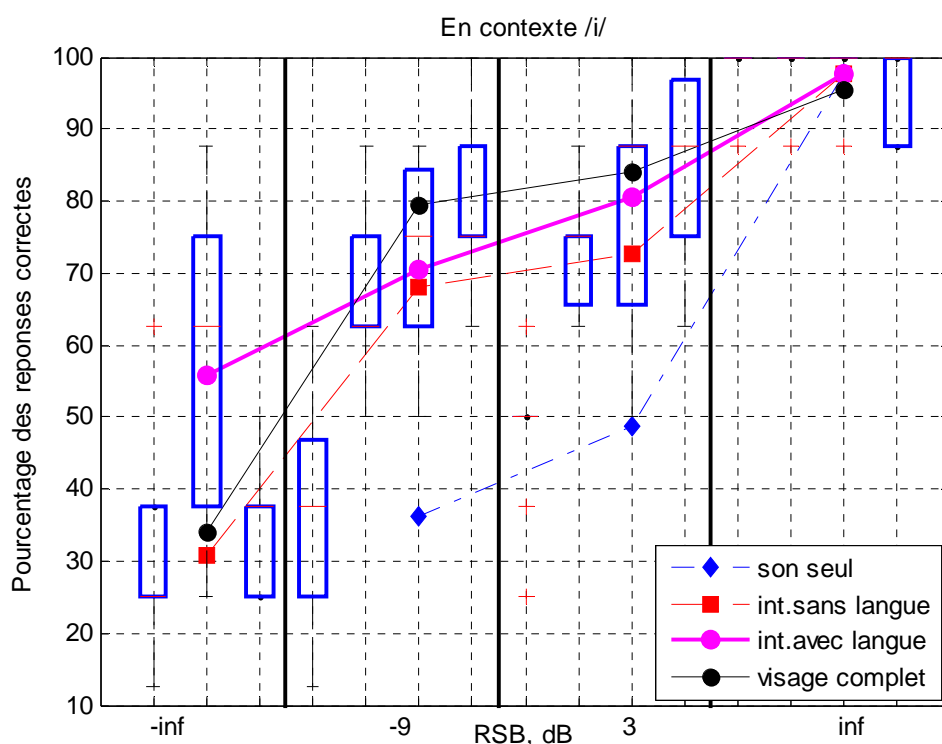


b)

Fig. E.2 – Scores d’intelligibilité du groupe I en fonction du bruit (a – en contexte /y/; b – en contexte /u/) pour les différentes conditions (de bas en haut, de gauche à droite dans chaque secteur) : audio seul; intérieur sans langue; intérieur avec langue; visage complet ; série de généralisation ; les barres dans chaque secteur représentent l’étendue des scores pour l’ensemble des sujets, les valeurs de médiane et des 2 quartiles pour ces scores

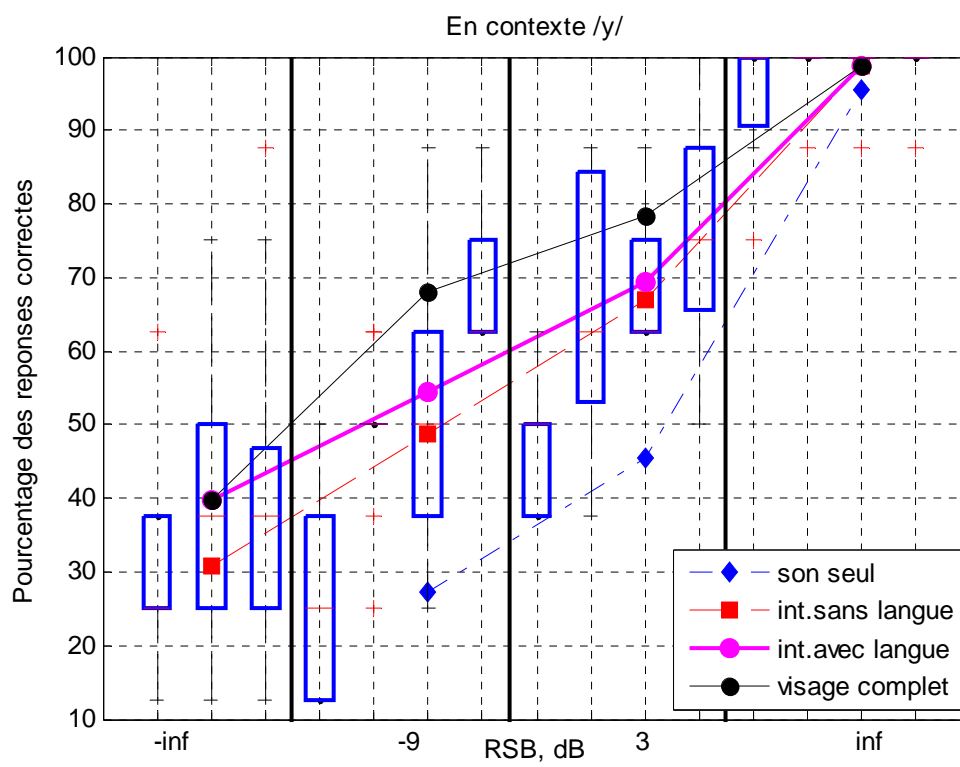


a)

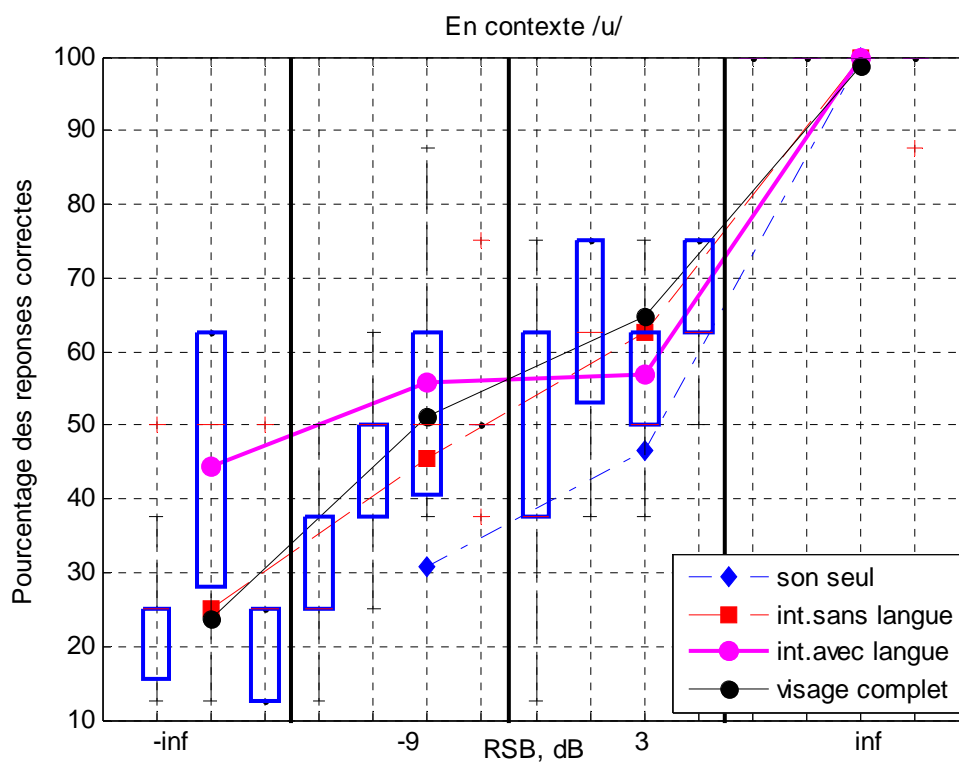


b)

Fig. E.3 – Scores d’intelligibilité du groupe II en fonction du bruit (a – en contexte /a/; b – en contexte /i/) pour les différentes conditions (de bas en haut, de gauche à droite dans chaque secteur) : audio seul; intérieur sans langue; intérieur avec langue; visage complet ; série de généralisation ; les barres dans chaque secteur représentent l’étendue des scores pour l’ensemble des sujets, les valeurs de médiane et des 2 quartiles pour ces scores



a)



b)

Fig. E.4 – Scores d’intelligibilité du groupe II en fonction du bruit (a – en contexte /y/; b – en contexte /u/) pour les différentes conditions (de bas en haut, de gauche à droite dans chaque secteur) : audio seul; intérieur sans langue; intérieur avec langue; visage complet ; série de généralisation ; les barres dans chaque secteur représentent l’étendue des scores pour l’ensemble des sujets, les valeurs de médiane et des 2 quartiles pour ces scores