

# Identification visuelle et catégorisation de consonnes en français québécois Visual Identification and Categorization of Consonants in Quebec French

par • by

Benoît Jutras, MOA, Jean-Pierre Gagné, PhD, Michel Picard, PhD et Julie Roy, MOA

Université de Montréal  
Montréal, Québec

## ABRÉGÉ

L'objectif de la présente étude est d'établir les catégories de consonnes qui forment les visèmes du français parlé au Québec. À cet effet, nous avons élaboré trois versions d'un test de reconnaissance de visèmes à partir des 17 consonnes du français parlé au Québec. Deux locutrices ont produit ces consonnes à l'intérieur d'une structure syllabique de type (a-C-a) conduisant à la création de deux tests. Chaque test est constitué de 85 éléments (5 présentations des 17 consonnes réparties de façon aléatoire). Quarante-six adultes ayant une acuité auditive normale ont participé à l'étude. Les résultats ont démontré que le nombre de visèmes varie d'une locutrice à l'autre. La production de la Locutrice 1 a permis de former six visèmes tandis que celles de la Locutrice 2 en a créé sept, dont un indépendant (/r/). Les visèmes /p, b, m/, /f, v/, /ʃ, ʒ/ et /s, z/ sont communs aux deux locutrices. La présente étude est une première étape vers l'élaboration d'un test clinique de reconnaissance de visèmes.

## ABSTRACT

The goal of this study is to categorise the consonants that form the visemes of the French spoken in Québec. To this end, we've developed three versions of a viseme recognition test based on the 17 consonants of Québec French. Two female speakers produced these consonants within a (a-C-a)-type syllabic structure, which led to the creation of two tests. Each test is composed of 85 elements (5 presentations of the 17 consonants, randomly distributed). Forty six adults with normal hearing participated in the study. The results showed that the number of visemes varied from one speaker to the other. Speaker One produced six visemes while Speaker Two produced seven, one being an independent (/r/). Visemes /p,b,m/, /f, v/, /ʃ, ʒ/ and /s, z/ were common to both speakers. This study is a first step towards the development of a clinical test for viseme recognition.

## MOTS CLÉ

perception visuelle de la parole • lecture labiale • évaluation

Les recherches sur la perception de la parole peuvent être regroupées, entre autres, en trois grandes catégories : les études qui traitent des aspects auditifs, celles qui portent sur les aspects visuels et celles qui s'intéressent à la fois aux aspects auditifs et visuels de la perception des différents éléments de la parole. La présente recherche s'inscrit dans le mouvement du second groupe d'études puisqu'elle vise à déterminer la catégorisation visuelle des consonnes en français québécois.

Il est connu que la vision joue un rôle important pour la reconnaissance de la parole lorsque l'information auditive est altérée par une situation d'écoute difficile. Si on prend l'exemple d'un individu ayant une acuité auditive normale qui entretient une conversation avec une autre personne dans un environnement sonore bruyant ou réverbérant. Les indices visuels donnés par le locuteur deviennent indispensables à cet individu pour le maintien de la conversation. Quant aux personnes ayant une atteinte auditive, ces indices sont souvent nécessaires à la compréhension d'un message verbal. Binnie, Jackson et Montgomery (1974) mentionnent que les individus ayant

une hypoacousie, plus spécifiquement ceux qui ont une perte auditive sévère ou profonde, ont recours principalement à la vision pour les aider à décoder l'information verbale. Ils ajoutent que les éléments acoustiques sont secondaires à la reconnaissance du message. Dans certains cas, la lecture labiale est le seul moyen dont les personnes disposent pour percevoir la parole (Erber, 1975; Gagné, 1994).

La lecture labiale est le thème générique qui désigne la perception visuelle de la parole. Jacobs (1982) rapporte que la lecture labiale est un processus actif. Cette habileté permet, selon elle, d'obtenir de l'information linguistique en regardant les mouvements des lèvres du locuteur ainsi que ceux de sa mâchoire, de sa musculature et de ses expressions faciales. Selon Jeffers et Barley (1971), la lecture labiale est l'habileté qui permet la perception, la discrimination et l'identification des mouvements individuels des sons, incluant les mouvements de transition et de fusion des sons. Les unités minimales sous-jacentes à une telle perception sont les visèmes ou les phonèmes visuels (Fisher, 1968). Fisher définit le visème comme étant toute unité minimale de la parole qui est perçue visuellement distincte

d'une autre unité. C'est un concept créé à partir de la perception visuelle des éléments de la parole et qui mène à l'élaboration de valeurs normatives tout comme dans le cas des phonèmes. Cependant, le nombre de phonèmes d'une langue est généralement supérieur à celui des visèmes. Cette différence s'explique par le fait que plusieurs éléments de la parole se distinguent acoustiquement, mais qu'ils se confondent visuellement (ex.: /p, b, m/, /f, v/, etc.). On regroupe les sons qui ont un *pattern* visuel similaire, les considérant comme formant une seule catégorie. Il existe des catégories qui peuvent ne contenir qu'un seul son de la parole, ce que Owens et Blazek (1985) appellent les visèmes indépendants.

O'Neill (1954) a démontré que la vision contribue à la discrimination de 57% des consonnes et de seulement 30% des voyelles. Ces résultats suggèrent que les consonnes sont des éléments critiques pour la perception visuelle de la parole. Walden, Prosek, Montgomery, Scherr et Jones (1977) soulignent que la première étape de l'entraînement à la lecture labiale, en utilisant la méthode analytique, consiste à consolider d'abord les habiletés de base impliquées dans la perception visuelle des consonnes avant de travailler les phrases, le contexte ou l'entraînement auditivo-visuel.

Les résultats de plusieurs études ont permis de regrouper en catégories les consonnes qui ont été reconnues visuellement comme ayant des traits similaires ou distinctifs (Binnie & coll., 1974; Binnie, Jackson & Montgomery, 1976; Erber, 1972; Fisher, 1968; Jackson, 1988; Walden & coll., 1977; Woodward & Barber, 1960). Cependant, le nombre de catégories et le nombre d'éléments formant un visème varient d'une recherche à l'autre. En anglais, Woodward et Barber (1960) ont trouvé quatre catégories de consonnes, Fisher (1968) ainsi que Binnie et coll. (1974) en ont rapporté cinq tandis que Walden et coll. (1977) ont trouvé six visèmes. Par ailleurs, Woodward et Barber (1960) ont formé une catégorie de consonnes composée de /p, b, m/ tandis que Fisher (1968) a regroupé les consonnes /p/, /b/, /m/ et /d/ en un seul visème.

Plusieurs facteurs peuvent expliquer les différences entre les études portant sur le nombre de visèmes et sur le nombre de consonnes incluses dans une catégorie. Lesner (1988) rapporte que l'un de ces facteurs est l'articulation propre à chaque locuteur. Plus l'articulation est claire, plus il est facile de reconnaître visuellement les éléments de la parole (Gagné, Masterson, Munhall, Bilida, & Querengesser, 1994; Gagné, Querengesser, Folkeard, Munhall, & Masterson, 1995). Cependant, un locuteur qui est intelligible dans le mode auditif ne le sera pas nécessairement en modalité visuelle (Gagné & coll., 1994, 1995). De plus, les effets environnementaux, tels que

l'éclairage, la distance et l'angle d'observation sont des variables qui peuvent influencer le nombre de catégories ou le nombre d'éléments par visème (Jackson, 1988). De leur côté, Owens et Blazek (1985) soulignent que les stimuli utilisés, les participants recrutés pour l'étude, la tâche demandée aux participants, la procédure d'application du test et les analyses statistiques comptent parmi d'autres facteurs qui peuvent expliquer l'inclusion ou l'exclusion d'une consonne dans une catégorie ou la formation de catégories. Les critères utilisés par les chercheurs pour définir, de façon opérationnelle, des catégories de phonèmes font également partie des éléments qui montrent une différence entre les résultats des études (Jackson, 1988).

Selon certains chercheurs (Hutton, 1960; Walden & coll. 1977; Walden, Erdman, Montgomery, Schwartz & Prosek, 1981), l'effet d'apprentissage exerce une influence sur les résultats des participants soumis à des tâches d'identification visuelle de consonnes. Les résultats des recherches ont démontré une augmentation du nombre de visèmes chez les participants ayant eu des sessions d'entraînement à la reconnaissance visuelle de consonnes.

#### **Les épreuves de reconnaissance visuelle de consonnes et la réadaptation audiolinguistique**

Binnie et coll. (1976) propose un test de dépistage en lecture labiale qui comporte neuf catégories de consonnes qui ont été formées à partir de résultats obtenus auprès de participants sans perte auditive. Ils mentionnent qu'une personne n'ayant pas reconnu visuellement certaines consonnes, devrait se voir offrir un programme d'exercices spécifiques adapté à ses besoins. Ce programme devrait contenir des exercices de discrimination et d'identification pour aider la personne à développer des habiletés de base en lecture labiale (Walden & coll., 1977). Il est entendu, par exemple, qu'un entraînement axé sur l'identification du lieu d'articulation n'est pas justifié auprès d'un individu qui reconnaît systématiquement les consonnes ayant en commun ce trait visuel (Binnie & coll., 1976). Lorsque l'individu a atteint les critères de reconnaissance visuelle de consonnes, il faut alors orienter les thérapies vers des activités d'entraînement qui visent à développer, de façon optimale, les indices linguistiques et contextuels, présents dans le message verbal (Gagné, 1994).

Gagné, Seewald et Stouffer (1987) ont proposé un protocole canadien d'évaluation de la lecture labiale en anglais. Ce protocole contient deux parties : l'épreuve de *Speech Perception in Noise (SPIN)* et celle de la reconnaissance visuelle des consonnes. Il a été expérimenté auprès de participants adultes sans perte auditive. Les auteurs en ont conclu que le test de reconnaissance visuelle de consonnes est approprié pour l'évaluation de la percep-



tion visuelle de la parole.

En français, il n'existe pas de test de reconnaissance visuelle de consonnes. Cette étude est donc une première démarche vers l'élaboration d'une épreuve valide, fiable et normalisée en français parlé au Québec<sup>1</sup>. Il s'agit d'une épreuve qui évalue l'habileté à identifier visuellement les consonnes chez des «labio-lecteurs» québécois dont la langue maternelle est le français. Dans ce sens, l'objectif spécifique de la présente étude est de déterminer quelles sont les consonnes qui forment les visèmes en français québécois.

### Méthodologie

Le déroulement de l'expérimentation s'est effectué en quatre étapes : (a) le choix des stimuli, (b) l'enregistrement du matériel, (c) le montage du test et (d) l'application du test de reconnaissance visuelle des consonnes en français québécois.

#### Le choix des stimuli

Le matériel linguistique est constitué de syllabes sans sens de type VCV. La voyelle «a» précède et suit invariablement la consonne. Cette voyelle a été choisie parce qu'elle permet de voir clairement les mouvements de la bouche et des muscles faciaux. Elle est également celle qui produit le moins de coarticulation parmi l'ensemble des voyelles utilisées en français. Les 17 consonnes prononcées par les locutrices représentent celles que l'on retrouve en français québécois, c'est-à-dire /p/, /b/, /m/, /f/, /v/, /ʃ/, /ʒ/, /t/, /d/, /n/, /l/, /s/, /z/, /k/, /g/, /ŋ/ et /r/.

#### L'enregistrement du matériel

**Les locutrices.** Deux personnes adultes de sexe féminin ont été choisies comme locutrices. La Locutrice 1 était une étudiante en audiologie à l'Université de Montréal et la Locutrice 2 est audiologiste dans un centre hospitalier à Montréal. Leur participation à l'étude s'est faite sur une base volontaire. Les critères de sélection des locutrices étaient : (a) avoir le français comme langue maternelle, (b) avoir des mouvements articulatoires facilitant la lecture labiale et (c) ne pas avoir de régionalismes lors de l'élocution. Deux expérimentateurs (J.-P. G. et J. R.) ont vu à ce que les locutrices rencontrent les deuxième et troisième critères.

**L'équipement.** L'enregistrement des stimuli a été effectué avec un caméscope Sony High 8, modèle CCDV-5000. Un moniteur Panasonic CT 2010Y a été placé devant et à la gauche de la locutrice pour lui permettre de corriger au besoin la position dans laquelle elle se trouvait pour l'enregistrement des syllabes. Celui-ci permettait ainsi à la locutrice de se centrer dans un cadre préalablement

établi à l'aide de marqueurs fixés à l'écran du moniteur. Ces marqueurs délimitaient le contour du visage, des épaules et la position des yeux.

L'enregistrement nécessitait également un microphone Atus AT-828 qui était fixé à 19 cm de la bouche de la locutrice; deux projecteurs lumineux qui avaient une orientation différente dans l'espace : un projecteur recouvert d'un filtre blanc placé directement derrière la caméra qui éclairait la locutrice et un autre situé à la droite de la locutrice dont le jet de lumière était dirigé vers le plafond pour uniformiser la lumière ambiante; deux ampoules halogènes de type JPD 120V-600 WD; trois lampes de 100 Watts dont deux se trouvaient derrière la locutrice éclairant le panneau dans un angle de 45° pour éliminer les ombrages de la tête et l'autre recouverte d'un filtre blanc, placée en face de la locutrice pour éclairer l'intérieur de la bouche et pour éliminer les ombrages sous le menton; un sonomètre Radio Shack pour mesurer le niveau de pression sonore du signal de référence au début de l'enregistrement et celui du bruit de fond dans le local d'enregistrement; un audiomètre portatif Peter pour émettre le signal de référence; un panneau gris vert érigé derrière la locutrice servant de toile de fond.

**La procédure d'enregistrement.** Les locutrices ont participé à une séance d'enregistrement sur bande vidéoscopique d'une durée d'environ deux heures. Elles devaient prononcer à cinq reprises chaque syllabe en utilisant la voix afin de favoriser une articulation la plus naturelle possible. Le cadrage employé était le gros plan axé sur le visage de la locutrice vu de face incluant la tête et la partie supérieure des épaules. Les locutrices avaient le cou dégagé et les épaules étaient vêtues de bleu marin. Elles portaient un maquillage qui favorisait la reconnaissance des mouvements articulatoires tout en ayant une apparence naturelle (Jacobs, 1982). Les cheveux de la Locutrice 1 ont été relevés et attachés de façon à ce que le visage et le cou soient entièrement visibles.

Préalablement à l'enregistrement des syllabes, l'expérimentateur a procédé à l'étalonnage des appareils. Le 0 du vumètre de la caméra correspondait à 84.5 dB SPL. Cette mesure a été obtenue dans le but éventuel d'effectuer l'expérience en modalité auditive et auditivo-visuelle. Ensuite, l'expérimentateur a procédé à la mesure du niveau de bruit ambiant présent dans le local d'enregistrement. Les niveaux trouvés se situaient entre 34 et 36 dB(A).

Lors des séances d'enregistrement, la locutrice prononçait clairement chaque syllabe à cinq reprises de façon consécutive. Elle devait fixer la caméra deux secondes avant l'élocution, puis produire verbalement la syllabe et garder la même position pour encore deux autres secondes. Pendant ce temps, elle ne pouvait pas cligner des yeux, ni bouger la tête ou les yeux ou effectuer des

mouvements distrayants. Elle pouvait prendre de courtes pauses entre la production des syllabes si elle en ressentait le besoin. L'expérimentateur surveillait la position de la locutrice, les syllabes produites, les mouvements imprévus et demandait d'apporter occasionnellement des corrections lorsque jugé nécessaire.

### **Le montage du test**

L'enregistrement a été visionné par deux expérimentateurs afin de déterminer laquelle parmi les cinq productions de la même syllabe devait être retenue pour le montage du test. Selon la convention établie entre les deux expérimentateurs, la troisième répétition était le stimulus retenu. Toutefois, si cette répétition s'avérait inacceptable par rapport aux critères d'enregistrement pré-établis, alors les deux expérimentateurs choisissaient, d'un commun accord, une autre syllabe parmi celles produites.

Un test a été créé pour chacune des locutrices. L'extrait retenu par les deux expérimentateurs de chacune des 17 consonnes a été copié cinq fois. Les deux tests comportent donc 85 éléments (5 présentations des 17 consonnes distribuées au hasard). Ils existent en trois versions et pour chaque version, les 85 syllabes sont réparties de façon aléatoire. Au début du test, il y a huit items de pratique.

Le montage des images a été conçu de la façon suivante : (a) pendant trois secondes, une fenêtre de couleur bleue indiquait le numéro de l'item; (b) la locutrice prononçait la syllabe; (c) pendant six secondes, une fenêtre de couleur rose apparaissait pour indiquer le numéro de la réponse. Le même scénario se répétait pour les 85 syllabes du test. Pour chaque version, la durée approximative du test est de 20 à 22 minutes.

### **L'application du test de reconnaissance visuelle des consonnes en français québécois**

**Les participants.** Quarante-deux femmes et quatre hommes ont participé sur une base volontaire à l'expérimentation. Ils étaient âgés entre 18 et 30 ans (moyenne de 20,9 ans). Tous les participants ont affirmé que le français était leur langue maternelle. De plus, ils ont tous réussi le test de dépistage auditif (seuils auditifs inférieurs à 25 dB HL pour les fréquences de 250, 500, 1000, 2000 et 4000 Hz) et un test de dépistage visuel (acuité visuelle d'au moins 20/40 avec ou sans verres correctifs). Les participants étaient des étudiants inscrits au premier ou au deuxième cycle à l'Université de Montréal. Ils provenaient des programmes d'orthophonie et d'audiologie (n = 28), d'ergothérapie (n = 6), de physiothérapie (n = 4), de sciences infirmières (n = 2), de génie (n = 2), de finances (n = 2), de biologie (n = 1) et d'études françaises (n = 1). Les participants ont reçu 25\$

pour leur participation à l'étude.

**Le déroulement.** Les 46 participants de l'étude ont formé dix groupes de deux à six personnes, selon la disponibilité de chacun. Ils ont participé à deux sessions de deux heures. Pour chacune des sessions, les participants étaient soumis au test de reconnaissance visuelle de consonnes et également à un test de reconnaissance visuelle de phrases. Il y avait un délai de deux à quatorze jours entre les deux sessions. Il est à noter que dans la présente étude, seules les données provenant de l'épreuve de reconnaissance visuelle de consonnes seront traitées.

Lors de la première session, les participants ont reçu une lettre expliquant le projet et ont signé le formulaire de consentement. Ensuite, ils ont été soumis aux tests de dépistage auditif et visuel. Par la suite, ils ont effectué la première partie du test comprenant trois des six versions prévues. Pour ce faire, les participants étaient assis à 1,25 mètre d'un moniteur, en demi-cercle de façon à ce que chacun puisse clairement voir l'écran. Ils devaient visionner les cassettes du test et répondre sur une feuille que l'expérimentateur leur avait préalablement distribuée. Cette feuille avait comme en-tête le choix des réponses possibles. À chaque item, la voyelle /a/ était déjà inscrite sur la feuille et les participants devaient seulement écrire la consonne qu'ils percevaient. Après les trois premières épreuves, une pause de 10 à 15 minutes était allouée. Huit items de familiarisation précédait les épreuves du test. Lors de la deuxième session, les participants ont eu droit à un bref rappel des instructions. Le déroulement de cette rencontre était semblable à la première, à l'exception que les participants n'ont pas été exposés aux items de familiarisation. Ils ont visionné les cinq dernières épreuves. Durant les sessions, l'expérimentateur surveillait les participants pour éviter toute possibilité de plagiat.

Tous les groupes ont passé les épreuves dans un ordre différent pour limiter l'effet d'apprentissage. Ils ont tous effectué les trois versions du test de consonnes pour chaque locutrice. Les participants n'avaient pas accès à la trame sonore et se basaient seulement sur les indices visuelles pour identifier les consonnes.

### **Traitement des données**

Nous avons traité les données en trois étapes. Premièrement, les réponses recueillies ont été compilées sous forme de scores de *reconnaissance de consonnes* pour chaque participant et pour l'ensemble des participants. Deuxièmement, les données de l'ensemble des participants ont permis de créer une matrice de confusion de consonnes pour chaque locutrice. Troisièmement, les matrices de confusion ont servi à l'établissement de catégories de consonnes (visèmes). Nous avons utilisé les règles de Walden et coll. (1977) pour déterminer la composition



des différentes catégories de consonnes. Plus spécifiquement, les consonnes qui étaient le plus souvent confondues entre elles ont été groupées jusqu'à ce qu'un minimum de 75% de bonnes réponses ait été atteint. Ce critère a également été utilisé par d'autres auteurs: Benguerel et Pichora-Fuller (1982), Gagné, Seewald et Hassan (1989), Gagné et coll. (1987), Owen et Blazek (1985), Walden et coll. (1981), Walden et coll. (1977).

### Analyse des données

Les figures 1 et 2 représentent les catégories de consonnes et la moyenne du pourcentage de bonnes réponses obtenues pour chaque visème pour la Locutrice 1 et la Locutrice 2, respectivement. Il est possible d'observer une différence dans le nombre de visèmes entre les deux locutrices. On dénombre six catégories de consonnes selon la production verbale de la Locutrice 1 et sept visèmes pour celle de la Locutrice 2. Pour cette dernière, une catégorie indépendante formée à partir de la consonne /r/ s'ajoute aux six autres visèmes.

Les catégories les mieux reconnues sont /p, b, m/; /f, v/ et /ʃ, ʒ/. Les consonnes produites plus à l'arrière de la cavité buccale sont les moins bien identifiées. La catégorie /s, z/ est la seule dont le pourcentage de reconnaissance diffère nettement entre les deux locutrices (loc. 1 : 87% et loc. 2 : 98,6%).

### Discussion

Les résultats de l'étude démontrent que le nombre de catégories de consonnes varient d'une locutrice à l'autre. Ces résultats vont dans le même sens que ceux des études de Kricos et Lesner (1982, 1985) qui concluent que l'intelligibilité visuelle et la production de visèmes changent en fonction du locuteur. Bien que les deux locutrices de la présente étude avaient été

choisies pour la clarté de leur articulation, il n'en demeure pas moins qu'un nombre inégal de visèmes a été créé à partir de leur production verbale. Ce phénomène pourrait s'expliquer selon deux facteurs. Le premier est relié à la production verbale des consonnes propre à chacune des

Figure 1. Moyenne (colonne) et écart-type (petit carré) du pourcentage de reconnaissance de visèmes en fonction des catégories de consonnes (n = 46) pour la Locutrice 1.

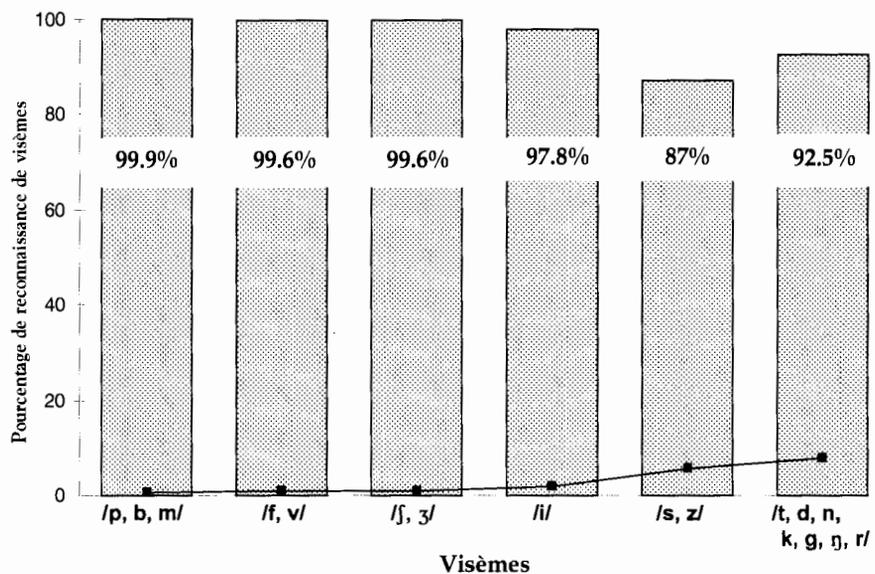
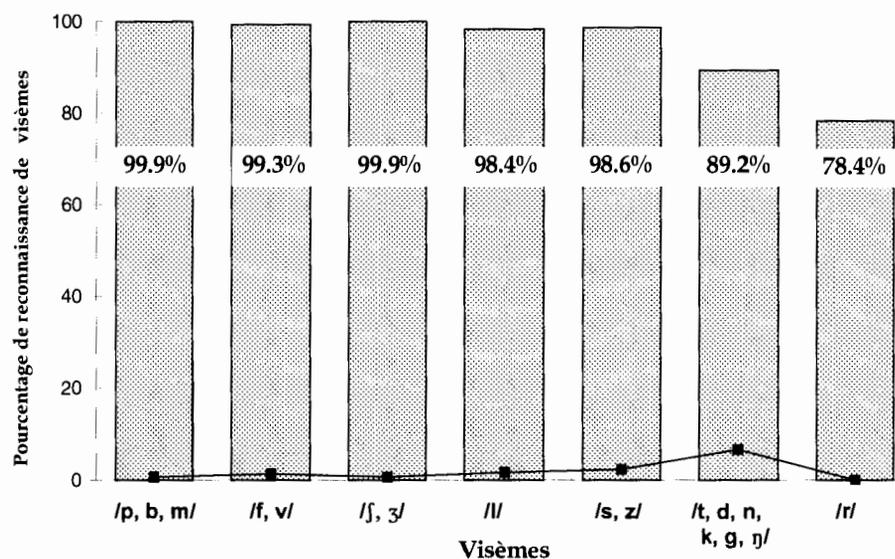


Figure 2. Moyenne (colonne) et écart-type (petit carré) du pourcentage de reconnaissance de visèmes en fonction des catégories de consonnes (n = 46) pour la Locutrice 2.



locutrices, tel que relevé par Lesner (1988). La Locutrice 2 articulait les syllabes avec une plus grande ouverture buccale ce qui fournissait aux participants des indices permettant une meilleure perception de la position de la langue. Les participants pouvaient alors associer le phonème /r/ à ce type de production. Quant au deuxième facteur, il découle d'une décision méthodologique et est relié à la définition du visème. On a déjà mentionné que le visème est un concept qui permet d'établir des «normes» à partir de données perceptives. Dans la présente étude, le critère normatif pour la formation de catégories a été établi à 75%. Le pourcentage de reconnaissance de la consonne /r/ était de 63% et de 78% lorsqu'elle était produite par la Locutrice 1 et 2, respectivement. Si le critère avait été plus sévère (ex. : 80%), il y aurait eu un nombre égal de catégories.

Par ailleurs, la composition d'une catégorie indépendante pour la consonne /r/ ne semble pas spécifique au français québécois. Selon l'étude de Binnie et coll. (1976), le /r/ peut former, à lui seul, un visème en anglais.

#### *Implications cliniques d'un test d'évaluation de certaines capacités en lecture labiale*

L'un des objectifs de la recherche sur la catégorisation de consonnes est de développer un test clinique qui permettrait d'évaluer la capacité de personnes malentendantes à identifier les visèmes en français québécois. Des valeurs normatives découlant de ce test peuvent servir de point de repère au clinicien pour l'élaboration d'un plan d'intervention. Binnie et coll. (1976) propose que les normes soient établies à partir de la moyenne des participants de l'étude normative moins un écart-type pour l'ensemble des réponses correctement identifiées. Le même calcul s'applique également pour chacune des catégories. Nous rapportons dans le Tableau 1 la moyenne et l'écart-type des données des 46 participants de la présente étude, recueillies à partir des consonnes énoncées par la Locutrice 1. Ces valeurs vont

servir de normes et les réponses du participant 7 seront utilisées comme exemple pour illustrer l'application des valeurs normatives. Rappelons que les données utilisées pour établir les normes proviennent d'étudiants universitaires. Ces données peuvent être potentiellement supérieures à celles que l'on retrouverait auprès de participants moins sélectionnés et plus représentatifs de la population en général. Par contre, elles démontrent qu'il est possible d'obtenir ces valeurs et qu'en situation d'entraînement à la lecture labiale, elles pourraient servir de cibles à atteindre.

Le participant 7 a un taux de réussite inférieur aux normes pour le pourcentage de reconnaissance de l'ensemble des visèmes (75,33%). Cependant, il peut identifier correctement les visèmes /p, b, m/, /f, v/, /ʃ, ʒ/, /s, z/ et /l/ avec un pourcentage qui dépasse les valeurs normatives. Par ailleurs, les résultats qu'il a obtenus pour le visème /t, d, n, k, g, ŋ, r/, se retrouvent en deçà des normes. Ces données suggèrent que le participant 7 pourrait bénéficier de sessions d'entraînement en lecture labiale. Les interventions devraient débiter d'abord par des séances de type analytique en mettant l'accent sur la discrimination de visèmes. À la fin des sessions, le participant 7 devrait arriver à mieux distinguer les consonnes du visème /t, d, n, k, g, ŋ, r/ des consonnes formant les autres visèmes.

#### **Conclusion**

La présente étude avait pour but d'établir les catégories de consonnes en français québécois à partir de leur identification visuelle. Les résultats ont montré que le nombre de visèmes varie en fonction du locuteur ce qui va dans le même sens que les conclusions des études Kricos et Lesner (1982, 1985) et de Lesner (1988). De plus, parmi les catégories formées, on retrouve les visèmes /p, b, m/, /f, v/, /ʃ, ʒ/ que Jackson (1988) considère comme étant des catégories «universelles», car elles se retrouvent dans plusieurs études (voir Jackson, 1988).

**Tableau 1. Moyenne des scores et écart-type obtenus à partir de la production de la Locutrice 1 pour chacun des visèmes et pour l'ensemble des visèmes. Il apparaît également les valeurs se situant à un écart-type de la moyenne pour chacune des données ainsi que les valeurs recueillies auprès du participant 7.**

Locutrice 1	Total	/p, b, m/	/f, v/	/ʃ, ʒ/	/l/	/s, z/	/t, d, n, k, g, ŋ, r/
Moyenne (M)	80,87	14,98	9,97	9,96	4,89	8,70	32,39
Écart-type (É-t)	3,46	0,15	0,22	0,19	0,31	1,58	3,15
M moins É-t	77,41	14,83	9,75	9,77	4,58	7,12	29,24
Participant 7	75,33	15	10	10	4,67	9,67	26



La catégorisation de consonnes en français québécois est l'une des étapes vers l'élaboration d'un test clinique de reconnaissance de visèmes. Il faudrait subséquemment vérifier l'équivalence des versions du test pour chacune des locutrices. Les étapes suivantes seraient d'évaluer la fidélité du test et de normaliser une telle épreuve auprès des personnes ayant une atteinte auditive.

#### Note

1. Dans le texte, nous utilisons le terme «français québécois» pour simplifier l'expression «français parlé au Québec».

**Prière d'envoyer toute correspondance à :** Jean-Pierre Gagné, École d'orthophonie et d'audiologie, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal (Québec), H3C 3J7.

Soumis : décembre 1996  
 Accepté : septembre 1997

#### Références

- Benguereel, A. P., & Pichora-Fuller, M. K. (1982). Coarticulation effects in lipreading. *Journal of Speech and Hearing Research, 25*, 600-607.
- Binnie, C. A., Jackson, P. L., & Montgomery, A. A. (1974). Auditory and visual contributions to the perception of consonants. *Journal of Speech and Hearing Research, 17*, 619-630.
- Binnie, C. A., Jackson, P. L., & Montgomery, A. A. (1976). Visual intelligibility of consonants: A lipreading screening test with implications for aural rehabilitation. *Journal of Speech and Hearing Research, 41*, 530-539.
- Erber, N. P. (1972). Auditory, visual, auditory-visual recognition of consonants by children with normal and impaired hearing. *Journal of Speech and Hearing Research, 15*, 413-422.
- Erber, N. P. (1975). Auditory-visual perception of speech. *Journal of Speech and Hearing Disorders, 40*, 481-492.
- Fisher, C. G. (1968). Confusions among visually perceived consonants. *Journal of Speech and Hearing Research, 11*, 796-804.
- Gagné, J.-P. (1994). Visual and audiovisual speech perception training: Basic and applied research needs. In J.-P. Gagné & N. Tye-Murray (Eds.), *Research in audiological rehabilitation: Current trends and future direction (supplement)*. *Journal of the Academy of Rehabilitative Audiology, 27*, 133-159.
- Gagné, J.-P., Masterson, V., Munhall, K. G., Bilida, N., & Querengesser, C. (1994). Across talker variability in auditory, visual, and audiovisual speech intelligibility for conversational and clear speech. *Journal of the Academy of Rehabilitative Audiology, 27*, 135-158.
- Gagné, J.-P., Querengesser, C., Folkeard, P., Munhall, K. G., & Masterson, V.M. (1995). Auditory, visual, and audiovisual speech intelligibility for sentence-length stimuli: An investigation of conversational and clear speech. *The Volta Review, 97*, 33-51.
- Gagné, J.-P., Seewald, R. C., & Hassan, R. A. (1989). A Canadian test to assess visual-consonant recognition abilities. Présentation effectuée dans le cadre du congrès annuel de l'Association canadienne des orthophonistes et des audiologistes, Toronto.
- Gagné, J.-P., Seewald, R. C., & Stouffer, J. L. (1987). The development of a Canadian speechreading test: A report of preliminary findings. Présentation effectuée dans le cadre du congrès annuel de l'Association des orthophonistes et des audiologistes de l'Ontario, Toronto.
- Hutton, C. (1960). A diagnostic approach to combined techniques in aural rehabilitation. *Journal of Speech and Hearing Disorders, 25*, 267-272.
- Jackson, P. L. (1988). The theoretical minimal unit for visual speech perception: Visemes and coarticulation. *The Volta Review, 90*, 99-115.
- Jacobs, M. A. (1982). Visual communication (speech-reading) for the severely and profoundly hearing-impaired young adult. In D. Smith, G. Walter, & R. Whitehead, *Deafness and communication: Assessment and training, 271-293*. Baltimore: Williams & Wilkins.
- Jeffers, J., & Barley, M. (1971). *Speechreading (Lipreading)*. Springfield, IL: Charles C. Thomas.
- Kricos, P. B., & Lesner, S. A. (1982). Differences in visual intelligibility across talkers. *The Volta Review, 84*, 219-225.
- Kricos, P. B., & Lesner, S. A. (1985). Effect of talker differences on the speechreading of hearing-impaired teenagers. *The Volta Review, 87*, 5-16.
- Lesner, S. (1988). The talker. *The Volta Review, 90*, 89-98.
- O'Neill, J. J. (1954). Contributions of the visual components of oral symbols to speech comprehension. *Journal of Speech and Hearing Disorders, 19*, 429-439.
- Owens, E., & Blazek, B. (1985). Visemes observed by hearing impaired and normal hearing adult viewers. *Journal of Speech and Hearing Research, 28*, 381-393.
- Walden, B. E., Erdman, S. A., Montgomery, A. A., Schwartz, D. M., & Prosek, R. A. (1981). Some effects of training on speech recognition by hearing-impaired adults. *Journal of Speech and Hearing Research, 24*, 207-216.
- Walden, B. E., Prosek, R. A., Montgomery, A. A., Scherr, C. K., & Jones, C. J. (1977). Effects of training on the visual recognition of consonants. *Journal of Speech and Hearing Research, 20*, 130-145.
- Woodward, M. F., & Barber, C. G. (1960). Phoneme perception in lipreading. *Journal of Speech and Hearing Research, 3*, 212-222.