

CORRÉLATS ACOUSTIQUES DE LA PERCEPTION DES VOYELLES PRODUITES PAR DES LOCUTEURS SOURDS

Lucie Ménard, Julie Chrétien, Robin Lachapelle, Isabelle Marleau

Laboratoire de phonétique, Département de linguistique, UQAM

RÉSUMÉ

L'objectif du présent article est d'étudier les corrélats acoustiques de la perception auditive de voyelles orales du français produites par des locuteurs sourds profonds congénitaux appareillés. Des répétitions des sept voyelles / i e ε a y u o / produites par deux locuteurs sourds ont été soumises lors d'un test d'identification auditive à un ensemble de 40 auditeurs francophones entendants. Les valeurs des deux premiers formants (F1 et F2) et de la fréquence fondamentale (F0) ont été extraites. Les résultats des tests perceptifs montrent que la distance tonotopique entre les formants et la fréquence fondamentale (F1-F0 et F2-F1) constituent les corrélats acoustiques du degré d'aperture perçu et du lieu d'articulation perçu. Les résultats de cette étude confirment les travaux précédents qui ont porté sur des voyelles synthétiques du français. Les corrélats acoustiques mis au jour sont interprétés en regard de la variabilité rencontrée lors de la production des voyelles par les locuteurs sourds profonds congénitaux.

MOTS-CLÉS Perception des voyelles, surdit , phon tique articulatoire

ACOUSTIC CORRELATES OF VOWELS PRODUCED BY DEAF SPEAKERS

*Lucie M nard, Julie Chr tien, Robin Lachapelle,
Isabelle Marleau*

ABSTRACT

The objective of this paper is to study the acoustic correlates of auditory perception of French oral vowels produced by congenitally profound deaf speakers with prosthesis. Repetitions of the seven vowels / i e ε a y u o / produced by two profoundly deaf speakers have been submitted as an auditory perceptual test to 40 Quebec French normal hearing listeners. Values of two formants (F1 and

CORRELACI N AC STICA DE LA PERCEPCI N DE VOCALES EMITIDAS POR LOCUTORES SORDOS

*Lucie M nard, Julie Chr tien, Robin Lachapelle,
Isabelle Marleau*

R SUMEN

El objetivo del presente art culo, consiste en estudiar la correlaci n ac stica de la percepci n auditiva de vocales orales del franc s, producidas por locutores con sordera profunda y cong nita, que usan pr tesis auditivas. Las repeticiones de las siete vocales / i e ε a y u o / producidas por dos locutores sordos fueron sometidas durante una

F2) and of the fundamental frequency (F0) have been extracted. Results of the perceptual test show that the tonotopic distance between formant frequencies and F0 (F1-F0 and F2-F1) can be considered as good correlates of perceived height degree and perceived place of articulation. This study confirms earlier work carried out on French synthesized vowels. Those acoustic correlates can be used to interpret the variability found in the produced data.

KEYWORDS Vowel production, deafness,
articulatory phonetics

prueba de identificación auditiva, a un conjunto de 40 oyentes francófonos de audición normal. Los valores de dos formantes (F1 y F2) y de la frecuencia fundamental (F0) fueron extraídos. Los resultados de las pruebas perceptivas demuestran que la distancia tonotópica entre las formantes y la frecuencia fundamental (F1-F0 y F2-F1) constituyen la correlación acústica del grado de apertura y del lugar de articulación percibidos. Los resultados de este estudio confirman los trabajos anteriores refiriéndose a vocales sintéticas del francés. La correlación acústica puede ser utilizada para interpretar la variabilidad encontrada en la producción de las vocales por los locutores con sordera profunda y congénita.

PALABRAS CLAVES Percepción de las vocales,
sordera, fonética articulatoria

INTRODUCTION

Plusieurs études ont démontré qu'une perte auditive sévère ou profonde influence considérablement l'intelligibilité des voyelles produites par les locuteurs sourds. Les outils permettant de restaurer les informations sensorielles auditives, comme la prothèse amplificatrice ou l'implant cochléaire, ont un impact positif sur l'intelligibilité des productions (Lane et al., 2005; Perkell et al., 2004; McCaffrey et Sussman, 1994). Les travaux visant à comparer, pour un même locuteur, les caractéristiques acoustiques des sons produits en condition appareillée ou non font état d'une importante variabilité interindividuelle. Les valeurs de formants (F1, F2) et de fréquence fondamentale (F0) sont globalement affectées, mais à des degrés divers d'un locuteur à l'autre. Peut-on alors décrire les productions de ces sujets sourds appareillés ou non selon d'autres dimensions acoustiques?

Cette question trouve écho dans le cadre des études sur la normalisation et l'invariance acoustique des voyelles perçues. En français, il a été montré qu'une combinaison linéaire des valeurs de F0 et des deux premiers formants (F1, F2) peut être corrélée à l'identité perçue de voyelles synthétiques sur les plans du degré d'aperture et du lieu d'articulation (Ménard et al., 2002). Ces paramètres acoustiques invariants permettent de définir un espace auditif au sein duquel le locuteur a pour tâche de produire les patrons acoustiques correspondant à la voyelle perçue. Ces patrons acoustiques, définis non plus en termes de valeurs formantiques indépendantes, mais plutôt en termes de distances tonotopiques (F1-F0, F2-F1) (Syrdal et Gopal, 1986; Lehiste et Meltzer, 1973), guident le locuteur, par exemple au cours de la croissance du conduit vocal (Ménard et al., 2007; Ménard et al., 2004; Ménard et al., 2002). Toutefois, la validité de ces paramètres acoustiques invariants n'a pas été évaluée en français sur des voyelles naturelles, qui comportent des variations de durée et d'intensité pouvant intervenir dans leur identification auditive. Ces variations sont encore plus importantes pour les voyelles produites par les locuteurs atteints de surdité profonde. L'objectif du présent article est d'étudier les corrélats acoustiques de la perception des traits de degré d'aperture et de lieu d'articulation pour des voyelles produites par des locuteurs sourds profonds, avec ou sans prothèse auditive amplificatrice. Nous tenterons par la suite d'interpréter les productions des locuteurs dans les deux conditions de port de prothèse à la lumière des résultats perceptifs. Cette étude fait partie d'un

projet plus vaste dont le but est d'étudier les relations production-perception en différentes conditions sensorielles.

VARIABILITÉ PRODUITE ET INVARIANCE PERÇUE

Les sources de variabilité dans la production de la parole sont très nombreuses. Si l'on se réfère au trapèze vocalique d'une langue comme le français, dans les dimensions F1 et F2 (en Hertz), des facteurs comme l'âge, le sexe et l'état du système perceptif des locuteurs influencent grandement la position des voyelles au sein de cet espace. Au résultat, un même timbre vocalique perçu correspond à des valeurs de formants produites très variables. La recherche des paramètres permettant de définir cette invariance perçue revêt donc une grande importance.

La présence ou non de rétroaction auditive comme source de variabilité produite

Au chapitre de la variabilité produite, l'état du système de perception auditive d'un locuteur influence considérablement les valeurs formantiques des voyelles qu'il produit. Les recherches ont en effet démontré que l'absence de rétroaction auditive peut modifier le signal acoustique des voyelles, voire les rendre inintelligibles, et ce, tant chez des sujets entendants que chez des sujets sourds appareillés.

En ce qui concerne les sujets entendants, Bell-Berti *et al.* (1979) ont montré, par une étude de l'activité musculaire de la langue chez des locuteurs de l'anglais articulant les voyelles [i e ε æ], que la variation des stratégies individuelles était corrélée à la variation des capacités perceptives des locuteurs. Dans le même ordre d'idées, Fox (1982) suggère que les différences acoustiques entre les paires de voyelles d'un individu à l'autre soient reliées aux différences perceptives. Jones et Munhall (2002) se sont intéressés au lien existant entre la perception auditive et le contrôle articulatoire. Les auteurs ont mesuré les productions des sujets entendants et ont démontré le lien existant entre le contrôle de la production vocalique et la rétroaction auditive, puisqu'une modification de la rétroaction auditive entraîne automatiquement une compensation articulatoire. De même, Perkell et al. (2004) ont déterminé que les entendants présentant les meilleures habiletés discriminatives produisent un contraste acoustique plus significatif entre deux

voyelles distinctes. Ce résultat confirme donc l'existence d'un lien entre la production et la perception de la parole.

Puisque perception auditive et production des voyelles sont intimement reliées, il y a lieu de croire qu'une altération du système de perception auditive comme la surdité profonde entraîne une variation des caractéristiques acoustiques et articulatoires des voyelles produites. Plusieurs études montrent en effet une variation plus importante de F0 chez les sujets sourds par rapport aux sujets entendants, alors que les variations de F1 varient de façon moindre (par exemple Angelocci et al., 1964; Monsen, 1976). Les valeurs de F2 seraient également plus élevées chez les sujets sourds (Angelocci et al., 1964; Stein, 1980) mais l'étendue de leurs variations serait réduite (Monsen et al., 1976). Ces variations acoustiques conduisent parfois à des erreurs de substitution de voyelles et à une diminution d'intelligibilité (McCaffrey et Sussman, 1994; Angelocci et al., 1964; Markides, 1983; Rubin, 1985; Abraham, 1989), mais un grand nombre de voyelles produites par des locuteurs sourds demeurent intelligibles malgré leur variabilité, suggérant en cela que l'invariance a été préservée.

Les corrélats acoustiques invariants des voyelles perçues

La recherche de corrélats acoustiques des traits perçus est d'une importance cruciale dans le domaine de la production-perception de la parole. Au plan perceptif, de tels corrélats peuvent permettre de normaliser les données interlocuteurs et ainsi de comparer des phénomènes produits par plusieurs locuteurs. Au plan de la production, la mise au jour de corrélats acoustiques reliés à la perception des catégories vocaliques permet de définir en termes acoustiques les régions cibles associées à une voyelle donnée. Ces régions cibles sont par la suite interprétées comme des gabarits qui guident la tâche de production d'un locuteur malgré les changements observés dans la morphologie de son système de production (pour le français, voir Ménard et al., 2007; Ménard et al., 2004) ou de son système perceptif (pour l'anglais, voir McCaffrey et Sussman, 1994). L'une des retombées cliniques des études sur l'invariance acoustique consiste en l'amélioration des technologies d'aide auditive et des méthodes de rééducation.

Pour la plupart, les études menées dans le cadre de cette approche¹ impliquent d'abord une transformation des fréquences selon l'échelle des Hertz en « bandes critiques », ou Bark. Par la suite, une différence entre les valeurs formantiques et la fréquence fondamentale est estimée (Syrdal et Gopal, 1986; Traunmüller, 1981). Le débat reste ouvert quant à l'effet des informations dynamiques (Bohn et Strange, 1995) dans la détermination des catégories vocaliques.

L'aperture perçue

Le rôle de la fréquence fondamentale (F0) dans le processus de normalisation pour l'identification de l'aperture a été confirmé, entre autres, par Traunmüller (1981) et Syrdal et Gopal (1986). Lors de l'analyse de la perception de voyelles synthétiques à un formant en bavarois, Traunmüller (1981) montre la forte corrélation entre l'aperture perçue et la valeur de la différence entre F1 et F0, en Bark. Les résultats révèlent que les frontières perceptives des catégories, pour les cinq degrés d'aperture, pour F0 inférieur à environ 350 Hz, sont situées à des valeurs de F1-F0 aux environs de 1,2 Bark, 2,2 Bark, 3,2 Bark et 6 Bark. Syrdal et Gopal (1986), après avoir réanalysé des voyelles naturelles, confirment le rôle de F0 dans la perception de la hauteur des dix voyelles de l'anglais américain. Les auteurs proposent un modèle du système de traitement auditif à deux composantes. Au premier stade a lieu la transformation des valeurs de fréquence en distances tonotopiques, en Bark. La seconde étape de traitement, de plus haut niveau, consiste en une classification binaire des distances tonotopiques, basée sur la valeur de 3 Bark. Ce dernier stade permet l'extraction des traits phonétiques. Les voyelles perçues hautes correspondent à F1- F0 inférieur à 3 Bark, et les voyelles non hautes, à des valeurs de F1-F0 supérieures à 3 Bark. Pour le français, Ménard et al. (2002), à partir d'un corpus de voyelles synthétiques, proposent que la dimension F1-F0, en Bark, soit reliée au degré d'aperture perçue, la frontière de 2 Bark permettant de distinguer les voyelles perçues fermées des voyelles perçues mi-fermées, et la frontière de 4 Bark permettant de distinguer les voyelles perçues mi-fermées des voyelles perçues mi-ouvertes et ouvertes.

¹ Notons que nous nous situons ici dans le cadre de l'invariance acoustique, par opposition à celui de l'invariance articulatoire.

Le lieu d'articulation perçu

Selon Syrdal et Gopal (1986) la différence entre F3 et F2 serait reliée à la perception de l'antéro-posteriorité des voyelles de l'anglais américain. Une bonne classification est opérée par ce paramètre, les voyelles antérieures correspondant à une valeur de F3-F2 inférieure à 3 Bark, et les voyelles postérieures, à une différence de F3-F2 supérieure à 3 Bark. Selon Fant (1983), le suédois distingue les voyelles antérieures et postérieures sur la base de la valeur de la différence entre F2 et F1. D'autres études proposent que la différence entre F2 et F0 soit un bon corrélat du lieu d'articulation perçu. Pour le français, Ménard et al. (2002) ont montré que la différence entre F2 et F1, en Bark, était significativement reliée au lieu d'articulation perçu (les voyelles perçues postérieures étant associées à des valeurs de F2-F1 inférieures à 5 Bark, les voyelles perçues antérieures correspondant à des valeurs de F2-F1 supérieures à 5 Bark).

La plupart des études sur les corrélats acoustiques invariants des voyelles ont été réalisées sur des productions de sujets entendants. McCaffrey et Sussman (1994) ont par ailleurs analysé les voyelles produites par des locuteurs anglophones sourds appareillés dans le cadre développé par Syrdal et Gopal (1986). Leur analyse a montré que les paramètres F1-F0 et F3-F2, en Bark, décrivaient avec justesse les voyelles produites et perçues par leurs sujets. Nous tenterons de prolonger ce travail en étudiant les voyelles du français.

MÉTHODOLOGIE

Locuteurs et corpus

Deux sujets sourds profonds de sexe masculin ont participé à l'expérience. Les sujets ont été recrutés par le biais du « Centre Recherche Interdisciplinaire en Réadaptation » (CRIR) de Montréal. Ils avaient une surdité profonde congénitale, portaient des prothèses auditives et avaient bénéficié (depuis leur enfance) d'une rééducation vocale leur permettant de s'exprimer oralement. Le sujet S1 était âgé de 22 ans et présentait un degré de surdité de 102 dB HL pour l'oreille gauche et de 95 dB HL pour l'oreille droite. Le sujet S2 était âgé de 29 ans et présentait un degré de surdité de 110 dB HL pour l'oreille gauche et de 90 dB HL pour l'oreille droite. Ces sujets avaient le français québécois comme langue maternelle et d'usage.

Les locuteurs avaient pour tâche de prononcer les voyelles orales /i e ε a y u o/ dans le contexte « V

comme pVpa », où V correspond à l'une des voyelles à l'étude. Seule la voyelle en contexte « pVpa » a été analysée. Ces sept voyelles permettaient d'obtenir des contrastes d'aperture (fermées /i y u/ vs. mi-fermées /e o/ vs. mi-ouverte /ε/ vs. ouverte /a/), de lieu d'articulation (antérieures /i y e ε/ vs. postérieures /u o a/) et d'arrondissement (non-arrondie /i/ vs. arrondie /y/), tout en maintenant une durée d'enregistrement relativement courte. Il est à noter que le contraste d'arrondissement ne sera pas étudié ici. Cinq répétitions de chacune des voyelles ont été obtenues (pour un total de 35 voyelles) et l'ensemble des voyelles a été produit en ordre aléatoire. Les enregistrements ont été effectués dans deux conditions : avec prothèse et sans prothèse. Les études présentées à la section précédente ont en effet montré les liens entre la présence de la rétroaction auditive et la production des voyelles. Les enregistrements dans la condition « sans prothèse » ont été obtenus au tout début de la session d'enregistrement, après que les sujets eurent retiré leur appareil auditif pendant quinze minutes. Les productions vocaliques des sujets ont été enregistrées dans une pièce isolée, à l'aide d'un ordinateur IBM Think Pad T 41, d'un micro unidirectionnel Sure et d'un préamplificateur.

Caractéristiques acoustiques des stimuli

Chacune des voyelles produites a été numérisée à un taux d'échantillonnage de 22050 Hz et analysée à l'aide du logiciel *Praat*. Une analyse par LPC (*Linear Predictive Coding*) a été effectuée afin d'extraire les valeurs des deux premiers formants (F1 et F2), en Hertz.² Le nombre de pôles variait de 12 à 16 et le signal était préamplifié à partir de 50 Hz. La fenêtre d'analyse était d'une largeur de 25 ms. Aucune mesure de référence n'était fournie à l'algorithme de détection des formants afin de ne pas contraindre l'analyse à des valeurs typiques, compte tenu du caractère variable des voyelles à l'étude. Afin de garantir l'exactitude des mesures, les valeurs de formants détectées par l'algorithme automatique LPC ont été superposées sur un spectrogramme à bande large. Lorsque la détection automatique ne correspondait pas à la

² La méthode LPC a été choisie par rapport à la méthode FFT (*Fast Fourier Transform*) puisqu'elle s'était avérée la méthode la plus fiable dans l'analyse de voyelles très variables produites par des enfants (Ménard, 2002). Puisque le corpus à l'étude est également variable, nous avons préféré utiliser cet algorithme.

représentation spectrographique, le nombre de pôles était réajusté et la procédure LPC était appliquée à nouveau. Les valeurs de F0 ont été extraites à l'aide d'un algorithme de détection automatique par autocorrélation. La durée des voyelles variait entre 185 ms et 305 ms. Un test T n'a révélé aucune différence significative de durée entre les deux locuteurs, suggérant ainsi une variation similaire entre les deux sujets. Il en est de

même de l'amplitude de chacune des voyelles, mesurée en valeurs RMS. Les valeurs de formants et de F0, en Hz, ont été transformées en Bark selon la formule proposée par Zwicker et Terhardt (1980) :

$$F_{\text{Bark}} = 13 \cdot \arctan(0.76 \cdot F_{\text{Hz}}) + 3.5 \cdot \arctan(F_{\text{Hz}}/7.5)^2.$$

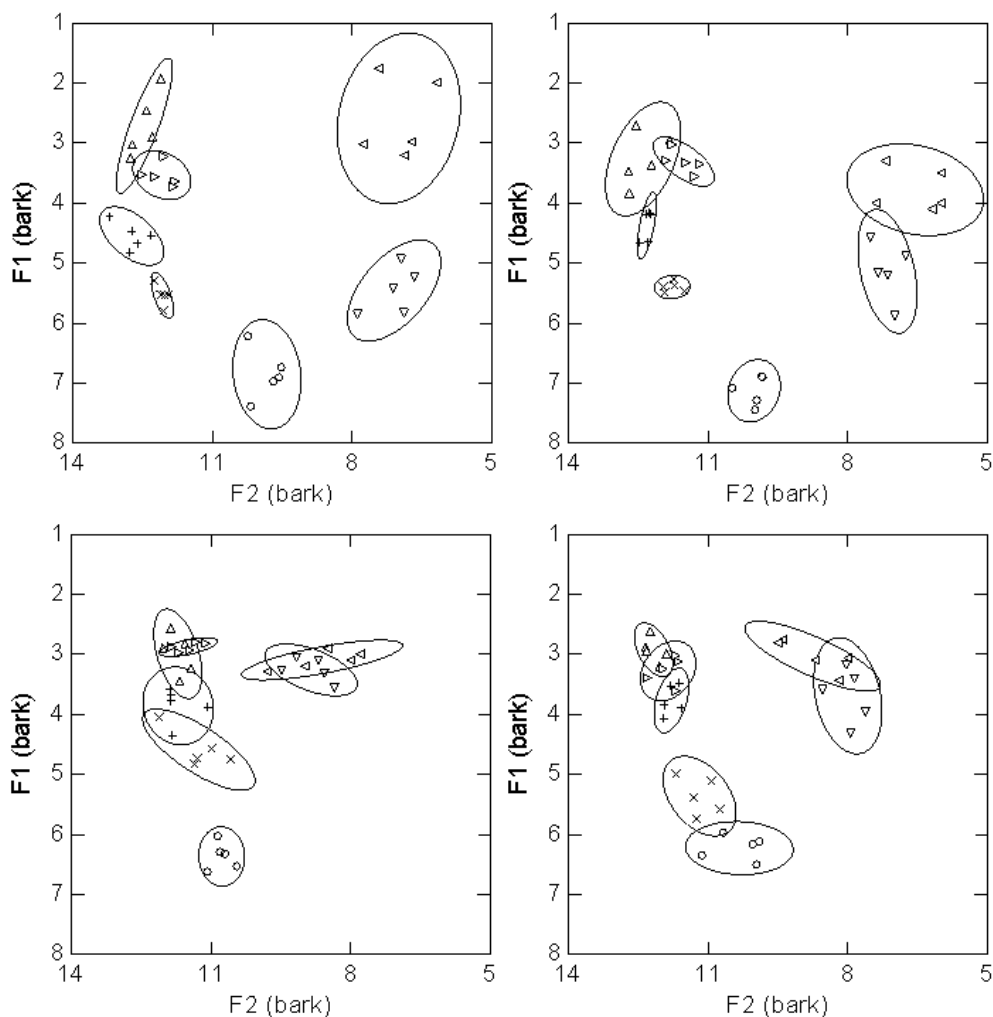


Figure 1 : Valeurs de F1 et de F2 pour les voyelles /i y u e ε a o/ produites par les deux sujets, dans les deux conditions de port de prothèse : S1 avec prothèse (haut gauche), S1 sans prothèse (haut droite), S2 avec prothèse (bas gauche), S2 sans prothèse (bas droite).

	S1	S2
/i/	180 (+4.8)	123 (+13)
/y/	181 (+4)	121 (+2)
/u/	187 (+14)	119 (+19)
/e/	150 (-1,5)	119 (-14)
/ɛ/	124 (-10)	113 (-2)
/a/	145 (+11)	108 (+17)
/o/	138 (-12)	112 (-8)

Tableau I : Valeurs moyennes de F0 (en Hertz) des voyelles produites par chacun des locuteurs dans la condition « avec prothèse ». Entre parenthèses figurent les pourcentages de variation par rapport à la condition « sans prothèse ».

Les valeurs de F1 et de F2, en Bark, des voyelles produites par les locuteurs S1 et S2, sont présentées dans la Figure I, dans les conditions « avec prothèse » et « sans prothèse ». Les ellipses de dispersion sont tracées à +1.5 écart type de la moyenne. Les valeurs de F0 dans la condition « avec prothèse » sont présentées dans le tableau I, et les pourcentages de variation des valeurs de la condition « sans prothèse » par rapport à la condition « avec prothèse » figurent entre parenthèses.³

En ce qui concerne le sujet S1 (Figure I), l'effet du port de la prothèse a une incidence sur les valeurs de formants pour certaines voyelles. Comme en témoigne la position des ellipses de dispersion, la valeur de F1 est plus élevée dans la condition « avec prothèse » que dans la condition « sans prothèse » pour la voyelle /u/. La taille réduite des ellipses de dispersion pour les voyelles /e/, /a/ et /u/ témoigne par ailleurs de la variabilité moins importante de ces catégories vocaliques dans la

condition « sans prothèse » par rapport à la condition « avec prothèse ». Les valeurs de F0, quant à elles (tableau I), varient d'une voyelle à l'autre et d'une condition de port de prothèse à l'autre, la voyelle /u/ étant associée à l'augmentation la plus marquée de la condition « sans prothèse » à la condition « avec prothèse ». Aucune différence significative de durée des voyelles ou d'amplitude RMS entre la condition « sans prothèse » et la condition « avec prothèse » n'a été relevée.

Pour ce qui est du locuteur S2 (Figure I), les valeurs de F1 suggérées par la position des ellipses de dispersion sont plus élevées seulement pour les voyelles /u/ et /y/ produites sans prothèse par rapport à ces mêmes voyelles produites avec prothèse. Concernant F2, les voyelles /o/ produites avec prothèse présentent un F2 plus élevé que leurs contreparties produites sans prothèse, contrairement aux voyelles /i/, /y/ et /u/, dont le F2 est plus élevé dans la condition « sans prothèse » que dans la condition « avec prothèse ». Conformément à ce qui avait été observé pour le locuteur S1, les valeurs de F0 produites par le locuteur S2 (tableau I) montrent également des

³ Compte tenu du petit nombre de sujets (n=2), les données ne sont pas soumises à des analyses statistiques et sont décrites qualitativement.

variations importantes. Les valeurs associées à /u/ correspondent à l'augmentation la plus importante entre les conditions « sans prothèse » et « avec prothèse ». Pour ce sujet comme pour le sujet S1, aucune différence significative de durée ou d'amplitude RMS n'a été relevée entre les conditions « sans prothèse » et « avec prothèse ».

Pour résumer, les valeurs de F1, de F2 et de F0 produites par les deux locuteurs varient en fonction de l'identité de la voyelle produite et du port ou non de la prothèse auditive, quoique dans une moindre mesure. L'effet de ces variations semble plus important pour le sujet S2 que pour le sujet S1. Le corpus ainsi obtenu représente un ensemble de voyelles dont les caractéristiques acoustiques varient considérablement. Ces résultats confirment les travaux antérieurs qui montrent que les résonances de F2 pour les voyelles produites par des sourds dévient souvent des valeurs normales, particulièrement pour les voyelles postérieures (par exemple Angelocci et al., 1964; Stein, 1980).

Test d'identification des voyelles

Une occurrence des 140 voyelles V produites en contexte « pVpa » (5 répétitions * 7 voyelles * 2 locuteurs * 2 conditions) par les locuteurs sourds ont été extraites de leur contexte produit et soumises, lors d'un test d'identification auditive, à 40 participants normo-entendants, locuteurs natifs du français québécois âgés de 18 à 40 ans. Tous les sujets avaient une audition normale et ne présentaient aucun trouble du langage. Le test de perception a été généré dans le logiciel *Praat*, et les stimuli y étaient présentés dans un ordre aléatoire⁴, via des casques d'écoute de haute qualité. La tâche des participants consistait à sélectionner à l'écran la voyelle correspondant au stimulus perçu dans la première syllabe du mot « pVpa » parmi l'ensemble des voyelles suivantes : /i y u e ø o ε œ ɔ a a ɔ̃/. Aucune limite de temps n'était imposée au participant⁵. Chaque voyelle était représentée, sur l'écran, par une case sur laquelle étaient inscrits la voyelle et un mot du français représentatif de cette voyelle. La voyelle nasale /ɔ̃/ a été incluse dans les choix de réponses conformément aux études

⁴ Nous nous sommes assurés que le test ne comportait pas de suite de stimuli constituée de plus de deux répétitions d'une même voyelle par un même locuteur.

⁵ Certains sujets peuvent avoir utilisé différentes stratégies d'identification, mais la durée totale du test était peu variable parmi les 40 sujets (environ 15 minutes).

antérieures (notamment McCaffrey et Sussman, 1994) qui ont noté une hypernasalisation de certaines voyelles produites par les sujets sourds, phénomène à l'origine de difficultés d'analyse acoustique. Les voyelles étaient disposées de la même façon sur l'écran pour tous les locuteurs (ordre du trapèze vocalique). Le test était d'une durée d'environ 15 minutes et incluait une courte phase de familiarisation.

Pour chaque stimulus, le taux d'identification correcte a été calculé. De plus, la voyelle perçue de façon dominante, c'est-à-dire la catégorie vocalique pour laquelle au moins 50% des auditeurs ont opté, a été identifiée. Ces taux d'identification ont été reliés aux paramètres acoustiques décrits à la section précédente.

RÉSULTATS

L'intelligibilité globale est d'abord présentée dans la section 4.1. Les corrélats acoustiques des voyelles font l'objet de la section 4.2.

Intelligibilité globale

Les résultats du test d'identification (matrices de confusion) sont présentés au Tableau II, pour chacun des sujets. Pour chacune des voyelles produites (lignes) figurent les pourcentages de réponses correspondants à chacune des voyelles perçues (colonnes). Les pourcentages sont présentés pour la condition « avec prothèse », et la différence entre les taux d'identification en condition « avec prothèse » par rapport à la condition « sans prothèse » apparaît entre parenthèses. Des valeurs positives reflètent une augmentation des taux d'identification des voyelles produites avec prothèse par rapport aux mêmes voyelles produites sans prothèse. Les taux d'identification des voyelles conformément à la catégorie produite (identification correcte) figurent en gras dans le tableau II et correspondent aux taux d'intelligibilité.

Un examen du tableau II révèle que l'intelligibilité varie d'un locuteur à l'autre et d'une voyelle à l'autre. Une ANOVA à mesures répétées incluant les facteurs « prothèse » et « voyelle » menée sur les taux d'identification correcte des voyelles produites par le locuteur S1 révèle un effet significatif de la variable « voyelle » uniquement ($F(6, 24)=21,73$; $p<0,05$). Conformément aux résultats de tests post-hoc, les voyelles /i/, /y/, /u/ et /a/ sont associées à des taux d'identification correcte significativement plus élevés que les voyelles /e/ et /ε/, qui, à leur tour, ont des taux d'identification correcte

Sujet S1	/i/	/y/	/u/	/e/	/ɛ/	/ɑ/	/o/	Autres
/i/	97.5 (-0.5)	0.5 (+0.5)	0.5 (0)	0 (-1)	0 (0)	1.5 (+1.5)	0 (-0.5)	0 (0)
/y/	1.5 (-0.5)	96.5 (+1.5)	0.5 (0)	0 (0)	1 (+0.5)	0.5 (+0.5)	0 (0)	0 (-2)
/u/	0.5 (-1)	1.5 (0)	97 (+3.5)	0 (-0.5)	0 (0)	0 (0)	0.5 (-1.5)	0.5 (-0.5)
/e/	0.5 (-0.5)	0.5 (-1)	0 (-0.5)	73.5 (+3)	23.5 (-1.5)	1 (+0.5)	0 (0)	1 (0)
/ɛ/	1.5 (+1)	0 (-1)	0.5 (+0.5)	24 (-0.5)	71.5 (+0.5)	0 (-1)	1 (1)	1.5 (+0.5)
/ɑ/	0.5 (+0.5)	0 (-0.5)	0 (-1.5)	0 (-0.5)	2 (+1)	89.5 (+5.5)	0 (0)	8 (-4.5)
/o/	0.5 (+0.5)	0.5 (0)	0 (-0.5)	0 (-0.5)	0.5 (+0.5)	0.5 (0)	43.5 (-15)	54.5 (+15)
Sujet S2	/i/	/y/	/u/	/e/	/ɛ/	/ɑ/	/o/	Autres
/i/	94 (-3.5)	1.5 (+1)	1.5 (+1)	0 (-0.5)	1.5 (+1.5)	0 (0)	0 (0)	1 (0)
/y/	31 (-16)	63 (+14)	0.5 (-0.5)	1 (0)	0 (-1)	0 (-0.5)	0.5 (+0.5)	4 (+3)
/u/	0.5 (0)	11.5 (+4.5)	85.5 (-3)	0 (0)	0 (0)	0.5 (-0.5)	0 (-0.5)	2 (-0.5)
/e/	4.5 (+0.5)	7.5 (+2)	0.5 (-0.5)	45.5 (-8.5)	6 (-3.5)	2 (+1.5)	0 (0)	34 (+8.5)
/ɛ/	3.5 (-1)	1 (-1.5)	1 (-0.5)	29 (-4)	10 (-13)	9 (+7)	0 (-0.5)	46.5 (+13.5)
/ɑ/	0.5 (-0.5)	0 (0)	0 (-1)	0 (0)	0.5 (-1)	84.5 (+16.5)	0 (0)	14.5 (-14)
/o/	1 (0)	3 (+1.5)	64 (-14)	0 (-1)	0.5 (+0.5)	0.5 (0)	9 (-4)	22 (+17)

Tableau II : Matrices de confusion issues du test d'identification perceptive pour les voyelles produites dans la condition « avec prothèse ». Les catégories des voyelles produites sont représentées par lignes, les catégories perçues sont représentées par colonnes. Entre parenthèses figurent les pourcentages de variation par rapport à la condition « sans prothèse ».

supérieurs à la voyelle /o/. Il est à noter que le port de la prothèse n'a pas d'effet significatif (en tant qu'effet principal ou en interaction avec la voyelle) sur les taux d'identification correcte pour ce sujet. En ce qui concerne le sujet S2, conformément à l'analyse menée pour le sujet S1, une ANOVA à mesures répétées avec les facteurs « prothèse » et « voyelle » a révélé un effet significatif de la variable « voyelle » ($F(6, 24)=58,34; p<0,05$) mais aucun effet de la variable « prothèse ». Des tests post-hoc ont montré que les voyelles /i/, /u/ et /a/ étaient associées à des taux d'identification correcte supérieurs aux voyelles /e/, /ɛ/ et /o/. Parmi ces dernières, la voyelle /e/ était significativement mieux identifiée que les voyelles /ɛ/ et /o/.

Corrélatés acoustiques des voyelles perçues

Les résultats présentés dans la section 4.1 suggèrent des variations importantes d'intelligibilité parmi les voyelles produites par les deux locuteurs sourds. Par exemple, la voyelle /o/ produite par le locuteur S2 a été perçue /o/ par 9% des auditeurs alors qu'elle a été identifiée /u/ par 64% d'entre

eux. Afin de déterminer les corrélatés acoustiques des voyelles perçues, nous avons identifié, pour chacune des voyelles produites, la voyelle perçue majoritairement, c'est-à-dire la voyelle perçue par au moins 50% des auditeurs (au moins 20 auditeurs sur 40). Aux résultats, 23 voyelles ont été majoritairement perçues /i/, 17 ont été majoritairement perçues /y/, 29 ont été majoritairement perçues /u/, 15 ont été majoritairement perçues /e/, 10 ont été majoritairement perçues /ɛ/, 5 ont été majoritairement perçues /o/ et 20 ont été majoritairement perçues /a/. Soulignons que les voyelles cardinales /i u a/ ont été associées à un plus grand nombre de voyelles majoritairement perçues que les voyelles périphériques. Les paramètres acoustiques F1-F0 et F2-F1 (en Bark) ont été calculés pour chacune des voyelles majoritairement perçues. Les graphiques de la figure II présentent les valeurs acoustiques de ces voyelles perçues par les auditeurs dans deux espaces : l'espace traditionnel F1 vs. F2, en Bark (gauche), et l'espace acoustico-auditif F1-F0 vs. F2-F1, en Bark (droite) (Ménard et al., 2002). L'identification perceptive des voyelles produites par

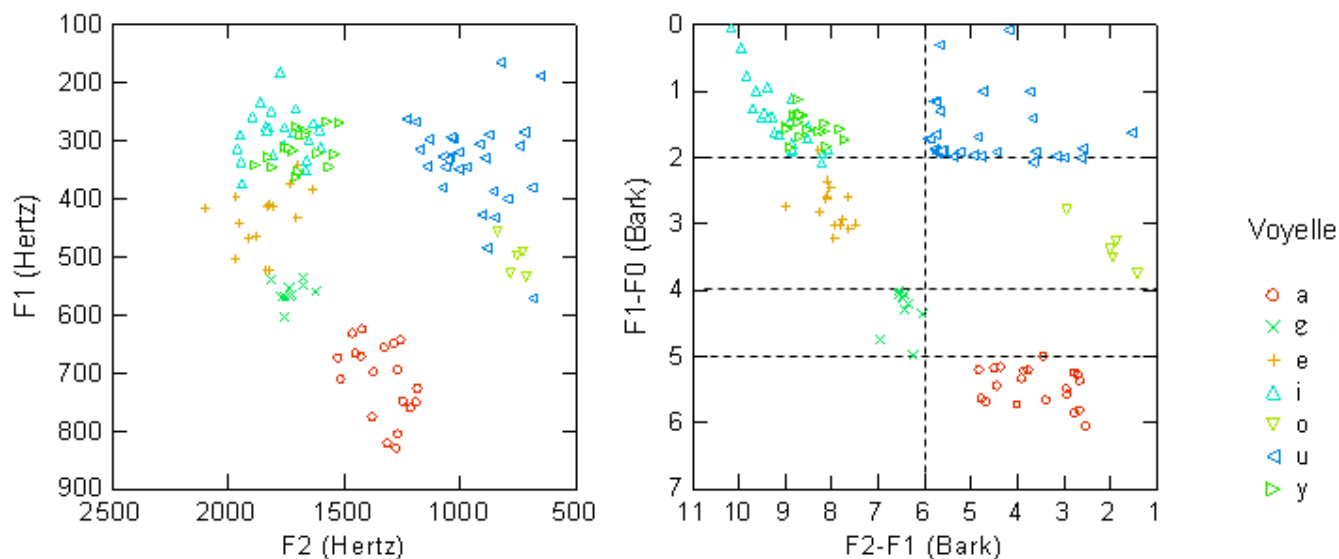


Figure II : Représentation des voyelles perçues majoritairement par les auditeurs dans les espaces F1 vs. F2, en Bark (gauche) et F1-F0 vs. F2-F1, en Bark (droite). Dans l'espace F1-F0 vs. F2-F1 sont représentées les frontières d'identification en traits pointillés : F1-F0=2 Bark, F1-F0=4 Bark, F1-F0=5 Bark et F2-F1=6 Bark.

les deux locuteurs et dans les deux conditions de prothèse (avec et sans prothèse) est représentée sur ces graphiques dans le but d'analyser l'ensemble de la variabilité soumise aux auditeurs. Les paramètres acoustiques qui permettront d'obtenir une classification optimale des voyelles perçues (distances interclasses maximales et distances intraclasses minimales) seront ceux qui pourront être interprétés comme corrélats acoustiques des traits perçus. Comme le montre la figure II, les paramètres acoustiques traditionnels F1 vs. F2 (en Bark) opèrent une bonne distinction des voyelles perçues selon le lieu d'articulation, l'ensemble des voyelles perçues postérieures /u o a/ étant associées à des valeurs de F2 inférieures aux voyelles perçues antérieures /i y e ε/. Cependant, les voyelles perçues mi-fermées /e/ correspondent parfois à des valeurs de F1 similaires à celles des voyelles perçues mi-ouvertes /ε/. Il en est de même pour les voyelles /u/ et /o/. Le paramètre F1 ne semble donc pas corrélé au degré d'aperture perçu.

En revanche, l'espace F1-F0 vs. F2-F1, en Bark (figure II, droite), permet de regrouper les différentes catégories de voyelles perçues en différentes régions. Les frontières d'identification perceptives sont tracées en traits pointillés sur la figure. D'abord, le degré d'aperture perçu est relié à la valeur de F1-F0 (en Bark). En effet, les voyelles perçues fermées /i y u/ sont reliées à des valeurs de F1-F0 inférieures à 2 Bark. Les voyelles perçues mi-fermées /e o/ correspondent à des valeurs de F1-F0 supérieures à 2 Bark et inférieures à 4 Bark. Les voyelles perçues mi-ouvertes /ε/ sont celles dont la valeur de F1-F0 se situe entre 4 et 5 Bark. Enfin, les voyelles perçues ouvertes /a/ sont associées à des valeurs de F1-F0 supérieures à 5 Bark. En ce qui concerne le lieu d'articulation perçu, le paramètre F2-F1 (en Bark) permet de différencier les voyelles perçues antérieures /i y e ε/, dont les valeurs de F2-F1 sont supérieures à 6 Bark, des voyelles perçues postérieures /u o a/, dont les valeurs de F2-F1 sont inférieures à 6 Bark. Afin de comparer les deux espaces acoustiques, des analyses linéaires discriminantes ont été menées. Dans un premier temps, les paramètres F1 et F2 (en Bark) ont été inclus dans l'analyse en tant que variables indépendantes, alors que les facteurs de classification correspondaient aux voyelles majoritairement perçues. Il est à noter que les voyelles /i/ et /y/ ont été regroupées pour la présente analyse, puisque le contraste d'arrondissement n'est pas à l'étude ici. Le taux de classification correcte obtenu était de 88%. En

revanche, lorsque les paramètres F1-F0 et F2-F1 (en Bark) étaient considérés comme variables indépendantes, le taux de classification correcte était de 98%. En somme, la combinaison des paramètres acoustiques F1, F2 et F0 permet d'obtenir un espace à 2 dimensions F1-F0 vs. F2-F1 au sein duquel les catégories vocaliques perçues selon les traits d'aperture et de lieu d'articulation occupent des régions distinctes.

DISCUSSION

Les résultats de la présente étude permettent de confirmer l'existence de paramètres acoustiques reliés de façon invariante aux traits d'aperture perçue et de lieu d'articulation perçu des voyelles du français. Ces résultats ont des implications importantes pour l'étude des voyelles produites par les locuteurs atteints de surdité.

Corrélat acoustiques invariants des voyelles perçues

Afin de générer un corpus de voyelles du français représentatif de la production de locuteurs sourds, deux sujets adultes atteints de surdité profonde ont produit des répétitions des sept voyelles orales /i y u e o ε a/. Deux conditions ont été suscitées : avec prothèse auditive amplificatrice et sans prothèse auditive amplificatrice. Les résultats du test perceptif révèlent que les différences tonotopiques entre les valeurs de formants et de F0 sont corrélées à la perception des catégories vocaliques. En effet, les différences de degré d'aperture perçue sont reliées à la valeur de F1-F0, en Bark, les frontières de 2 Bark, 4 Bark et 5 Bark permettant de différencier respectivement les voyelles perçues fermées des voyelles perçues mi-fermées, les voyelles perçues mi-fermées des voyelles perçues mi-ouvertes et les voyelles perçues mi-ouvertes des voyelles perçues ouvertes. Le lieu d'articulation perçu est relié au paramètre F2-F1, en Bark, les voyelles perçues postérieures étant associées à des valeurs de F2-F1 inférieures à 6 Bark, alors que les voyelles perçues antérieures correspondent à des valeurs de F2-F1 supérieures à 6 Bark. Ces données confirment les résultats que nous avons proposés pour l'identification de voyelles synthétiques en français (Ménard *et al.*, 2002). Les voyelles naturelles, et plus particulièrement les voyelles produites par les locuteurs atteints de surdité profonde, comportant d'autres dimensions pouvant intervenir dans leur identification (comme la durée et l'intensité, par exemple), la validité des paramètres acoustico-

perceptifs F1-F0 et F2-F1 ne pouvait être déduite des études antérieures. Le taux de classification correcte très élevé (98%) obtenu à l'aide de ces paramètres permet par ailleurs d'en confirmer la validité. Il semble donc que la combinaison des deux premiers formants et de F0 définissent de façon plus adéquate que les deux premiers formants seuls les régions-cibles associées aux buts acoustiques des voyelles /i y u e o ε a/. Ces résultats font écho aux études menées dans le cadre de la normalisation acoustique, notamment en anglais (entre autres Syrdal et Gopal, 1986; McCaffrey et Sussman, 1994).

Une réinterprétation des voyelles produites par les locuteurs sourds

La présente étude constitue le premier volet d'un projet plus étendu visant à étudier les stratégies articulatoires exploitées par les locuteurs atteints de privations sensorielles afin de produire les catégories vocaliques. Ces stratégies seront étudiées à la lumière des corrélats acoustiques des traits perçus. Bien que seulement deux sujets aient participé aux enregistrements, ils représentaient une variation acoustique suffisamment importante pour mener à bien le test perceptif. Les paramètres acoustico-perceptifs relevés dans notre test peuvent être utilisés afin d'interpréter les données de production de ces deux locuteurs, à titre exploratoire.

Malgré cette variation des paramètres acoustiques F1 et F2, les taux d'intelligibilité présentés dans le tableau II n'ont pas varié de façon significative entre les conditions « avec prothèse » et « sans prothèse ». Comment expliquer cette apparente contradiction? Puisque la perception des voyelles est corrélée non pas aux valeurs brutes de F1 ou F2 mais bien aux paramètres F1-F0 et F2-F1, il apparaît que les locuteurs aient été en mesure de compenser les variations de l'un des paramètres acoustiques (par exemple F2) par une variation équivalente du paramètre complémentaire (par exemple F1) afin d'obtenir des valeurs combinées de ces deux paramètres (par exemple F2-F1) conformes aux zones associées à la perception correcte (par exemple une valeur de F2-F1 supérieure à 6 Bark pour les voyelles antérieures). Par exemple, les voyelles /u/ produites par le sujet S1, en condition sans prothèse, correspondent à des valeurs de F1 significativement plus élevées que dans la condition avec prothèse (comparaison des ellipses bleues et rouges associées à /u/ dans la figure 1). Or, le taux d'identification correcte de

cette voyelle demeure très élevé, comme en témoigne le tableau II (variation minimale de 97% à 93,5% entre les conditions avec prothèse et sans prothèse). Ce taux élevé s'explique par une augmentation de la valeur de F0 pour la voyelle /u/ produite sans prothèse, par rapport à la même voyelle produite avec prothèse. L'augmentation parallèle de F0 et de F1 permet donc de maintenir une valeur de F1-F0 inférieure à 2 Bark, assurant ainsi la perception correcte de cette voyelle fermée.⁶

Soulignons que, chez les deux sujets à l'étude, les voyelles cardinales ont été mieux identifiées que les voyelles intermédiaires par les participants : les sujets sourds ont donc eu moins de difficultés à produire les premières. À cet égard, Rubin (1985) a aussi constaté une plus grande variabilité acoustique chez les voyelles intermédiaires que chez les voyelles cardinales dans un plan acoustique F1-F2. Aussi, l'instabilité des voyelles centrales chez le sujet S2 fait écho aux observations de McCaffrey et Sussman (1994) concernant la variabilité acoustique de celles-ci et la difficulté à faire des contrastes distincts entre elles. La stabilité des voyelles cardinales s'explique également dans l'espace F1-F0 vs. F2-F1. En effet, ces voyelles impliquent des valeurs de F1-F0 inférieures à 2 Bark pour /i/ et /u/ et des valeurs supérieures à 5 Bark pour /a/. Au contraire, les voyelles intermédiaires doivent être produites avec des valeurs de F1-F0 inférieures et supérieures à des frontières, ce qui nécessite un contrôle plus fin des valeurs de F1 et de F0. L'effet plafond semble donc favoriser la stabilité des voyelles cardinales /i a u/ au détriment des voyelles intermédiaires /e ε o/ dans l'espace F1 vs. F2 comme dans l'espace acoustico-perceptif F1-F0 vs. F2-F1.

REMERCIEMENTS

Nous remercions les sujets pour leur grande patience. Cette étude a bénéficié du soutien financier du Conseil de Recherches en Sciences Humaines du Canada et du Conseil de Recherches en Sciences Naturelles et en Génie du Canada.

⁶ Les locuteurs peuvent aussi avoir renforcé la variation naturelle de fréquence fondamentale intrinsèque des voyelles (les voyelles hautes ayant une F0 plus élevée que les voyelles basses, puisque la tension de la langue vers le palais exerce une pression sur les muscles du larynx, tendant ainsi les cordes vocales). La sensation proprioceptive de tension musculaire serait donc ici associée à la cible vocalique.

RÉFÉRENCES

- Abraham, S. 1989. Using a phonological framework to describe speech error of orally trained, hearing-impaired school-agers. *Journal of Speech and Hearing Disorders*. Vol. 54 : 600-609.
- Angelocci, A. A., G. A. Kopp and A. Holbrook. 1964. The vowel formant of deaf and normal-hearing eleven- to fourteen-year-old boys. *Journal of Speech and Hearing Disorders*. Vol. 29 : 156-170.
- Bell-Berti, F., L.J. Raphael, D.B. Pisoni, et J.R. Sawush. 1979. Some relationships between speech production and perception. *Phonetica*, 36, 373-383.
- Bohn, O. S., et Strange W. 1995. Discrimination of coarticulated German vowels in the silent-center paradigm : 'Target' spectral information non needed. *Proc. Int. Congr. Phon. Sc.*, 2, 270-273.
- Fant, G. 1983. Feature analysis of Swedish vowels—A revisit. *KTH, Speech Transmission Laboratory-Quarterly Progress and Status Report 2-3*, 1-19.
- Fox, R.A. 1982. Individual variation in the perception of vowels, *Phonetica*, 39, 1-22.
- Jones, J. A. and K. G. Munhall. 2002. The role of auditory feedback during phonation : studies of mandarin tone production. *Journal of Phonetics*. 30, 303-320.
- Lane, H., F. H. Guenther, M. L. Matthies, L. Ménard, J. S. Perkell, E. Stockmann, M. Tiede, J. Vick and M. Zandipour. 2005. Effects of bite blocks and hearing status on vowel production. *Journal of the Acoustical Society of America*. 118, 1636-1646.
- Lehiste, I. & Meltzer, D. 1973. Vowel and speaker identification in natural and synthetic speech. *Language and Speech*, 16, 356-364.
- Markides, A. 1983. *The speech of hearing-impaired children*. Manchester : Manchester University Press.
- McCaffrey, H. A. and H. M. Sussman. 1994. An investigation of Vowel, Organization in Speakers With Severe and Profound Hearing Loss. *Journal of Speech and Hearing Research*. 37, 938-951.
- Ménard, L., Schwartz, J.-L., Boë, L.-J. et Aubin, J. 2007. Production-perception relationships during vocal tract growth for French vowels: analysis of real data and simulations with an articulatory model. *Journal of Phonetics*. 35, 1-19.
- Ménard, L., Schwartz, J.-L., Boë, L.-J., 2004. Role of vocal tract morphology in speech development: Perceptual targets and sensori-motor maps for synthesized French vowels from birth to adulthood. *Journal of Speech, Language, and Hearing Research*. 47, 1059-1080.
- Ménard, L., Schwartz, J.-L., Boë, L.-J., Kandel, S., Vallée, N., 2002. Auditory normalization of French vowels synthesized by an articulatory model simulating growth from birth to adulthood. *Journal of the Acoustical Society of America*. 111, 1892-1905.
- Monsen, R. B. 1976. Normal and reduced phonological space: The production of English vowels by deaf adolescents. *Journal of Phonetics*, 4, 189-198.
- Perkell, J. S., F. H. Guenther, H. Lane, M. L. Matthies, E. Stockmann, M. Tiede and M. Zandipour. 2004. The distinctness of the speakers' production of vowel contrasts is related to their discrimination of the contrast. *Acoustical Society of America*. Vol. 116, No. 4, October 2004 : 2338-2344
- Rubin, J. A. 1985. *Static and dynamic information in vowels produced by the hearing impaired*. Bloomington, Indiana : Indiana University Linguistics Club.
- Smith, C. R. 1975. Residual hearing and speech production of deaf children. *Journal of Speech and Hearing Research*. Vol. 18.
- Stein, D. M. 1980. A study of articulatory characteristics of deaf talkers. *Dissertation Abstract Internationnal*. Vol. 40, October 1980.
- Syrdal, A. K. et Gopal, H. S. 1986. A perceptual model of vowel recognition based on the auditory representation of American English vowel, *Journal of the Acoustical Society of America*. 79, 1086-1100.
- Traunmüller, H. 1981. Perceptual dimension of openness in vowels, *Journal of the Acoustical Society of America*. 69, 1465-1475.
- Zwicker, E. et Terhardt, E. 1980. Analytical expressions for critical band rate and critical bandwidth as a function of frequency. *Journal of the Acoustical Society of America*. 68, 1523-1525.