… la perception du genre …

Parmi les différentes dimensions de l'identité véhiculées par la voix, le genre (c'est à dire le sexe du locuteur, qu'il soit biologique et/ou identitaire[1]) est une facette que nous pensons tous être capables de déterminer. Et pourtant, d'une part il nous arrive de nous tromper (notamment au téléphone), et d'autre part, nombreux sont nos patients surtout les femmes qui viennent consulter justement parce que leurs interlocuteurs font fréquemment des erreurs. Dans le dernier numéro de Journal of Voice, Marylou Gelfer & Quinn Bennett [1] ont publié un article épatant à propos des indices sur lesquels nous nous appuyons pour fonder notre jugement.

Évidemment, dans la thématique de recherche qui est la mienne, « la voix et l'identité », les mécanismes de production vocale, leur résultante sonore, et les indices perceptifs qui sous-tendent la détermination du genre sont un domaine d'étude passionnant et font l'objet de mémoires d'orthophonie réalisés actuellement sous ma direction. Ces choses-là peuvent paraître simples, et pourtant, la question est loin d'être triviale…

INTRODUCTION

Il existe de nombreuses différences biologiques, anatomiques, comportementales, entre les hommes et les femmes, qui sont susceptibles d'affecter la production de la voix, et même les plus jeunes enfants sont capables de distinguer les locuteurs masculins des locuteurs féminins [2]. Pourtant, les vocologistes, qu'ils soient scientifiques ou cliniciens, peinent à déterminer les attributs spécifiques qui permettent à un auditeur de distinguer le genre d'une voix.

Nous savons tous que la hauteur de la voix revêt un caractère d'importance dans cette tâche-là (les hommes ont une voix plus grave que les femmes), mais même lorsque les locuteurs modifient leur fréquence usuelle, la catégorisation des auditeurs ne change pas nécessairement. Nous savons également que les résonances du conduit vocal, par l'intermédiaire des formants des voyelles, influence la perception du genre, mais la manière dont ces caractéristiques interagissent pour nous permettre de prendre notre décision reste floue.

LA QUESTION DE LA « HAUTEUR »

Le terme de « hauteur de la voix » fait référence au caractère aigu ou grave d'une voix, telle qu'elle est perçue par des auditeurs. La mesure acoustique de la « hauteur de la voix » est la fréquence usuelle (que l'on retrouve dans les publications internationales sous l'appellation « *speaking fundamental frequency* » ou SFF), ou la fréquence moyenne en hertz (Hz) d'un signal de parole. Classiquement, on retrouve dans la littérature des valeurs d'environ 118Hz pour les hommes (la#1) et 205Hz (sol#2) pour les femmes, donc une différence moyenne d'environ 10 demi-tons, un peu moins d'une octave (pour info, une octave est composée de 12 demi-

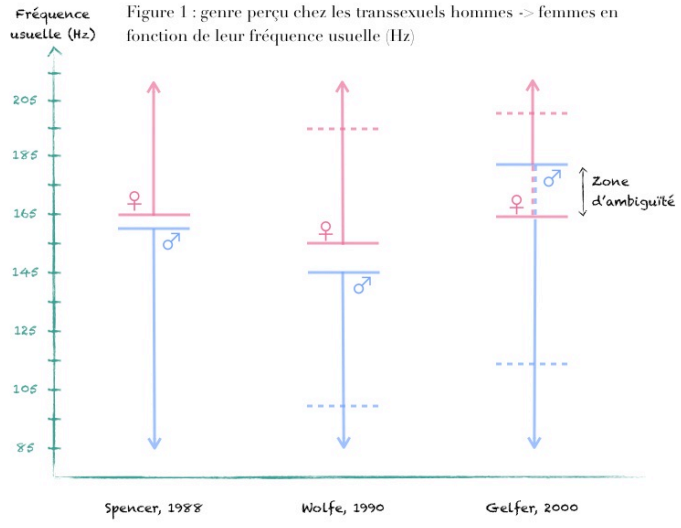
[1] Le sexe biologique est celui qui correspond à l'anatomie de l'individu, le sexe identitaire est celui qui correspond au genre auquel il s'identifie (qui peut être son genre biologique ou celui du sexe opposé comme dans le cas des transsexuels).

tons sur l'échelle musicale chromatique occidentale) [3]. Nous parlons évidemment ici de voix parlée et pas de voix chantée.

Évidemment, comme toujours, s'intéresser à des valeurs moyennes est une démarche normative un peu réductrice qui donne une idée générale mais ne tient pas compte des variations individuelles qui sont dans ce domaine relativement importantes. Dans une étude publiée par Hollien & Jackson [4], on retrouve des valeurs de 90Hz (fa1) pour la voix masculine la plus grave et 165Hz (mi2) pour la voix masculine la plus aigue, et chez les femmes 165Hz pour la voix la plus grave et 294Hz (ré3) pour la voix la plus aiguë, le tout en parole spontanée. Ainsi, Gelfer & Bennett [1]notent que malgré des valeurs moyennes franchement distinctes entre les hommes et les femmes, il est possible d'observer une zone de recouvrement entre les genres aux environs de 165Hz.

Cette valeur de 165Hz (mi2) est particulièrement intéressante et mérite qu'on s'y attarde. Les recherches publiées sur la voix des transsexuels homme -> femme suggèrent qu'une fréquence usuelle située dans cette zone pourrait représenter une « frontière » importante pour l'indentification du genre basée sur la seule écoute de la voix (figure 1). En 1988, Spencer [5] étudie la perception du genre d'après les voix de 8 transsexuels et montre que parmi eux, les quatre individus identifiés comme des femmes dans au moins 72% des cas (genre identitaire mais non biologique) ont tous une fréquence usuelle d'au moins 165Hz. En revanche, les quatre autres sujets, qui présentent tous une fréquence inférieure ou égale à 160Hz, sont systématiquement perçus comme des hommes par les 46 auditeurs participant à l'étude. Dans un travail similaire incluant 20 transsexuels mais seulement 10 auditeurs, Wolfe [6] trouve des résultats sensiblement différents indiquant que les 9 sujets de son corpus perçus comme des femmes ont une fréquence usuelle comprise entre 155 et 195Hz et les 11 sujets perçus comme des hommes entre 97 (!) et 145Hz. Enfin, dans une étude précédente à propos de 15 sujets transsexuels, Gelfer [7] note que seulement 3 sujets de son corpus sont systématiquement ou presque (≥ 90% des fois) comme des femmes. Elle relève chez ces 3 sujets des valeurs de fréquence usuelle comprises entre 164 et 199Hz. Mais elle note aussi que 10 sujets de son corpus, perçus systématiquement comme des hommes, montrent des valeurs comprises entre 112 (pas surprenant) et 181Hz (donc bien au delà de cette fameuse zone des 165Hz). Elle trouve une zone comprise entre 164 et 181Hz, au sein de laquelle certains sujets sont perçus comme des femmes (genre identitaire) et d'autres comme des hommes (genre biologique). Elle note cependant, à l'instar de Spencer [5], qu'aucun sujet présentant une fréquence inférieure ou égale à 164Hz n'est perçu comme une femme.

Figure 1 : genre perçu chez les transsexuels hommes -> femmes en fonction de leur fréquence usuelle (Hz)



Si on convertit les « Hertz » (qui donnent une mesure physique de la fréquence) en « demi-tons » (qui en donnent une mesure perceptive), alors Spencer [5] identifie le demi-ton #40 (mi2) comme le seuil de la voix féminine, alors que Wolfe [6] le situe plutôt au niveau du demi-ton #39 (ré#2). Concernant le plafond des voix masculines, pour Spencer il se place au demi-ton #39, et pour Wolfe au demi-ton #38 (ré2). Pour Gelfer [7], la fréquence la plus basse à partir de laquelle une voix de transsexuel homme -> femme est identifiée comme féminine se situe au demi-ton #40. Au total, ces résultats suggèrent l'étendue de fréquences comprise entre 145 et 165Hz pourrait être une région « frontière », une zone d'ambiguïté, dans laquelle certaines voix sont perçues comme masculines et d'autres comme féminines à valeurs de fréquence égales ou similaires. Notons que cela représente un intervalle de seulement 1 ton, c'est à dire la différence de hauteur entre deux touches blanches d'un piano (précisément le ré2 et le mi2), donc une différence relativement subtile.

Gelfer & Bennett [1] reconnaissent certaines limites dans les travaux cités ci-dessus en faveur de l'existence d'une région frontière : peu de sujets sont identifiés comme des femmes (3 dans l'article de Gelfer & Schofield,

4 dans celui de Spencer, les données de Wolfe n'étant pas précisées), et l'effectif variable des jurys d'écoute mis en place (10-46) pourrait amoindrir la fiabilité de l'emplacement des fréquences limites de la zone frontière. Pourtant, si cette zone d'ambiguïté existe réellement, alors d'autres indices que la fréquence usuelle doivent probablement être utilisés par les auditeurs pour identifier le genre lorsque la hauteur de leur voix se situe entre 145 et 165Hz. Il est alors vraisemblable que les fréquences des formants des voyelles prennent dans ce cas le pas sur la fréquence fondamentale usuelle pour permettre aux auditeurs de trancher.

LA QUESTION DES « FORMANTS » DES VOYELLES

L'emplacement des formants des voyelles est considéré depuis longtemps comme un critère saillant dans la perception du genre, même si le poids de leur contribution dans la décision finale des auditeurs est difficile à déterminer. Dans la littérature, des données normatives ont été publiées notamment par Peterson & Barney [8] et plus récemment par Hillenbrand [9], qui ont montré que les voyelles prononcées par les femmes ont des formants placés plu hauts que celles prononcées par des hommes, probablement en raison d'un conduit vocal présumé plus long chez les hommes. Quoi qu'il en soit, il est difficile de déterminer réellement les différences et les similitudes formantiques entre les genres : ni Peterson & Barney ni Hillenbrand, ne donnent les extrêmes ou les écarts-types des valeurs qu'ils ont relevées, pourtant les normes qu'ils avancent dans leurs résultats respectifs sont franchement différentes. Notamment, pour la voyelle /i/, Peterson & Barney situent le premier formant (F1) aux environs de 310Hz pour les femmes, ce qui est moins élevé que la valeur relevée par Hillenbrand pour le même formant chez les hommes (342Hz) ! Il faut noter que les moyens techniques mis en place dans ces deux études étaient franchement différents, donnant sans surprise des résultats qui le sont tout autant. Pour Gelfer & Bennett [1], de nouvelles recherches sur les normes des formants des voyelles sont clairement nécessaires, mais quoi qu'il en soit, si l'on prend ces deux études séparément, alors on peut tout de même leur trouver un point commun : les formants des femmes sont en effet, comme attendu, toujours plus élevés que ceux des hommes (tableau 1).

Fréquence des formants	♂ [Peterson & Barney, 1952]	♀ [Peterson & Barney, 1952]	♂ [Hillenbrand, 1995]	♀ [Hillenbrand, 1995]
/i/	F1 : 270 F2 : 2290 F3 : 3010	F1 : 310 F2 : 2790 F3 : 3310	F1 : 342 F2 : 2322 F3 : 3000	F1 : 437 F2 : 2761 F3 : 3372
/a/	F1 : 730 F2 : 1090 F3 : 2440	F1 : 850 F2 : 1220 F3 : 2810	F1 : 768 F2 : 1333 F3 : 2522	F1 : 936 F2 : 1551 F3 : 2815
/u/	F1 : 300 F2 : 870 F3 : 2240	F1 : 370 F2 : 950 F3 : 2670	F1 : 378 F2 : 997 F3 : 2343	F1 : 459 F2 : 1105 F3 : 2735

Tableau 1 : Valeurs normatives de l'emplacement des formants (Hz) en fonction du genre pour les voyelles /i/, /a/, /u/, d'après Peterson & Barney [8], et Hillenbrand [9].

Cette différence hommes/femmes étant systématique malgré l'hétérogénéité des résultats pris globalement, tout laisse à penser qu'il s'agit d'une réalité : les formants des voyelles sont placés plus hauts chez les femmes que chez les hommes. Ainsi, en l'absence d'une fréquence usuelle appropriée au genre, ou en cas d'incohérence entre la répartition formantique et la hauteur de la voix, on peut légitimement penser que l'emplacement des formants des voyelles devienne un indice encore plus important pour l'identification du genre. Par exemple, Mount & Salmon [10] décrivent une étude longitudinale à propos du cas d'un sujet transsexuel homme -> femme de 63 ans, proposant de suivre les progrès réalisés après un an de thérapie vocale. Durant cette période, la voix du sujet passe de 110 à environ 200Hz (du la1 au sol2, soit 10 demi-tons, presqu'une octave). Malgré cette augmentation considérable de la hauteur de sa voix, les auteurs notent que le sujet continuait à être identifié comme un homme au téléphone, jusqu'à ce qu'il parvienne à élever le second formant de ses voyelles (/i/, /a/, et /u/). Malheureusement, cette conclusion était présentée comme anecdotique et n'a pas été confirmée expérimentalement. Néanmoins, cela suggère que l'élévation des formants des voyelles pourrait jouer un rôle majeur dans la perception du genre pour les sujets transsexuels hommes -> femmes, en

particulier lorsque, comme ici, la hauteur de la voix n'est pas cohérente avec la répartition formantique. Dans une étude plus récente, Gelfer & Schofield [7] confirment que les sujets transsexuels en effet identifiés comme des femmes, ont toujours des formants plus élevés que ceux persistant à être perçus comme des hommes.

Il faut néanmoins noter la difficulté expérimentale à déterminer la contribution relative des deux indices décrits précédemment (hauteur de la voix et répartition formantique) en utilisant des enregistrements de parole naturelle (par opposition à de la parole de synthèse). Il est possible que les contraintes physiques de production induisent une interaction entre la hauteur de la voix et la répartition des voyelles, ce qui veut dire que nous ne pouvons pas vraiment modifier la hauteur de notre voix sans modifier notre manière d'articuler les voyelles, donc sans impacter leurs formants. Il est alors difficile de considérer ces deux éléments comme indépendants dans la parole naturelle et les auteurs se sont naturellement tournées vers une phonation « en condition de laboratoire », utilisant des moyens artificiels d'émission vocale afin de faire varier isolément la fréquence usuelle et les formants, afin de pouvoir étudier leur impact respectif sur l'identification du genre.

FRÉQUENCE USUELLE ET RÉPARTITION FORMANTIQUE : COMMENT FAIRE LA PART DES CHOSES ?

Une des premières études utilisant des moyens artificiels pour explorer la part relative de l'impact de ces deux éléments a été présentée par Coleman en 1976 [11]. Dans ce travail, 20 hommes et 20 femmes devaient lire un texte dans trois conditions différentes : avec leur voix naturelle, en utilisant un laryngophone[2] réglé à une fréquence de 120Hz (valeur typique masculine), puis à 240Hz (hauteur féminine). Cette méthode est théoriquement supposée permettre au sujet de conserver ses habitudes articulatoires (donc formantiques) tout en modifiant la hauteur de sa voix et ainsi d'obtenir une parole masculine avec une voix féminine et inversement. Coleman montre ainsi que les sujets utilisant le laryngophone réglé à une hauteur conforme à leur genre sont correctement identifiés dans presque 100% des cas. Lorsque les femmes utilisent le laryngophone à 120Hz, elles sont perçues comme féminines dans seulement 33% des cas, mais concernant les hommes, même avec un laryngophone à 240Hz, ils continuent à être identifiés comme tels dans 70% des cas. Coleman interprète ces résultats comme la preuve que les formants des voyelles ont un effet sur la perception du genre, précisant que la différence observées entre les hommes et les femmes pourrait être reliée 1/ au caractère très artificiel de la résultante sonore lorsque le laryngophone est réglé à 240Hz, 2/ au manque d'indices prosodiques dans la parole électrolaryngée, et 3/ à l'introduction de résonnances produites par le dispositif lui-même.

Fréquence du laryngophone	120Hz		240Hz	
Sujets perçus comme :	♂	♀	♂	♀
♂	100%		70%	
♀		33%		100%

Tableau 2 : Reconnaissance du genre avec laryngophone d'après Coleman [11]

Pour réduire ces artéfacts liés au laryngophone, Whiteside [12] propose d'utiliser des voyelles synthétiques, manipulées par informatique. Il conçoit donc différents stimuli avec des caractéristiques fréquentielles et formantiques congruentes pour chaque genre, et des caracté-ristiques incongruentes. Les résultats qu'il obtient sont simi-laires à l'étude de Coleman, avec toujours une identification des femmes un peu en deçà de celle des hommes (tableau 3). Mais ce qu'il convient de souligner c'est que les auditeurs in-clus dans le jury d'écoute se sont basés davantage sur la fréquence que sur les formants puisque les résultats montrent que la reconnaissance du genre suit la hauteur de la voix quelle que soit la répartition formantique.

	Formants ♂	Formants ♀
SFF ♂	97% homme	93% homme
SFF ♀	75% femme	85% femme

Tableau 3 : Reconnaissance du genre d'après Whiteside [12]

[2] Un laryngophone est un vibrateur électrique que le sujet place au niveau de son cou et qui produit la source sonore à la place du larynx. On voit assez fréquemment des patients laryngectomisés utiliser ce genre de dispositif.

Enfin, Gelfer & Mikos [13] ont repris le travail de Whiteside en ajoutant au corpus d'enregistrement des voix de transsexuels homme -> femme. Ils ont donc recruté 10 hommes, 10 femmes et 10 transsexuels, tous appariés en âge et en taille et dont ils ont analysé la production des voyelles. Ils ont ensuite synthétisé deux variations de chaque voyelle : une avec une fréquence à 120Hz et une autre à 240Hz. Les résultats montrent que lorsque les formants issus des échantillons masculins et féminins sont associés à leurs fréquences respectives appropriées, parviennent à identifier leur genre de manière plutôt précise (tableau 4). Concernant les sujets transsexuels, lorsque les formants sont associés à la fréquence de 240Hz, ils sont identifiés comme des femmes dans 84% des cas, mais les scores chutent lorsqu'ils sont associés à la fréquence de 120Hz (seulement 16% d'identification en femme). Gelfer & Mikos, comme Whiteside, tranchent donc en faveur d'un poids relatif plus important accordé à la fréquence dans la détermination du genre : les hommes ont la voix grave et les femmes ont la voix aigue. D'accord. Mais est-ce vraiment si simple ?

Fréquence de la voyelle	120Hz		240Hz	
Sujets perçus comme :	♂	♀	♂	♀
♂	84%		19%	
♀		20%		73%
⚥		16%		84%

Tableau 4 : reconnaissance du genre d'après Gelfer & Mikos [13]

Chez nos patients transsexuels justement, lorsque l'on travaille à la féminisation de leur voix (et pour reprendre d'observation mentionnée par Mount & Salomon plus haut), si on ne fait rien d'autre que d'augmenter la fréquence, la voix devient au mieux un peu étrange, et au pire franchement caricaturale. Cette conviction clinique ne va donc pas dans le sens d'un poids aussi exclusif de la hauteur de la voix. De nombreux autres aspects, articulatoires et prosodiques notamment, sont mis en place pour un résultat crédible et naturel. A plus y réfléchir, il est probable que l'étude de voyelles isolées (comme dans les travaux de Whiteside et de Gelfer & Mikos), gommant ces aspects articulatoires et prosodiques, ne permet pas une vue d'ensemble du phénomène. Hillenbrand & Clark [14] utilisent non pas de simples voyelles, mais des phrases. Ils enregistrent les voix de 25 hommes et 25 femmes à partir desquelles ils créent 4 échantillons différents : 1/ voix naturelle ; 2/ transposition de la fréquence (vers les aigus pour les hommes, vers les graves pour les femmes) ; 3/ transposition des formants (idem) ; et 4/ transposition de la fréquence et des formants (idem). Puis ils présentent les stimuli obtenus à un jury composé de 21 auditeurs. Pour les voix naturelles, le genre est correctement identifié dans 99,6% des cas. Lorsqu'une seule variable est modifiée, les résultats sont bien plus mitigés (tableau 5), mais lorsque les deux variables sont modifiées, alors 82% des hommes sont identifiées en femmes, et inversement. D'après les résultats de cette étude il semble donc que sur des échantillons de parole articulée (phrases), comprenant donc tous les aspects du discours (hauteur, articulation donc formants, prosodie, etc.), le jury fait davantage d'erreurs lorsque on modifie seulement la fréquence que lorsqu'on modifie seulement les formants, et la bascule significative d'un genre à l'autre nécessite la modification des deux : à ce moment là, 82% des hommes sont perçus en femmes et inversement.

	Voix naturelle		Modification de la fréquence		Modification des formants		Modification des deux critères	
	♂	♀	♂	♀	♂	♀	♂	♀
Identifiés ♂	99,6%		65,7%		81,1%		18,1%	
Identifiés ♀		99,6%		80,9%		88,3%		17,9%

Tableau 5 : Reconnaissance du genre sur une phrase d'après Hillenbrand & Clark [14]

Dans une seconde partie du même travail, Hillenbrand & Clark [14] proposent la même démarche sur un seul mot « *hVd* » (ou V représente la voyelle, en l'occurrence /a/, /i/, /u/, ou /ae/). Ils observent des résultats sensiblement identiques sur les voix naturelles avec une très bonne identification du genre (97%), et un taux d'erreurs encore plus significatif lorsque les deux critères sont modifiés puisque dans ce cas seuls 19% des hommes sont perçus comme tels, et 8% pour les voix de femmes (tableau 6). Mais concernant la part relative de

ces deux critères lorsqu'ils sont modifiés séparément, on observe une augmentation de l'écart, avec des erreurs d'identification beaucoup plus fréquentes lorsque seule la fréquence est modifiée que lorsqu'on change la répartition formantique. Ainsi, il faut comprendre que lorsqu'on modifie un seul paramètre, l'impact de la fréquence est plus important que celui de la répartition formantique. Mais ! L'importance de l'influence des formants est loin d'être négligeable puisque sa modification permet significativement de faire basculer l'identification vers l'autre genre lorsque les deux critères sont modifiés.

	Voix naturelle		Modification de la fréquence		Modification des formants		Modification des deux critères	
	♂	♀	♂	♀	♂	♀	♂	♀
Identifiés ♂	97,5%		44%		79%		19%	
Identifiés ♀		97%		32%		79%		8%

Tableau 6 : Reconnaissance du genre sur une syllabe d'après Hillenbrand & Clark [14]

Ainsi, pris tous ensemble, les travaux de Hillenbrand & Clark [14], Gelfer & Mikos [13], et Whiteside [12], suggèrent que les critères de fréquence et de répartition formantique sont tous les deux pris en considération par les auditeurs dans la détermination du genre. La fréquence serait un critère plus saillant lorsque les informations fréquentielles et formantiques sont incongruentes, en particulier lorsque l'étude porte sur des voyelles isolées. Cependant, comme évoqué plus haut, les valeurs normatives de l'emplacement des formants en fonction du genre sont mal définies. Il est ainsi impossible de savoir quel est le recouvrement entre les genres auquel on pourrait s'attendre : les valeurs entre les genres sont elles bien distinctes ou certaines valeurs pourraient-elle être observées indifféremment au sein des deux genres ?

COMMENT ALLER PLUS LOIN DANS L'ÉTUDE DE LA PERCEPTION DU GENRE ?

À ce stade, les différents travaux de recherche ont toujours comparé les formants masculins et féminins dans le contexte de hauteur de voix franchement dichotomiques, séparées de 0.75 à 1 octave. Mais que se passerait-il si on étudiant l'importance des informations formantiques à partir de voix dont la hauteur n'est ni clairement masculine, ni clairement féminine ? Les recherches citées précédemment dans le paragraphe consacré à l'influence de la hauteur, suggéraient que dans la gamme de fréquences comprises entre 145 et 165Hz, les voix sont susceptibles d'être perçues quelquefois comme masculines et quelquefois comme féminines. Pour un individu dont la fréquence usuelle est située dans cette zone d'ambiguïté, est-ce que l'emplacement des formants ne deviendrait pas un marqueur du genre davantage saillant ? Et inversement, est-ce que les valeurs fréquentielles ne prendraient pas aussi plus d'importance si ce sont les informations formantiques qui sont ambiguës ?

Une explication courante à propos des différences formantiques observées entre les hommes et les femmes tient à l'idée anatomique que le conduit vocal est plus long chez les hommes (puisqu'ils sont la plupart du temps plus grands). On peut donc imaginer que chez des hommes et des femmes sensiblement de la même taille, ces différences devraient s'estomper et les informations formantiques devenir ambiguës. Dans ce cas, il est probable que le jury utilise davantage les informations fréquentielles pour reconnaître le genre. La poursuite du même raisonnement conduit à penser que les hommes les plus grands et les femmes les plus petites devraient montrer des différences formantiques maximales. Gelfer & Bennett [1] supposent donc que lorsque la différence de taille entre les individus est très importante, alors l'emplacement des formants devient un critère très saillant pour la détermination du genre, plus que la fréquence.

Pour explorer ces hypothèses, ils recrutent 30 étudiants âgés de 20 à 35 ans d'après les critères suivants : des hommes « grands » (entre 187 et 193cm, n=8), des femmes petites (entre 152 et 157cm, n=8), et des hommes et des femmes de taille similaire (entre 162 et 172cm, appariés à 2cm près , n=7 pour chaque sexe, soit 14 en tout). Ils enregistrent leur voix, et procèdent à une modification synthétique de leur fréquence usuelle, de telle sorte qu'ils obtiennent échantillons pour chaque participant : la voix naturelle, un échantillon à 116Hz (hauteur masculine typique), un échantillon à 207Hz (hauteur féminine typique), et trois échantillons dans la zone de fréquences ambiguë, donc à 145, 155 et 165Hz.

À l'aide de ces stimuli, ils entendent répondre à 4 questions : 1/ existe-t-il une répartition formantique typique des genres ? et si oui, de quel ordre est la potentielle zone de recouvrement ? 2/ quel est le degré de pertinence de l'identification du genre chez les hommes grands et les femmes petites lorsque leur répartition formantique naturelle est conservée mais que l'on fait varier la hauteur de leur voix ? 3/ quel est le degré de pertinence de la perception du genre chez les hommes et les femmes de taille identique dans les mêmes conditions de variation de hauteur ? et enfin 4/ quelle est la contribution des résonances du tractus vocal (donc de la répartition formantique) dans l'identification du genre ?

Concernant la question 1/, les valeurs normatives confirment que d'une manière générale, les femmes possèdent une répartition formantique placée plus haut que celle des hommes pour toutes les voyelles. Si on s'intéresse maintenant aux résultats par groupe, on constate que cela reste vrai même lorsque les hommes et les femmes sont de taille similaire. Si on compare les deux groupes d'hommes, les formants des grands sont plus bas que ceux des hommes de taille moyenne. En revanche, aucune différence significative n'apparaît entre les deux groupes de femmes. Si on regarde maintenant les écarts à la moyenne pour les hommes d'une part et les femmes d'autre part, on note un recouvrement important entre les genres : dans tous les cas, la plus haute valeur formantique des hommes est largement au dessus de la plus basse valeur relevée chez les femmes. Cela suggère que dans certains cas individuels, les formants d'un homme pourraient être les mêmes, voire plus élevés que ceux de certaines femmes. Les auteurs ne font pas d'analyse spécifique de ces cas particuliers et donc les questions soulevées par ce constat la perception du genre était-elle pertinente chez eux ? leurs voix étaient-elles perçues comme ambiguës ? la hauteur de leur voix était-elle, elle aussi ambiguë ? si non, la fréquence permettait-elle de trancher ? restent posées.

Si on s'intéresse maintenant aux résultats de l'identification du genre chez les hommes grands et les femmes petites (question 2/), donc entre les deux groupes sensés avoir le maximum de différences formantiques, voici ce que l'on obtient (figure 2). Notons que les auteurs considèrent comme constamment performantes les évaluations correctes à plus de 80%. Ainsi, quelle que soit la fréquence, les femmes petites sont toujours correctement identifiées (100% dans leur zone de fréquence spécifique (207Hz) mais toujours plus de 86% dans la zone spécifique masculine). En ce qui concerne les hommes grands, leur identification est performante de 116Hz à 165Hz (qui est je vous le rappelle la zone frontière), puis la courbe se brise et le taux de d'identification correcte à 207Hz n'est plus que de 53%, c'est à dire au niveau du hasard. En d'autres termes, à 207Hz, on a autant de chance de déterminer le genre d'un homme grand par tirage au sort qu'en écoutant sa voix.

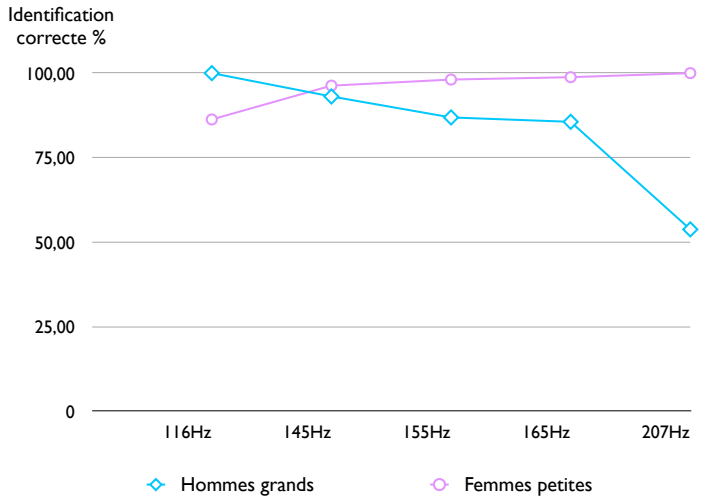


Figure 2 : pourcentage d'identification correcte du genre chez les hommes grands et les femmes petites en fonction de la hauteur de la voix. D'après Gelfer et Bennett [1]

Lorsqu'on compare ces deux groupes, il semble donc que dans le contexte de différences formantiques maximales, la perception du genre des hommes grands est davantage perturbée par des informations incongruentes de hauteur de voix que celle des femmes. Rappelons que l'hypothèse de base était que dans ces deux groupes, l'identification du genre serait basée principalement sur l'emplacement des formants et serait performante quelle que soit la hauteur de la voix. On note ici que cette hypothèse est validée pour les femmes petites, mais aussi même si dans une moindre mesure pour les hommes grands, puisque nous n'assistons pas à une franche bascule en faveur de l'identification au genre opposé : si la hauteur de la voix avait une véritable importance, du moins du niveau de

celle attendue, alors à 207Hz, le taux de réponse correcte chez les hommes grands serait de moins de 20% pour signer une identification massive en femme.

Concernant l'identification du genre chez les hommes et les femmes de taille similaire (question 3/), l'hypothèse était que ces derniers ayant *a priori* des conduits vocaux de taille comparable, leur répartition devrait être similaire et l'identification du genre devrait donc dépendre fortement de la hauteur de leur voix. En réalité, les résultats montrent (figure 3) exactement la même tendance que précédemment, avec une identification correcte systématique du genre des femmes moyennes, quelle que soit la hauteur de leur voix, et des performances moindres concernant les hommes moyens, avec une identification qui passe en dessous du seuil d'identification correcte cette fois dès 165Hz (75%) et seulement de 44,3% à 207Hz (ce qui reste relativement proche du hasard). Concernant la question 3, l'hypothèse est donc cette fois davantage pertinente pour les hommes que pour les femmes, mais en l'absence d'une franche bascule du taux d'erreur (qui devrait être aux environs de 20% pour parler d'identification en femme), elle n'est pas validée.

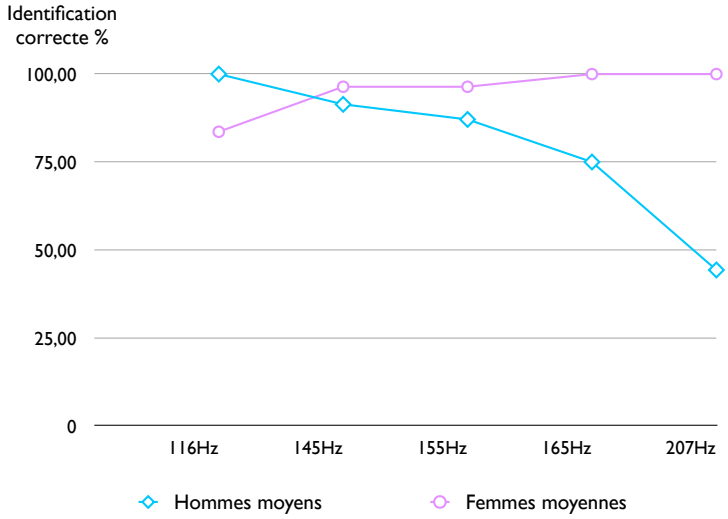


Figure 3 : pourcentage d'identification correcte du genre chez les hommes et les femmes de taille similaire en fonction de la hauteur de la voix. D'après Gelfer et Bennett [1]

Enfin, concernant la contribution spécifique de la répartition formantique dans la tâche d'identification du genre, les auteurs proposent d'utiliser l'outil statistique « *Pearson Product-Moment Correlations* » avec un taux de significativité fixé à p=0,0038. Je vous fais grâce des détails de la procédure statistique ;-) Toutes les corrélations entre l'identification correcte du genre et les valeurs d'emplacement des formants sont négatives et statistiquement significatives (ce qui veut dire qu'il y a moins de 3% de chances que ces résultats soient dus au hasard) : plus le pourcentage de réponses « identification masculine » est élevé, plus les formants sont bas, et inversement. Cela suggère en d'autres termes que les résonances du tractus vocal ont une influence significative sur le jugement perceptif du genre.

Donc l'emplacement des formants a une importance significative dans notre prise de décision, mais pourtant on ne retrouve pas de tendance nette lorsqu'on regarde les tableaux 2 et 3. Alors comment interpréter ces résultats ? Dans la discussion, les auteurs reconnaissent que malgré le soin apporté au recrutement de leurs sujets, seuls trois individus (un homme et deux femmes) ont réellement montré des valeurs formantiques comparables, ce qui veut dire d'une part que malgré les recouvrements observés dans les valeurs normatives il n'est pas si facile que ça de trouver des similarités entre les genres, et d'autre part que si le contrôle de la taille des individus n'a pas permis de retrouver des répartitions formantiques similaires, alors ce n'est pas surprenant que la question 3/ n'ait pas trouvé de réponse valide : ce n'est pas que chez les hommes et les femmes moyens l'importance de la hauteur de la voix n'est pas significative, c'est surtout que dans ces deux groupes là, les valeurs formantiques n'étaient pas réellement similaires comme espéré. Il faut donc considérer que la question 3/ n'a juste pas trouvé de réponse…

BREF… QUE RETENIR DE TOUT ÇA ?

*T*ous les travaux de recherche publiés sur ce thème vont dans le sens de montrer une participation des phénomènes de résonance dans l'identification du genre. C'est donc bien qu'il doit s'y passer quelque chose ! L'immense difficulté pour les chercheurs réside dans la possibilité de créer des échantillons de voix, qui « sonnent » naturels, tout en créant des conditions qui sont tout sauf naturelles (les formants d'un genre, la hauteur de l'autre)… Il est donc probable que les limitations qui freinent la recherche dans ce domaine soient d'ordre essentiellement technique, et le temps, faisons lui confiance, aura raison de ces difficultés-là. Mais vous l'aurez compris, ce n'est pas vraiment cette question-là qui m'intéresse…

Revenons donc à notre intuition clinique. Nos patientes dysphoniques, en particulier celles qui sont porteuses d'un œdème de Reinke, se plaignent continuellement d'être appelées « *Monsieur* » au téléphone. Certes, la hauteur de leur voix est largement aggravée, mais pourquoi diable leurs formants seraient-ils affectés ? Il n'existe aucun argument physiologique qui pourrait être mis en avant en faveur de cette idée-là. Pourtant on les appelle fréquemment « *Monsieur* », c'est une réalité, et moi même, avec mes presque vingt ans d'expérience et mes connaissances sur la chose, il m'arrive de commettre l'erreur. D'un autre côté, nos patients transsexuels peinent à adopter les caractéristiques vocales féminines et donc se plaignent de la même chose… S'il est relativement aisé de modifier la hauteur d'une voix, c'est beaucoup plus difficile d'avoir une action consciente sur la position des résonateurs pour modifier la répartition formantique. On peut essayer d'élever le larynx pour raccourcir le tractus vocal, on peut aussi antérioriser l'articulation pour élever le premier formant, mais on ne sait pas trop quoi faire de plus… Ainsi nous sommes face ici précisément à deux cas de figure clinique qui illustrent bien l'article que nous venons de résumer : d'un côté des patientes qui ont des formants congruents avec leur genre mais une fréquence incongruente, et de l'autre des sujets transsexuels qui parviennent à adopter une fréquence congruente avec leur genre identitaire (féminin) mais peinent à modifier l'emplacement de leurs formants qui demeure confruent avec leur genre biologique (masculin). Les unes ont les bons formants mais la mauvaise fréquence, les autres la bonne fréquence mais les mauvais formants. Et dans les deux cas, le résultat est le même : on les identifie comme des hommes… Nous l'avons vu précédemment, Hillenbrand & Clark pensent que la bascule de l'identification du genre ne peut se faire que lorsque les caractéristiques de hauteur de voix et d'emplacement formantique sont congruentes. Pourtant, ce n'est le cas d'aucun des exemples cliniques développés ici…

Loin de moi l'idée de nier la participation de la fréquence et des formants : cela joue, c'est indéniable, mais ça ne peut pas tout expliquer. L'an dernier, forts de nos intuitions de cliniciens, et cédant à la pression de nos patients qui cherchent sinon à comprendre, du moins à trouver des solutions pour que leur genre soit perçu dans leur voix, nous avons lancé un plan expérimental sur ce thème [15]. Il existe de très nombreux travaux publiés sur le thème des valeurs normatives de hauteur de la voix en fonction de la sévérité de la dysphonie. Tous sont d'accord : plus la dysphonie est sévère plus la voix des femmes s'aggrave et plus la fréquence des hommes s'élève (tableau 7).

Fréquence moy (e-t)	Voix normale (grade 0)	Dysphonie légère (grade 1)	Dysphonie moyenne (grade 2)	Dysphonie sévère (grade 3)
♂	122 (24)	131 (27)	130 (33)	167 (32)
♀	218 (28)	204 (34)	195 (37)	168 (57)

Nous avons donc sélectionné 198 voix dans notre base de données (18 femmes et 23 hommes à la voix normale, 50 femmes et 18 hommes grade 1, 41 femmes et 14 hommes grade 2, et 23 femmes et 11 hommes grade 3), que nous avons fait écouter à 29 auditeurs naïfs (c'est à dire non spécialistes de la voix). Sur les 4640 réponses recueillies, nous avons observé un taux d'erreur de… 1,3%… Évidemment, sur l'ensemble de notre population, tous les sujets ne se plaignaient pas d'erreurs sur leur genre, et qui plus est tous ne montraient pas une modification significative de la hauteur de leur voix. Mais même si on ne prend que les grades 3, dont la fréquence appartient à la zone d'ambiguïté, seuls 1,7% d'entre eux ont montré des erreurs d'identification du genre…

Tout ça pour dire que l'identification du genre est une tâche que nous pratiquons quotidiennement et que nous réussissons plutôt bien, quelles que soient les contextes (dysphonie, manipulation de la hauteur, emplacement des formants). J'en viens donc à ma conclusion… et si c'était ailleurs que ça se passe ? J'ai plein d'idées sur la question. La prosodie, en particulier, le rythme particulier de notre parole, sa mélodie, voilà des indices qui forcément- participent à notre faculté de discrimination. J'ai plein d'idées disais-je mais je suis aussi preneur des vôtres ! La recherche dans ce domaine se poursuit dans le monde, à ma modeste mesure d'autres expérimentations sont aussi en cours, mais il reste aujourd'hui difficile d'apporter d'autre conclusion que le gimmick habituel que l'on retrouve en fin de toute publication ou presque : *further studies are needed*.

La suite, donc, au prochain numéro !

BIBLIOGRAPHIE

1. PAUSEWANG GELFER M & BENNETT QE (2013) Speaking fundamental frequency and vowel formant frequencies: effect on perception of gender. *Journal of Voice* 27(5): 556-566
2. PERRY TL OHDE RN, ASHMEAD DH (2001) The acoustic bases for gender identification from children's voices. *Journal of the Acoustic Society of America* 109(6): 2988-2998
3. BAKEN RJ (1996) *Clinical Measurement of Speech and Voice*. San Diego (USA): Singular Publishing Group
4. HOLLIEN H & JACKSON B (1972) Normative data on the speaking fundamental frequency and chronologic age in males. *Journal of Phonetics* 1: 117-120
5. SPENCER L (1988) Speech characteristics of male-to-female transsexuals: a perceptual and acoustic study. *Folia Phoniatrica* 40: 31-42
6. WOLFE V, RATUSNIK D, SMITH FG & NORTHROP G (1990) Intonation and fundamental frequency in male-to-female transsexuals. *Journal of Speech & Hearing Disorders* 55: 43-50
7. PAUSEWANG GELFER M & SCHOFIELD KJ (2000) Comparison of acoustic and perceptual measures of voice in male-to-female percieved as female versus those percieved as male. *Journal of Voice* 14(1): 22-33
8. PETERSON GE & BARNEY HL (1952) Control methods used in a study of the vowels. *Journal of the Acoustic Society of America* 24: 175-184
9. HILLENBRAND J, GETTY LA & CLARK MJ (1995) Acoustic characteristics of american english vowels. *Journal of the Acoustic Society of America* 97: 3099-3111
10. MOUNT K & SALMON S (1988) Changing the vocal characteristics of a postoperative transsexual patient: a longitudinal study. *Journal of Communication Disorders* 21: 229-238
11. COLEMAN R (1976) A comparison of the contributions of two voice quality characteristics to the perception of maleness and femaleness in the voice. *Journal of Speech & Hearing Research* 19: 168-180
12. WHITESIDE SP (1998) The identification of a speaker's sex fron syntheszed vowels. *Perceptual & Motor Skills* 87: 595-600
13. PAUSEWANG GELFER M & MIKOS VA (2005) The relative contributions of speaking fundamental frequency and formant frequencies to gender identification based on isolated vowels. *Journal of Voice* 19: 544-554
14. HILLENBRAND J & CLARK MJ (2009) The role of f0 and formant frequencies in distinguishing the voices of men and women. *Attention Perception Psychophysics* 71: 1150-1166
15. MORACCHINI MC, PASCAL A & REVIS J (2013) *Identification perceptive du genre par un jury à partir de voix normales et dysphoniques*. Mémoire de fin d'études d'Orthophonie, Centre de Formation en orthophonie de Marseille. Marseille (France): Aix-Marseille Université