



Différences mélodiques et spectrales entre sexes comparées chez les locuteurs coréens et français

Dayeon Yoon¹ Nicolas Audibert¹ Cécile Fougeron¹

(1) Laboratoire de Phonétique et Phonologie (UMR7018 CNRS/Univ. Sorbonne-Nouvelle), 19 rue des Bernardins, 75005 Paris, France

{dayeon.yoon, nicolas.audibert, cecile.fougeron}@sorbonne-nouvelle.fr

RÉSUMÉ

Cette étude examine la variation entre langues des différences entre sexes dans la production de la parole. Un ensemble de paramètres acoustiques liés à la fréquence fondamentale et sa variation, à la forme du spectre et au degré de périodicité sont analysés sur les productions de 70 hommes et femmes coréens et français. Les résultats montrent un effet du sexe sur la plupart des paramètres avec des valeurs plus élevées chez les femmes (f_0 , $H1^*-H2^*$, $H1^*-A1^*$, $H1^*-A2^*$, et HNR) excepté une valeur plus importante de $H2kHz^*-H5kHz^*$ chez les hommes, ce qui indique une voix perçue comme plus soufflée chez les femmes. Pour $H1^*-A1^*$, $H1^*-A2^*$, $H1^*-A3^*$, les différences entre sexes ne sont significatives que dans le groupe français. Ces résultats suggèrent qu'au-delà du dimorphisme sexuel dans les deux langues, les différences entre voix d'hommes et voix de femmes seraient moins marquées en coréen qu'en français.

ABSTRACT

Sex differences in melody and spectrum compared in Korean and French speakers

This study examines the cross-language variation of sex differences in speech production. A set of acoustic parameters related to the fundamental frequency and its variation, to the spectral shape and to the degree of periodicity are analyzed on the productions of 70 Korean and French men and women. The results show an effect of sex on most parameters with higher values in women (f_0 , $H1^*-H2^*$, $H1^*-A1^*$, $H1^*-A2^*$, and HNR) except for a higher value of $H2kHz^*-H5kHz^*$ in men, which indicates a voice perceived as breathier in women. For $H1^*-A1^*$, $H1^*-A2^*$, $H1^*-A3^*$, the differences between sexes are significant only in the French group. These results suggest that beyond the sexual dimorphism in the two languages, the differences between male voices and female voices are less marked in Korean than in French.

MOTS-CLÉS : sexe, variation inter-langues, voix, fréquence fondamentale, spectre

KEYWORDS: sex, cross-language variation, voice, fundamental frequency, spectrum

1 Introduction

Les différences vocales entre les sexes se basent d'une part sur des différences anatomiques et physiologiques. Les différences morphologiques entre hommes et femmes au niveau des plis vocaux, par exemple, engendrent des aspects vocaux spécifiques au sexe. Par exemple, Boë et al. (1975) ont montré que la fréquence fondamentale (f_0) est un des corrélats acoustiques les plus discriminants entre les hommes et les femmes. Les plis vocaux moins longs et moins épais chez les femmes que

chez les hommes permettent un nombre plus élevé de vibrations dans la voix de femmes (207 Hz) que dans la voix d'hommes (118 Hz).

Le dimorphisme sexuel au niveau des plis vocaux a également une influence importante sur les différences entre les sexes dans les corrélats acoustiques liés à la source vocale. À titre d'exemple, les mesures des différences d'amplitude harmonique comme H1-H2 (différence d'amplitude entre premier et deuxième harmoniques (« H »)), H1-A1, H1-A2, H1-A3, (différence d'amplitude entre premier harmonique (« H ») et premier/deuxième/troisième formant (« A »), respectivement) etc., sont souvent utilisées pour examiner des différences entre la voix d'hommes et la voix de femmes (Hanson, Chuang, 1999 ; Henton, Bladon, 1985 ; Iseli et al., 2007 ; Klatt, Klatt, 1990 ; Narra et al., 2015). Les mesures des différences d'amplitude entre harmoniques sont supposées dépendre du degré de fermeture de la glotte, qui est lui-même influencé par l'épaisseur des plis vocaux. Comme les plis vocaux sont relativement moins épais chez les femmes que chez les hommes, l'adduction incomplète des plis vocaux se trouve plus fréquente dans la phonation des femmes (Titze, 1989). Par conséquent, les différences H1-H2, H1-A1, H1-A2, H1-A3 ont été rapportées être plus grandes chez les femmes que chez les hommes (Iseli et al., 2007; Narra et al., 2015). La fermeture incomplète des plis vocaux permet au flux d'air de toujours passer entre les plis vocaux en permettant aux voix de femmes d'être perçues comme plus soufflées que les voix d'hommes (Södersten et al., 1990). Cette caractéristique des plis vocaux est reflétée par d'autres paramètres acoustiques tels que le HNR (rapport harmonique/bruit) et le CPP (Cepstral Peak Prominence). Ils sont souvent utilisés pour l'évaluation de la dysphonie (Hillenbrand, Houde, 1996), mais aussi pour la comparaison de la voix en fonction de l'âge et du sexe dans les voix non-pathologiques (Chen et al., 2010 ; Choi, Choi, 2016).

Un autre aspect contribuant à la différence vocale entre hommes et femmes est la langue parlée. Autrement dit, la voix d'hommes et la voix de femmes sont construites dans la communauté linguistique. Ce phénomène est illustré notamment par le cas des bilingues, qui utilisent le même appareil vocal dans les deux langues et permettent donc une comparaison directe dans laquelle la dimension morphologique est parfaitement contrôlée. Pépiot et Arnold (2020) ont montré que les femmes bilingues français-anglais présentent une valeur de la f_0 plus élevée dans la phrase en français qu'en anglais alors que leur variation de la f_0 est plus grande en anglais qu'en français ; en revanche, chez les hommes bilingues français-anglais, il n'y a aucune différence significative de f_0 entre les langues alors que la variation de la f_0 est plus grande en français qu'en anglais. Des différences de valeurs de H1-H2 entre sexes ont été aussi observées dans leur étude. Dans les deux langues, les valeurs H1-H2 ont été plus élevées chez les femmes, ce qui suggère que les voix des femmes sont plus soufflées que les hommes. De plus, les femmes ont montré une valeur plus élevée en anglais alors que les hommes ont présenté une valeur moins élevée en anglais : les femmes bilingues français-anglais présentent donc une voix plus soufflée quand elles parlent en anglais, et inversement pour les hommes. Quant aux bilingues anglais-coréen étudiés par Cheng (2019), la f_0 est plus élevée dans l'interview en coréen qu'en anglais indépendamment du sexe. En plus, pour la modulation de la f_0 (l'écart-type de la f_0 en demi-ton), il y a une différence en fonction de la langue parlée entre les sexes : chez les hommes, la f_0 est plus variée en coréen; chez les femmes, la f_0 est plus variée en anglais.

Enfin, toutes les études examinées ci-dessus mettent en évidence une interaction entre le sexe et la langue parlée du locuteur, ce qui suggère que le dimorphisme ne serait pas marqué acoustiquement de la même façon dans les différentes langues étudiées. Dans la même ligne que ces recherches, le premier objectif de la présente étude est de vérifier l'effet du sexe sur les paramètres acoustiques liés à la phonation. Le deuxième objectif de cette étude est d'explorer si les différences entre les sexes dans les variables acoustiques examinées varient en fonction de la langue parlée. Comme langue parlée, le coréen et le français ont été examinés.

2 Méthode

2.1 Participants

Un total de 70 locuteurs qui n'ont jamais souffert d'une maladie liée aux plis vocaux ont été examinés : 17 hommes et 20 femmes coréanophones et 16 hommes et 17 femmes francophones, tous âgés de 18 à 42 ans (l'âge moyen de 28,5 ans ($\sigma = 3,9$) chez les H et de 27,3 ans ($\sigma = 4,9$) chez les F dans le groupe CR ; de 30 ans ($\sigma = 6,4$) chez les H et de 24,3 ans ($\sigma = 5,1$) chez les F dans le groupe FR).

2.2 Procédure d'enregistrement

Les participants ont été enregistrés dans des espaces calmes tels qu'une salle de cours, une salle de réunion ou leur domicile. Un micro-casque dynamique unidirectionnel Shure SM10A et une carte son Studio 24c de PreSonus ont été utilisés pour l'enregistrement. Les voix des participants ont été enregistrées à une fréquence d'échantillonnage de 44,1 kHz, numérisées avec une résolution de 16 bits et sauvegardées au format WAV. Le microphone a été placé à un angle de 45 degrés et à 7 à 8 centimètres de la bouche du locuteur.

2.3 Données expérimentales

Les locuteurs ont reçu l'instruction de lire une phrase dans laquelle une séquence cible « mama, mamama, mamamama » est intégrée. Les Coréens ont lu une phrase « 오늘 오후에 마마, 마마마, 마마마마, 셋이 공원에 가기로 했다 » /ontul ohue mama, mamama, mamamama s'heei koŋw'ane kagiŋo heŋ'a/, dont le sens correspond à la phrase « Cet après-midi, mama, mamama et mamamama, nous irons au parc de l'île Sainte-Ida » lue par les Français. Pour une lecture la plus naturelle possible de la séquence cible « mama, mamama, mamamama », les locuteurs ont été invités à imaginer des noms de personnes à deux syllabes ('미나'[mina] pour les CR, 'Marie' pour les FR), trois syllabes ('영식이'[yŋŋeigi] pour les CR, 'Amélie' pour les FR), et quatre syllabes ('가희 누나'[kahi nuna] pour les CR, 'Isabella' pour les FR), respectivement.¹ Puis, les phrases ont été manuellement segmentés dans Praat (Boersma, Weenink, 2022) pour analyser les occurrences de la voyelle /a/.

2.4 Analyses des données

Les données ont été analysées basé sur les paramètres acoustiques qui peuvent être groupés en trois catégories : f0 (Hz) et sa modulation (demi-ton), forme du spectre (H1*-H2*, H2*-H4*, H4*-H2kHz*, H2kHz*-H5kHz, H1*-A1*, H1*-A2*, H1*-A3*), et degré de périodicité (CPP et HNR).² Les valeurs

¹ La plupart des prénoms coréens ont deux syllabes. Mais, si la dernière syllabe se termine avec une syllabe fermée, [i] est ajouté à la fin pour en faire une syllabe ouverte, comme '영식이'[yŋŋeig + i]. Et, si un homme ou une femme plus jeune parle d'une femme plus âgée que lui ou elle dans une relation étroite, on ajoute '누나'[nuna] ou '언니'[ŋŋni], respectivement, à la fin du nom comme '가희 누나'[kahi + nuna] ou '가희 언니'[kahi + ŋŋni].

² Les mesures corrigées, représentées par l'astérisque, ont été utilisées à cause d'éventuelles différences des formants dans la voyelle /a/ entre le coréen (F1 et F2 moyens de 638 Hz et de 1333 Hz chez les hommes ; de 825 Hz et de 1688 Hz chez les femmes) et le français (F1 et F2 moyens de 602 Hz et de 1476 Hz chez les hommes ; de 708 Hz et de 1705 Hz chez les femmes) (Yoon, Kang, 2014; Kahn et al., 2011). Dans VoiceSauce, l'algorithme développé par Iseli et al. (2007) est utilisé.

acoustiques ont été estimées dans le logiciel VoiceSauce (Shue et al., 2009). Elles ont été mesurées sur chaque intervalle de 5 ms dans la séquence cible. Préalablement à ces mesures, tous les enregistrements ont été sous-échantillonnés à 16 kHz.

Tout d'abord, la f_0 a été mesurée basée sur l'algorithme STRAIGHT (Kawahara et al., 1999) utilisé par défaut dans le logiciel. La plage pour la détection de f_0 a été définies entre 70 Hz et 300 Hz pour les hommes et entre 100 Hz et 600 Hz pour les femmes. Pour la variation de la f_0 , les valeurs de la f_0 ont été converties en demi-tons pour normaliser les différences de hauteur mélodique entre hommes et femmes. L'écart-type de la f_0 en demi-tons a été calculé pour chaque locuteur.

Pour la forme du spectre, les paramètres de l'amplitude harmonique $H1^*-H2^*$, $H2^*-H4^*$, $H4^*-H2kHz^*$, $H2kHz^*-H5kHz$, $H1^*-A1^*$, $H1^*-A2^*$, et $H1^*-A3^*$ ont été mesurés avec Snack Sound Toolkit (Sjölander, 2004) utilisé par défaut par VoiceSauce. $H1^*-H2^*$ est un indicateur de la longueur relative de la phase d'ouverture de l'impulsion glottale (quotient d'ouverture) (Hanson, 1997). Donc, plus la valeur de $H1^*-H2^*$ est élevée, plus une voix est perçue comme soufflée. Selon Kreiman et Garellek (2011), une valeur importante de $H2^*-H4^*$ peut également refléter une voix perçue comme soufflée. Cependant, lorsque $H1^*-H2^*$ et $H2^*-H4^*$ covarient dans le sens opposé, c'est-à-dire quand $H1^*-H2^*$ est plus élevé que dans la voix modale alors que $H2^*-H4^*$ est moins élevé, et vice versa, les deux composantes peuvent s'annuler et la voix peut être perçue comme modale. $H4^*-H2kHz^*$ est une mesure de la pente spectrale dans une bande de fréquence plus élevée entre la quatrième harmonique et l'harmonique autour de 2000 Hz. $H2kHz^*-H5kHz$ est une mesure de la pente spectrale harmonique qui ne dépend pas de la f_0 (Johnson et al., 2020). $H1^*-A1^*$ correspond à la largeur de bande du F1. C'est un indicateur de la présence d'une fente glottique postérieure, fréquente dans la voix soufflée. $H1^*-A2^*$ est un indicateur de la pente spectrale aux fréquences médianes des formants. $H1^*-A3^*$ représente la pente spectrale aux fréquences de formants supérieures. Cela reflète une fermeture brusque des plis vocaux (Stevens, Hanson, 1995).

Quant au degré de périodicité, le HNR et le CPP ont été mesurés. Le HNR a été estimé par l'algorithme de Krom (de Krom, 1993) entre 0-500 Hz, entre 0-1500 Hz, entre 0-2500 Hz et entre 0-3500 Hz. Comme le HNR est le rapport harmonique/bruit, une valeur moins élevée indique typiquement une voix perçue plus soufflée avec relativement plus de bruit. Le CPP (Cepstral Peak Prominence) est une mesure de l'amplitude de pic cepstral normalisé pour l'amplitude globale (Hillenbrand et al., 1994). L'idée est qu'un signal périodique doit montrer une structure harmonique bien définie et, par conséquent, un pic cepstral plus proéminent qu'un signal moins périodique. Dans VoiceSauce, le CPP a été détecté à l'aide de l'algorithme décrit par Hillenbrand et al. (1994).

Enfin, un modèle linéaire à effets mixtes a été exécuté en utilisant la fonction lmer dans le package lme4 pour R (Bates et al., 2015) pour examiner l'effet du sexe et de la langue parlée sur les paramètres acoustiques : le sexe et la langue parlée des locuteurs en tant que facteurs fixes et le locuteur en tant qu'ordonnée à l'origine (intercept) aléatoire ($VD \sim \text{sexe} * \text{langue} + (1|\text{locuteur})$, VD désignant la variable dépendante). Les effets des facteurs sur la modulation de la F_0 ont été évalués par un modèle linéaire simple ($VD \sim \text{sexe} * \text{langue}$) en raison du nombre restreint de valeurs (une par locuteur) ne permettant pas au modèle mixte de converger. Les contrastes entre chaque groupe ont été estimées avec une correction de Tukey.

3 Résultats

La table 1 récapitule les effets des facteurs et interactions ainsi que les comparaisons post-hoc. Les comparaisons entre sexes et langues réalisées dans chacun des modèles sont illustrées par la figure 1.

Variable	Effets			Différence H-F	
	Sexe	Langue	Interaction	CR	FR
f0 (Hz)	$\chi^2(1) = 117.59$ $p < .001^{***}$	$\chi^2(1) = 0.62$ $p = .431$	$\chi^2(1) = 0.75$ $p = .386$	H < F $p < .001^{***}$	H < F $p < .001^{***}$
Modulation f0 (demi-ton)	F(1, 66) = 0.8 $p = .375$	F(1, 66) = 0.002 $p = .966$	F(1, 66) = 1.43 $p = .237$	H < F $p = .998$	H > F $p = .455$
H1*-H2*	$\chi^2(1) = 32.89$ $p < .001^{***}$	$\chi^2(1) = 0.64$ $p = .424$	$\chi^2(1) = 2.51$ $p = .113$	H < F $p = .001^{**}$	H < F $p < .001^{***}$
H2*-H4*	$\chi^2(1) = 1.93$ $p = .165$	$\chi^2(1) = 0.04$ $p = .841$	$\chi^2(1) = 0.03$ $p = .867$	H > F $p = .803$	H > F $p = .701$
H4*-H2kHz*	$\chi^2(1) = 2.02$ $p = .156$	$\chi^2(1) = 1.52$ $p = .218$	$\chi^2(1) = 3.54$ $p = .06$	H > F $p = .995$	H < F $p = .079$
H2kHz*-H5kHz	$\chi^2(1) = 23.93$. $p < .001^{***}$	$\chi^2(1) = 0.63$ $p = .428$	$\chi^2(1) = 1.35$ $p = .245$	H > F $p = .009^{**}$	H > F $p < .001^{***}$
H1*-A1*	$\chi^2(1) = 11.02$ $p < .001^{***}$	$\chi^2(1) = 2.22$ $p = .136$	$\chi^2(1) = 2.7$ $p = .101$	H < F $p = .487$	H < F $p = .002^{**}$
H1*-A2*	$\chi^2(1) = 5.64$ $p = .018^*$	$\chi^2(1) = 0.32$ $p = .574$	$\chi^2(1) = 4.94$ $p = .026^*$	H < F $p = .993$	H < F $p = .004^{**}$
H1*-A3*	$\chi^2(1) = 2.19$ $p = .139$	$\chi^2(1) = 0.72$ $p = .397$	$\chi^2(1) = 4.44$ $p = .035^*$	H > F $p = .985$	H < F $p = .044^*$
CPP	$\chi^2(1) = 0.88$ $p = .347$	$\chi^2(1) = 6.27$ $p = .012^*$	$\chi^2(1) = 0.55$ $p = .46$	H > F $p = .629$	H > F $p = .999$
HNR 0-500 Hz	$\chi^2(1) = 19.28$ $p < .001^{***}$	$\chi^2(1) = 7.88$ $p = .005^{**}$	$\chi^2(1) = 0.23$ $p = .628$	H < F $p = .011^*$	H < F $p = .002^{**}$
HNR 0-1500 Hz	$\chi^2(1) = 24.17$ $p < .001^{***}$	$\chi^2(1) = 5.48$ $p = .019^*$	$\chi^2(1) = 1.01$ $p = .314$	H < F $p = .007^{**}$	H < F $p < .001^{***}$
HNR 0-2500 Hz	$\chi^2(1) = 19.14$ $p < .001^{***}$	$\chi^2(1) = 6.96$ $p = .008^{**}$	$\chi^2(1) = 0.63$ $p = .426$	H < F $p = .021^*$	H < F $p < .001^{***}$
HNR 0-3500 Hz	$\chi^2(1) = 20.17$ $p < .001^{***}$	$\chi^2(1) = 6.35$ $p = .012^*$	$\chi^2(1) = 0.39$ $p = .534$	H < F $p = .011^*$	H < F $p < .001^{***}$

TABLE 1 : Significativité des effets principaux et interactions évalués dans les modèles linéaires mixtes appliqués à chacun des descripteurs, ainsi que de la différence entre hommes et femmes pour chacune des langues

4 Discussion

Dans cette étude, nous avons examiné 14 paramètres acoustiques liés à la voix chez les locuteurs coréens et français pour approfondir notre compréhension de l'effet du sexe et sa variation en fonction de la langue. Nos résultats montrent que la différence entre voix d'hommes et voix de femmes est conforme aux résultats attendus en raison du dimorphisme sexuel pour la majorité des variables liées à la source vocale prises en compte, mais pour certaines variables la différence entre sexes n'est observée que dans l'une des deux langues ou est plus importante dans l'une des langues.

Tout d'abord, pour H1*-A1*, H1*-A2* et H1*-A3*, la différence entre sexes apparaît uniquement chez les Français. L'effet du sexe a été trouvé dans H1*-A1* et H1*-A2* et l'interaction entre le sexe et la langue a été observée dans H1*-A2* et H1*-A3*. Cependant, d'après le test post-hoc, les différences entre sexes ne sont significatives que dans le groupe français avec une valeur plus importante chez les femmes. Puisque H1*-A1* reflète le quotient ouvert et que H1*-A2* et H1*-A3* représentent la fermeture glottique abrupte, nos résultats suggèrent que les voix des femmes françaises seraient plus soufflées que celles des hommes français. Ce sont des résultats allant dans le même sens que les études précédentes au niveau de l'effet du sexe (Hanson, Chuang, 1999 ; Iseli et al., 2007).

Avec une présence d'une fente glottique postérieure qui est fréquente chez les femmes, les valeurs $H1^*-A1^*$, $H1^*-A2^*$, $H1^*-A3^*$ sont attendues plus grandes dans le groupe féminin. En revanche, les voix des hommes et des femmes coréens n'étaient pas significativement différentes les unes les autres dans ces trois paramètres, ce qui montre des résultats différents des autres langues (Hanson, Chuang, 1999 ; Narra et al., 2015). Enfin, chez les Coréens, la pente spectrale ne semble pas différencier les voix d'hommes et de femmes, à l'exception des mesures $H1^*-H2^*$ et $H2kHz^*-H5kHz$.

Puis, pour f_0 (Hz), $H1^*-H2^*$, et $H2kHz^*-H5kHz$, la différence entre sexes ne varie pas en fonction de la langue. La f_0 moyenne (Hz) est plus élevée chez les femmes que chez les hommes dans les deux langues. Ce résultat n'est pas surprenant comme la f_0 est un des facteurs les plus discriminants entre la voix d'hommes et la voix de femmes. À part quelques exemples comme le dialecte Wu étudié par Rose (1991), la voix d'hommes et la voix de femmes se différencient par la f_0 basé sur le dimorphisme sexuel au niveau des plis vocaux. $H1^*-H2^*$ est également plus élevée chez les femmes que chez les hommes dans les deux langues. D'après les résultats déjà publiés (Henton, Bladon, 1985 ; Hanson, Chuang, 1999 ; Narra et al., 2015 ; Iseli et al., 2007 ; Klatt, Klatt, 1990), il existe dans $H1^*-H2^*$ une différence d'environ 3,1 dB à 8,4 dB entre les hommes et les femmes. Dans notre étude, la différence en $H1^*-H2^*$ entre sexes a été de 4,39 dB chez les Coréens et de 7,15 dB chez les Français. Les valeurs plus élevées chez les femmes que chez les hommes, indépendamment de la langue, suggèrent que la voix des femmes peut être perçue relativement plus soufflée que celle des hommes. Pour $H2kHz^*-H5kHz$, les valeurs sont plus élevées chez les hommes que chez les femmes dans les deux langues. Ce résultat est cohérent avec les résultats de Klatt et Klatt (1990) et Mendoza et al. (1996), selon lesquels la voix des femmes montre plus de bruits dans les hautes fréquences.

Ensuite, pour le HNR, la différence entre les sexes apparaît dans les deux langues, mais la valeur est plus élevée chez les Coréens dans les deux groupes de sexe. Quant à la différence entre sexes, la valeur HNR est plus grande chez les femmes. Ce sont des résultats inattendus. Comme la voix de femmes est considérée comme plus soufflée que la voix d'hommes en raison des différences morphologiques entre sexes au niveau des plis vocaux, nous nous attendions à ce que la valeur de l'HNR soit plus grande chez les hommes que chez les femmes. Chen et al. (2010) a montré le même résultat inattendu du HNR que cette étude, mais n'ont pas pu expliquer ce résultat.

Pour le CPP, seul un effet de la langue a été observé, avec une valeur plus grande chez les hommes et femmes coréens. Nous attendions que les hommes montrent une valeur plus importante de CPP que les femmes car il existe plus de composantes harmoniques dans la voix d'hommes et leur conduit vocal est plus long que celui des femmes (Choi et Choi, 2016). Ce résultat inattendu pourrait être lié à l'influence de la tâche de production mise en évidence en coréen par Choi et Choi (2016) avec une absence d'effet en lecture.

Ni la différence selon le sexe ni la variation selon la langue n'ont été observées dans la modulation f_0 , $H2^*-H4^*$, ni $H4^*-H2kHz^*$. Pour la modulation de la f_0 , cela peut être lié au matériel linguistique de l'expérience d'après Traunmüller et Eriksson (1994) : la différence de modulation de la f_0 est plus grande entre les sexes dans un discours spontané que dans la lecture de phrases. La littérature sur d'éventuelles différences entre sexes pour $H2^*-H4^*$ et $H4^*-H2kHz^*$ est par ailleurs limitée, mais on peut s'interroger sur la pertinence de ces descripteurs pour étudier la différence de voix entre les sexes et sa variation en fonction de la langue.

Dans l'ensemble, nos résultats indiquent que la voix des femmes serait plus soufflée par rapport à la voix des hommes chez les Coréens et les Français. Comme Södersten et al. (1990) l'ont montré, cela peut être dû aux différences physiologiques entre les hommes et les femmes : la fente glottique postérieure, la principale cause de voix soufflée, est observée plus fréquemment chez les femmes que

chez les hommes. Par conséquent, chez les femmes, le taux de contact des plis vocaux est inférieur à celui des hommes, qui provoque un flux d'air important fuyant à travers la glotte. Cela augmente le quotient ouvert chez les femmes lorsque les plis vocaux vibrent.

Nous constatons également que la différence de voix entre les hommes et les femmes francophones est plus marquée que celle entre les hommes et les femmes coréens. Cela suggère qu'il existe des caractéristiques de voix qui sont privilégiées chez les hommes et les femmes en fonction de la langue (et la culture). Par exemple, selon van Bezooijen (1995), une faible différence de f_0 entre hommes et femmes est préférée chez les Néerlandais alors qu'une grande différence est préférée chez les Japonais. D'après son interprétation, ces résultats sont dus à la tendance égalitaire de la culture néerlandaise qui considère les différences entre les sexes comme moins importantes que la culture japonaise. De ce point de vue, il est possible que les francophones dans notre étude perçoivent les voix des femmes comme psychologiquement plus attirantes quand elles sont plus soufflées par rapport à celles des hommes. Il est probable également que les caractéristiques phonétiques fortement utilisées pour distinguer la voix d'hommes de la voix de femmes puissent être différentes entre les Coréens et les Français. Même si les auditeurs dans de nombreuses langues ont tendance à dépendre le plus de la f_0 dans l'évaluation de l'attractivité des voix d'hommes et de femmes (van Bezooijen, 1995), Šebesta et al. (2017) ont constaté que l'attractivité masculine dans la population namibienne est évaluée par à quel point la voix est soufflée plutôt que par à quel point la f_0 est élevée. Chez les locuteurs coréens, il pourrait y avoir d'autres facteurs considérés comme plus importants que les mesures de qualité vocale prises en compte dans notre étude dans la distinction des voix masculines et féminines.

De plus, en ce qui concerne la périodicité et le bruit de la voix, il est confirmé que les voix des locuteurs coréens seraient plus périodiques que celles des locuteurs français. Cela peut être dû au fait que les locuteurs coréens modulent la qualité de la voix selon la situation de communication. D'après Idemaru et al. (2019), les locuteurs coréens ont une f_0 plus basse et une voix plus périodique dans un contexte où ils doivent être polis, par exemple pour s'adresser à leur professeur. Comme l'expérimentatrice et la plupart des participants de notre expérience ne se connaissaient pas beaucoup, il est possible que les locuteurs coréens aient eu recours à une voix plus périodique que celle des Français, ce qui pourrait expliquer les valeurs de HNR plus élevées relevées pour les Coréens, à la fois pour les hommes et les femmes.

Cette étude confirme enfin que les différences de voix entre hommes et femmes dépendent en partie de la langue parlée. Cependant, il sera nécessaire de l'approfondir en examinant des productions d'hommes et de femmes français et coréens sur un matériel linguistique plus varié afin de prendre les effets possibles de la tâche ou de la voyelle. Dans cette étude, une seule voyelle /a/ dans une phrase lue a été analysée. En effet, si Choi et Choi (2016) ont mis en évidence un effet significatif du sexe sur les valeurs de CPP pour toutes les voyelles coréennes /a/, /ɛ/, /i/, /o/, /u/, /ɯ/, /ʌ/, cet effet était dépendant de la voyelle et n'était pas observé pour /a/ lors de la lecture de phrases longues. Cela suggère qu'en coréen, la différence homme-femme de degré de périodicité de la voix peut varier non seulement selon la voyelle mais aussi en fonction du type de production. Narra et al. (2015) ont également observé chez les locuteurs malais des différences entre sexes sur H1-H2, H1-A1, H1-A2, et H1-A3 de magnitude variable en fonction de la voyelle avec un effet plus fort pour /a/. Iseli et al. (2007) ont identifié également des différences entre sexes sur f_0 et H1-H2 qui varient en fonction de la voyelle en anglais avec un effet plus fort sur /æ/. La variation de l'effet du sexe sur les paramètres de voix en fonction de la voyelle semble donc dépendante de la langue. Il sera donc nécessaire dans une étude ultérieure d'étendre les résultats de cette étude à des tâches et voyelles plus variées.

En outre, une étude perceptive nous permettra de comprendre si les paramètres acoustiques qui distinguent les voix d'hommes et de femmes chez les Coréens et Français sont exploités également

dans la représentation de la voix d'hommes et de femmes chez les auditeurs masculins et féminins des mêmes langues. Selon Smith & Patterson (2005), la f_0 et les formants jouent un rôle important dans la distinction entre sexes à la fois dans la production de la parole mais aussi dans le processus de perception. Si nous arrivons aussi à trouver des variables acoustiques liées à la qualité vocale que les locuteurs et les auditeurs considèrent ensemble comme importantes, cela nous rapprochera de la compréhension de la représentation dans chaque langue des voix d'hommes et des voix de femmes.

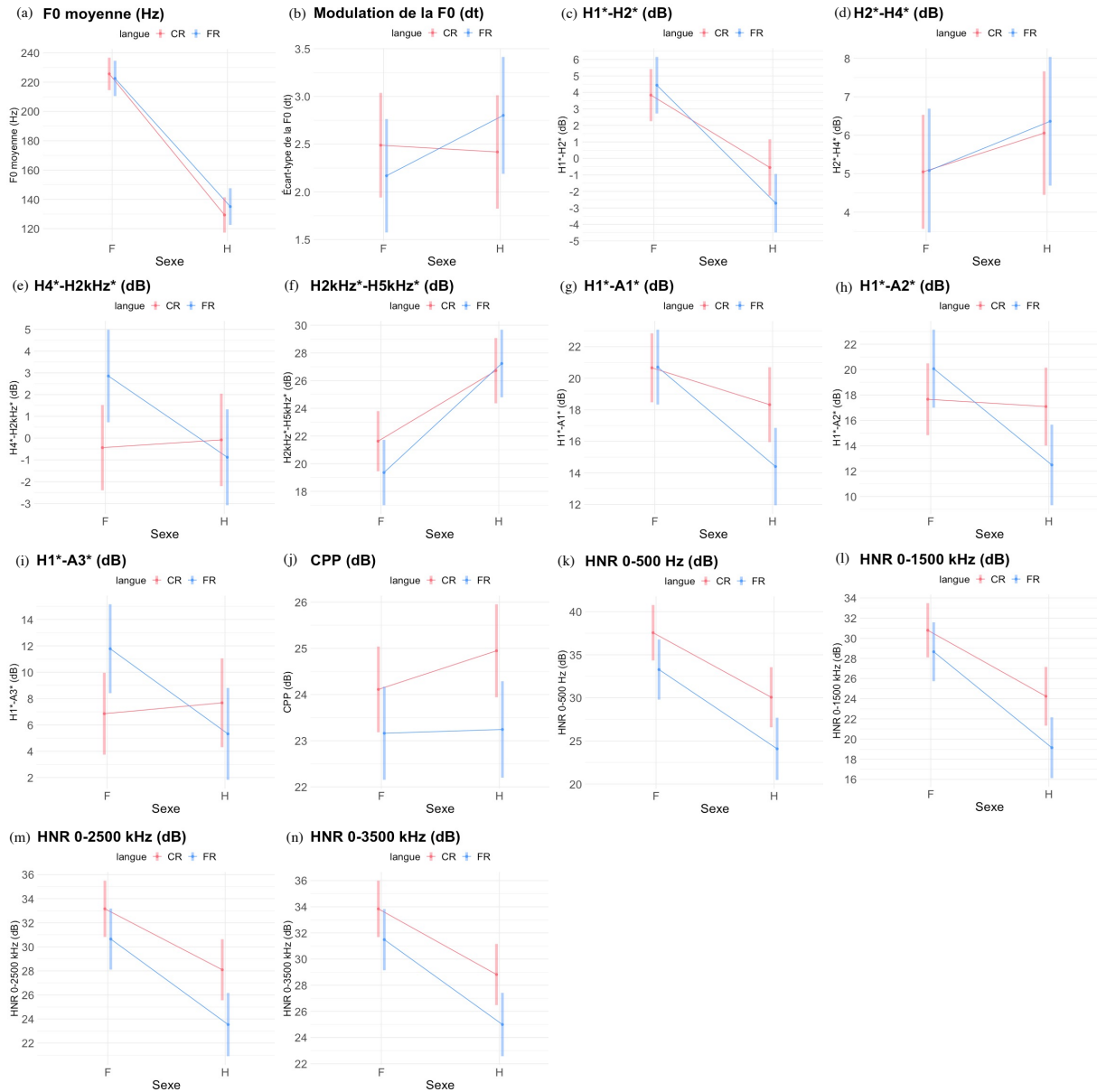


FIGURE 1 : Comparaisons entre sexes (F : femme ; H : homme) et entre langues parlées (CR : coréen ; FR : français) pour les paramètres acoustiques liés à la source vocale : (a) f_0 (Hz), (b) modulation de f_0 (demi-tons), (c) $H1^*-H2^*$, (d) $H2^*-H4^*$, (e) $H4^*-H2k^*$, (f) $H2kHz^*-H5kHz^*$, (g) $H1^*-A1^*$, (h) $H1^*-A2^*$, (i) $H1^*-A3^*$, (j) HNR 0-500 Hz, (k) HNR 0-1500 Hz, (l) HNR 0-2500 Hz, (m) HNR 0-3500 Hz, (n) CPP. Mesures en dB sauf pour (a) et (b).

Références

- BATES D., MÄCHLER M., BOLKER B., WALKER S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1), 1–48.
- BOË L.-J., CONTINI M., RAKOTOFIRINGA H. (1975). Étude statistique de la fréquence laryngienne. *Phonetica* 32(1), 1-23.
- BOERSMA P., WEENINK D. (2022). Praat: doing phonetics by computer [Computer program]. Version 6.2.07, récupérée le 28 janvier 2022 from <http://www.praat.org/>.
- CHEN G., FENG X., SHUE, Y.-L., ALWAN A. (2010). On using voice source measures in automatic gender classification of children's speech. Actes de 11th Annual Conference of the International Speech Communication Association (INTERSPEECH 2010), Makuhari, Chiba, Japan, 26-30 septembre 2010, 673-676.
- CHENG A. (2019). Cross-linguistic f0 differences in bilingual speakers of English and Korean. *The Journal of the Acoustical Society of America* 127(2), EL67-EL73.
- CHOI S., CHOI C. (2016). The Effect of Gender and Speech Task on Cepstral- and Spectral-Measures of Korean Normal Speakers. *Audiology and Speech Research* 12(3), 157-163.
- DE KROM G. (1993). A Cepstrum-Based Technique for Determining a Harmonics-to-Noise Ratio in Speech Signals. *Journal of Speech and Hearing Research* 36, 254-266.
- IDEMARU K., WINTER B., BROWN L. (2019). Cross-cultural multimodal politeness: The phonetics of Japanese deferential speech in comparison to Korean. *Intercultural Pragmatics* 16(5), 517-555.
- ISELI M., SHUE Y., ALWAN A. (2007). Age, sex, and vowel dependencies of acoustic measures related to the voice source. *The Journal of the Acoustical Society of America* 121, 2283-2295.
- JOHNSON K., BABEL M., FUHRMAN R. (2020). Bilingual acoustic voice variation is similarly structured across languages.
- KAHN J., AUDIBERT N., BONASTRE J. F., ROSSATO S. (2011). Inter and Intra-Speaker Variability in French: An Analysis of Oral Vowels and Its Implication for Automatic Speaker Verification. Actes de 17th International Congress of Phonetic Sciences (ICPhS XVII) Hong Kong, China, 17-21 août 2011, 1002-1005.
- KAWAHARA H., MASUDA-KATSUSE I., DE CHEVEIGNÉ A. (1999). Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds. *Speech Communication* 27, 187-207.
- KLATT D., KLATT L. (1990). Analysis, synthesis, and perception of voice quality variations. *Journal of Acoustical Society of America* 87(2), 820-857.
- KREIMAN J., GARELLEK M. (2011). Perceptual importance of the voice source spectrum from H2 to 2 kHz. *Journal of the Acoustical Society of America* 130(4), 2570-2570.
- MENDOZA E., VALENCIA N., MUÑOZ J., TRUJILLO H. (1996). Differences in voice quality between men and women: Use of the long-term average spectrum (LTAS). *Journal of Voice* 10(1), 59-66.
- NARRA M., DODDERI T., ANU C., VARGHESE S., DATTATREYA T. (2015). Harmonic Amplitude Measures to Note Gender Difference. *Advances in Life Science and Technology* 31, 17-24.

- HANSON H. (1997). Glottal characteristics of female speakers: acoustic correlates. *Journal of the Acoustical Society of America* 101(1), 466–481.
- HANSON H., CHUANG E. (1999). Glottal characteristics of male speakers: Acoustic correlates and comparison with female data. *The Journal of the Acoustical Society of America* 106(2), 1064-1077.
- HENTON C., BLADON R. (1985). Breathiness in normal female speech: inefficiency versus desirability. *Language and Communication* 5(3), 221-227.
- HILLENBRAND J., CLEVELAND R., ERICKSON R. (1994). Acoustic Correlates of Breathy Vocal Quality. *Journal of Speech and Hearing Research* 37, 769-778.
- HILLENBRAND J., HOUDE R. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech and Hearing Research* 39(2), 311-321.
- PÉPIOT E., ARNOLD A. (2020). Cross-Gender Differences in English/French Bilingual Speakers: A Multiparametric Study. *Perceptual and Motor Skills* 128(1), 153-177.
- ROSE P. (1991). How effective are long term mean and standard deviation as normalisation parameters for tonal fundamental frequency? *Speech Communication* 10, 229-247.
- ŠEBESTA P., KLEISNER K., TUREČEK P., KOČNAR T., AKOKO R., TŘEBICKÝ V., HAVLÍČEK J. (2017). Voices of Africa: acoustic predictors of human male vocal attractiveness. *Animal Behaviour*, 127, 205–211.
- SHUE Y.-L., KEATING P., VICENIK C., YU K., (2009). VoiceSauce: A Program for Voice Analysis. *Actes de 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong, China, 17-21 août 2011*, 1846-1849.
- SJÖLANDER K. (2004). Snack sound toolkit. KTH Stockholm, Sweden. <http://www.speech.kth.se/snack>.
- SMITH D., PATTERSON R. (2005). The interaction of glottal-pulse rate and vocal-tract length in judgements of speaker size, sex, and age. *The Journal of the Acoustical Society of America* 118(5), 3177-3186.
- Södersten M., Lindestad P.-Å. (1990). Glottal Closure and Perceived Breathiness During Phonation in Normally Speaking Subjects. *Journal of Speech Language and Hearing Research*, 33(3), 601-611.
- STEVENS K., HANSON H. (1995). Classification of glottal vibration from acoustic measurements. Dans O. Fujimura, M. Hirano (éds.), *Vocal fold physiology: Voice quality control*. San Diego, CA : Singular Publishing Group Inc. 146-170.
- TITZE I. R. (1989). Physiologic and acoustic differences between male and female voices. *Journal of the Acoustical Society of America* 85(4), 1699-1707.
- TRAUNMÜLLER H., ERIKSSON A. (1993). The frequency range of the voice fundamental in the speech of male and female adults. Unpublished manuscript. Stockholm.
- VAN BEZOOIJEN R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*, 38(3), 253–265.
- YOON T., KANG Y. (2014). Monophthong Analysis on a Large-scale Speech Corpus of Read-Style Korean. *Phonetics and Speech Sciences* 6(3), 139-145.