



HAL
open science

Variation interculturelle de la perception du spectre masculin-féminin : indexation de la voix genrée en France et aux Etats-Unis

Maria Candea, Leann Brown

► To cite this version:

Maria Candea, Leann Brown. Variation interculturelle de la perception du spectre masculin-féminin : indexation de la voix genrée en France et aux Etats-Unis. Vinay Swamy; Louisa Mackenzie. Devenir non-binaire en français contemporain, Le Manuscrit, 2022, 9782304052428. hal-03573714

HAL Id: hal-03573714

<https://hal.science/hal-03573714>

Submitted on 14 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Titre: Variation interculturelle de la perception du spectre masculin-féminin : indexation de la voix genrée en France et aux Etats-Unis¹

Maria CANDEA, LeAnn BROWN

Mots-clés: masculinité, féminité, non-binarité, perception, français, anglais américain, indexicalité

Abstract

Expanding a gender paradigm beyond binary categories requires new perceptual tools, which may be implicit or explicit. In this paper we test the use of a masculinity-femininity scale with an “overlap” zone in the middle that is not explicitly delimited and ask for listeners’ gender assessments based on short speech samples produced by cis women, cis men and non-binary people from the USA. As gender is a sociocultural artefact and the social “footprint” of gender non-conforming or non-binary individuals is greater in North America than in France (for example), how listeners use the scale potentially indicates 1) something about how they view gender and 2) something about the cues they use to index gender. Statistical analyses suggest that French and American listeners share the same indexations for the vocal cues in their masculinity-femininity assessments, but there is evidence that there are some interesting differences, particularly concerning the strategic use of the entire scale versus the endpoints.

Introduction

La linguistique considère fondamentalement les langues comme des systèmes de signes qui permettent aux humains de produire du sens. La sociolinguistique prend appui sur ce postulat et s’intéresse tout particulièrement aux variations à tous les niveaux que connaît la langue à travers les actes singuliers de son utilisation, en contexte, par des individus singuliers. Plus précisément, la sociolinguistique s’intéresse à l’organisation, la catégorisation et la hiérarchisation sociale de toutes les variations dans une langue et cela peut concerner les phrases, les mots, les parties de mots, et jusqu’à la microvariation des sons et des voix produites. Son objectif est de saisir ce que les individus font avec le langage en contexte, non seulement comment ils produisent du sens mais également comment ils perçoivent le sens produit par autrui.

La recherche sociophonétique que nous présentons ici se consacre à l’étude de la variation dans le domaine de la production et de la perception de la voix, car la voix fait partie des ressources sémiotiques utilisées par les êtres humains.

¹ Ce travail a été réalisé grâce à un financement ANR [projet NoBiPho, ANR-18-CE26-0006-01]. Les stimuli proviennent du *Corpus Gender Spectrum Speech*, réalisé grâce au financement Amidex [SIRL/GSS https://hdl.handle.net/11403/gender_spectrum_speech].

A la différence de la morphologie, de la syntaxe ou des discours produits dans une langue, la voix que nous émettons est en partie dépendante de paramètres individuels physiologiques : longueur du conduit vocal, taille des plis vocaux (fréquemment nommés « cordes vocales »), forme des cavités de résonance. C'est ce qui explique que nous reconnaissons une personne au timbre de sa voix et que nous avons des représentations convergentes de ce qu'est une « voix féminine » et une « voix masculine ». Cela induit l'illusion d'une voix qui serait strictement déterminée par le corps (comme la couleur des yeux) et l'illusion d'un dimorphisme de l'appareil phonatoire humain (féminin/masculin).

Or, contrairement à ce qui est souvent sous-entendu dans les discours en phonétique (voir Arnold 2016 pour une analyse critique), l'appareil phonatoire humain ne présente pas de dimorphisme : la variation des tailles et des formes se distribue sur un continuum, et les voix aussi. C'est ce qui explique que nous ne pouvons pas toujours deviner le genre d'une personne inconnue à partir de sa seule voix. En outre, la production de la voix n'est pas uniquement déterminée par la physiologie : c'est aussi une pratique culturelle. Les stéréotypes partagés dans une langue-culture incitent les locutaires² à ajuster leurs pratiques vocales en fonction des attentes de leur entourage. C'est ce qui explique que les personnes bilingues, socialisées dans deux cultures différentes, pratiquent des voix différentes (voir par exemple Benoist-Lucy & Pillot-Loiseau 2013 pour une étude de cas sur la pratique du *vocal fry* en anglais américain et en français de France). Il en est de même pour le fait de genrer sa voix : selon les cultures et les époques, nous mobilisons notre appareil phonatoire diversement pour rendre notre genre lisible. Ce processus est d'autant plus explicite et facile à observer dans le cas des personnes qui effectuent une transition de genre : le manuel de Mills & Stoneham (2017) rassemble des techniques disponibles pour les personnes trans et non-binaires qui souhaitent féminiser, masculiniser ou neutraliser leur voix.

La présente étude, partie intégrante d'un projet plus vaste portant sur la production et la perception de la voix genrée ou neutre, constitue une exploration des indices que les auditaires³ peuvent percevoir pour catégoriser une voix sur une échelle des genres. Le projet inclut des enregistrements de lectures et entretiens avec des personnes (femmes et hommes) cisgenres et non-binaires, anglophones vivant aux Etats-Unis ou francophones vivant en France. Les personnes non-binaires sont provisoirement définies ici comme des individus qui ne s'identifient jamais ou qui ne s'identifient pas exclusivement comme des femmes/femelles/féminines ou des hommes/mâles/masculins (Brown,

² Dans cet article nous avons choisi d'employer les formes néologiques « auditaires » et « locutaires » proposées par Alpheratz (voir interview dans ce volume) pour neutraliser l'opposition genrée des formes « auditeurs/auditrices » et « locuteurs/locutrices » et rendre ces noms épicènes.

³ Voir note 2.

Candea, Reid-Collins, Abbou, & German 2019 ; Skinner 2017). Des extraits de ces enregistrements recueillis sont inclus dans des protocoles de perception de la parole pour vérifier, dans une démarche interculturelle, de quelle manière la voix produit du sens dans des interactions suscitées de manière expérimentale.

Selon Podesva & Callier, “[the fact that] voice quality patterns vary widely across individuals and communities suggests that voice quality can serve as an important resource in the construction of identity” (2015 : 174). Parmi les personnes non-binaires de notre corpus, certaines utilisent en effet la voix comme ressource pour performer leur non-binarité de genre, ou leur non-appartenance au genre féminin ou masculin, tandis que d’autres n’utilisent pas leur voix mais d’autres ressources sémiotiques. Notre intérêt s’est porté principalement ici sur la possibilité de tester empiriquement la possibilité de la validation sociale d’une voix non-binaire. Autrement dit : puisque des personnes non-binaires ajustent leurs pratiques vocales de manière à se rendre lisibles comme personnes non-binaires, notamment – pour des personnes AFAB⁴ – en diminuant les valeurs des paramètres qui indexent la féminité (Brown & Pillot-Loiseau, soumis), ces pratiques ont-elles une chance d’être interprétées en ce sens par des personnes du tout venant ?

Les stimuli présentés dans cette étude proviennent de personnes nord-américaines⁵.

L’objectif poursuivi ici n’est pas seulement d’identifier s’il est possible de parler de catégorisation non-binaire des voix humaines, mais également de tester le poids du facteur interculturel en comparant les résultats perceptifs recueillis auprès d’échantillons de personnes vivant en France et aux États-Unis.

Hypothèses et cadres théoriques

Nous faisons l’hypothèse que les différences culturelles entre la France et les Etats-Unis en matière de visibilité non-binaire et de sensibilisation/acceptation sociale peuvent entraîner des évaluations genrées différentes lors de l’écoute de brefs extraits d’enregistrements de voix humaines. Bien entendu, les deux groupes d’auditeurs sollicités utilisent culturellement la catégorisation binaire (femmes et hommes) pour identifier le genre d’un être humain; cependant, nous avons l’impression d’un véritable changement de fond dans le paysage états-unien en matière de saillance de la multiplication des catégories de genre compte tenu de l’augmentation des options de genre qu’il est

⁴ Assigned Female At Birth, « assignée fille à la naissance ».

⁵ Le groupe français du projet en cours a été enregistré plus tard et les études sont en phase initiale.

possible de déclarer officiellement pour les certificats de naissance et les cartes d'identité dans plusieurs Etats du pays, l'augmentation des personnes qui s'identifient comme non-binaires (Nolan, Kuhner & Dy 2019; James, Herman, Rankin, Keisling, Mottet, & Anafi 2016), parmi lesquelles des célébrités, comme Miley Cyrus, Jonathon Van Ness, Asia Kate Dillon, Elliot Page et Sam Smith, ainsi que la présence croissante de personnes non-binaires dans les médias grand public et dans tous les domaines de la création audiovisuelle. Le contexte semble pour le moment différent en France, malgré une légère augmentation des dossiers dans la presse écrite et audiovisuelle consacrés à la non-binarité et malgré l'apparition de quelques rares célébrités ouvertement non-binaires dans le paysage musical, comme Chris de *Christine and the Queens*.

L'hypothèse contraire serait que malgré les quelques différences culturelles entre la France et les Etats-Unis en matière de notoriété de la notion de non-binarité du genre, les réponses obtenues seront rigoureusement identiques car elles reposeront sur des modèles de genre, cognitifs et indexicaux, identiques ou très proches.

Pour tester nos hypothèses, nous avons mené une expérience de perception comparée de brefs extraits audio de voix ; celles-ci doivent être catégorisées sur une échelle de masculinité-féminité. Le sous-corpus présenté ici est constitué d'enregistrements de lectures produites en anglais américain par des personnes vivant aux Etats-Unis ; ces lectures sont écoutées et évaluées par deux groupes de quelques dizaines d'auditeurs vivant aux Etats-Unis et en France.

Notre projet tente d'isoler ce qui pourrait éventuellement constituer un indice saillant de non-binarité de genre (ou de non-conformité au système binaire du genre) dans la perception des voix humaines. Des résultats convergents pourraient être interprétés au prisme de la théorie de l'indexicalité (Ochs 1992 ; Eckert 2008 ; Silverstein 2003) qui suggère que les variantes de prononciation peuvent être indexées ou associées à une ou plusieurs significations comme des postures, des attitudes, des identités, des affiliations.

Les protocoles expérimentaux pour répondre aux questions que nous nous posons peuvent mobiliser la catégorisation explicite des personnes : il s'agit d'associer un genre féminin, masculin ou non-conformiste à des personnes en se fondant uniquement sur l'écoute de leur voix. Ce type de demande explicite est pratiqué, par exemple, par Lee, Dutton & Ram (2010) et nous l'envisageons dans une autre opération de notre projet. D'autres protocoles peuvent demander des évaluations de la « masculinité » et de la « féminité » (Coleman 1973 ; Munson 2007), ou de dire s'il s'agit de voix d'hommes « moins masculines » ou « plus masculines » (Avery & Liss 1996). Enfin, d'autres protocoles

encore peuvent mobiliser la catégorisation implicite, comme c'est le cas dans la présente étude, où l'interface de test propose l'utilisation d'une échelle continue à 15 points entre deux pôles (féminin et masculin), comportant *de facto* une zone médiane, non délimitée explicitement et non focalisée dans la question posée. La façon dont les personnes sollicitées comme auditaires vont utiliser les 15 points de l'échelle va nous donner des informations à la fois sur la façon dont elles construisent leurs catégories de genre et sur les marqueurs vocaux dont elles se servent pour indexer le genre.

La littérature disponible montre qu'il y a des marqueurs phonétiques et vocaux qui indexent le genre – y compris non-binaire pour certaines études – et aussi l'orientation sexuelle. C'est le cas principalement de la valeur moyenne de fréquence fondamentale (F0), ce qui correspond à la hauteur de la voix mesurée en Hertz, (Coleman 1973; Linville, 1998; Oates & Dacakis 1983; Smyth, Jacobs, & Rogers 2003), la variance en valeur absolue de la fréquence fondamentale (Davies et al. 2015; Hancock et al. 2014; McConnell-Ginet 1978), une combinaison des paramètres de l'intonation (Schmid & Bradley 2019, Corwin 2019) ainsi que la variation de pente spectrale (Mendoza et al. 1996; Klatt & Klatt 1990). Nous détaillons ces paramètres ci-dessous.

Bien qu'il y ait des tendances globales communes à toutes les cultures, des différences importantes de valeur ont été attestées dans la littérature (Johnson 2005). Par exemple, pour l'anglais la fourchette moyenne des voix masculines se situe entre 100-110 Hz, et celle des voix féminines entre 200-220 Hz (Simpson 2009). Pour le français, Pépiot (2014) avance une moyenne à 133 Hz pour les hommes français, et à 234 Hz pour les femmes françaises, qui constituent des valeurs plus hautes que celles qu'il calcule en comparaison pour les Américains du même groupe d'âge (119 Hz pour les hommes, 210 Hz pour les femmes).

La variance de la fréquence fondamentale peut être mesurée de différentes manières. Pépiot (2014), par exemple, calcule la plage de variation F0 max-F0 min, en convertissant les Hertz en demi-tons en raison de la plus grande pertinence perceptive de cette mesure : selon ses calculs, les Françaises présentent la plus grande variance de hauteur en demi-tons (l'écart entre les valeurs les plus basses et les valeurs les plus hautes est très important), tandis que les trois autres groupes (hommes français, femmes et hommes états-uniens) présentent des valeurs de variance de hauteur très similaires, moins amples. L'analyse statistique indique que la variance de hauteur des femmes françaises est significativement plus importante que celle des hommes français. Aucune différence statistique n'a été constatée pour les Américain-es selon le genre pour les données mesurées en demi-tons.

La pente spectrale mesure la richesse du timbre. Dans le logiciel Praat, la pente spectrale mesure la différence d'énergie (dB) entre la bande de fréquences élevées (2000-4000 Hz) et la bande de basses fréquences (0-2000 Hz). Ces bandes sont visibles dans le spectre moyenné à long terme (*long term average spectrum* - désormais LTAS) de la voix, c'est-à-dire le calcul de l'énergie moyenne dans la voix sur une longue période phonétique, plus d'une minute de parole par exemple. Une pente très négative signifie que la voix a peu d'énergie dans la bande supérieure par rapport à la bande inférieure. Si la pente est faiblement négative, cela signifie que l'amplitude de la bande des fréquences supérieures égale presque celle des fréquences basses : la voix a alors un timbre plus riche, plus brillant. Différentes études ont montré une corrélation positive entre le degré de « richesse » ou « brillance » du timbre perçu et l'importance du renforcement d'énergie entre 2000 et 4000 Hz pour la voix parlée (Smith et al., 2005) et chantée (Omori, Kacker, Carroll, Riley, & Blaugrund 1996; Pillot & Vaissière 2007).

Selon une étude en préparation (Brown & Pillot-Loiseau *soumis*) les personnes non-binaires de notre corpus présentent en moyenne une pente spectrale moins négative que les personnes cisgenres dans la bande de fréquence 2000-4000 Hz, donc avec une énergie plus forte et plus uniforme, ce qui montre la production d'une qualité de voix plus brillante, plus claire, d'un timbre plus riche. Nous faisons l'hypothèse que la richesse du timbre de la voix pourrait être un bon candidat pour devenir une marque exploitable dans le système indexical du genre. Comme elle n'est pas mobilisée pour indexer le masculin ou le féminin, la voix brillante pourrait peut-être indexer la non-binarité ou la non-conformité de genre. Cette suggestion est confortée par les conseils qui sont donnés explicitement aux personnes non binaires qui souhaitent travailler leur voix pour la « dégenrer » (Mills & Stoneham 2017 : 71) et sont incitées à modifier leur résonance pour obtenir un timbre plus riche.

D'autres marqueurs peuvent se voir insérer dans des systèmes indexicaux différents selon les langues-cultures. Par exemple, la phonation en *creaky voice* (voix craquée, laissant vibrer ses plis vocaux de façon irrégulière et peu tendue, rappelant le coassement d'un crapaud) a été décrite comme étant un index de masculinité au Royaume-Uni (Henton & Bladon 1988), un index associé aux jeunes femmes aux Etats-Unis (Yuasa 2010), et sans aucune capacité à indexer le genre en France (Benoist-Lucy & Pillot-Loiseau 2013).

C'est la raison pour laquelle nous avons pensé qu'il était pertinent de mener le même test, avec les mêmes échantillons de voix, auprès d'un public français et états-unien ; nous nous sommes intéressées à trois paramètres acoustiques : les valeurs de la F0 (la fréquence fondamentale), l'amplitude de sa plage de variation mélodique et la valeur de pente spectrale (ou richesse du timbre).

Méthode expérimentale et matériaux

Les extraits de voix, d'une durée de 3 à 4 secondes, proviennent de 20 personnes enregistrées dans le cadre du *Gender Spectrum Speech Corpus* (Brown et al. 2019)⁶ parmi lesquelles : 8 femmes cisgenres, 7 hommes cisgenres et 5 personnes non-binaires. Le corpus comporte des lectures et des entretiens. Dans cette étude, nous avons utilisé des échantillons de la partie « lecture » : le contenu des extraits audio donnés comme stimuli est toujours le même pour favoriser la focalisation de l'attention des auditeurs sur la voix. Dans nos stimuli toutes les personnes lisent la même phrase (*But that only makes them even more endearing to their owners*) au sein d'un paragraphe ; la qualité de l'enregistrement est toujours la même, sans bruits ni chevauchement de parole, et l'intensité a été normalisée à 70dB pour ne pas introduire de disparités entre les stimuli (détails dans Brown & Pillot-Loiseau *soumis*).

L'interface de l'enquête a été diffusée en ligne, en français et en anglais, dans le même format. Les participants ont été invités à donner leur consentement pour l'utilisation de leurs réponses, à lire les instructions et à évaluer deux extraits d'entraînement pour se familiariser avec la tâche. Les extraits étaient programmés pour s'afficher en ordre aléatoire et les auditeurs devaient évaluer chaque voix en cochant une position sur une échelle. L'extrémité gauche de l'échelle était marquée comme « très masculine », l'extrémité droite comme « très féminine », et la position centrale de l'échelle était marquée « 0 ». Ultérieurement nous avons attribué aux cinq points de l'extrémité gauche de l'échelle le codage comme *zone masculine*, et aux cinq points de l'extrémité droite de l'échelle le codage *zone féminine*. Le point zéro ainsi que les deux points situés d'un côté ou de l'autre du zéro constituent une zone codée comme *zone médiane*.

Le protocole permet de poser plusieurs questions. Outre le fait de tester si les résultats obtenus auprès d'un échantillon d'auditeurs des Etats-Unis sont similaires ou différents de ceux obtenus auprès d'un échantillon de France, cela permet d'observer les stratégies de réponse de ces auditeurs confrontés à l'échelle de 15 points divisée en 3 zones et de déterminer, notamment, s'ils utilisent avec prédilection les zones extrêmes, correspondant à un marquage binaire du genre, ou bien s'ils utilisent également la *zone médiane*. Enfin, au cas où la zone médiane de l'échelle est sélectionnée pour évaluer certaines voix, le protocole permet également de tester si cela arrive plus souvent pour les

⁶ LeAnn Brown, Maria Candea, James Sneed German, Oriana Reid-Collins, Tim Mahrt, Julie Abbou, Aleksandra Chikulaeva, *Gender Spectrum Speech* https://hdl.handle.net/11403/gender_spectrum_speech.

voix produites par les personnes non-binaires ou non, et de déterminer s'il est éventuellement possible de constater l'émergence d'une correspondance entre une auto-identification comme personne non-binaire, la production d'une voix non genrée ou peu genrée et la catégorisation perceptive de cette voix dans la zone médiane de l'échelle du genre.

Pour la France, les participants ont été recrutés via Twitter et Facebook⁷. Le recrutement des participants états-uniens a été fait par la plateforme Amazon Mechanical Turks contre une petite compensation pécuniaire. Une liste de questions démographiques a été présentée à la fin de la tâche expérimentale, demandant des données sur l'âge, le genre, la langue maternelle et l'éducation des participants.

La composition des deux échantillons était très similaire pour le critère de l'âge : 62% des 47 Français, et 60% des 40 Américains qui ont répondu à l'enquête étaient âgés de 18 à 40 ans. L'échantillon français comportait plus de femmes (68%) que l'échantillon américain (50%). Les deux échantillons présentaient également une certaine différence en termes de niveaux d'éducation des participants, car 51% des Français déclarent détenir au moins un diplôme de master contre 7,5% des Américains, et en parallèle, 25% des Américains déclarent la fin du lycée comme dernier diplôme alors que seulement 2% des Français sont dans ce cas. Néanmoins, un pourcentage similaire dans les deux groupes déclare comme dernier diplôme le degré immédiatement supérieur au baccalauréat : 45% des Français contre 47,5% des Américains.

Les mesures de la F0 en valeur absolue, sa variance et la pente spectrale ont été calculées avec le logiciel Praat (Boersma & Weenink 2019) en Hz ainsi qu'en demi-tons par la seconde autrice (L. Brown). La pente spectrale a été calculée avec le même logiciel à l'aide d'un script pour le LTAS avec correction de hauteur (Pettrossi, Audibert, & Crevier-Buchman 2017).

Plutôt que de traiter ces valeurs comme continues, elles ont été catégorisées en niveaux fondés sur la distribution des valeurs et sur les indexations déjà attestées dans la littérature. La F0 a été divisée en trois niveaux sur la base d'une interprétation des différentes fourchettes données pour les voix masculines, féminines et neutres ou ambiguës attestées dans la littérature (Adler et al. 2012 ; Davies

⁷ En particulier à partir du compte de la première autrice (M. Candea) de l'article ainsi que de celui de Clémence Bobkiewicz, étudiante en master de sciences du langage à l'Université Sorbonne nouvelle, dans le cadre de sa recherche de première année. Les messages ont été postés en direction de communautés très larges de profils divers et non dans les réseaux amicaux ou intimes.

Papp & Antoni 2015 ; Hancock, Colton, & Douglas 2014 ; Irwig, Childs, & Hancock 2017)⁸. La plage de variance de la F0 ainsi que les valeurs de pente spectrale ont été divisées en deux niveaux (variation mélodique forte, variation mélodique faible ; pente spectrale forte, pente spectrale faible) en prenant leur valeur médiane comme repère de séparation dans la mesure où les études antérieures mentionnent seulement des oppositions binaires.

Résultats et discussion

Les résultats globaux montrent que le genre des locutaires produit des effets sur les jugements exprimés, alors que ces jugements s'opèrent par la simple écoute d'un extrait audio de 3 à 4 secondes. Les extraits voix d'hommes cisgenres ont globalement été associées avec des réponses classées dans la zone masculine de l'échelle et marginalement avec des réponses en zone médiane ou féminine. Les extraits des voix de femmes cisgenres ont globalement été associées avec des réponses classées dans la zone féminine de l'échelle, marginalement en zone médiane et jamais en zone masculine. Les extraits des voix des personnes non-binaires (qui sont dans notre échantillon des personnes assignées femmes à la naissance) ont été associées avec des réponses dispersées dans les trois zones de l'échelle, féminine, médiane et masculine, même si les réponses en zone féminine étaient relativement plus nombreuses.

Si on rassemble toutes les réponses regroupées par zones de l'échelle, il n'y a pas d'effet global de la langue-culture des auditeurs de France ou des Etats-Unis (le test de Mann-Whitney U indique une probabilité élevée que les différences soient dues au hasard⁹). Nous observons donc globalement une bonne convergence lorsqu'il s'agit de distinguer chacun de ces trois groupes (Figure 1). Dans le détail, nous avons néanmoins identifié quelques divergences.

⁸ 145-185 Hz constitue la zone médiane (6.49-10.20 en demi-tons), la plage haute commence au-dessus de 185 Hz et la plage basse se situe en dessous de 145 Hz.

⁹ Z=-1,232, p=0,218.

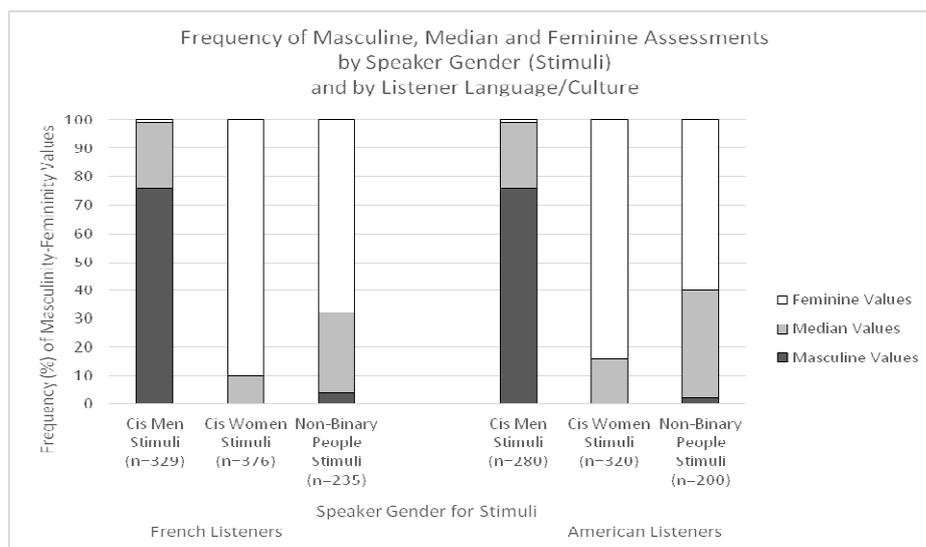


Figure 1 : Fréquence des évaluations situées en zone masculine, féminine ou médiane ventilées par genre des locutaires et langue-culture des auditeurs

Notre question de recherche principale visait à établir si les auditeurs américain-es et français-es utilisent la *zone médiane* de l'échelle de masculinité-féminité (la zone contenant le point « zéro » et les deux positions situées à sa gauche et à sa droite) de la même manière. « De la même manière » doit se comprendre d'abord dans l'absolu, et ensuite en particulier pour évaluer les stimuli produits par les hommes cisgenres « H », les stimuli produits par les femmes cisgenres « F » et les stimuli produits par les personnes non-binaires « NB ».

Dans l'absolu, tous stimuli confondus, nous avons décelé deux stratégies de réponse possibles. La « stratégie binaire » consiste à ne jamais utiliser la zone médiane de l'échelle et toujours évaluer les voix en sélectionnant une des 5 positions des zones masculine ou féminine. Selon notre interprétation, cette stratégie laisse penser que l'utilisation d'un modèle binaire de genre incite les auditeurs à favoriser un choix clair pour chaque voix, même lorsque les voix sont acoustiquement ambiguës. La stratégie que nous appelons « non-binaire » consiste à utiliser toute l'échelle, avec ses trois zones, pour situer les évaluations des différentes voix entendues. Selon notre interprétation, cette stratégie montre que non seulement les auditeurs acceptent aisément l'idée d'une échelle continue mais en plus ils utilisent volontiers les positions situées autour de zéro qui constitue le point neutre, le point de bascule ni féminin ni masculin ; cela laisse penser que les auditeurs sont capables de mobiliser un modèle de genre qui peut permettre une certaine forme, même implicite, de sortie de la stricte binarité.

Nos résultats indiquent que la majorité des auditeurs des deux groupes linguistiques (87%) utilisent toute l'échelle, la zone médiane tout comme les zones des deux extrémités (stratégie non-binaire). Une minorité néanmoins, (13%), n'utilise jamais la zone médiane (stratégie binaire), et la majorité parmi eux appartient au groupe français (64%, alors que les auditeurs français constituent 54% du nombre total d'auditeurs). En raison de la petite taille de ce groupe, nous n'avons pas exploré sa ventilation par groupes sociaux tels que l'âge et l'éducation. La ventilation par genre des auditeurs n'a donné aucun résultat significatif.

Des analyses statistiques par équations d'estimation généralisées (GEE) avec des matrices échangeables, en considérant chaque participant comme une mesure répétée, ont été menées sur chaque type de stimulus séparément pour tester les régularités dans les réponses et la probabilité qu'elles soient dues au hasard ou qu'elles soient significatives. Nous avons effectué trois régressions logistiques binaires pour les trois catégories de stimuli constituées¹⁰ dans le respect de l'auto-identification des locuteurs enregistrés.

Cette analyse¹¹ indique si la langue-culture de l'auditeur ou le paramètre acoustique pris en compte a un effet sur les évaluations perceptives, ainsi que s'il y a une interaction significative entre la langue-culture de l'auditeur et chaque paramètre acoustique considéré. S'il y a des interactions bidirectionnelles significatives, cela indiquera des différences de réponses selon la langue-culture des personnes qui ont répondu ; s'il n'y a pas d'interactions bidirectionnelles significatives mais que le paramètre acoustique (F0, variance mélodique, pente spectrale) est significatif, cela suggèrera une indexation commune de ces paramètres acoustiques dans le champ de la masculinité-féminité à travers les deux langues-cultures étudiées, français de France et anglais des États-Unis.

L'analyse de régression logistique binaire montre une différence entre les réponses données par les Américain-es et celles données par les Français-es dans les évaluations des voix produites par les locutrices (femmes cisgenres) et les personnes non-binaires. Dans les deux cas, les Américain-es ont accordé plus d'évaluations en zone médiane de l'échelle que les Français.es¹². Cela veut dire, en

¹⁰ Nous avons comparé les évaluations par des réponses cochées en zone médiane par rapport à celles cochées en zone masculine de l'échelle, ensuite par rapport à celles cochées en zone féminine, et les évaluations cochées en zone masculine par rapport à celles en zone féminine.

¹¹ Pour rappel, les prédicteurs statistiques retenus pour les analyses factorielles sont : la langue-culture des auditeurs (2 niveaux France / Etats-Unis), la moyenne de la F0 (3 niveaux : bas, moyen, haut), la variance de la plage de hauteur (2 niveaux : variation forte, faible) et la valeur de pente spectrale (2 niveaux : pente négative forte, faible).

¹² Pour les voix de femmes, les valeurs américaines sont : moy=0,16, erreur-type=0,02, et les valeurs françaises : moy=0,10, erreur-type=0,02, ($p<0,05$). Pour les voix des personnes non-binaires, les valeurs américaines sont : moy=0,39, erreur-type=0,03 et françaises moy=0,29, erreur-type=0,03 ($p<0,05$).

résumé, que la tendance légèrement plus forte des Français-es à utiliser la stratégie binaire, et donc à éviter de donner des évaluations dans la zone médiane, s'est surtout manifestée pour évaluer les voix des femmes cis-genres et des personnes AFAB non-binaires. Il pourrait s'agir ici d'une différence culturellement marquée, même si cela demande à être vérifié sur des échantillons plus grands.

Nous avons souhaité affiner les analyses pour tenter de savoir quels seraient les paramètres acoustiques qui ont pu influencer les réponses données, parmi les trois paramètres acoustiques annoncés : valeurs de la fréquence fondamentale, F0 ; amplitude de la plage de variation mélodique ; valeur de pente spectrale).

Aucun effet significatif de la F0 n'a été mesuré pour les évaluations des voix des femmes en zone médiane ou en zone féminine. Le facteur qui a eu un effet sur les évaluations des voix produites par les femmes cisgenres, a été la variance de la plage mélodique ($p < 0,001$) : plus la plage est ample, plus les voix sont mélodieuses, moins elles sont associées à la zone médiane¹³ par rapport à la zone féminine, et plus la plage est faible, plus les voix sont monotones, et plus elles sont catégorisées en zone médiane¹⁴ par rapport à la zone féminine. Cette tendance, comme indiqué plus haut, est plus prononcée dans l'échantillon des auditeurs états-uniens qui montrent une plus grande propension générale à utiliser la zone médiane dans leurs évaluations.

En ce qui concerne la F0 étudiée comme facteur, celle-ci a un effet significatif pour les évaluations des voix produites par les personnes non-binaires ($p < 0,001$). Une F0 plus basse favorise les évaluations plutôt en zone médiane qu'en zone féminine¹⁵.

Contrairement à nos prédictions, nous n'avons pas pu mettre en évidence un effet global de la pente spectrale pour les évaluations données aux voix produites par les personnes non-binaires ; cette absence de résultat qui peut être dû à la petite taille de notre échantillon. Il en est de même pour les voix produites par les femmes cisgenres, mais ce dernier résultat était attendu.

En ce qui concerne les hommes cisgenres, les analyses statistiques montrent néanmoins une interaction intéressante dans les évaluations de leurs voix, de façon générale, sans aucune différence selon la langue-culture des auditeurs. Les voix qui ont une F0 en plage moyenne ont toujours été, comme nous l'avons dit, davantage associées à la zone médiane de l'échelle du genre que celles qui ont une F0 basse, mais cet écart est significativement net lorsque les voix ont un timbre plus riche, plus

¹³ Moyennes marginales estimées: 0,08, erreur-type=0,06.

¹⁴ Moyennes marginales estimées: 0,47, erreur-type=0,01.

¹⁵ Pour la F0 basse moy=0,94, erreur-type=0,03; pour la F0 moyenne moy=0,26, SE=0,03; pour la F0 haute moy=0,02, erreur-type=0,03.

clair, avec une pente spectrale plus faible¹⁶. Cela nous incite à proposer de garder l'attention des futures analyses sur le comportement de ce paramètre qui pourrait constituer un indice de non-conformité ou non-binarité du genre.

S'il est difficile de trouver des tendances robustes pour l'indexation des paramètres acoustiques avec la zone médiane de l'échelle du genre (à part la F0 qui est un paramètre stable et connu) il est possible que ce soit dû au fait que nous avons groupé les données selon la catégorie de genre auto-identifiée par les personnes enregistrées (femmes, homme, non-binaire), ce qui correspond à la perception habituelle de la voix des personnes dans la vie quotidienne : en dehors des protocoles expérimentaux, nous connaissons le genre des gens avec qui nous interagissons. Mais dans ce protocole, seule la voix était disponible pour l'évaluation.

Or, les personnes de genre fluide ou non binaire utilisent leur agentivité sémiotique (Corwin 2017) de manières diverses et dynamiques, et la voix n'est pas investie de la même valeur sémiotique pour rendre intelligible son genre.

Nous pouvons prendre comme exemple trois personnes non-binaires de notre échantillon qui ont reçu des jugements différents sur la base de leur voix : les locutaires SE02, SE051 et SE018. Les deux premiers prennent de la testostérone, ce qui a comme effet d'abaisser les valeurs de leur fréquence fondamentale (leur voix est rendue plus grave). SE02 a reçu des évaluations très proches de la position « zéro », au milieu de l'échelle du genre, et même au milieu de la zone médiane. Cela correspond à la façon dont iel se genre, comme « non-binary and gender-neutral » ; sa voix fait partie des ressources sémiotiques qu'iel mobilise pour performer son genre, et iel affirme que celle-ci est perçue comme neutre ou androgyne, selon les retours spontanés reçus. SE051, malgré une F0 basse, reçoit des évaluations situées au tout début de la zone féminine de l'échelle et exprime son souhait d'avoir une voix perçue comme celle d'un « effeminate gay man ». Enfin, SE018 ne prend pas de testostérone et déclare ne pas se soucier de modifier sa voix, d'une quelconque manière et ne pas y investir du temps ; ses façons de performer la fluidité du genre passent par des ressources visuelles, parfois perçues comme très féminines et parfois comme très masculines. Dans notre protocole, sa voix a obtenu des jugements placés à l'extrémité de la zone féminine ; plus féminine, en moyenne, que la plupart des voix des femmes cisgenres de notre échantillon.

¹⁶ Tests *post hoc* Bonferroni séquentiel, $p < 0,001$ et $p < 0,01$.

Il est important de se rappeler que ce n'est pas parce que les humains ont la capacité de produire une voix peu genrée que les personnes non-binaires vont obligatoirement choisir de produire une voix peu genrée.

Conclusion

Nous avons voulu savoir si les auditeurs français-es et américain-es utilisent de façon convergente ou différente les paramètres acoustiques dans leurs évaluations de la masculinité-féminité des voix humaines, et si les deux échantillons de personnes sollicités s'emparent avec les mêmes stratégies d'une échelle de genre comportant 15 points (affichée à l'horizontale : une zone médiane située entre une zone masculine à gauche et une zone féminine à droite).

Nos résultats indiquent que la majorité des auditeurs utilisent la stratégie non-binaire (utilisation du point central/neutre de l'échelle de masculinité-féminité), ce qui permet de penser la possibilité de l'émergence dans l'imaginaire collectif d'une nouvelle catégorie de voix (neutres, ou agenres), que ce soit en contexte nord-américain ou en contexte français. Parmi les auditeurs utilisant la stratégie "binaire" (qui ne s'emparent jamais de la zone médiane dans leurs réponses) la majorité sont Français, ce qui indique néanmoins des différences interlinguistiques/culturelles potentielles dans l'utilisation d'une telle échelle. Cela devra être confirmé par le volet symétrique de cette étude dans lequel nous allons demander à des Américain-es et des Françaises d'évaluer des voix françaises. Le sous-groupe de Français-es qui a systématiquement évité d'utiliser la zone médiane de l'échelle est trop faible pour être étudié ici, mais il mériterait une étude sociologique qualitative approfondie.

La tendance des auditeurs de France à privilégier légèrement les points finaux de l'échelle plutôt que la zone médiane se manifeste, dans nos données, particulièrement pour les voix produites par des femmes cisgenres et par des personnes non-binaires ; elle peut refléter une perception plus binaire de la masculinité et de la féminité, ou la présence d'indices linguistiques et notamment vocaux qui sont interprétés comme plus genrés par les auditeurs français-es par rapport à leurs homologues états-uniens.

En revanche, les auditeurs des USA évaluent plus souvent les voix des femmes cisgenres et des personnes non-binaires dans la zone médiane que dans la zone féminine, ce qui pourrait indiquer un schéma de genre plus nuancé, peut-être plus proche de briser le cadre strictement binaire, peut-être en lien avec la plus forte exposition à la culture non-binaire plus émergente dans les médias grand public aux Etats-unis qu'en France.

Concernant les évaluations selon les paramètres acoustiques des extraits audio, nos résultats suggèrent que les auditeurs français-es et américain-es indexent les indices vocaux disponibles de manière similaire et cohérente. Cela suggère que les deux échantillons d'auditeurs partagent le même système indexical, et donc une évolution convergente serait très facile à imaginer. Le système indexical du genre semble plus stable chez les auditeurs américain-es ; ces derniers associent de façon plus nette les voix qui présentent une faible ampleur de variation mélodique avec la zone médiane de l'échelle du genre. La taille de notre échantillon ne permet pas de montée en généralité mais encourage à explorer davantage l'association de plusieurs paramètres pour étudier l'effet des combinaisons d'indices, en particulier pour mieux explorer le rôle de la pente spectrale (LTAS).

Nous ne perdons pas de vue que la signification activée par un indice peut tout à fait correspondre à l'intention des locutaires qui la produisent ou bien y échapper ; l'étude de leurs discours sur leur propre voix est en cours et elle constitue un complément indispensable. Par ailleurs, une enquête qualitative semble également nécessaire pour une meilleure compréhension des stratégies des auditeurs et du degré de conscience de leurs choix. Pour le moment, notre étude exploratoire montre que tous les éléments sont potentiellement disponibles pour imaginer aisément l'émergence d'une voix non-genrée qui pourrait être une ressource pour les personnes non-binaires souhaitant rendre leur identité de genre intelligible par la voix. Les paramètres acoustiques mobilisables en production de la parole ont toutes les chances d'être correctement activés à la perception. Le système indexical est phonétiquement et cognitivement disponible ; son activation dépendra de facteurs socio-politiques, et d'ailleurs rien n'empêcherait que la production d'une voix non genrée ou peu genrée puisse devenir une ressource accessible et souhaitable de façon plus large, y compris pour des personnes binaires.

Références

- Adler, Richard K., Sandy Hirsch et Michelle Mordaunt. (Eds). *Voice and communication therapy for the transgender/transsexual client: a comprehensive clinical guide*. (2012). Plural Publishing: San Diego, USA.
- Arnold, Aron. « Idéologies de genre et construction des savoirs en sciences phonétiques », *GLAD!* [En ligne], 01 | 2016. (2016). <http://journals.openedition.org/glad/117> ; DOI : <https://doi.org/10.4000/glad.117>.
- Avery, Jack D. et Julie M. Liss. "Acoustic characteristics of less-masculine-sounding male speech." *Journal of the Acoustical Society of America* 99, no. 6 (June 1996): 3738-3748.
- Benoist-Lucy, Agathe et Claire Pillot-Loiseau. "The Influence of language and speech task upon creaky voice use among six young American women learning French." *Interspeech 2013*, (2013). Lyon, France: 2395-2399. fahal-00862349f
- Boersma, Paul et David Weenink. PRAAT: Doing phonetics by computer. Version 6.0.49. (2019). <http://www.praat.org>.

- Brown, LeAnn, Maria Candea, Oriana Reid-Collins, Julie Abbou, et James Sneed German. (2019). *Non-binary speakers & mixed methods: Interpreting multiplex gender identity in language research*. Lavender Languages and Linguistics 26. Gotenburg, Sweden.
- Brown, LeAnn et Claire Pillot-Loiseau (*soumis*). Voice as non-binary gender identity resource: The roles of spectral slope and fundamental frequency.
- Coleman, Ronald O. "A comparison of the contributions of two vocal characteristics to the perception of maleness and femaleness in the voice." *STL-Quarterly Progress and Status Report* 14(2-3) (1973): 13-22.
- Corwin, Anna I., Language and gender variance: constructing gender beyond the male/female binary. *Electronic Journal of Human Sexuality* 12 (2009).
- Corwin, Anna. Emerging genders: semiotic agency and the performance of gender among genderqueer individuals. *Gender and Language* 11 (2017): 255–277.
- Davies, Shelagh, Viktória G. Papp et Christella Antoni. "Voice and Communication Change for Gender Nonconforming Individuals: Giving Voice to the Person Inside." *International Journal of Transgenderism* 16, no. 3 (2015): 117-159, DOI: 10.1080/15532739.2015.1075931
- Eckert, Penelope. "Variation and the indexical field." *Journal of Sociolinguistics* 12, no. 4 (2008): 453–476.
- Hancock, Adrienne B., Julianne Krissing et Kelly Owen. "Intonation and Gender Perception: Applications for Transgender Speakers." *Journal of Voice*, 28, no. 2 (2014): 203-209.
- Henton, Caroline et Anthony Bladon. Creak as a sociophonetic marker. In L. Hyman & C. Li (Eds.), *Language, Speech, and Mind* (pp. 3–29). (1988). London: Routledge.
- Irwig, Michael S., Childs, K., et Hancock, A. B. (2017). Effects of testosterone on the transgender male voice. *Andrology* 5:107–112.
- James, Sandy E., Jody L. Herman, Susan Rankin, Mara Keisling, Lisa Mottet, et Ma'ayan Anafi. (2016). *Executive Summary of the Report of the 2015 U.S. Transgender Survey*. Washington, DC: National Center for Transgender Equality.
- Klatt, Dennis M. et Laura C. Klatt. "Analysis, synthesis, and perception of voice quality variation among female and male talkers." *Journal of the Acoustical Society of America* 87, (1990): 820-857.
- Lee, Chao-Young, Lauren Dutton et Gayatri Ram. "The role of speaker gender identification in relative fundamental frequency height estimation from multispeaker, brief speech segments." *Journal of the Acoustical Society of America* 128, no.1 (2010): 384-388.
- Linville, Sue Ellen. "Acoustic correlates of perceived versus actual sexual orientation in men's speech." *Folia Phoniatrica et Logopaedica* 50 (1998): 35-48.
- Limesurvey GmbH. / LimeSurvey: An Open Source survey tool /LimeSurvey GmbH, Hamburg, Germany. URL <http://www.limesurvey.org>
- McConnell-Ginet, Sally. "Intonation in a Man's World." *Signs* 3, no. 3 (1978):541-559.
- Mendoza, Elvira, Nieves Valencia, Juana Mufioz, et Humberto Trujillo. "Differences in voice quality between men and women: use of the long-term average spectrum (LTAS)." *Journal of Voice*, 10, (1996): 59-66.
- Mills, Matthew et Gillie Stoneham. *The Voice Book for Trans and Non-Binary People: A Practical Guide to Creating and Sustaining Authentic Voice and Communication*. (2017) Jessica Kingsley Publishers.
- Munson, Benjamin. "The acoustic correlates of perceived masculinity, perceived femininity, and perceived sexual orientation." *Language and Speech* 50, no. 1 (2007): 125-142
- Nolan, Ian T., Christopher J., Kuhner et Geolani W. Dy. Demographic and temporal trends in transgender identities and gender confirming surgery. (2019). *Translational Andrology and Urology* 8, 184-190 doi: [10.21037/tau.2019.04.09](https://doi.org/10.21037/tau.2019.04.09)

- Oates, Jennifer M. et Georgia Dacakis. "Speech pathology considerations in the management of transsexualism – a review." *British Journal of Disordered Communication* 18, no. 3 (1983): 139-151.
- Ochs, Elinor. "Indexing Gender." In Alessandro Duranti et Charles Goodwin, *Rethinking Context. Language as an interactive phenomenon*. (1992). Cambridge university Press.
- Omori, Koichi, Ashutosh Kacker, Linda M. Carroll, William D. Riley, et Stanley M. Blaugrund. Singing power ratio: quantitative evaluation of singing voice quality. *Journal of Voice*, 10, (1996): 228-235.
- Pépiot, Ewan. "Male and female speech: a study of mean f0, f0 range, phonation type and speech rate in Parisian French and American English speakers." *Speech Prosody* 7, Dublin, Ireland (2014): 305-309. ffhalshs-00999332
- Pettirossi, Amelia, Nicolas Audibert, et Lise Crevier-Buchman. « Le spectre moyen à long-terme corrigé en F0 (pitch-corrected LTAS) : une mesure robuste de la voix dysphonique sur des énoncés de nature variable. » *7èmes Journées de phonétique clinique (JPC7)*, (2017). Paris, France.
- Pillot, Claire et Jacqueline Vaissière. Spectral correlates of carrying power in speech and western lyrical singing according to acoustic and phonetic factors, *15th International Congress of Acoustics*, (2007). Madrid, 1-7 septembre 2007, 1-6.
- Podesva, Robert J. et Patrick Callier. Voice Quality and Identity. *Annual Review of Applied Linguistics*, 35 (2015), pp. 173–194. doi: 10.1017/S0267190514000270
- Schmid, Maxwell et Evan Bradley. "Vocal pitch and intonation characteristics of those who are gender non-binary." (2019). *ICPhS Conference*. Melbourne, Australia.
- Skinner, Ashton. *Transgender Experiences Beyond the Binary: A Phenomenological Study of Arizonans with Non-Binary Gender Identities*. (2017). Masters thesis, Arizona State University.
- Silverstein, Michael. "Indexical order and the dialects of sociolinguistic life." *Language and Communication* 23, (2003): 193-229.
- Simpson, Adrian P. "Phonetic differences between male and female speech." *Language and Linguistics Compass* 3, (2009): 621–640.
- Smith, Cara G., Eileen M. Finnegan et Michael P. Karnell. Resonant voice: Spectral and nasendoscopic analysis. *Journal of Voice*, 19, (2005): 607-622.
- Smyth, Ron, Greg Jacobs, et Henry Rogers. "Male voices and perceived sexual orientation: an experiential and theoretical approach." *Language in Society* 32, (2003): 329–350.
- Yuasa, Ikuko Patricia. Creaky Voice: A New Feminine Voice Quality for Young Urban-Oriented Upwardly Mobile American Women? *American Speech*, 85(3), (2010): 315-337.