



HAL
open science

Évolution de la voix humaine : le rôle de la sélection sexuelle

Alexandre Suire

► **To cite this version:**

Alexandre Suire. Évolution de la voix humaine : le rôle de la sélection sexuelle. Biodiversité et Ecologie. Université de Montpellier, 2019. Français. NNT : . tel-02434154

HAL Id: tel-02434154

<https://theses.hal.science/tel-02434154>

Submitted on 9 Jan 2020

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.

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Sciences de l'Évolution et de la Biodiversité

Au sein de l'école doctorale GAIA – Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau

Et de l'unité de recherche Institut des Sciences de l'Évolution de Montpellier – UMR 5554

ÉVOLUTION DE LA VOIX HUMAINE : LE ROLE DE LA SELECTION SEXUELLE

Présentée par Alexandre SUIRE

Le 15 novembre 2019

Sous la direction de Melissa BARKAT-DEFRADAS
et Michel RAYMOND

Devant le jury composé de

Jacqui SHYKOFF

Directrice de recherche, Université Paris-Sud – Orsay

Rapporteuse

Dan DEDIU

Chercheur, Université Lyon 2 – Lyon

Rapporteur

Jean-Baptiste ANDRE

Chargé de recherche, Institut Jean Nicod – Paris

Examineur

Julien RENOULT

Chargé de recherche, Centre d'Écologie Fonctionnelle et Évolutive – Montpellier

Examineur

Didier DEMOLIN

Professeur, Université Sorbonne Nouvelle – Paris

Examineur

Evelyne HEYER

Professeure, Muséum National d'Histoire Naturelle – Paris

Présidente de jury



UNIVERSITÉ
DE MONTPELLIER

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Sciences de l'Évolution et de la Biodiversité

Au sein de l'école doctorale GAIA – Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau

Et de l'unité de recherche Institut des Sciences de l'Évolution de Montpellier – UMR 5554

ÉVOLUTION DE LA VOIX HUMAINE : LE ROLE DE LA SELECTION SEXUELLE

Présentée par **Alexandre SUIRE**

Le 15 novembre 2019

Sous la direction de **Melissa BARKAT-DEFRADAS**
et **Michel RAYMOND**

Devant le jury composé de

Jacqui SHYKOFF

Directrice de recherche, Université Paris-Sud – Orsay

Rapporteure

Dan DEDIU

Chercheur, Université Lyon 2 – Lyon

Rapporteur

Jean-Baptiste ANDRE

Chargé de recherche, Institut Jean Nicod – Paris

Examineur

Julien RENOULT

Chargé de recherche, Centre d'Écologie Fonctionnelle et Évolutive – Montpellier

Examineur

Didier DEMOLIN

Professeur, Université Sorbonne Nouvelle – Paris

Examineur

Evelyne HEYER

Professeure, Muséum National d'Histoire Naturelle – Paris

Présidente de jury

Sommaire

I. Contexte	1
II. Physio-anatomie de la parole	12
II.1. Production de la parole	12
i. La soufflerie	15
ii. La source	15
iii. Le filtre	18
iv. Production des voyelles	23
v. Production des consonnes	27
vi. La qualité vocale	29
vii. Ontogénie de la voix	34
viii. Système auditif	36
II.2. Sources de variation	39
i. Taille corporelle	39
ii. Profils hormonaux	41
iii. Configuration corporelle	45
iv. Force physique	46
v. Indices développementaux	47
vi. Âge	49
vii. Variation inter-linguistique	50
viii. Variation temporelle	51
ix. Facteurs environnementaux	51
II.3. Conclusion	52
III. La sélection sexuelle	55
III.1. Mécanisme général	55
i. Compétition intrasexuelle	56
ii. Compétition intersexuelle	57
III.2. Conclusion	58
IV. Évolution des préférences vocales	59
IV.1. Les préférences masculines et féminines	59
i. Préférences pour la hauteur vocale (F0)	69
ii. Préférences pour le timbre (Formants)	70
iii. Préférences pour le bruit (HNR et jitter)	71
iv. Préférences pour les variations de l'intonation (F0-SD)	71
v. Préférences pour le débit	72
IV.2. Sources de variation des préférences	72
i. Cycle menstruel	72
ii. Environnement socio-culturel	75
IV.3. Limites des préférences	76
IV.4. Autres composantes de l'attractivité vocale	77
IV.4. Conclusion	79
Article 1	81
Article 2	125
V. Évolution du dimorphisme vocal	159
V.1. Poids relatif des pressions de sélection sexuelle dans les deux sexes	159
V.2. La voix : un signal « honnête » ?	160
V.3. La voix : un biais sensoriel ?	163

V.4. La voix : un signal « redondant » ?	165
V.5. Conclusion	166
Article 3	171
VI. Succès copulatoire et reproducteur	203
VI.1. Populations non-industrialisées.....	203
VI.2. Populations industrialisées	204
VI.3. Conclusion.....	207
Article 4.....	209
VII. La modulation vocale	221
VII.1. État de l'art.....	222
i. Dominance vocale.....	222
ii. Attractivité vocale.....	224
V.II. Avantages et coûts de la modulation vocale	228
VII.3. Modulation vocale et émergence du langage.....	230
V.II.4. Conclusion	232
Étude exploratoire 1.....	235
Étude exploratoire 2.....	249
VIII. Symbolisme phonétique et sélection sexuelle.....	265
VIII.1. Principe du symbolisme phonétique.....	265
VIII.2. La théorie « code-fréquence »	267
VIII.3. Dimorphisme sexuel et caractéristiques phonétiques des prénoms.....	268
VIII.4. Vers un autre corpus onomastique : le cas des Pokémon.....	273
VIII.5. Conclusion.....	275
Article 5	277
Article 6	299
Article 7	313
IX. Conclusion générale.....	345
Bibliographie.....	351
Remerciements.....	363

I. Contexte

« Ce regard allègre animait tout alentour, dans l'ombre, d'une joie jeunette, d'un entrain minime mais pur comme nous n'en avons plus à notre disposition, sa voix cassée quand elle vociférait reprenait guillerette les mots quand elle voulait bien parler comme tout le monde et vous les faisait alors sautiller, phrases et sentences, caracoler et tout, et rebondir vivantes tout drôlement comme les gens pouvaient le faire avec leur voix et les choses autour d'eux au temps encore où ne pas savoir se débrouiller à raconter et chanter tour à tour, bien habilement, passait pour niais, honteux, et maladif. »

Louis-Ferdinand Céline – *Voyage au bout de la nuit*, pages 253-254.

Si la teneur du *Voyage* s'en tient à une description absurde et atroce de l'existence humaine, Céline dresse pendant un bref instant un portrait lumineux de la mère Henrouille. Au milieu d'une *Nuit* où rien ne luit pour l'antihéros du roman, cet extrait dessine les caractéristiques vocales d'un personnage qui nous permet d'introduire l'objet d'étude du présent travail : caractériser finement la façon dont chacun s'emploie à utiliser sa voix en fonction du contexte, indépendamment du message linguistique véhiculé.

En effet, si elle apparaît en premier lieu comme le support privilégié de la communication verbale, la voix n'est pas un simple outil de médiation de l'information sémantique ; elle constitue également un objet identitaire et social dont le rôle est fondamental au sein des rapports humains. La voix est en effet un instrument par lequel le sujet parlant partage avec autrui son état mental, émotionnel et affectif via une infinité de subtiles nuances, pouvant exprimer tour à tour la douleur, la passion, l'ivresse, la fureur, l'exaltation, la peur, la joie, le doute, la conviction et même... le sourire (Tartter, 1980). Elle peut également signaler l'intérêt ou une intention particulière, dont les variations mélodiques peuvent complètement altérer le sens d'une phrase : nous pouvons souffler un « *je t'aime* » avec la tendresse la plus sincère comme nous pouvons l'exulter de façon cinglante et sarcastique. Enfin et surtout, les

caractéristiques acoustiques propres à la voix, telles que sa hauteur (grave ou aiguë), son articulation (tendue ou relâchée) et son rythme (lent ou rapide) résultent d'un tissu complexe de facteurs biologiques, sociaux et culturels délivrant de précieuses informations sur le locuteur telles que le sexe, l'âge, la santé, l'origine socio-culturelle ou encore l'orientation sexuelle.

Si les sciences humaines et sociales se sont quelque peu attachées à décrire son rapport avec la personnalité (Scherer, 1978), son rôle en tant que marqueur social et culturel (Laver & Trudgill, 1979 ; van Bezooijen, 1995), en tant que vecteur d'émotions (Banse & Scherer, 1996) et plus globalement sa fonction dans la communication interpersonnelle (Apple, Streeter, & Krauss, 1979), elle est restée dans ce paradigme un objet d'étude essentiellement descriptif jusqu'au début des années 2000, en ce sens où peu d'études se sont interrogées sur sa fonction évolutive.

Pourtant, Darwin (1890) se questionne longuement sur le rôle de la voix humaine en tant que vecteur d'émotions et d'intentions par analogie aux vocalisations animales :

« *The sexes of many animals incessantly call for each other during the breeding-season; and in not a few cases, the male endeavours thus to charm or excite the female. This, indeed, seems to have been the primeval use and means of development of the voice [...]. When male animals utter sounds in order to please the females, they would naturally employ those which are sweet to the ears of the species; and it appears that the same sounds are often pleasing to widely different animals, owing to the similarity of their nervous systems, as we ourselves perceive in the singing of birds and even in the chirping of certain tree-frogs giving us pleasure. On the other hand, sounds produced in order to strike terror into an enemy, would naturally be harsh or displeasing.* » *The expression of the emotions in man and animals*, pages 90-96.

Darwin avance ainsi l'idée que la forme des productions sonores, c'est-à-dire leur dimensions acoustiques et perceptives (sons graves ou aigus, plaisants ou non), sont empreintes d'esthétisme et véhiculent en elles-mêmes – au-delà de leur fonction communicative première

– une intention. S’il remarque que les vocalisations peuvent servir à projeter une impression d’hostilité et d’agression pour repousser un éventuel adversaire, il note surtout l’importance de celles-ci lors des saisons d’accouplements, lorsqu’il s’agit d’attirer et courtiser des femelles. En partant de ces observations, Darwin explique, en s’appuyant sur le mécanisme d’évolution qu’il nomme « *sélection sexuelle* » et décrit dans un ouvrage antérieur (1871) , que certaines formes sonores des vocalisations produites par nos ancêtres auraient évolué dans le but de séduire les conspécifiques. Dans ce contexte, il apparaît légitime de se demander dans quelle(s) mesure(s) et de quelle(s) manière(s) la voix humaine peut jouer un rôle dans le cadre de la recherche d’un partenaire, et plus généralement au sein des stratégies d’accouplements.

Dans un article majeur relatif à la communication vocale chez les animaux, Morton formalise en 1977 une théorie qu’il appela « *motivation-structural rules* ». Il suggère que les propriétés physiques des signaux acoustiques (i.e. sons de basses ou hautes fréquences), sont « *motivées* », c’est-à-dire qu’elles portent en elles-mêmes une information sur les intentions et/ou l’état émotionnel de l’émetteur. Son argument repose sur l’observation d’une convergence fonctionnelle des signaux acoustiques chez un certain nombre d’oiseaux et de mammifères, dont les exemples d’espèces sont aussi variés en termes d’écologie, de système d’appariement et d’histoire évolutive : les individus utilisent des sons de basses fréquences (i.e. sons graves) pour signaler l’hostilité, la menace et l’agression, et des sons de hautes fréquences (i.e. sons aigus) pour signaler la peur, l’apaisement ou l’amicalité. Ainsi, une hirondelle, un macaque rhésus ou encore un sanglier produiront respectivement, par exemple, des pépiements secs et intenses, des grognements et des rugissements lorsqu’ils cherchent à menacer un autre individu, et des gémissements, couinements ou plus généralement des cris perçants lorsqu’ils se soumettent.

Afin d’expliquer ce phénomène d’un point de vue évolutif, l’auteur émet l’hypothèse que les vocalisations de basses et hautes fréquences auraient été sélectionnées en raison de la

relation directe qui existerait entre la fréquence d'un son et la taille corporelle de l'émetteur, en notant que cette dernière représente la clé de voûte des affrontements liés à l'acquisition des ressources. En effet, selon une supposée relation allométrique entre la taille corporelle et ce qu'il nomme les « *membranes vibratoires* » de la production vocale, les vocalisations de basses fréquences sont produites par des individus plus corpulents, tandis que celles situées dans les hautes fréquences sont produites par des individus plus petits et fins. Morton (1977) suggère que cette association signal acoustique-taille corporelle aurait été conservée au cours de l'évolution, se substituant peu à peu à la taille corporelle comme unique facteur déterminant dans l'acquisition des ressources :

« The argument is that genes that promote fighting, where larger body size is favored, will tend to be replaced by genes involved in the communication of motivation if the communication genes are as effective as fighting in acquiring the resources in question. Genes promoting communication will be as effective as those promoting fighting if the communication system replaces size as the main determinant of which gene carrier obtains the resources. This seems to lead inevitably to low harsh sounds as the adaptive sound structure in hostile contexts, since lower sounds represent larger size. » page 865.

Bien que son argument apparaisse comme discutable, Morton (1977) suggère que les vocalisations autorisent une « *projection* » perceptive de la taille corporelle, fonction particulièrement favorable dans de nombreux contextes de rivalité, permettant ainsi de deviner l'issue d'une confrontation sans avoir à l'engager pour éviter les coûts d'un combat réel. Toutefois, certaines zones d'ombre restent à investiguer : les vocalisations de basses et hautes fréquences servent-elles à exploiter efficacement un simple biais cognitif de perception (e.g. sons graves systématiquement associés à une perception de masse imposante), lequel aurait été sélectionné pour les raisons précédemment évoquées, ou bien sont-elles contraintes par les caractéristiques physio-anatomiques de l'émetteur, limitant ainsi la plage des vocalisations

possibles, reflète donc « *honnête* » de sa taille corporelle ? De même, dans quelle mesure ces vocalisations sont-elles modulables et dans les faits, toutes les espèces en sont-elles capables ? Ces questionnements sur les liens existant entre signaux acoustiques et taille corporelle constitueront par la suite une des pierres angulaires de la recherche en bioacoustique animale et, plus globalement, dans la compréhension de l'émergence et l'évolution des systèmes de communication vocaux (Fitch & Hauser, 2003 ; Taylor & Reby, 2010).

Toutefois, ce n'est qu'un peu plus tard que Ohala (1984) propose une explication évolutive quant au rôle fonctionnel de la voix humaine en formalisant le « *frequency-code hypothesis* » (ou théorie « *code-fréquence* »). Sur la base du travail de Morton (1977), sa théorie suggère que l'utilisation de la hauteur vocale de l'espèce humaine (i.e. voix graves ou aiguës) présente des similitudes d'associations sons-sens identiques à celles observées chez les autres espèces animales. Mais il suggère également que certaines de ces associations sont universelles au niveau translinguistique, notamment dans l'intonation de certains types d'énoncés, la communication de l'affect, et même dans la composition sonore de certains mots. Selon l'auteur, cette capacité serait « *innée* » en raison de son incidence à travers les diverses langues du monde (i.e. universaux linguistiques), laquelle ne peut être due ni au hasard ni à l'emprunt linguistique.

Ohala (1984) note par exemple que l'utilisation de la hauteur vocale suit le même schéma intonatif selon la nature grammaticale des phrases énoncées dans une très grande majorité des langues naturelles. Typiquement, lorsqu'il s'agit de poser des questions, les individus ont généralement tendance à employer un motif mélodique particulier, lequel consiste en une hauteur vocale ascendante et plus élevée en fin d'énoncé. De même, l'utilisation de la hauteur vocale est largement utilisée à travers les langues du monde pour marquer et communiquer son affect, ses intentions et/ou ses émotions : lorsqu'il s'agit de signaler la déférence, la politesse, la soumission et le manque de confiance, les locuteurs auront tendance

à utiliser un registre vocal plus aigu, et à l'inverse, ils emploieront un registre plus grave pour signaler l'autorité, l'agression, la confiance ou la menace.

Au-delà du contexte des vocalisations interpersonnelles, il convient de remarquer (et nous y reviendrons ultérieurement) que Ohala (1984) étend sa théorie « *code-fréquence* » aux unités langagières supérieures (i.e. sons constitutifs des mots) en s'appuyant sur le principe du symbolisme phonétique, qui stipule qu'il existe un lien naturel entre le signifié (i.e. l'image mentale) et le signifiant (i.e. la forme sonore) des signes linguistiques. Il notera à ce sujet que, dans de nombreuses langues, les mots désignant la « *petitesse* » (le concept et toutes les notions relatives) semblent attester une plus grande occurrence de voyelles et de consonnes de hautes fréquences (i.e. sons aigus), quant, à l'inverse, les mots désignant la « *grandeur* » et/ou la « *largeur* » exhibent une plus grande proportion de voyelles et de consonnes produites dans les basses fréquences (i.e. sons graves). L'auteur note quelques exemples, dont l'anglais (« *teeny* », « *wee* », « *itsy bitsy* » contre « *large* », « *humongous* »), le français (« *petit* », « *fin* », « *mince* » contre « *grand* », « *large* », « *gros* »), le grec (« *mikros* » contre « *makros* ») ou encore le japonais (« *tchiisai* » contre « *ooki* »).

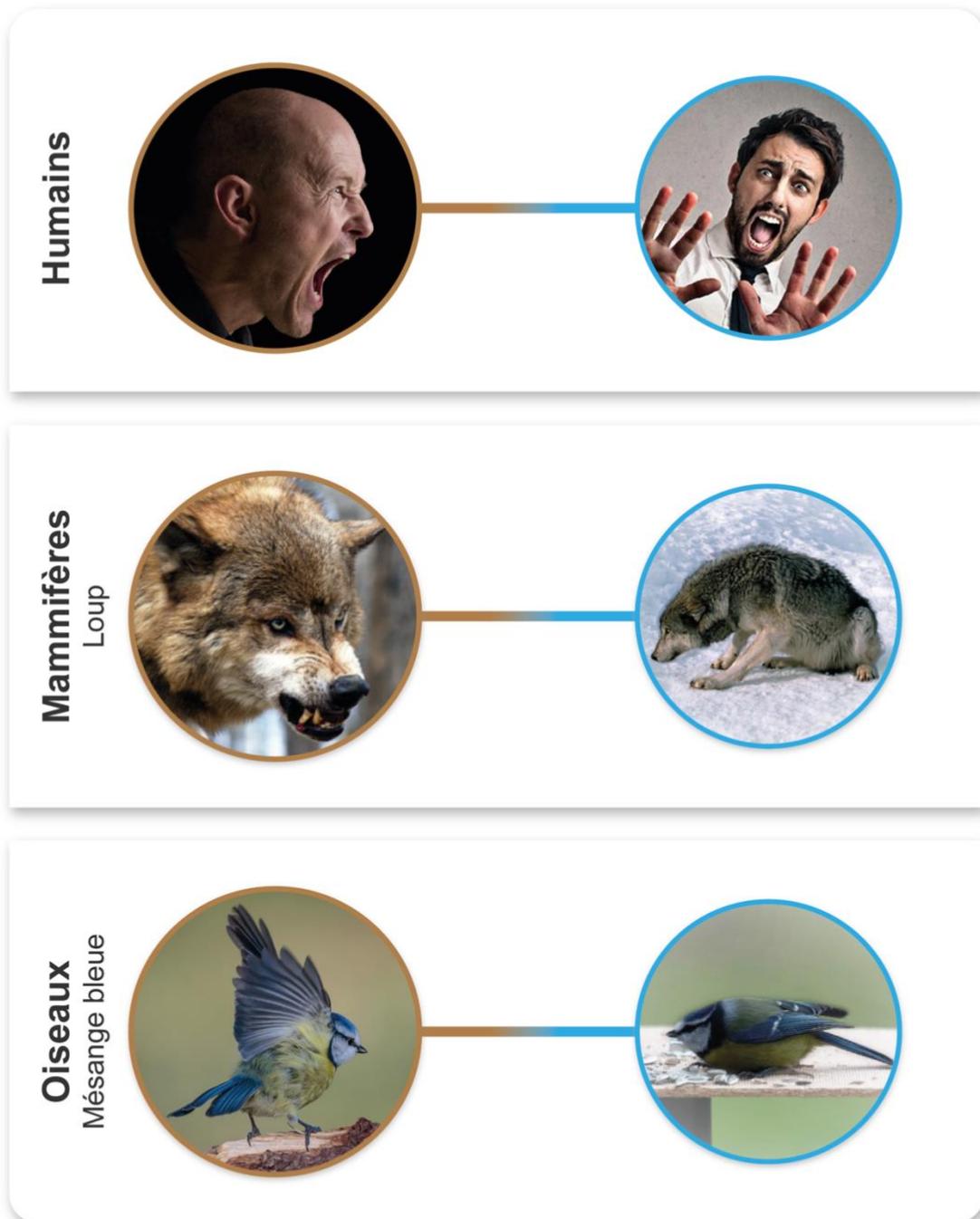
Finalement, l'auteur avance le même argument que Morton (1977) pour expliquer la supposée universalité de ces associations sons-sens, illustrées dans la Figure 1 :

« *As Morton [1977] points out, however, the F_0^1 of voice can also indirectly convey an impression of the size of the signaler since F_0 , other things being equal, is inversely related to the mass of the vibrating membrane (vocal cords in mammals, syrinx in birds) which, in turn, is correlated with overall body mass [...]. To give the impression of being large and dangerous, then, an antagonist should produce a vocalization as rough and as low in F_0 as possible. On*

¹ La hauteur vocale, exprimée ici en termes de fréquence fondamentale et correspondant au premier harmonique d'un son complexe. L'ensemble des termes acoustiques sera décrit dans la partie suivante.

the other hand, to seem small and nonthreatening a vocalization which is tone-like and high in F0 is called for. » page 5.

Figure 1. Illustration de la théorie « code-fréquence » de Ohala (1984), construite à partir des idées de Morton (1977).



Basses fréquences :

signaler la menace ou l'autorité
paraître imposant

Hautes fréquences :

signaler la soumission ou l'amicalité
paraître petit

Enfin et surtout, pour souligner le caractère « *inné* » de l'utilisation fonctionnelle de la hauteur vocale, Ohala (1984) propose de la considérer en regard du dimorphisme sexuel de la voix humaine. En constatant notamment les différences anatomiques liées à la taille du larynx entre les hommes et les femmes (15 à 20% supérieure chez les premiers) et qui se produisent lors de la puberté, il en conclut que le développement d'une voix relativement grave chez les hommes a évolué pour renforcer la perception de dominance et d'agression dans le contexte de la compétition pour la reproduction. En faisant un parallèle avec l'apparition des bois des cerfs mâles – absents chez les femelles – et qui se développent lorsqu'ils atteignent la maturité sexuelle, la voix mue pour « *s'assombrir* » d'environ une octave chez le jeune adolescent pubère. La voix est, au même titre que la pilosité ou la répartition des graisses sur le corps, un caractère sexuel secondaire qui apparaît à la puberté. Ainsi, Ohala (1984) fut le premier à établir le rôle fonctionnel du dimorphisme vocal en évoquant directement le rôle de la sélection sexuelle :

« I think the enlargement of the vocal apparatus also occurs to enhance aggressive displays – the acoustic component. Males, by their role in the family unit and the fact that they compete for the favors of the female – i.e, they are subject to what Darwin called sexual selection – would be the ones to develop such deviations from the 'norm'. However, they would only need these aggressive decorations when they are ready to compete and retain the favors of a female, that is, at the time of sexual maturity. » page 14.

De manière générale, le mécanisme de sélection sexuelle, tel que défini par Darwin (1871), constitue effectivement une force évolutive majeure dans l'évolution des traits sexuellement dimorphique et plus généralement des stratégies liées à l'accès à la reproduction. Dans ce contexte, elle a permis à la recherche en bioacoustique animale d'interpréter d'un point de vue évolutif la fonction des signaux acoustiques dans les contextes d'accouplements et d'affrontements. Les études se sont ainsi focalisées à comprendre les liens entre les signaux

acoustiques et la taille corporelle (Bowling, Garcia, Dunn, Ruprecht, Stewart, Frommolt & Fitch, 2017 ; Garcia, Herbst, Bowling, Dunn, & Fitch, 2017), la dominance sociale (Bartholomew & Collias, 1962 ; Fischer, Kitchen, Seyfarth, & Cheney, 2004), les systèmes d'appariements (Puts, Hill, Bailey, Walker, Rendall, Wheatley, Welling, Dawood, Cárdenas, Buriss, Jablonski, Shriver, Weiss, Lameira, Apicella, Owren, Barelli, Glenn & Ramos-Fernandez, 2016) ou encore leurs modulations selon les différents types d'habitats (Waser & Brown, 1986). Dans différents contextes de compétition pour la reproduction, elles se sont aussi attachées à décrire diverses adaptations anatomiques, comme celles autorisant une production de vocalisations dans une plage de fréquence acoustique non-prédite par la taille corporelle (i.e. « *acoustic size exaggeration* », Charlton & Reby, 2016) ou d'autres permettant d'augmenter la portée des appels sur de longues distance pour attirer des femelles et signaler la territorialité (Mitani & Stuht, 1998 ; Wich & Nunn, 2002). Ces modifications peuvent être liée à la hauteur vocale, telles que la présence particulière de gras dans les plis vocaux chez les tigres et les lions (Klemuk, Riede, Walsh, & Titze, 2011), la présence d'un jeu additionnel de vibrateurs laryngés avec l'intervention de « *bandes ventriculaires* » chez le koala (Charlton, Frey, McKinnon, Fritsch, Fitch, & Reby, 2013), de larynx hypertrophiés chez les hypsignathes (Bradbury, 1977) ; ou à des modifications affectant le timbre, comme un élargissement de l'os hyoïde chez les singes hurleurs (Dunn, Halenar, Davies, Cristobal-Azkarate, Reby, Sykes, Dengg, Fitch, & Knapp, 2015), l'utilisation de larynx mobiles chez le cerf rouge (Fitch & Reby, 2001) et la gazelle à goitre (Frey, Volodin, Volodina, Soldatova, & Juldachev, 2011), la présence de résonateurs vocaux supplémentaires tels que des sacs vocaux chez le tétras des armoises (Krakauer, Tyrrell, Lehmann, Losin, Goller & Patricelli, 2009) et certaines espèces d'anoures (Starnberger, Preininger, & Hödl, 2014), ou encore les proboscis nasaux chez les éléphants de mers (Bartholomew & Collias, 1962) et les nasiques (Koda, Murai, Tuuga, Goossens, Nathan, Stark, Ramirez, Sha, Osman, Sipangkui, Seino, & Mastuda, 2018).

Pourtant, ce n'est qu'au début des années 2000 que la compréhension du dimorphisme vocal et du rôle de la voix humaine sera abordée du point de vue de la sélection sexuelle, lorsque Collins (2000) évoque l'idée d'un possible effet du choix du partenaire comme force évolutive majeure dans le façonnement des préférences vocales. Plus tard, Puts, Gaulin & Verdolini (2006) avancent l'idée selon laquelle la voix masculine aurait également évolué vers les graves via la compétition intrasexuelle. Depuis ces travaux fondateurs, la dernière décennie a vu se multiplier les études sur l'évolution de la voix sous le prisme de la sélection sexuelle (Puts, Doll, & Hill, 2014), et c'est également dans ce paradigme que s'inscrit ce travail de thèse.

Dans ce contexte, il s'agira de décrire dans un premier temps les mécanismes liés à la production de la parole, puis d'introduire le mécanisme de la sélection sexuelle. Nous verrons ensuite comment ce principe permet d'aborder d'un point de vue évolutif la question des préférences vocales, mais aussi la relation entre les composantes acoustiques des voix et diverses caractéristiques biologiques et sociales des locuteurs. Plus généralement, la sélection sexuelle nous permettra de comprendre le comportement vocal dans les contextes interactionnels de séduction et de compétition. Sous une perspective comparative, nous nous intéresserons également aux vocalisations compétitives d'un primate non-humain. Enfin, nous nous intéresserons à d'autres domaines du langage en montrant que certaines unités de plus haut niveau – typiquement les noms propres – peuvent également prétendre au statut d'ornement actif dans le contexte de la sélection sexuelle. A chacune de ces parties, nous tenterons d'établir une revue exhaustive et critique de la littérature afin de dégager les questions scientifiques que nous nous sommes posés.

II. Physio-anatomie de la parole

II.1. Production de la parole

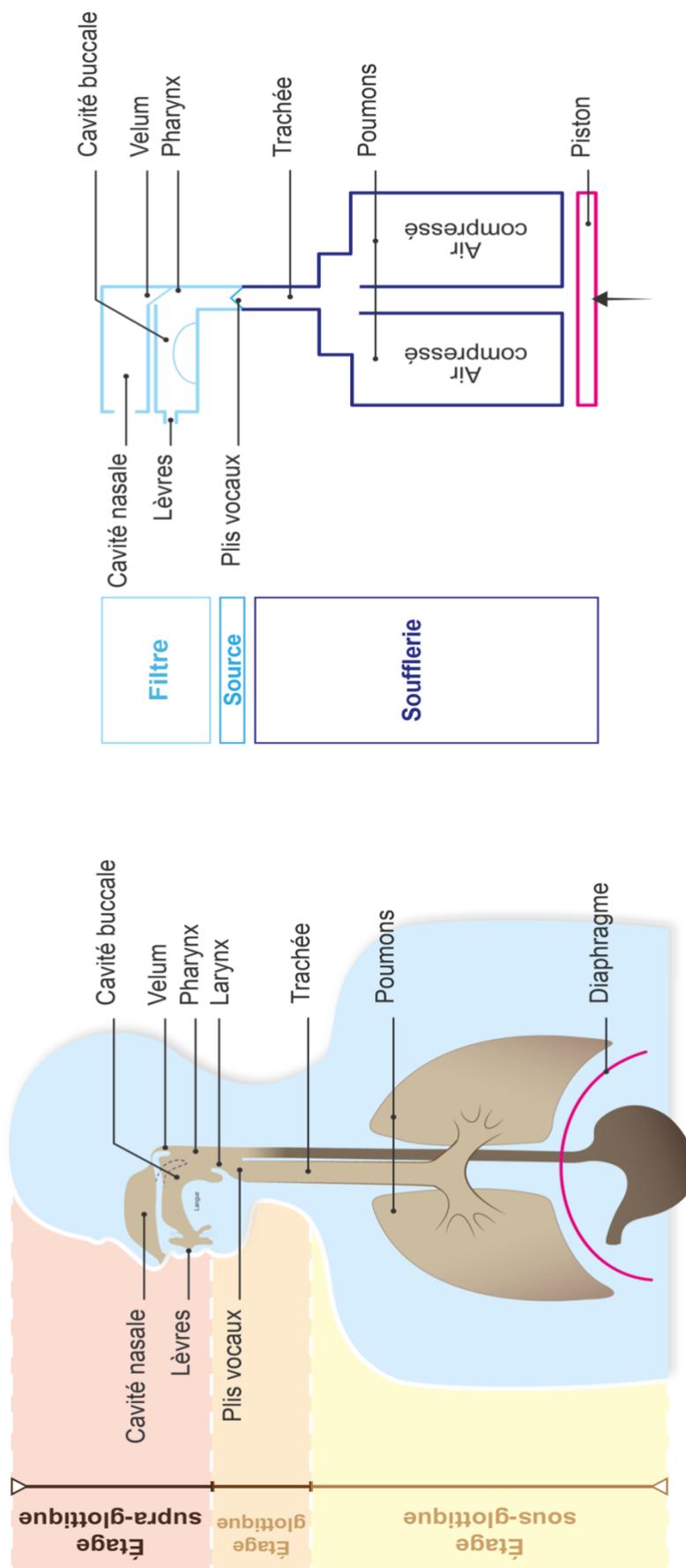
Si le terme de « *langage* » renvoie à la faculté cognitive à acquérir une langue naturelle dont chaque humain exempt de pathologie cérébrale est doté à la naissance, celle de « *langue* » renvoie au système linguistique abstrait (i.e. le code), connu et partagé par l'ensemble des membres d'un groupe et qui permet à ses derniers de communiquer entre eux, et la notion de « *parole* » désigne l'utilisation concrète de la langue que possède chaque individu (Saussure, 1916). Cette dernière correspond à la manière dont la voix est employée pour exprimer la langue au niveau individuel. Elle prend ainsi en compte la hauteur, le timbre, l'articulation, la prononciation, l'accent, le rythme ou encore l'intonation. C'est ce domaine linguistique que nous avons étudié.

Dans ce contexte, la phonation correspond au phénomène de production physio-anatomique des sons de la parole. L'appareil phonatoire (ou appareil vocal) correspond à l'ensemble des organes impliqués dans ce processus. Il fait intervenir trois structures anatomiques distinctes dont le rôle premier n'est pas spécifiquement dédié à la parole : (1) l'appareil respiratoire (initialement dédié à la respiration), (2) le larynx et (3) les résonateurs (nécessaires à la déglutition et à la mastication).

Chacune de ces structures joue un rôle particulier dans la production de la parole comme modélisé à travers la théorie « *source-filtre* » définie par Fant (1960). Ce modèle classique de la production de la parole pose que la « *source* » des sons provient du souffle pulmonaire lequel fait mécaniquement vibrer les plis vocaux logés dans le larynx produisant ainsi une onde sonore complexe. A l'issue de son passage dans la trachée, celle-ci parvient au « *filtre* », constitué de différents résonateurs et comprenant d'une part les cavités supra-laryngées (i.e. les cavités pharyngale, orale et nasale) et d'autre part les muscles articulateurs qui jouent le rôle d'amplificateur (ou d'atténuateur) de certaines fréquences (i.e. la langue et les lèvres). En

d'autres termes, la source correspond à la transformation de l'air en son et la configuration articulatoire du filtre détermine la mise en forme de ce son. Cette théorie suppose également une indépendance entre la « *source* » et le « *filtre* » : les propriétés de l'un peuvent être modifiées sans influencer celles de l'autre (Fitch & Hauser, 1995). Nous nous attacherons ainsi à décrire l'anatomie ainsi que les mécanismes articulatoires et acoustiques de la production de la parole, lesquels sont illustrés dans la Figure 2 qui nous servira de canevas tout au long de ce chapitre.

Figure 2. Anatomie des organes de la parole (gauche) et sa représentation mécanique (droite).



i. La soufflerie

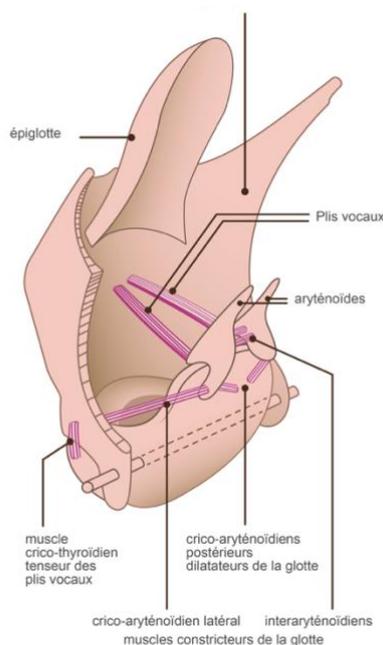
L'air est la matière première de la voix. Il fournit l'énergie nécessaire à la production d'un son. A cette fin, l'appareil respiratoire (étage sous-glottique) agit comme une soufflerie. Celle-ci est constituée de trois parties : le diaphragme, les poumons et la trachée.

Le diaphragme est le muscle inspirateur principal situé en dessous des poumons, séparant le thorax de l'abdomen. Il est le principal moteur d'entrée et sortie de l'air, dont la mécanique s'apparente à celui d'un piston ; son rôle physiologique est donc fondamental : sa contraction permet de créer une dépression dans la cavité thoracique afin de permettre l'entrée d'air via la trachée dans les voies respiratoires (bronches et poumons) ; son relâchement, à l'inverse, fait diminuer la capacité volumique des poumons en expirant l'air, qui reprend le chemin inverse pour aborder le larynx.

ii. La source

Le larynx (étage glottique) est le générateur des sons de la parole. Il est composé des plis vocaux, de l'épiglotte et des cartilages qui leur servent de support et les protègent (Figure 3).

Figure 3. Anatomie du larynx.



L'armature du larynx est constituée de trois cartilages : le cartilage cricoïde, dernier anneau de la trachée, les cartilages aryténoïdes, sur lesquels s'insèrent les plis vocaux, et le cartilage thyroïde, dont la partie antérieure correspond à la pomme d'Adam (présente seulement chez les hommes). Chaque pli vocal est constitué d'une superposition de deux muscles et d'un ligament vocal. Chaque ligament va s'insérer sur la partie antérieure du cartilage thyroïde à l'apophyse des cartilages aryténoïde, longés par des muscles sur lesquels on peut agir ; le tout est recouvert d'une muqueuse qui constitue la partie vibrante produisant le son. La glotte correspond à l'espace entre les deux plis vocaux. Enfin, l'épiglotte, qui surmonte les plis vocaux et la glotte, agit comme un clapet lors de la déglutition pour empêcher un passage des aliments dans la trachée vers les poumons.

Lors de la production d'un son, l'accolement et la tension des plis vocaux va constituer une résistance à l'air expulsé des poumons, entraînant une élévation de la pression en amont de la glotte (i.e. pression sous-glottique) : l'air, en essayant de sortir de la glotte, fera alors vibrer les plis vocaux. Cette vibration correspond à un éloignement et rapprochement périodique (i.e. cyclique) des muscles superposés des plis vocaux, comme une sorte de battement par analogie aux hanches d'un hautbois. Lorsqu'elles sont en vibration, il y a production d'un son qui est dit voisé ou sonore. Par exemple, les /a/ de « *paranoïa* », le /v/ de « *vodka* » ou encore le /b/ de « *bistrot* ». Par contraste, on dit qu'un son est non-voisé ou sourd lorsqu'il n'y a pas vibration des plis vocaux. Par exemple, le /s/ de « *sanglot* », ou encore le /ʃ/ de « *chanvre* ».

Du point de vue physio-anatomique, c'est le nombre de vibrations (ou d'oscillations) des plis vocaux par seconde qui détermine la hauteur vocale. Plus précisément, l'onde sonore produite au niveau de la glotte correspond à un signal périodique complexe, qui par l'application de la transformée de Fourier, peut-être décomposé en une somme de composantes simples, les harmoniques, dont les fréquences sont des multiples entiers de la fréquence fondamentale (correspondant elle-même au premier harmonique, ou H1). Cette fréquence

fondamentale (notée F0, et exprimée en Hz) correspond perceptivement à la hauteur vocale (ou hauteur tonale) dont la valeur fréquentielle (basse ou élevée), correspond au plan perceptif à un son « *grave* » ou « *aigu* ».

De manière générale, la vitesse de vibration des plis vocaux dépend de leur longueur, de leur épaisseur, de leur tension et de la pression sous-glottique exercée. La fréquence fondamentale est inversement proportionnelle à la longueur des plis vocaux et proportionnelle à la racine carrée du rapport entre la tension subie par les plis vocaux et leur densité (Titze, Riede, & Mau, 2016). Elle peut être obtenue à partir de l'équation suivante :

$$F0 = \frac{1}{2L} \sqrt{\frac{\sigma}{\rho}}$$

Où F0 est la fréquence fondamentale (Hz), L la longueur des plis vocaux (cm), σ la tension des plis vocaux (force par unité de surface) et ρ la densité du tissu (1.02 g/cm³). Dans ce contexte, un son relativement aigu sera produit par des plis vocaux plus courts, moins denses et plus tendus, ce qui produit une vitesse de vibration plus rapide (haut F0). A l'inverse, un son relativement grave sera produit par des plis vocaux plus longs, plus denses et moins tendus, aboutissant à une vitesse de vibration plus lente (bas F0).

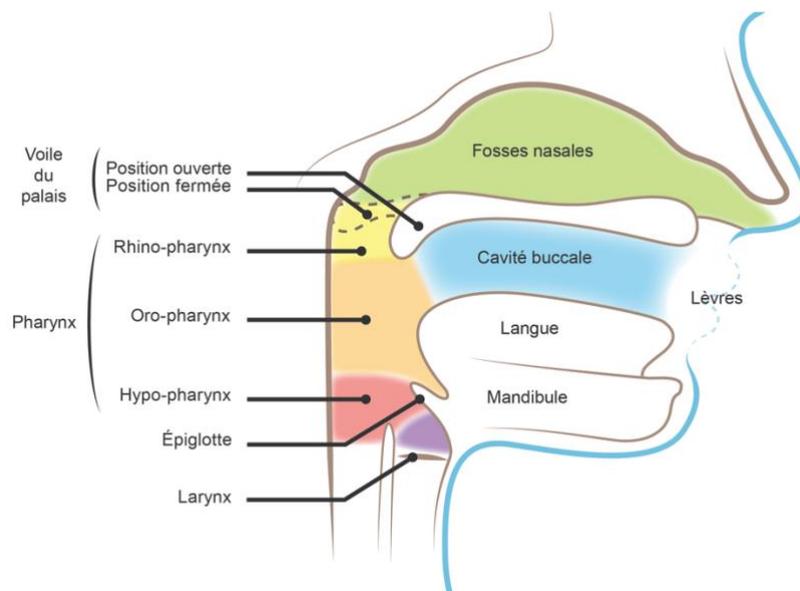
Enfin, et de manière importante, les hommes présentent des plis vocaux plus longs et plus épais que ceux des femmes à l'âge adulte, de l'ordre de 17 à 25 mm chez les premiers et de 12 à 17 mm chez les seconds (Titze, 1994). Ces différences sexuelles entraînent, du moins chez des locuteurs français, une différence de fréquence fondamentale variant entre 100 et 150 Hz chez les hommes, et entre 140 et 240 Hz chez les femmes (Vaissière, 2015)¹. Les hommes possèdent ainsi des hauteurs vocales plutôt grave par rapport aux femmes.

¹ Pour les valeurs moyennes des différents paramètres acoustiques que nous présentons au cours de ce chapitre, nous avons donné, lorsque cela est possible, les valeurs observées chez des locuteurs français. Parfois, certaines valeurs ne sont pas bien connues ou même indisponibles chez ceux-ci. Dans ce contexte, nous indiquons dans quelle langue/culture elles ont été mesurées.

iii. Le filtre

L'onde glottique produite au niveau du larynx va ensuite être « filtrée » par différents éléments appelés résonateurs et situés à l'étage supra-glottique. L'ensemble constitue le conduit vocal et est formé des cavités pharyngale, buccale et nasale (Figure 4). L'articulation des différents sons de la parole se réalise via les changements de forme et de volume de ces différentes cavités.

Figure 4. Anatomie de l'étage supra-glottique. Les résonateurs comprennent le pharynx (dont l'hypo-pharynx, l'oro-pharynx et le rhino-pharynx), la cavité buccale (contenant les articulateurs) et les fosses nasales.



Trois résonateurs constituent le « filtre » et participent à la modulation de l'onde glottique :

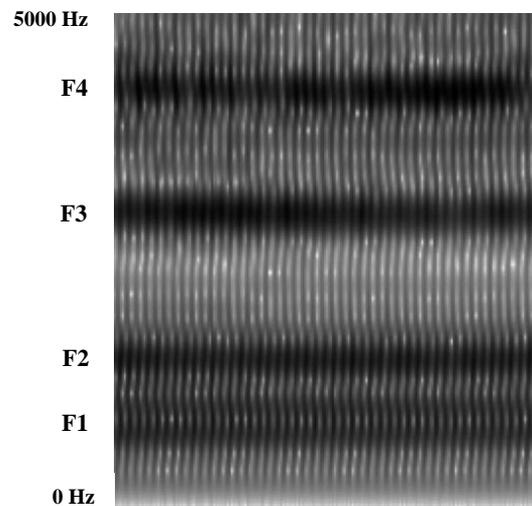
- La cavité pharyngale, située juste au-dessus du larynx et correspondant au pharynx ou « carrefour aéro-digestif ». Ce dernier est un conduit musculaire capable de modifier son diamètre sous l'effet de muscles constricteurs. Son volume est également capable de varier verticalement sous la dépendance des mouvements du larynx. Cette cavité débouche ensuite sur la cavité buccale et peut communiquer avec la cavité nasale.
- La cavité buccale (ou orale), correspondant à la bouche et séparée de la cavité nasale par une cloison appelée le palais. Dans cette cavité se trouvent les articulateurs principaux de la parole que sont la langue, les lèvres et les dents. C'est la partie la plus

importante du conduit vocal car sa forme et sa taille peuvent varier selon les positions relatives du palais par rapport à la langue, de la langue elle-même et des lèvres, et qui sont en grande partie responsable de la production des différents sons de la parole (i.e. voyelles et consonnes).

- La cavité nasale, correspondant aux deux fosses nasales séparées par une cloison verticale médiane et qui sont recouvertes de muqueuses. Cette cavité peut être couplée aux deux précédentes cavités via l'abaissement du voile du palais (ou velum) lors de la phonation.

Au niveau acoustique, l'onde sonore produite au niveau de la glotte va être « *mise en forme* » par les résonateurs, qui vont amplifier (ou atténuer) certaines fréquences, c'est-à-dire en renforcer ou en affaiblir certaines selon la configuration et le volume des résonateurs : on dit de ces derniers qu'ils « *ajustent* » le timbre du son via le phénomène de résonance. Plus précisément, on entend par résonance la modification de la qualité de l'onde glottique qui résulte de l'enrichissement et de l'appauvrissement de certains de ses harmoniques. Lors de la production des voyelles et de certaines consonnes voisées, ces fréquences renforcées sont appelées formants ou fréquences de résonance. Les formants correspondent ainsi à une concentration d'énergie dans une bande de fréquence particulière selon la configuration et la position des organes articulateurs. Ils sont notés F1, F2, F3, F4... etc. (Figure 5). Les formants correspondent ainsi au corrélat acoustique du timbre, et constitue ce que l'on appelle la « *couleur* » d'un son (Vaissière, 2015). Théoriquement, il existe un nombre infini de formants, mais seuls les trois premiers sont nécessaires pour caractériser le timbre d'un son vocalique. Ceux supérieures au F3 correspondent à des caractéristiques individuelles. D'une façon générale, nous savons que plus la cavité est grande, plus grave est le son résultant.

Figure 5. Spectrogramme de la voyelle /a/ comme dans « *doctorat* ». La fréquence est représentée verticalement et le temps horizontalement. A chaque fréquence est associée une énergie (ou intensité), représentée par les bandes noires.



Les formants peuvent être calculés à partir de l'équation suivante (Reby & McComb, 2003) :

$$F_i = \frac{(2i - 1)c}{4LCV},$$

Où F_i correspond au formant i , c la vitesse du son dans l'air (approximé à 350 mètres par seconde pour le conduit vocal d'un mammifère), et LCV est la longueur du conduit vocal (Titze, 1994). Comme la distance entre les formants est la même lorsque le conduit vocal est ouvert ou fermé au niveau de la glotte, l'espacement des formants, c'est-à-dire la distance entre chaque formant adjacent, est aussi utilisé comme mesure générale du timbre.

L'espace entre deux formants consécutifs dans le spectre des fréquences est donc une constante et est donné par :

$$\Delta F = F_{i+1} - F_i = \frac{c}{2LCV}$$

En remplaçant $\frac{c}{2LCV}$ par ΔF dans la première équation, nous obtenons :

$$F_i = \frac{(2i - 1)}{2} \Delta F$$

Nous pouvons ainsi obtenir une estimation de la longueur du conduit vocal à partir de l'équation suivante :

$$LCV = \frac{c}{2\Delta F}$$

Nous remarquons ainsi que ΔF (c'est-à-dire l'espacement entre les formants) et la longueur du conduit vocal sont inversement corrélés : plus le conduit est allongé et grand, plus les formants sont proches les uns des autres. Par ailleurs, il est à noter que plusieurs mesures de l'espacement des formants autre que ΔF ont été proposées. Nous en présenterons deux ici, les plus communément utilisés en bioacoustique animale (humains compris) et que nous avons étudiés au cours de cette thèse.

La première est la « *dispersion des formants* » (i.e. Df), et a été proposée par Fitch (1997). Elle est définie comme la somme des distances adjacentes entre deux formants consécutifs, divisée par le nombre de formants mesurés. Elle est obtenue à partir de l'équation suivante :

$$Df = \frac{\sum_{i=1}^{N-1} F_{i+1} - F_i}{N - 1},$$

Où Df est la dispersion des formants (Hz), F_i la fréquence du formant i (Hz) et N le nombre de formants mesurés.

La deuxième mesure est la « *position des formants* » (i.e. Pf), et a été proposée par Puts, Apicella & Cardenas (2012). Elle est définie comme la somme des moyennes des formants standardisés pour les N formants mesurés, où la standardisation est réalisée à partir des moyennes non pondérées et des écarts-types combinés des deux sexes. Elle est obtenue à partir de l'équation suivante :

$$Pf = \frac{\sum_{i=1}^N F'_i}{N},$$

Où F'_i est le $i^{\text{ème}}$ formant standardisé et N le nombre de formant mesuré.

Cette dernière mesure a été proposée en réponse aux supposées inconsistances dans les associations entre Df (la dispersion des formants) et les caractéristiques physiques, physiologiques et comportementales d'un individu, que nous présenterons dans les chapitres ultérieurs. Par exemple, si nous mesurons 4 formants, Puts et al. (2012) argumentent que Df ne correspond finalement qu'à une mesure de la distance entre le 1^{er} et le 4^{ème} formant (F1 et F4), et ne contient en somme aucune information sur les 2^{ème} et 3^{ème} formants (F2 et F4). De plus, la variance associée au F4 est beaucoup plus élevée que celle du F1 et, comme les formants ne sont pas standardisés, la mesure sera particulièrement biaisée par la mesure du F4. Dans ce contexte, Df porte de l'information sur l'espacement, mais non pas sur la position même des formants.

Généralement, si les timbres présentent une distribution des formants situés dans les basses fréquences, ils sont plutôt rapprochés et nous parlons de timbre « *graves* » (ou « *compact* »). A l'inverse, si des timbres présentent une distribution des formants situés dans les hautes fréquences, ils sont plutôt éloignés les uns des autres et nous parlons de timbres « *aigus* » (ou « *diffus* »).

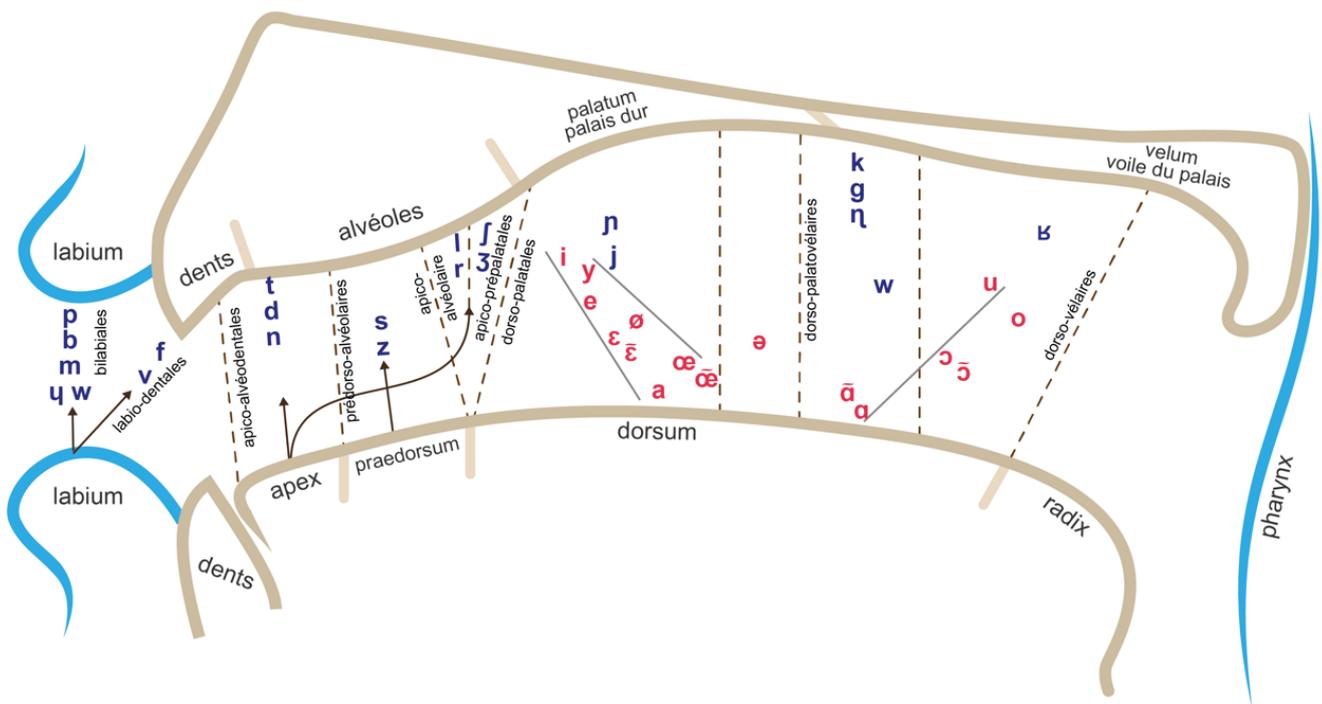
Enfin, et de manière importante, les hommes possèdent un conduit vocal plus long que les femmes à l'âge adulte : la distance moyenne entre la glotte et les lèvres est de 17 à 18 cm chez les hommes, tandis qu'elle est de l'ordre de 14.5 cm chez les femmes (Fitch & Giedd, 1999). Les hommes présentent ainsi des timbres dont les formants sont situés dans les basses fréquences et plutôt rapprochés les uns des autres, à l'inverse des femmes qui présentent des timbres aux formants situés dans les hautes fréquences et plutôt éloignées les uns des autres. Pour les locuteurs du français, les hommes exhibent en moyenne des timbres présentant des F1, F2, et F3 respectivement dans les 390 Hz, 1500 Hz et 2600 Hz, avec des dispersions d'environ 1105 Hz. Par contraste, les femmes exhibent en moyenne des timbres présentant des F1, F2, et F3 respectivement dans les 400 Hz, 1900 Hz et 3100 Hz, avec des dispersions d'environ 1350

Hz (Pépiot, 2013). Les hommes exhibent ainsi en moyenne des timbres plutôt graves par rapport aux femmes.

iv. Production des voyelles

Tous les sons de la parole humaine et leurs lieux d'articulation dans la cavité orale sont représentés dans la Figure 6, qui nous servira de canevas pour illustrer leurs productions.

Figure 6. Points d'articulation des sons du français dans la cavité orale. Les voyelles sont représentées en rouge et les consonnes en bleu.



Dans le cas des voyelles, chacune est produite selon une configuration particulière des articulateurs de la cavité orale et présentera ainsi une structure formantique qui lui est propre. La connaissance des deux (voire des trois) premiers formants permet d'identifier toutes les voyelles. Selon les mécanismes articulatoires que nous avons décrits, les voyelles peuvent être représentées classiquement sur un axe en deux ou éventuellement trois dimensions (Figure 7).

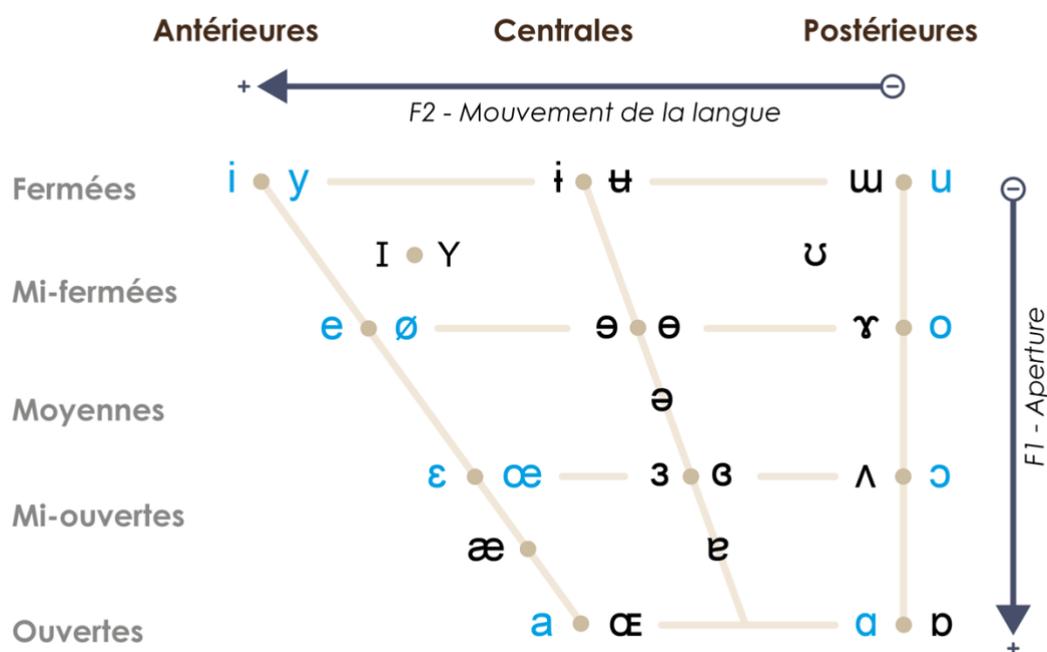
Chaque axe correspond à une position particulière d'un des articulateurs de la cavité orale :

- Le premier formant (F1) correspond à l'aperture de la bouche et chaque voyelle est définie selon le degré d'ouverture de la cavité orale. Nous parlons ainsi de voyelle fermée lorsque ce degré est minimal (i.e. bouche presque fermée), comme la voyelle /u/ de « *saoul* » ou la voyelle /i/ de « *whisky* ». A l'inverse, lorsque le degré est maximal (i.e. bouche ouverte), nous parlons de voyelle ouverte, comme le /a/ de « *pâte* » ou le /a/ de « *patte* » (du moins chez les locuteurs faisant la distinction entre « *pâte* » et « *patte* », comme les locuteurs lyonnais). Entre ces deux degrés, nous pouvons définir les voyelles semi-fermées, telle que /o/ de « *cocaïne* » ou le /e/ de « *tequila* », et les voyelles semi-ouvertes telle que /ɔ/ de « *porto* » ou le /ɛ/ de « *bière* ». Il convient de noter que plus le degré d'ouverture est grand, plus la fréquence du F1 est élevée.
- Le deuxième formant (F2) correspond au point d'articulation, et plus précisément à la position de la langue dans la cavité orale. Lorsque la langue est située à l'avant de la cavité, la voyelle est dite antérieure, telle que le /i/ de « *philosophie* » ou le /y/ de « *ubuesque* ». Lorsque la langue est située à l'arrière de la cavité orale, la voyelle est dite postérieure, comme c'est le cas pour la voyelle /u/ de « *fou* ». Une position centrale de la langue (i.e. ni à l'avant ni à l'arrière) conduit à la production d'une voyelle dite centrale, comme le /a/ de « *absinthe* » ou le /ə/ de « *dangereux* ». Entre la position antérieure et centrale, nous pouvons observer les voyelles mi-fermées comme le /e/ de

« *saké* » ou les voyelles mi-ouvertes comme le /ɛ/ de « *infernal* ». Plus la langue est en position avancée dans la cavité, plus la fréquence du F2 est élevée.

- Le troisième formant (F3) correspond à la configuration des lèvres et notamment à l'arrondissement de celles-ci lors de la production d'un son. Lorsque les lèvres sont protruses, nous parlons de voyelle arrondie ; quand elles ne le sont pas la voyelle est dite non-arrondie. Quant à l'inverse de l'arrondissement, les lèvres sont étirées au maximum, nous parlons de voyelle étirée (e.g. /i/ ou /e/). Ainsi les voyelles /y/ de « *bureaucratie* » ; /u/ de « *fou* » ; /o/ de « *sottise* » ; /ɔ/ de « *rhum* » ; /ø/ de « *vœu* » et /œ/ de « *jeunesse* » sont dites arrondies. La valeur du F3 contribue à enrichir le spectre des sons concernés dans les graves dans la mesure où elle est liée à un allongement du conduit phonatoire.

Figure 7. Représentation des voyelles sur deux axes selon la manière dont elles sont produites, ici l'aperture (ou degré d'ouverture) de la cavité orale, et le mouvement (ou position) de la langue dans celle-ci. Une troisième dimension peut être représentée par la protrusion des lèvres. Pour chaque paire de voyelles, la voyelle à gauche est non-arrondie et à droite se situe sa version arrondie. Les voyelles du français sont en bleu.



La classification des voyelles orales est donc établie sur chacune des dimensions suivante : l'aperture (i.e. voyelle fermée ou ouverte), le lieu d'articulation (i.e. voyelle antérieure ou postérieure) et l'arrondissement (i.e. voyelle arrondie ou non-arrondie et/ou étirée). Ainsi, la voyelle /u/ est définie comme une voyelle postérieure, fermée et arrondie, la voyelle /a/ comme une voyelle centrale, ouverte et non-arrondie et la voyelle /i/ comme une voyelle antérieure, fermée et étirée.

Notons qu'il existe également dans la langue française des voyelles nasales. Leur production correspond à un passage de l'air simultané par la cavité orale et le nez. La production de ce type de sons entraîne une résonance particulière dans les fosses nasales et le couplage des deux cavités de résonance (orale et nasale) se fait grâce à l'abaissement du voile du palais. Il existe en français quatre voyelles nasalisées, ayant chacune un pendant oralisé : le /ẽ/ de « *important* » par rapport au /ɛ/ de « *thèse* », le /œ̃/ de « *inopportun* » par rapport au /œ/ de « *jeune* »¹, le /ɔ̃/ de « *bon* » par rapport au /ɔ/ de « *sortie* » et le /ã/ de « *antécédent* » par rapport au /ɑ/ de « *pâtauger* ». La différence entre chaque paire réside essentiellement dans la position relevée contre abaissée du voile du palais.

Enfin, nous pouvons aussi définir une classification des voyelles plus ou moins graves et aiguës selon la fréquence et la distance entre les deux ou trois premiers formants. Si l'on s'en tient à une classification des sons vocaliques en termes d'acuité, les voyelles du français s'échelonne selon la distribution suivante, de la plus grave à la plus aiguë :

/u/ > /o/ > /ɔ/ > /ɑ/ > /a/ > /œ/ > /ø/ > /y/ > /ɛ/ > /e/ > /i/

Concernant les voyelles nasales, les descriptions de Delvaux & Metens (2002) ont montré un rapprochement des deux premiers formants ainsi que l'existence d'anti-formants témoignant d'une résonance supplémentaire au niveau des fosses nasales.

¹ Souvent en français « parisien », le /œ̃/ est remplacé par /ẽ/.

Notons enfin que toutes les voyelles présentent ce que l'on appelle une fréquence fondamentale intrinsèque (IF0, ou hauteur intrinsèque), qui correspond au F0 propre à la production d'une voyelle particulière, indépendamment des facteurs liés à la variabilité des locuteurs.

v. Production des consonnes

Si les voyelles sont produites avec la vibration des plis vocaux et un échappement libre de l'air dans le conduit supra-laryngé, les consonnes présentent un profil articulatoire et spectral beaucoup plus variables (Meynadier, 2013). D'une façon générale, les consonnes sont définies comme des bruits, dont la production est caractérisée selon deux dimensions : 1) le mode d'articulation, qui dépend de la façon dont l'air s'écoule du conduit vocal à la sortie des cavités supra-laryngées, celui-ci pouvant s'échapper de façon continu ou pouvant être expulsé brutalement après que le conduit soit totalement obstrué, et 2) le lieu d'articulation, qui définit l'endroit où a lieu l'obstruction et/ou le rétrécissement du conduit vocal. Comme pour les voyelles, le lieu d'articulation correspond généralement à la position de la langue dans la cavité orale et à la manière dont elle va se placer dans celle-ci pour obstruer le passage de l'air.

De manière générale, les consonnes produites avec un écoulement continu de l'air sont produites dans de plus hautes fréquences que celles produites après un relâchement soudain de ce dernier. Ceci est dû à leur mode de production : l'écoulement de l'air provoque du bruit dans les hautes fréquences, par opposition à celles produites avec un bruit impulsionnel après accumulation de l'air en amont de la constriction, augmentant la pression aérodynamique à l'intérieur du conduit vocal. De même, toute consonne voisée est produite dans de plus basses fréquences, et donc perçues comme plus grave que sa version non voisée. Ceci s'explique par le fait que la vibration des plis vocaux diminuent la pression supra-glottique, une part de l'énergie étant dévolue à la vibration des plis vocaux.

Pour les modes articulatoires, nous distinguons en français les types suivants :

- Les occlusives, produites en deux temps avec 1) fermeture complète de la cavité orale en un point et 2) séparation des articulateurs avec un soudain relâchement de l'air. En français, nous distinguons le /p/ de « *post-doc* », le /b/ de « *bière* » (forme voisée de /p/), le /t/ de « *thèse* », le /d/ de « *doctorat* » (forme voisée de /t/), le /k/ de « *cannabis* » et le /g/ de « *galère* » (forme voisée de /k/).
- Les constrictives, produites par constriction ou rétrécissement du conduit vocal, conduisant à une turbulence du flux d'air à la sortie de ce dernier. En français, nous distinguons le /f/ de « *fumer* », le /v/ de « *vodka* » (forme voisée de /f/), le /s/ de « *stupéfiant* », le /z/ de « *zut* » (forme voisée de /s/), le /ʃ/ de « *malchance* », le /ʒ/ de « *jeunesse* » (forme voisée de /ʃ/), et le /ʁ/ de « *rhum* » (forme voisée de /ʁ/, absente du système phonologique français).
- Les occlusives nasales, produites de la même façon que les occlusives orales auquel se surimpose un abaissement du palais, avec un écoulement de l'air simultané par les cavités orales et nasales. En français, les consonnes nasales sont toutes voisées et nous distinguons le /m/ de « *marijuana* », le /n/ de « *sniffer* », le /ɲ/ de « *gnôle* » et le /ŋ/ de « *zapping* ».
- Les latérales, produites par un écoulement de l'air qui s'échappe par les côtés de la langue. Il en existe une seule en français, noté /l/ comme dans « *laideur* ».
- Les semi-consonnes ou semi-voyelles, correspondant à des sons qui sont phonétiquement similaires aux voyelles mais qui fonctionnent en tant que frontière au noyau de la syllabe (qui elle est nécessairement une voyelle). En d'autres termes, ce sont des voyelles non-syllabique, c'est-à-dire que leur point d'articulation varie au cours de l'articulation avec la voyelle syllabique qui suit. Nous distinguons le /j/ de « *fioriture* », le /ɥ/ de « *fuire* » et le /w/ de « *moi* ».

Pour les lieux d'articulations, nous distinguons en français les type suivants :

- Les bilabiales : les deux lèvres sont en contacts telles que [p b m].
- Les labiodentales : la lèvre inférieure s'appuie contre les incisives supérieures telles que [f v].
- Les alvéolaires : la pointe de la langue s'appuie contre les alvéoles dentaires telles que [l t d n s z].
- Les post-alvéolaires : resserrement entre l'avant de la langue et les alvéoles, dont la pointe de la langue peut se trouver derrière les incisives inférieures telle que /ʃ/, ou derrière les alvéoles telle que /ʒ/.
- Les vélares : le dos de la langue est en contact avec le palais telle que [k g ŋ].
- Les uvulaires : la racine de la langue est rapprochée de la luette, ce qui peut provoquer une constriction /ʁ/.
- Les palatales : le milieu ou le dos de la langue s'appuie contre le palais rigide telles que [ɲ j].

vi. La qualité vocale

Nous avons jusqu'ici décrit les paramètres acoustiques considérés comme les plus pertinents pour la description de la parole humaine. Toutefois, la voix d'un individu peut être caractérisée au-delà de sa hauteur et de son timbre. Comme nous le verrons au cours de cette thèse, d'autres composantes acoustiques d'une voix, notamment le bruit et la prosodie, sont peu étudiés alors même qu'ils peuvent jouer un rôle relativement important pour l'attractivité vocale. Dans ce contexte, il convient de définir la notion complexe de qualité vocale pour nous permettre d'introduire d'autres éléments constitutifs de la parole humaine.

Bien qu'il existe de multiples définitions et qu'aucun consensus n'existe à ce jour pour définir la qualité vocale, nous nous appuyerons sur la définition donnée par Trask (2012). Selon l'auteur, la qualité vocale correspond à la coloration auditive caractéristique de la voix d'un

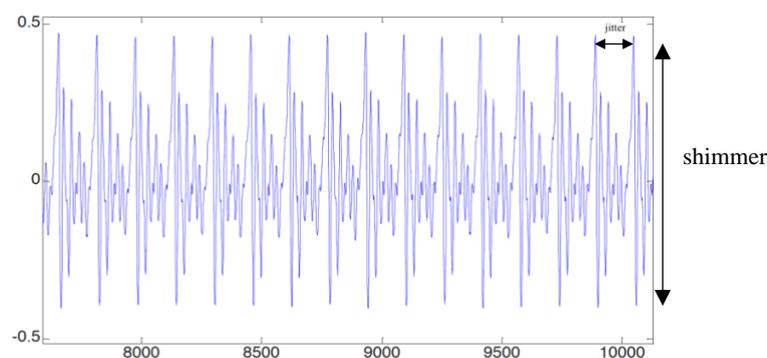
individu, laquelle est dérivée d'un ensemble de caractéristiques vocales propre à chaque locuteur. Elle englobe ainsi la hauteur, le timbre, mais également l'intensité, l'articulation, le bruit et des éléments prosodiques tels que l'intonation et le débit.

a. Le bruit

L'étude des dysphonies a permis de mettre en évidence d'autres critères utiles pour la caractérisation de la voix au-delà de son caractère grave ou aiguë. Nous pensons particulièrement aux voix « *soufflées* » et/ou « *rauques* » dont la spécificité réside dans la quantification du bruit dans le signal sonore.

Parmi eux, le jitter et le shimmer sont deux paramètres acoustiques mesurant la « *perturbation* » de la fréquence fondamentale. Si un signal sonore est produit par vibration des plis vocaux et par l'oscillation cyclique de celles-ci, le jitter correspond aux perturbations dans la régularité des vibrations des plis vocaux d'un cycle à l'autre, tandis que le shimmer correspond à l'amplitude de cette variation (Figure 8, Teixeira, Oliveira, & Lopes, 2013). Le jitter est principalement lié à un mauvais contrôle des plis vocaux tandis que le shimmer dépend de la réduction de la résistance glottale lors de la phonation. Les patients avec des voix présentant des pathologies vocales (e.g. lésions des plis vocaux) en présentent des valeurs élevées.

Figure 8. La vibration des plis vocaux oscille cycliquement, et chaque cycle correspond à ce que l'on appelle une pulsation glottique. Le jitter correspond ainsi aux perturbations d'une oscillation à l'autre, et le shimmer à la perturbation de l'amplitude du cycle. Figure emprunté de Teixeira et al. (2013).



Il existe de nombreuses mesures du jitter, mais nous considérerons ici le jitter local (ou jitter ratio), communément utilisé pour décrire les voix pathologiques. C'est une mesure des perturbations à « *court terme* » du F0. Il s'agit de la moyenne de toutes les différences, en valeur absolue, entre les durées de deux périodes consécutives d'un signal. On divise ensuite cette moyenne par sa durée. Le résultat est donc un rapport, exprimé ici en %. Le jitter peut être considéré comme le corrélat acoustique de la « *raucité vocale* » (Hillenbrand, 1988 ; Rabinov, Kreiman, Gerratt, & Bielamowicz, 1995 ; Wendahl, 1966)¹. Le shimmer ne sera pas inclus dans nos travaux car il semblerait que le jitter et le shimmer ne soient pas des dimensions simultanément utiles à considérer en tant qu'indices de la qualité vocale perçus par des auditeurs. En effet, bien qu'ils soient concordants sur le niveau global du bruit contenue dans une voix, les auditeurs sont fortement discordants sur les valeurs qui y contribuent (Kreiman & Gerratt, 2005). De plus, ces deux mesures sont fortement corrélées (Hillenbrand, 1987) et donc non indépendantes.

Hormis le jitter et le shimmer, un autre paramètre acoustique lié au bruit est tout aussi pertinent pour décrire la qualité vocale. Celui-ci est défini comme le rapport harmoniques-bruit (HNR pour « *harmonics-to-noise ratio* »), qui représente un ratio entre des composantes périodiques, ici les harmoniques, et une composante non-périodique, le bruit. Le premier résulte de la vibration des plis vocaux (pour rappel, les harmoniques sont des multiples entiers du F0, qui sont ensuite filtrées dans les cavités supra-glottiques), et le second résulte du bruit au niveau de la glotte. Le HNR s'exprime en décibels (dB) et reflète « *l'efficacité* » de la production de la parole : plus le souffle est transformé avec succès en énergie pour la vibration des plis vocaux, plus le HNR sera élevé. La voix sera alors perçue comme plus riche en harmoniques. A l'inverse, si ce ratio est peu élevé, le souffle ne se transforme pas en énergie et passe via la fente

¹ Des voix présentant des valeurs en deçà de 0.5% ou au-delà de 1% sur des voyelles tenues (i.e. production d'une voyelle le plus longtemps possible par un locuteur) sont considérées comme pathologiques (Teixeira et al., 2013). Toutefois comme nous le verrons, avec des locuteurs ne présentant pas de pathologie vocale et sur de la parole spontanée, les valeurs de jitter deviennent bien plus élevées.

glottique, ce qui augmente la composante du bruit. Nous parlons alors de fuite glottique. Dans ce contexte, le HNR peut être considéré comme le corrélat acoustique du « *souffle vocal* » (Hillenbrand & Houde, 1996 ; Krom, 1995)¹.

Enfin, et de manière importante, les hommes semblent exhiber des voix ayant des valeurs de jitters plus élevées que les femmes, du moins chez des locuteurs canadiens (Goy, Fernandes, Pichora-Fuller, & van Lieshout, 2013), et dans une méta-analyse incluant des locuteurs canadiens, écossais et allemands (Pisanski, Jones, Fink, O'Connor, DeBruine, Röder, & Feinberg, 2016). De plus, Il a été montré que les hommes exhibent également des voix avec des HNR moins élevés, entre 11 et 13 dB, contre 13 et 15 dB chez les femmes, du moins chez des locuteurs japonais et anglais (Heffernan, 2004). Il semblerait ainsi que les hommes ont des voix plus rauques et plus soufflées que les femmes.

b. La prosodie

Si les éléments segmentaux de la parole se réfèrent à l'aspect articulatoire et acoustique des sons, les éléments supra-segmentaux relèvent de la prosodie. Comme la notion de qualité vocale, il faut d'emblée noter que de multiples définitions existent et qu'aucun consensus définitif n'existe à ce jour. Toutefois, nous retiendrons la définition de Di Cristo (2000), qui nous semblent appropriée dans le contexte de ce travail de thèse :

« La prosodie est une branche de la linguistique consacrée à la description (phonétique) et à la représentation formelle (aspect phonologique) des éléments de l'expression orale tels que les accents, les tons, l'intonation et la quantité, dont la manifestation concrète dans la production de la parole, est associée aux variations de la fréquence fondamentale (F0), de la durée et de l'intensité (paramètres prosodiques physiques). Ces variations étant perçues par

¹ Les voix sont considérées comme pathologiques lorsqu'elles présentent des valeurs de HNR inférieures à 7 dB (Teixeira et al., 2013).

l'auditeur comme des changements de hauteur (ou de mélodie), de longueur et de sonie (paramètres prosodiques subjectifs). »

En d'autres termes, la prosodie désigne les phénomènes liés à la variation de l'intonation portée par l'évolution à moyen et court terme du F0, de l'intensité et du rythme de la parole¹. Dans le cadre de cette thèse, nous nous sommes intéressés à la micro-intonation et au débit ou rythme de la parole.

La micro-intonation correspond aux variations de la fréquence fondamentale dans la chaîne parlée (noté F0-SD, SD pour « *standard-deviation* » en anglais, ou déviation standard en français). D'une façon générale, les micro-variations de l'intonation au sein d'un énoncé permettent de percevoir les intentions et/ou l'état émotionnel d'un individu (Banse & Scherer, 1996). Au niveau perceptuel, les voix sont ainsi perçues comme « *monotones* » et/ou « *froides* » lorsque le F0-SD est bas, et plutôt comme « *dynamiques* », « *enjouées* », « *vivantes* » et/ou « *expressives* » lorsqu'il est élevé (Hodges-Simeon, Gaulin, & Puts, 2011). Le débit de parole (rythme ou « *speaking rate* » en anglais) est lui défini comme le nombre de syllabes par seconde dans la chaîne parlée, incluant les pauses (i.e. moments de silence). Perceptuellement, le F0-SD peut être assimilé à la mélodie en musique, tandis que le rythme reflète le tempo.

Enfin, et de manière importante, dans une méta-analyse incluant des locuteurs anglophones et slaves, il a été montré que les hommes présentent des voix avec des F0-SD d'environ 15 Hz, contre 30 Hz chez les femmes (Pisanski et al., 2016). Concernant le débit, une

¹ Les phénomènes accentuels relèvent eux aussi de la prosodie de la langue. Ils englobent les différents types d'accents toniques qui peuvent – selon les langues – relever du niveau phonétique (comme c'est le cas par exemple en français) que du niveau phonologique (comme dans toutes les langues dites accentuelles, comme l'anglais ou l'espagnol). Dans les langues accentuelles, l'accent tonique (noté ['] avant la syllabe concernée) peut, au même titre que les phonèmes, modifier le sens (plus précisément la catégorie grammaticale) d'un mot. Par exemple, en anglais, la séquence ['prɪzənt] accentuée sur la première syllabe correspond au nom ou à l'adjectif « (*un*) *présent* » quand la séquence [prɪ'zænt] accentuée sur la seconde syllabe signifie « *présenter* » (verbe). Le rôle fonctionnel de l'accent en français n'est pas comparable dans la mesure où celui-ci est fixe (toujours situé sur la dernière syllabe du mot) à moins d'un effet stylistique particulier (i.e. accent d'emphase). Nous ne nous attarderons pas plus longuement sur les aspects accentuels ici, néanmoins, il convient de garder en mémoire ces éléments pour la suite de notre propos (cf section VIII.3.), que contrairement à ce qui est observé en anglais les noms où la syllabe préminente est située à gauche, en français la syllabe accentuée est à droite (i.e. l'accent tonique à droite).

étude a montré que les hommes français et américains parlent significativement plus vite que les femmes (Pépiot, 2013).

vii. Ontogénie de la voix

Après avoir décrit l'ensemble des mécanismes de la production de la parole et les différents critères acoustiques et perceptuels qui en découlent, nous allons rapidement décrire comment l'appareil vocal se développe.

Tout comme chez les chimpanzés adultes, chez le nouveau-né humain le larynx est placé très haut dans le cou en comparaison à ce qui est observé chez l'adulte. Conséquemment, les voies respiratoires et digestives sont séparées chez les primates non-humains et le nourrisson, ce qui leur permet de respirer et de déglutir simultanément. Néanmoins, si la position haute du larynx présente certains avantages (i.e. les singes et les bébés peuvent respirer par le nez tout en continuant de s'alimenter), elle limite aussi sévèrement le répertoire des vocalisations (Lieberman, Harris, Wolff, & Russell, 1971). Lors des premiers mois de la vie, le larynx va progressivement descendre dans le cou, jusqu'à un changement majeur entre la deuxième et la troisième année où les voies respiratoires et digestives s'inter-sectionnent dans le pharynx pour créer une unique voie. Cette première descente du larynx permet d'allonger le conduit vocal, entraînant une modification conséquente du répertoire vocalique que peut produire l'enfant. La longueur des cordes vocales à la naissance mesure environ 4.5 à 5 mm, jusqu'à atteindre la taille de 8 mm environ à 5 ans, avec des voix présentant des F0 autour des 300 Hz. De manière générale, il semblerait y avoir très peu de différences au niveau de l'anatomie vocale et donc acoustique (du moins concernant le F0 et les formants) entre la voix d'un garçon et celle d'une fille d'un même âge à la préadolescence (Simpson, 2009 ; Lieberman, McCarthy, Hiiemae, & Palmer, 2001).

A la puberté, l'effet des androgènes lié à la maturité sexuelle va induire des modifications sexuellement dimorphiques, résultant en une différence notable des paramètres

acoustiques entre les deux sexes. En effet, à la puberté, les taux d'androgènes chez les garçons provoquent le développement des caractères sexuels secondaires, dont l'appareil vocal (Richmond & Rogol, 2007). Plus particulièrement sous l'influence de la dihydrotestostérone – un métabolite de la testostérone –, les plis vocaux s'allongent d'environ 60% par rapport à leur longueur initiale et vont s'épaissir de manière plus ou moins permanente au cours de la vie adulte en fonction de différents facteurs environnementaux. Chez les filles, l'effet de la puberté sur la configuration de l'appareil phonatoire est moindre. Les plis vocaux vont eux aussi connaître un allongement significatif mais celui-ci est d'environ 30% (Jenkins, 1998). Rappelons qu'à l'âge adulte, la longueur des plis vocaux est de 17 à 25 mm chez l'homme et de 12 à 17 mm chez la femme (Titze, 1994).

En parallèle et simultanément à l'influence des hormones, une seconde descente du larynx a lieu. Une étude de résonance magnétique chez des participants âgés de 2.8 ans à 25 ans a révélé que la longueur moyenne du conduit vocal diverge significativement après 15 ans entre les deux sexes : les mâles post-pubères possèdent en moyenne un conduit vocal long de plus de 17 mm comparé à celui des femmes (Fitch & Giedd, 1999). A ce même âge, nous observons également une croissance du cartilage thyroïdien (i.e. la pomme d'Adam) particulièrement proéminente chez les hommes. Cette seconde descente du larynx va également – toujours chez les hommes – diminuer les fréquences des formants et leur espacement. Rappelons ainsi qu'après la puberté, les hommes possèdent un conduit vocal plus long : la distance moyenne entre la glotte et les lèvres étant de l'ordre de 17 à 18 cm chez les hommes, contre environ 14.5 cm chez les femmes. Les androgènes sont ainsi considérés comme le mécanisme proximal à l'origine des différences anatomiques et acoustiques entre la voix des hommes et des femmes.

De manière intéressante, deux récentes études longitudinales ont suggéré que les différences interindividuelles pour la fréquence fondamentale sont déterminées très précocement, c'est-à-dire bien avant la puberté. Dans la première étude, il a été montré que le

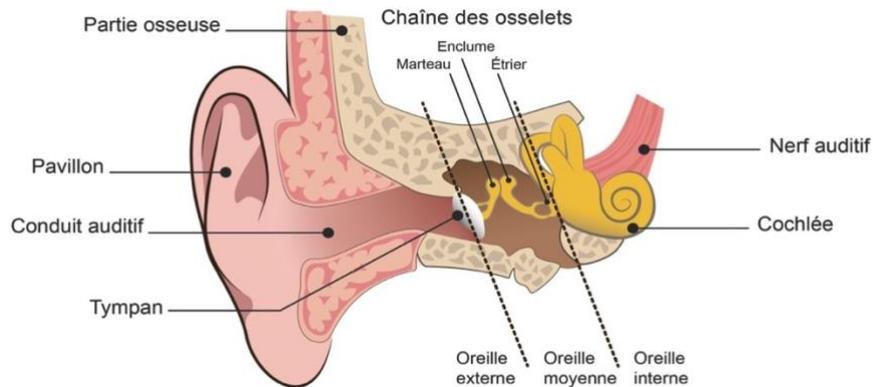
F0 des cris d'enfants de 4 mois prédit le F0 de leurs vocalisations verbales à 5 ans, avec 41% de variance interindividuelle expliquée à cet âge (Levrero, Mathevon, Pisanski, Gustafsson, & Reby, 2018). Les auteurs montrent également que le ratio 2D:4D (ratio entre la longueur de l'index et la longueur de l'annulaire, respectivement influencées par les œstrogènes et la testostérone), connu pour être un proxy de l'exposition aux hormones sexuelles prénatales, est positivement corrélé au F0 à 4 mois et à 5 ans. Dans la seconde étude, il semblerait que, au moins chez les garçons, le F0 à l'âge de 7 ans prédit 64% de la variance du F0 à l'âge adulte (Fouquet, Pisanski, Mathevon, & Reby, 2016). Bien que les tailles d'échantillons soient discutables dans les deux études (6 filles et 9 garçons dans la première, 10 hommes dans la seconde), ces résultats suggèrent qu'une proportion substantielle de la variation interindividuelle pour le F0 émerge très tôt au cours de la vie et est peut-être même déterminée avant la naissance.

viii. Système auditif

Si nous avons jusqu'ici décrit les mécanismes de production de la parole, nous allons maintenant la décrire d'un point de vue perceptif. En effet, l'intérêt de la communication vocale réside dans la production de signaux acoustiques par un émetteur à l'intention d'un ou plusieurs auditeurs. La parole doit ainsi s'étudier à la fois d'un point de vue de la production que de la perception.

Le fonctionnement de l'oreille humaine sollicite trois parties du système auditif (Figure 9).

Figure 9. Schéma de l'appareil auditif.



Chaque partie a un rôle particulier :

- L'oreille externe composée du pavillon et du conduit auditif externe. Elle a pour principal rôle de collecter et d'envoyer les ondes sonores via le conduit auditif jusqu'au tympan. L'oreille externe sert également de protection à l'oreille moyenne via la présence de cérumen. Le pavillon joue un rôle dans la localisation des sons et l'équilibre binaural.
- L'oreille moyenne est composée du tympan et d'osselets. Elle permet de transformer les vibrations « aériennes » en vibrations « solidiennes » pour les transmettre à l'oreille interne. Une fois l'onde sonore arrivée au tympan, cette dernière le fait vibrer. Cette vibration actionne la chaîne de trois osselets : le marteau, l'enclume et l'étrier. Ces derniers vont agir comme un « cric » de voiture : ils augmentent l'amplitude de la vibration. L'oreille moyenne sert également de protection à l'oreille interne, notamment en limitant l'amplification des sons trop puissants (au-delà de 80 dB).
- L'oreille interne, remplie de liquide d'où partent les nerfs vers le cerveau. Une fois le son transmis à l'oreille interne, cette dernière doit le transformer en signal nerveux interprétable par le cerveau. L'étrier va venir taper sur l'oreille interne remplie de liquide, la vibration se propage alors jusqu'à la cochlée qui possède des cellules

sensorielles et va capter ces vibrations pour les transformer en signaux nerveux, ensuite transmis au cerveau via le canal auditif.

Le F0 et les formants sont considérés comme les caractéristiques acoustiques les plus saillantes de la voix humaine. Toutefois, il est à noter que la fréquence d'un son n'est pas perçue de manière linéaire par celle-ci. La notion de « *différence juste notable* » (JND, « *just noticeable difference* » en anglais) ou seuil de discrimination, est donc une notion centrale à comprendre. Elle correspond à la plus petite différence entre deux stimulus à partir de laquelle un individu est capable de les différencier. Due à la variabilité inter- et intra-individuelle, l'établissement des seuils de discriminations repose sur une mesure statistique du seuil différentiel obtenu empiriquement au cours de multiples essais pour un même individu.

Les seuils de discrimination peuvent ainsi varier selon plusieurs paramètres liées à l'étude : voix naturelle ou synthétisée, voyelles isolées ou en contexte consonantique, si les auditeurs sont entraînés ou non, etc. Toutefois, sur de la parole naturelle (stimuli vocaux consistant en la prononciation de plusieurs mots), il semblerait que ces seuils soient de l'ordre de 5 à 6 % pour le F0 et les formants de locuteurs hommes et femmes (généralement de l'ordre des 3 Hz, Pisanski & Rendall, 2011 ; Puts et al., 2006). Toutefois, il semblerait nécessaire de multiplier ces seuils différentiels par 2, voire 3 fois par rapport aux valeurs de références pour obtenir des résultats fiables entre participants (Pisanski & Rendall, 2011)¹.

Concernant les paramètres de jitter et de HNR, il n'existe pas à notre connaissance d'étude ayant établi des seuils de discriminations basées sur de la parole naturelle. Toutefois, Hillenbrand (1988) rapporte que pour des voyelles tenues synthétisées, les perceptions de « *raucité* » et de « *souffle* » ne sont pas linéaires. La perception de « *raucité* » suit une relation

¹ Pour le F0, diverses échelles psycho-acoustiques de transformation de la hauteur d'un son en unité perceptible ont été proposées, comme les semi-tons, les Mels, les Barks et les ERB (« *Equivalent Rectangular Bandwidth* »), cette dernière étant la plus couramment utilisée. Les ERB constituent une échelle semi-logarithmique qui contrôle pour la diminution de la saillance perceptuelle du F0 lorsqu'elle augmente. Cette échelle semble mieux rendre compte de la perception qu'ont les individus de la hauteur d'un son qu'une échelle linéaire ou les autres échelles psycho-acoustiques (Nolan, 2003).

polynomiale du troisième degré : elle augmente linéairement jusque des valeurs de jitter à 2%, au-delà de laquelle elle augmente significativement moins pour ensuite se tasser vers l'extrémité de la distribution. Quant à la perception de « *souffle* », elle diminue avec l'augmentation du HNR, mais ne semblent pas augmenter en-dessous des -6 dB et ne diminue plus au-delà des 21 dB.

Concernant la micro-intonation (F0-SD), bien qu'il n'existe pas à notre connaissance d'étude ayant établi des seuils de discriminations, nous pouvons faire l'hypothèse qu'elle se situerait dans les mêmes plages de valeurs que la hauteur vocale puisqu'elle n'en est finalement que sa variation. Enfin, pour ce qui concerne la perception de la variation de débit dans la parole naturelle, la différence juste notable se situerait autour des 5%, indépendamment du sexe des locuteurs (Quené, 2007).

II.2. Sources de variation

Après avoir introduit et décrit l'ensemble des paramètres acoustiques présentés dans cette thèse, nous avons procédé à une description exhaustive de l'ensemble des facteurs ayant une influence sur ceux-ci.

i. Taille corporelle

Morton (1977) & Ohala (1984) ont tous deux suggéré que les propriétés physiques et donc oscillatoires des plis vocaux devraient être directement reliées à la taille corporelle. Toutefois, les études ayant tenté d'établir des corrélations entre celle-ci et divers paramètres vocaux ont rapporté des résultats variables. En effet, Pisanski, Fraccaro, Tigue, O'Connor, Röder, Andrews, Find, DeBruine, Jones & Feinberg (2014) ont montré que 80% des études rapportent une absence de corrélation avec la hauteur vocale (F0), tandis que 50% des études rapportent une absence de corrélation avec le timbre (formants), sujets hommes et femmes compris.

Par exemple, Cartei, Bond & Reby (2014) ont montré chez des hommes anglophones que le F0 et le ΔF (i.e. l'espace formantique, définissant le caractère plus ou moins grave ou aigüe du timbre, tout comme le Df, la dispersion des formants, et Pf, la position des formants) sont négativement corrélés à la taille. Ces résultats corroborent ceux de Kempe, Puts & Cárdenas (2013) et Puts, Apicella & Cardenas (2012), lesquels ont également trouvé une corrélation négative entre la taille, le F0 et le Pf dans une population d'étudiant américains. Chez les Hadza (une population de chasseurs-cueilleurs de Tanzanie), Puts et al. (2012) ont aussi rapporté une corrélation négative entre la taille et le Df, mais une absence de corrélation avec le F0. Chez les Tsimane (une population d'horticulteurs de Namibie), Hodges-Simeon, Gurven & Gaulin (2015) ont rapporté une corrélation négative entre le Pf et l'indice de masse corporelle, mais notent également une absence de corrélation avec le F0. A l'inverse, Bruckert, Lienard, Lacroix, Kreutzer & Leboucher (2006) n'ont trouvé aucune corrélation entre le F0, le Df et la taille dans une population d'étudiants francophones.

Si ces résultats contrastés peuvent s'expliquer par une relative différence en termes de tailles d'échantillon et de méthodologies d'analyse, l'étude la plus robuste concernant le lien entre le F0, les formants et la taille corporelle reste celle de Pisanski et al. (2014). Dans cette méta-analyse incluant 39 études, les auteurs ont montré que le F0 prédit moins de 2% de la variation en taille corporelle au sein de chaque sexe, tandis que l'analyse des formants, ainsi que les différentes mesures liées à leurs distances adjacentes (i.e. ΔF , Df et Pf), expliquent quant à elles jusqu'à 10% de la variation. Une étude supplémentaire avec un échantillon conséquent de 700 hommes et 700 femmes ont confirmé ces résultats (Pisanski et al., 2016). Il est toutefois à noter qu'il a été trouvé de modestes corrélations négatives entre le F0 et la taille chez les hommes, et le F0 et le poids chez les femmes (les coefficient de corrélation variant entre -0.1 et -0.2, Jost, Fuchs, Loeffler, Thiery, Kratzsch, Berger, & Engel, 2018 ; Pisanski et al., 2014/2016).

Les résultats hétérogènes précédemment mentionnés peuvent s'expliquer par le fait que les plis vocaux sont des tissus mous dont la taille et l'épaisseur varient en fonction de différents facteurs tels que les taux d'androgènes à la puberté et à l'âge adulte. C'est pourquoi les propriétés oscillatoires des plis vocaux n'entretiennent pas de relations allométriques avec la taille corporelle. En revanche, il a été montré que les dimensions et la longueur du conduit vocal dépendent directement de la taille du crâne et de la taille des individus entre et au sein de chaque sexe (Fitch & Giedd, 1999 ; Fitch, 1997), ce qui explique que des conduits vocaux plus longs produisant des formants relativement bas ou moins espacés soient positivement corrélés à la taille corporelle. Enfin et surtout, Pisanski et al. (2014/2016) montrent via des analyses de puissances que pour obtenir des estimations statistiquement fiable de la taille corporelle via le F0, un échantillon d'au moins 618 hommes et 2140 femmes est nécessaire, et via les formants, un échantillon d'au moins 99 hommes et 164 femmes est nécessaire.

Notons également que la micro-intonation (F0-SD), la raucité (jitter) et le souffle vocal (HNR) ne sont pas corrélés à la taille ou au poids (Hodges-Simeon et al., 2015 ; Pisanski et al., 2016 ; Puts et al., 2012), puisque ce sont des composantes de la qualité vocale indirectement liées au mécanisme physique d'oscillations des plis vocaux. Enfin, notons qu'il n'existe aucun lien entre la taille corporelle et le débit de parole, celui-ci dépendant d'une part des capacités cognitives des individus et, d'autre part, de facteurs stylistiques indépendants des relations allométriques observées au niveau anatomique.

ii. Profils hormonaux

Dans les deux sexes, des récepteurs aux androgènes, aux œstrogènes et à la progestérone ont été identifiés dans la muqueuse du pharynx et dans l'épithélium des plis vocaux (Amir & Biron-Shental, 2004 ; Newman, Butler, Hammond, & Gray, 2000 ; Voelter, Kleinsasser, Joa, Nowack, Martínez, Hagen, & Voelker, 2008 ; toutefois voir Schneider, Cohen, Stani, Kolbus, Rudas, Horvat, & van Trotsenburg, 2007). Ces observations suggèrent que les paramètres

vocaux pourraient être influencés par les niveaux hormonaux circulants à l'âge adulte, indépendamment des pics observés à la puberté. Il a ainsi été suggéré qu'une concentration relativement élevée de testostérone permettrait un épaissement et un allongement des plis vocaux, ce qui aurait pour effet de diminuer le F0. Chez les femmes, cela dépendrait des variations hormonales au cours du cycle menstruel (Kadokia, Carlson, & Sataloff, 2013). Au cours de la phase folliculaire, les taux d'œstrogènes sont élevés tandis que ceux de la progestérone restent bas. Ceci provoque un œdème aux niveaux des plis vocaux avec une augmentation de l'afflux sanguin. Les polysaccharides se décomposent et se lient plus facilement avec l'eau, favorisant l'accumulation de fluide dans les plis vocaux. Ceci provoque une vibration plus rapide de ceux-ci, résultant en de plus hauts F0. Au cours de la phase lutéale, les taux de progestérone dépassent les taux d'œstrogènes, favorisant la mue de l'épithélium laryngé et rendant les sécrétions glandulaires plus visqueuses, conduisant à une diminution du F0.

Chez les hommes, le lien entre taux de testostérone circulant et paramètres vocaux ne fait pourtant pas consensus¹. D'une part, un ensemble d'étude a rapporté des relations négatives entre les deux. Parmi elles, Dabbs & Mallinger (1999) ont trouvé une corrélation négative entre le F0 et la testostérone dans les populations anglophones, tout comme Kempe et al. (2013) et Puts et al. (2012). En utilisant des analyses de piste, Hodges-Simeon et al. (2015) ont montré chez les Tsimane que le F0 est négativement prédit par le taux de testostérone, lui-même positivement influencé par l'âge et la condition énergétique (i.e. un proxy de la taille corporelle). Ainsi, au sein des sujets masculins de ces études, ceux dont les voix présentent des hauteurs vocales relativement plus graves, affichent de plus hauts taux de testostérone. Notons également que si Cartei et al. (2014) et Evans, Neave, Wakelin & Hamilton (2008) ont observé

¹ La majeure partie de la testostérone circulant dans le sang est liée à deux protéines, l'albumine et la globuline liant les hormones sexuelles. Seule une petite quantité de testostérone est libre (non liée aux protéines) et est disponible pour pénétrer dans les tissus. C'est celle-ci qui est généralement mesurée dans les études que nous citons.

une corrélation négative entre le F0 et la testostérone dans des populations anglo-saxonnes, les premiers n'ont observé aucune relation avec le ΔF , tandis que les seconds ont observé une corrélation négative avec le Df, ce dernier résultat étant également observé par Bruckert et al. (2006) dans un échantillon de locuteurs francophones. De manière intéressante, Puts et al. (2016) ont aussi montré, dans un échantillon d'étudiants américains, que l'interaction entre la testostérone et le cortisol (une hormone liée au stress, la glycémie et à la croissance osseuse) prédit le F0 de telle manière que la testostérone est négativement corrélée au F0 seulement chez les hommes présentant des taux de cortisol relativement peu élevés.

A l'inverse, un nombre similaire d'études n'ont trouvé aucune relation entre taux de testostérone circulant et paramètres vocaux. Par exemple, Bruckert et al. (2006) n'ont établi aucun lien avec le F0 et le F0-SD. De même, Skrinda, Krama, Kecko, Moore, Kaasik, Meija, Lietuvietis, Rantala, & Krams (2014) n'ont trouvé aucune relation entre le F0, les formants et la testostérone dans une population de lettons ; pas plus que Arnocky, Hodges-Simeon, Ouellette & Albert (2018) avec le F0, le F0-SD et le Df chez des étudiants américains. Si ces résultats divergents peuvent être éventuellement liés à des tailles d'échantillons variables, l'étude la plus conséquente à ce jour, comprenant 1106 locuteurs allemands, n'a observé aucune corrélation entre le F0 et le taux de testostérone circulant (Jost et al., 2018). Toutefois, les auteurs ont trouvé chez les hommes une corrélation négative entre le F0 et le sulfate de déhydroépiandrostérone, une hormone constituant un réservoir important d'hormones stéroïdiennes permettant aux tissus de la transformer en androgènes ou en œstrogènes selon les besoins locaux. Enfin, rappelons que la divergence de l'ensemble de ces résultats peut s'expliquer par le fait que le niveau de testostérone circulant peut être confondu par de nombreux facteurs, comme la variation diurnale (Dabbs, 1990), l'âge (Vermeulen, Goemaere, & Kaufman, 1999), le statut social (Eisenegger, Haushofer, & Fehr, 2011) ou encore la santé

(Araujo, Dixon, Suarez, Murad, Guey, & Wittert, 2011), facteurs qui ne sont pas (ou rarement) contrôlés dans les études citées.

Chez les femmes, ce sont surtout les œstrogènes et la progestérone qui ont été étudiées. Parmi les œstrogènes, l'estradiol ou œstradiol est un dérivé naturel du métabolisme du cholestérol (via la testostérone) qui est nécessaire au maintien de la fertilité et des caractères sexuels secondaires chez les femmes (Thornhill & Gangestad, 2008). Toutefois, si l'hypothèse d'une corrélation positive entre œstrogènes et F0 a souvent été citée et prise pour acquise (e.g. Collins & Missing, 2003 ; Feinberg, Jones, DeBruine, Moore, Law Smith, Cornwell, Tiddeman, Boothroyd, & Perrett, 2005 ; Jones, Feinberg, DeBruine, Little, & Vukovic, 2010; O'Connor, Fraccaro, Pisanski, Tigue, & Feinberg, 2013 ; Pipitone & Gallup, 2008), les deux seules études ayant tenté de quantifier ce lien n'en ont trouvé aucun (Jost et al., 2018 ; Puts, Bailey, Cárdenas, Burriss, Welling, Wheatley, & Dawood, 2013). Toutefois, notons que si Puts et al. (2013) ne trouve pas de liens entre le F0, le Pf et le jitter avec l'estradiol et la progestérone dans un échantillon de plus de 200 femmes américaines, ces deux hormones prédisent toutefois de manière modeste l'attractivité vocale des femmes. Enfin, notons également que Dabbs & Malingier (1999) et Puts et al. (2016) n'ont rapporté aucun lien entre la testostérone et le F0 chez celles-ci.

Remarquons toutefois que suite à la ménopause, les niveaux d'œstrogènes et de progestérone chutent tandis que les taux de testostérone restent constants, ce qui entraîne généralement une diminution du F0 d'environ 10 à 35 Hz en moyenne (Abitbol, Abitbol, & Abitbol, 1999 ; Amir & Biron-Shental, 2004). Dans une étude longitudinale, il a également été noté qu'une baisse du F0, du F0-SD et de l'étendue du F0 (i.e. valeurs maximale et minimale de la hauteur vocale affichée par un individu, respectivement F0max et F0min) surviennent après chaque grossesse et restent stables après une année post-partum, pour ensuite revenir à des niveaux équivalents à ceux observés avant la grossesse (Pisanski, Bhardwaj, & Reby, 2018).

Ceci peut s'expliquer par le fait que les niveaux moyens d'œstrogènes, de progestérone et de cortisol chutent remarquablement dans les mois suivant l'accouchement, ce qui soulignerait l'action de ces hormones sur les récepteurs des hormones stéroïdiennes situés dans les cellules épithéliales des plis vocaux.

En conclusion, notons qu'il n'existe pas à notre connaissance d'étude (hormis Puts et al., 2013 pour les femmes) ayant cherché à apprécier le lien entre le HNR, le jitter et le débit de parole avec les différentes hormones que nous avons citées, bien qu'il ait été récemment suggéré que ces dernières influenceraient le HNR et le jitter (Pisanski et al., 2016).

iii. Configuration corporelle

Hormis la taille corporelle, il a aussi été suggéré que les paramètres vocaux pouvaient être liés à diverses formes du corps. En effet, comme l'appareil vocal et la configuration corporelle sont tous deux influencés par les androgènes et les œstrogènes, ils pourraient covarier ensemble (Pisanski et al., 2016). A notre connaissance, ces études sont rares et la plus robuste ayant tenté d'établir ces relations a étudié 27 paramètres vocaux et 8 mesures du corps sur un échantillon de locuteurs canadiens, écossais et allemands, comprenant au total 700 hommes et 700 femmes (Pisanski et al., 2016).

Chez les hommes, il a été montré que le F0max, le F0-SD, le jitter, le F1 (i.e. premier formant correspondant au degré d'ouverture de la mandibule), le F3 (i.e. troisième formant, correspondant à l'arrondissement des lèvres et donc à la longueur du conduit vocal) et le Pf sont négativement corrélés à la circonférence des hanches, avec les deux premières mesures expliquant le plus de variance (~11%). De plus, le F0max, le F1, le F3 et le Pf sont en même temps négativement corrélés à la circonférence de la poitrine. Le F0min est négativement corrélé au ratio poitrine-hanche, tandis que le HNR et le shimmer y sont positivement corrélés, ces deux derniers expliquant le plus de variance (respectivement 4.4% et 5.7%). Enfin, le Pf et

le ratio taille-hanche sont positivement corrélés. Chez les femmes, le F0, le F0-SD, le HNR et toutes les mesures de formants (ainsi que celles évaluant leurs positions adjacentes) sont négativement corrélés à la circonférence des hanches, tandis que le F0, le F0max, le F0-SD, le jitter, le F3, le F4 (i.e. formant liée aux caractéristiques interindividuelles), le Df et le ΔF sont négativement corrélés à la taille. Pour la circonférence de la poitrine, elle est négativement corrélée au F0, au F0max, au F0-SD, au jitter et à toutes les mesures de formants (ainsi que celles évaluant leurs positions adjacentes). Pour le ratio taille-hanche, le F0-SD, le jitter, le shimmer, et la LCV (i.e longueur du conduit vocal, inféré à partir des formants) y sont négativement corrélés, tandis que le HNR, le Fn (i.e. la moyenne des formants) et le Pf y sont positivement corrélés, le HNR et le Fn expliquant le plus de variation (9% et 10.8% respectivement), suivi du LCV, du jitter et du shimmer, chacun expliquant 7.3% de la variation. Enfin le F0max et le jitter sont négativement corrélés à la circonférence poitrine-hanche, et positivement avec le Fn.

De manière importante, soulignons que dans les deux sexes, l'ensemble de ces corrélations sont relativement faibles, variant dans l'absolu entre ~ 0.10 et ~ 0.33 . De plus amples études sont nécessaires pour comprendre les mécanismes sous-jacents liant les paramètres vocaux à diverses formes du corps.

iv. Force physique

S'il existe une relation positive entre la taille corporelle et la force physique, il a été suggéré que cette dernière pourrait être liée à différents paramètres vocaux (Sell, Bryant, Cosmides, Tooby, Sznycer, von Rueden, Krauss, & Gurven, 2010). Mais comme pour les aspects de la configuration corporelle, le mécanisme liant les deux restent relativement obscur, même si des médiateurs hormonaux y seraient à l'origine (Hodges-Simeon, Gurven, Puts, & Gaulin, 2014).

Dans une étude incluant des hommes américains et Hadzas, il a effectivement été montré chez les premiers que ceux ayant une plus grande force physique affichent un F0-SD et un Pf relativement plus bas ; tandis que chez les seconds, ceux avec une plus grande force physique affichent un F0 plus bas (Puts et al., 2012). Notons aussi que dans les deux groupes, le Df ne prédit pas la force physique. Ces résultats sont plutôt soutenus par ceux de Hogdes-Simeon et al. (2015), qui ont montré chez les hommes Tsimanes que la force physique est un prédicteur important du F0-SD et du Pf. A l'inverse, Sell et al. (2010) n'ont trouvé aucune relation significative entre la force physique, le F0 et le Df chez des hommes et femmes américains mais aussi chez des Tsimanes et des Andins. Notons tout de même que paradoxalement, les auteurs rapportent une corrélation perceptive entre ces différents facteurs (i.e. les sujets semblent plus ou moins capables au-delà de la chance d'associer correctement une voix avec la force physique correspondant à ce locuteur). Plus récemment, Han, Wang, Fasolt, Hahn, Holzleitner, Lao, DeBruine, Feinberg, & Jones (2018) ont montré une absence de corrélations dans les deux sexes entre la force physique et le F0, le F0-SD, le Df, le Pf et la LCV dans une population d'anglais et de chinois.

v. Indices développementaux

Il a été suggéré que les paramètres vocaux pourraient également refléter des qualités développementales. En effet, la variation de la condition physique et mentale à la puberté (Harries, Walker, Williams, Hawkins, & Hughes, 1997 ; Hodges-Simeon et al., 2014) et l'influence de l'environnement intra-utérin (Levrero et al., 2018) pourraient affecter le développement physiologique de l'appareil vocal.

D'une part, il a été proposé que l'indice de Manning, ou ratio 2D:4D, censé refléter le taux relatif de testostérone auquel le fœtus a été exposé au moment de la formation de la main, serait un bon candidat comme indice du développement. Dans ce contexte, il a été suggéré que les voix relativement graves pourraient être associées à des valeurs de ratios relativement plus

petites, donc appartenant à des individus ayant été exposés à de plus hauts taux de testostérone durant la vie fœtale. Toutefois, Evans et al. (2008) ont rapporté une absence de corrélation entre ce ratio, le F0 et le Df chez des hommes canadiens, de même que Hughes, Harrison & Gallup (2002) et Hughes, Pastizzo & Gallup (2008) n'ont rapporté aucun lien avec l'attractivité vocale dans les deux sexes dans une population américaine¹. En revanche, Levrero et al. (2018) ont montré que ce ratio est positivement corrélé au F0 à l'âge de 4 mois et 5 ans, mais seulement lorsque mesuré sur la main droite. Notons toutefois que de récentes méta-analyses ont remis en question la pertinence de ce ratio en tant qu'indice d'exposition à la testostérone fœtale (Berenbaum, Bryk, Nowak, Quigley, & Moffat, 2009 ; Hampson & Sankar, 2012 ; Voracek, 2014).

Un autre indice lié à la qualité du développement est l'asymétrie fluctuante. Celle-ci peut se définir comme les différences entre les côtés droit et gauche d'un organisme à symétrie bilatérale, dont la précision peut être influencée par des facteurs internes (i.e. variation génétique) et externes (i.e. influence environnementale). Dans ce contexte, il a été suggéré qu'un développement relativement instable (i.e. développement de mauvaise qualité) affecterait la symétrie des plis vocaux, la forme du conduit vocal et, conséquemment, les propriétés acoustiques qui en découlent. Hughes et al. (2002) ont effectivement rapporté que les voix de sujets arborant une plus grande symétrie bilatérale (mesurée sur les coudes, les poignets et les mains) ont été considérées comme plus attractives par les membres des deux sexes que les voix des sujets présentant des traits asymétriques. Ces résultats sont soutenus par ceux de Hill, Cárdenas, Wheatley, Welling, Burriss, Claes, Apicella, McDaniel, Little, Shriver, & Puts (2017) qui ont trouvé une corrélation négative entre l'asymétrie faciale et l'attractivité vocale (i.e. plus grande symétrie du visage = plus attractif). De même, Abend, Pflüger, Koppensteiner,

¹ Il a été suggéré que les individus présentant des indices liés à un meilleur développement (mesuré soit via le ratio 2D:4D ou l'asymétrie fluctuante) devraient être perçus comme plus attractif.

Coquerelle, & Grammer (2015) ont trouvé une relation négative entre l'attractivité vocale des femmes et l'asymétrie fluctuante des membres supérieurs et inférieurs du corps, de la tête et du visage. Toutefois, si Hughes et al. (2008) ont rapporté les mêmes résultats perceptifs avec une mesure de l'asymétrie corporelle, cette étude reste la seule n'ayant objectivement pas trouvé de liens direct avec le F0, le F0-SD, le F0max, le F0min, l'intensité, le jitter, le shimmer et le HNR.

vi. Âge

Un certain nombre d'études a montré que divers paramètres vocaux varient avec l'âge. Par exemple, une légère augmentation du F0 chez les hommes et une diminution de celle-ci chez les femmes a été observé avec le vieillissement. Chez les hommes, le F0 augmenterait vers l'âge de 50 ans (donnant lieu à une voix plus aiguë), ce qui serait possiblement dû à une diminution de la testostérone. A l'inverse, pour les femmes, le F0 diminue après la ménopause, ce qui se manifeste par une aggravation significative de la voix. Plus précisément, chez les hommes, le F0 chute d'environ 10 Hz jusqu'à environ 50 ans, puis commence à augmenter progressivement d'environ 35 Hz ; chez les femmes, le F0 continue de diminuer avec l'âge ou reste constant jusqu'à la ménopause, après quoi il diminue de 10 à 35 Hz (Decoster & Debruyne, 1997 ; Hollien & Shipp, 1972 ; McGlone & Hollien, 1963 ; Sataloff, Caputo Rosen, Hawkshaw, & Spiegel, 1997).

Pour ce qui est du F0-SD et du jitter, ces derniers semblent augmenter dans les deux sexes, mais pas le HNR (Brown, Morris, & Michel, 1989 ; Goy, Fernandes, Pichora-Fuller, & van Lieshout, 2013). Il a également été rapporté une diminution de la fréquence des formants et du débit tant chez les hommes que chez les femmes, laquelle serait vraisemblablement due aux processus cognitifs plus lents et aux changements neuromusculaires (i.e. rigidité cartilagineuse) impactant la précision du mouvement des articulateurs (Torre & Barlow, 2009).

vii. Variation inter-linguistique

Les langues ne se parlent pas dans le même registre de hauteur. En effet, cette dernière peut considérablement varier d'une langue à l'autre, et ce pour chacun des deux sexes.

Ceci peut tout d'abord s'expliquer par l'effet de la langue et de son système phonologique et prosodique. Par exemple, une large littérature a montré que les individus en cours d'acquisition d'une deuxième langue changent de registre tonal lorsqu'ils parlent dans cette langue. Cela a été démontré chez des locuteurs francophones apprenant l'allemand (Zimmerer, Jügler, Andreeva, Möbius, & Trouvain, 2014), tout comme des suédois apprenant le russe (Ullakonoja, 2007) ou encore des gallois apprenant l'anglais (Ordin & Mennen, 2017).

Mais il a également été suggéré que, au-delà de possibles variations physio-anatomiques entre populations parfois dues à l'environnement (Everett, Blasi, & Roberts, 2015), une part de la prosodie des langues est « *apprise* » (Simpson, 2009). Cela laisse entendre que le F0 peut être également sensible aux influences socio-culturelles. En effet, Traunmüller & Eriksson (1995) ont rapporté dans une description largement étayée d'exemples qu'il existe une large variabilité du F0 entre les langues. Par exemple chez les hommes, pour des mêmes conditions d'enregistrement (parole lue, spontanée, voyelle tenue, etc.) et appariés en âge, les hommes bulgares et polonais présentent un F0 nettement plus élevé que les hommes allemands et anglais britanniques (160 Hz et 163 Hz contre 118 Hz et 128 Hz respectivement, Andreeva, Demenko, Mobius, Zimmerer, Jugler, & Oleskowicz-Popiel, 2014). D'autres études multilingues ont montré que les anglophones britanniques affichent des F0 plus faibles que des locuteurs du chinois mandarin (107 Hz et 129 Hz respectivement, Keating & Kuo, 2012) et des différences de F0 entre l'anglais américain, l'espagnol et le japonais ont également été mises en évidence (105.6 Hz, 124.6 Hz et 130.6 Hz respectivement, Hanley, Snidecor, & Ringel, 1966). Du côté des femmes, et toujours dans des conditions expérimentales similaires, les américaines présentent un F0 inférieur à celui des femmes japonaises (211 Hz contre 224 Hz respectivement,

Yamazawa & Hollien, 1992), tandis que les femmes bulgares et polonaises présentent un F0 plus élevé que les femmes allemandes et anglaises (272 Hz et 266 Hz contre 210 Hz et 217 Hz respectivement, Andreeva et al., 2014). A titre comparatif, d'autres exemples inclus ceux de Rose (1991) qui a montré que dans les dialectes chinois du Wù, le F0 chez les hommes et les femmes sont respectivement de 170 Hz et 187 Hz.

viii. Variation temporelle

De manière intéressante, deux études suggèrent qu'il y a eu un changement moyen du F0 dans les deux sexes au cours du temps. Dans une première étude, des enregistrements de 28 femmes entre 18 et 25 ans en 1945 ont été comparés à des enregistrements d'un groupe de femmes similaire en 1993 : les résultats montrent une différence de F0 d'environ 23 Hz, les voix étant significativement plus graves dans le deuxième groupe (Pemberton, McCormack, & Russell, 1998). Dans une seconde étude, il a été montré que le F0 des hommes français journalistes avaient également baissé au cours des cinquante dernières années (de Mareüil, Rilliard, & Allauzen, 2012). Toutefois, il est difficile d'établir les raisons liées à ces changements, même si l'hypothèse la plus probable serait des changements sociétaux, le tabagisme ou encore les contraceptifs hormonaux qui affectent la qualité vocale.

ix. Facteurs environnementaux

D'autres facteurs environnementaux peuvent influencer la qualité vocale. Par exemple, la consommation élevée de tabac entraîne une baisse significative du F0. Sorensen & Horii (1982) ont rapporté une différence de hauteur vocale d'environ 10 Hz entre des hommes fumeurs et non-fumeurs, les premiers ayant ainsi une voix plus grave. Les auteurs observent la même tendance chez les femmes, avec toutefois une différence plus faible, de l'ordre des 5 Hz. Ce phénomène s'expliquerait par l'épaississement des plis vocaux induit par la consommation régulière de tabac, observé notamment par Auerbach, Hammond & Garfinkel (1970). Enfin, il semblerait que la tessiture vocale (i.e. étendue des sons, du plus grave au plus aigu pouvant être

produite par un locuteur) serait influencée par la température et l'humidité de l'habitat écologique considéré, ce qui pourrait expliquer la distribution des langues à tons autour de l'équateur, la production des tons nécessitant un contrôle averti des plis vocaux lui-même favorisé par une hydratation optimale (Everett et al., 2015).

II.3. Conclusion

Nous avons décrit dans ce chapitre introductif les grandes composantes acoustiques de la voix humaine dont :

1. Le F0, ou fréquence fondamentale, correspondant aux vibrations des plis vocaux donnant naissance au flux laryngé. Le F0 constitue le corrélat acoustique de la hauteur vocale.
2. Les formants (notamment les trois premiers), correspondant aux résonances du conduit vocal, autant de multiples entiers du fondamental constituant les harmoniques d'une voix. Ce sont les composantes acoustiques les plus saillantes des sons vocaliques qui permettent d'en déterminer le timbre. Le premier formant renvoie au degré d'ouverture de la mandibule et permet d'opposer les voyelles ouvertes (F1 élevé) aux voyelles fermées (F1 bas). Le deuxième formant dépend de la position de la langue dans la bouche et permet d'opposer les voyelles antérieures lorsque la langue est située à l'avant de la cavité buccale (F2 élevé), et les voyelles postérieures lorsque la langue est ramassée à l'arrière de la bouche (F2 bas). Le troisième formant correspond à l'arrondissement des lèvres lequel est corrélé à un allongement du conduit vocal et donc à un assombrissement du timbre. Plus la voyelle est arrondie, plus le F3 est bas.

Toutefois, la voix ne se caractérise pas uniquement avec ces caractéristiques saillantes que sont le F0 et les résonances vocales. Mais d'autres critères secondaires peuvent également intervenir dans la caractérisation de la qualité vocale. En effet, le jitter (corrélat acoustique de la raucité induit par l'irrégularité des cycles de vibrations des plis vocaux), le HNR (corrélat du

souffle vocal généré par la présence d'une fuite glottique due à un accolement imparfait des plis vocaux lors de la phonation), la micro-intonation (F0-SD), dont les modulations confèrent aux voix des contours plats (i.e. voix monotones) ou plus mélodieux (i.e. voix dynamiques) et le débit de parole (i.e. lents ou rapides) constituent aussi des paramètres d'intérêt. Nous avons aussi souligné l'existence d'un dimorphisme vocal entre les deux sexes et énuméré de nombreuses sources influençant la variation des paramètres acoustiques.

L'objet du chapitre suivant sera de décrire le mécanisme de sélection sexuelle afin d'inscrire dans un cadre évolutif l'ensemble des concepts et des mécanismes que nous avons jusque-là décrits.

III. La sélection sexuelle

III.1. Mécanisme général

Si nous comprenons facilement que les différences morpho-anatomiques, physiologiques et comportementales entre espèces procèdent d'adaptations à leurs milieux et résultent d'un processus de sélection naturelle, certaines caractéristiques ne semblent pas répondre à l'exigence d'un tel mécanisme (Darwin, 1871). Un exemple classique concerne la queue du paon bleu (*Pavo cristatus*), ornée de plumes iridescentes présentes uniquement chez les mâles. Darwin (1871) a remarqué que ce trait qu'il jugeait « *extravagant* » posait un problème à sa théorie de l'évolution par sélection naturelle. En effet, la queue du paon mâle constitue un encombrement majeur qui limite fortement ses déplacements dans les airs, sans ajouter le fait que la couleur des plumes attire plus facilement les prédateurs. Darwin (1871) remarqua également que ces ornements sont absents chez les femelles et les juvéniles, l'amenant à conclure qu'il n'y a probablement aucun avantage à les arborer en termes de survie. En effet, si les plumes de la queue du paon étaient d'une quelconque utilité contre les prédateurs, les femelles et les jeunes en seraient également dotées. Darwin (1871) propose alors que le tri des caractéristiques morphologiques pouvait s'effectuer non pas uniquement en fonction des avantages qu'elles procurent aux individus en termes de survie, mais en fonction de ce qu'elles peuvent apporter en termes d'accès à des partenaires sexuels. Il nomme ce processus sélection sexuelle et le distingue de celui de sélection naturelle. Il prit soin de restreindre ce mode de sélection aux caractères sexuels secondaires, et fut le premier à relier ce mécanisme au dimorphisme sexuel qu'il observa chez de nombreuses espèces, comme le plumage des oiseaux de paradis ou les bois des cervidés, observés uniquement chez les mâles.

Ainsi, la sélection sexuelle peut être définie comme le processus par lequel certaines caractéristiques sont sélectionnées en raison des avantages qu'elles confèrent aux individus qui les possèdent dans l'accès aux partenaires sexuels. Elle peut être considérée comme une

variante de la sélection naturelle (Cézilly & Allainé, 2010). En ce sens, trois conditions doivent être simultanément remplies afin qu'elle puisse opérer :

- Il doit exister une variation pour un trait entre individus de même sexe.
- Des variantes particulières de ce trait doivent influencer la capacité de certains individus à accéder à un nombre plus ou moins grand de partenaires sexuels.
- Le trait et sa variation doivent être héritables.

Darwin (1871) a également suggéré qu'il existait plusieurs moyens d'accéder à des partenaires sexuels, pouvant mener à l'évolution de caractéristiques distinctes : la sélection intrasexuelle et la sélection intersexuelle. Il est généralement admis que le dimorphisme sexuel résulte de l'action de l'un de ces deux processus.

i. Compétition intrasexuelle

Darwin (1871) formula l'idée que lorsque la compétition implique un affrontement physique entre les individus d'un même sexe, la sélection sexuelle devrait favoriser toute caractéristique renforçant la force physique ou l'endurance des individus, ou dans une moindre mesure une caractéristique diminuant celles des adversaires. Ce mécanisme est nommé sélection ou compétition intra-sexuelle.

Cette compétition aboutit généralement à l'évolution d'« armes » au sens large, c'est-à-dire des attributs offensifs tels que les bois chez les cervidés ou à un développement renforcé d'un trait dans l'un des deux sexes, comme les cornes chez les aurochs. Ces attributs peuvent également être défensifs tels que la crinière des lions. Toutefois, elle peut aussi aboutir à un dimorphisme relativement important en termes de force physique ou de taille corporelle, comme c'est le cas chez les gorilles et les éléphants de mer. De manière générale, cette compétition est généralement observée chez les mâles d'une espèce et lorsqu'elle est particulièrement intense, les enjeux sont cruciaux car seuls les vainqueurs accèdent aux

femelles, regroupées en harem comme chez les cerfs rouges, les gorilles, les éléphants de mers ou les lions.

ii. Compétition intersexuelle

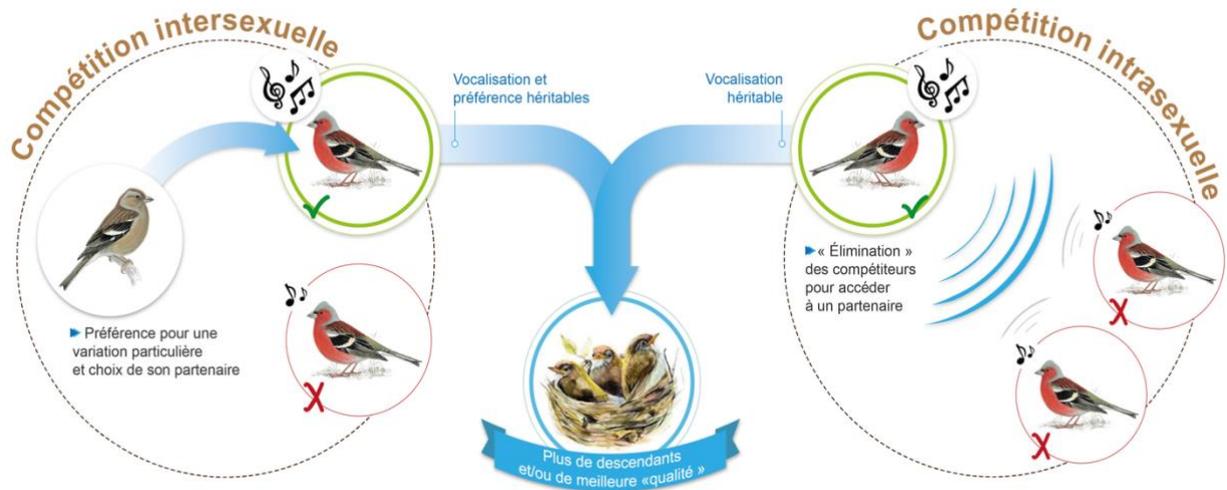
Darwin (1871) formula également l'idée que les femelles pouvaient aussi agir en tant qu'arbitre de la compétition entre les mâles, en favorisant ceux pourvu de certains attributs telles que les plumes extravagantes des oiseaux de paradis. Cette forme de compétition s'apparente au choix exercé par les individus d'un des deux sexes. Ce mécanisme correspond à ce que l'on nomme la sélection ou la compétition intersexuelle, et plus couramment au choix de partenaire.

Globalement, la qualité phénotypique est variable entre les individus. Puisqu'ils ne sont pas équivalents, le fait de choisir, plutôt que de prendre au hasard, peut permettre à un individu d'augmenter sa valeur reproductive (Geary, Vigil, & Byrd-Craven, 2004). Dans ce contexte, le choix de partenaire peut être défini comme une stratégie augmentant la probabilité de s'accoupler avec certains individus par rapport à d'autres (Halliday, 1983). Dans nombre d'espèces, c'est généralement le choix (ou les préférences) des femelles qui appliquera une pression sélective sur les mâles. Toutefois, il existe quelques espèces où les rôles sexuels sont inversés et où le mâle exerce un choix comme chez le jacana noir, les phalaropes et les hippocampes.

Le choix de partenaire aboutit généralement à l'évolution d'ornements, qui ont été bien étudiés chez les oiseaux, comme les plumes extravagantes du paon, les chants particulièrement remarquables chez l'alouette des champs ou encore les parades chez les oiseaux de paradis.

Les mécanismes de compétition intrasexuelle et intersexuelle sont illustrés dans la Figure 10.

Figure 10. Mécanismes du choix de partenaire (gauche) et de la compétition entre individus de même sexe (droite), en utilisant les vocalisations comme exemple. Dans le cas du choix d'un partenaire, la femelle va préférer certains chants plutôt que d'autres. Notons également que dans ce cadre, la vocalisation et la préférence doivent être hérissables. Dans le cas de la compétition intrasexuelle, le mâle possédant des vocalisations capables de repousser les compétiteurs est celui qui accédera aux partenaires sexuels.



III.2. Conclusion

Après une description succincte de la parole humaine et des deux grands mécanismes de la sélection sexuelle (i.e. compétition intra- et intersexuelle), nous allons maintenant pouvoir intégrer et considérer l'évolution de la voix humaine sous différents aspects : 1) les préférences vocales dans les deux sexes, 2) leurs interprétations évolutives, 3) les liens entre paramètres vocaux et succès copulatoire et reproducteur, 4) la flexibilité du comportement vocal, et enfin 5) une perspective originale sur des unités supérieures du langage.

IV. Évolution des préférences vocales

IV.1. Les préférences masculines et féminines

Il convient en premier lieu de définir ce qu'est une préférence et à quoi se réfère la notion d'attractivité d'un point de vue de l'évolution. La préférence (sexuelle ou non) peut être définie comme le mécanisme psychologique et comportemental permettant d'augmenter la valeur sélective lors d'un choix dans les conditions sociales, culturelles et écologiques dans laquelle elle a été sélectionnée. Une préférence est sélectionnée lorsqu'elle augmente la probabilité qu'un individu interagisse avec un stimulus d'une façon qui va augmenter la fréquence des allèles liés à cette préférence (Bovet, 2014). Il est à noter que les préférences ne sont pas des mécanismes généraux adaptatifs qui s'appliqueraient à tous les types de stimuli : si nous évaluons les potentiels partenaires sexuels en utilisant les mêmes mécanismes par lesquels nous évaluons la nourriture, nous trouverions la viande ou le sucre comme sexuellement attractifs (Sugiyama, 2015). Dans le cas du choix d'un partenaire, la sélection va favoriser les mécanismes permettant d'identifier la valeur sélective des différents partenaires potentiels via un ou plusieurs traits tels que le visage, le physique ou la voix. Dans ce cas-ci, nous définissons l'attractivité comme la propension à être attiré par une certaine valeur qu'un trait spécifique peut prendre le long d'un continuum. Notons également que les préférences sont, du moins en partie, des mécanismes inconscients. Un individu peut exprimer une préférence sans avoir conscience des facteurs l'ayant provoquée. Même dans le cas où un individu se rend compte des facteurs ayant influencé son choix, il ne sera pas pour autant conscient du lien entre le trait préféré et la propriété dont il est l'indice.

Prenons un exemple pour illustrer ces concepts. Les femmes peuvent être attirées par des voix graves (bas F0) plutôt qu'aiguës (haut F0) chez les hommes, parmi l'ensemble des voix exhibées chez ceux-ci. Les femmes peuvent se rendre compte qu'elles préfèrent des voix graves plutôt qu'aiguës, mais n'auront pas conscience qu'une telle voix peut refléter la qualité

phénotypique (et donc en partie la qualité génotypique sous-jacente héritable) de l'homme en question en tant que partenaire sexuel (i.e. haute force physique, système immunitaire de meilleure qualité, etc.). Ainsi, il y aura sélection de cette préférence chez les femmes si elle est héritable et qu'elle augmente leurs succès reproducteur.

Dans l'espèce humaine, la valeur sélective d'un partenaire comprend des qualités phénotypiques comme la santé, la fertilité, l'âge, l'intelligence, le statut social, les compétences parentales et la capacité ou la volonté à investir dans les descendants (Buss, 1989 ; Geary et al., 2004 ; Sugiyama, 2015). Ces indices transparaissent sur le visage, le physique et/ou la voix.

Toutefois, les préférences masculines et féminines pour le sexe opposé vont être différentes car certains traits associés à une haute qualité en tant que partenaire chez un homme vont différer de ceux associés à une haute qualité de partenaire pour une femme : les hommes et les femmes ne vont donc pas accorder le même poids aux différents critères lors de la recherche d'un partenaire. Ceci s'explique par le fait que l'investissement parental – le temps et l'effort investi dans la production des gamètes, les soins, la protection et l'éducation des enfants – varie considérablement entre les sexes (Trivers, 1972). En effet, les coûts physiologiques sont plus importants chez les femmes, avec la fertilisation interne, la placentation, la gestation et la lactation tandis que les hommes n'apportent que la contribution du sperme pour la reproduction. De plus, dans toutes les sociétés actuelles (comme pour toutes les espèces de mammifères) l'investissement des hommes dans les soins parentaux est moindre. Cette asymétrie dans l'investissement parental fait que la compétition intrasexuelle est plus importante pour les hommes, et que le choix ou la discrimination est moins présente chez ceux-ci (pour une discussion complète de ces aspects, consulter Sugiyama, 2015). Plus généralement, ces différences s'expliquent par des stratégies reproductives distinctes entre les hommes et les femmes. Rappelons qu'il existe certaines espèces où les rôles sont inversés (e.g. l'hippocampe ou le jacana noir).

D'une part, les hommes vont accorder plus de poids à l'attractivité physique de leur partenaire que les femmes, car le potentiel reproductif d'une femme est étroitement lié à sa condition physiologique : puisque le succès reproducteur des hommes est en grande partie déterminé par la fertilité du ou des partenaires auxquels ils accèdent, les mécanismes psychologiques permettant de détecter les indices de fertilité d'une femme vont être sélectionnés (Bovet, 2014). De plus, cette fertilité est étroitement liée à l'âge chez les femmes, avec un pic autour des 25 ans suivi d'une diminution après 30 ans, devenant ensuite nulle autour des 50 ans lors de la ménopause. Les hommes devraient ainsi posséder des mécanismes pour détecter et préférer des qualités vocales liées à une relative jeunesse associée à une fertilité perçue optimale. En revanche, les femmes vont accorder plus de poids à certains facteurs sociaux et auront tendance à préférer des hommes présentant un statut social plus élevé, en mesure de fournir des ressources importantes. Cette capacité est elle-même associée à une dominance sociale et physique élevées. Si le succès reproducteur des hommes est directement lié à la fertilité de leur partenaire, celui des femmes dépend donc plutôt du statut social de leur partenaire. Dans ce contexte, les femmes devraient ainsi posséder des mécanismes pour détecter et préférer des qualités vocales corrélées ou liées aux critères de masculinité reflétant des caractéristiques sociales avantageuses pour elles (Puts, Bailey, & Reno, 2015 ; Sugiyama, 2015).

Dans ce contexte, nous avons cherché à établir une liste exhaustive des études s'étant intéressées à l'attractivité et aux préférences des hommes et des femmes au niveau vocal (Tableau 1). Nous n'avons conservé que les études ayant clairement identifié les corrélats acoustiques liés à l'attractivité. Dans cette synthèse, nous pouvons d'emblée noter que la hauteur vocale (F0) et le timbre (formants) sont les paramètres les plus étudiés. De plus, les locuteurs anglais sont surreprésentés par rapport aux autres langues. Du point de vue méthodologique, il apparaît que le nombre et le type de stimuli vocaux utilisés sont relativement

variables (voyelles isolées, parole lue ou spontanée, etc.). De même, le nombre de juges est tout à fait hétérogène. Au plan des analyses effectuées par les différents auteurs, nous constatons deux types d'études : les études corrélationnelles, qui cherchent à établir une simple corrélation entre paramètres acoustique et attractivité (i.e. échelles de Likert) et les études expérimentales, lesquelles tentent d'établir des relations causales entre ceux-ci (i.e. choix forcé). L'ensemble de ces travaux permet de dégager les lignes générales quant à la question des préférences vocales.

Auteurs / année de publication	Langue(s) étudiée(s)	Nombre de stimuli	Nombre de juges	Type de stimulus	Méthode et nombre de stimuli jugés	Paramètres acoustiques étudiés et sens des corrélations
Tuomi & Fisher 1979	Anglais canadien (« <i>sexiness</i> ») ₁	10 ♀ 5 ♂	10 ♀ 10 ♂	Phrase naturelle	Échelle de Likert – Tous noté	- F0 bas + attractif ♂/♀ (n.s.)
Zuckerman & Miyake 1993	Anglais américain	62 ♀ 48 ♂	8 ♀ 9 ♂	Lecture naturelle	Échelle de Likert – Tous noté	- F0 bas + attractif ♂ - F0 maximale bas + attractif ♂ - amplitude basse + attractive ♂ - Moins de pause au total + attractif ♂ - tout n.s. ♀
Oguchi & Kikuchi 1997	Japonais (Expérience 1)	4 ♂	25 ♀	Phrase lue naturelle	Échelle de Likert- Tous noté	- F0 bas attractif ♂/♀ - F0-SD bas + attractif ♂/♀
	Japonais (Expérience 2)	8 ♀ 8 ♂	42 ♀ 20 ♂	Phrase lue naturelle	Échelle de Likert – Tous noté	- F0 bas attractif ♂/♀ - F0-SD bas + attractif ♂/♀
Collins (2000)	Néerlandais	34 ♂	54 ♀	Voyelles naturelles	Échelle de Likert – entre 10 et 14 stimuli	- Harmoniques bas + attractif ♂ - Espacement des harmoniques faible + attractif ♂
Collins & Missing (2003)	Anglais britannique	30 ♀	30 ♂	Voyelles naturelles	Échelle de Likert – 10 stimuli	- Harmoniques élevées + attractif ♀ - Formants élevés + attractif ♀ - Df élevée + attractif ♀
Feinberg et al. (2005)	Anglais canadien	10 ♂	68 ♀	Voyelles manipulées	Échelle de Likert – 10 stimuli	- F0 bas + attractif ♂ - Df faible et longueur tube vocal plus grand + attractif ♂ (n.s.)

Bruckert et al. (2006)	Français (« <i>pleasantness</i> ») ²	26 ♂	102 ♀	Voyelles naturelles	Échelle de Likert – 6 stimuli	- F0 bas + attractif ♂ - F0-SD élevé + attractif ♂
Riding et al. 2006	Anglais américain	9 ♂	54 ♀	Parole spontanée puis manipulée	Échelle de Likert – 11 stimuli	- F0 élevée - attractif ♂ - F0-SD n.s. ♂
Saxton et al. 2006	Anglais britannique et écossais	12 ♂	40 ♀ 7- 10 ans 40 ♀ 12- 15 ans 40 ♀ 20- 34 ans	Nombres naturels	Choix forcé – 6 ou 12 stimuli	- F0 bas + attractif ♂ pour 12-15 et 20-34 ans
Feinberg et al. 2008	Anglais canadien (expérience 1)	123 ♀	10 ♂	Voyelles naturelles	Échelle de Likert – 61 ou 62 stimuli	- F0 élevée + attractif ♀
	Anglais canadien (expérience 2)	15 ♀	263 ♀ 342 ♂	Voyelles manipulées	Choix forcé – 15 paires	- F0 élevée + attractif ♀
Hughes et al. 2008		31 ♀ 40 ♂	50 ♀ 51 ♂	Nombre (1 à 10)	Échelle de Likert – Chaque voix noté entre 13 et 15 juges	- minF0 bas + attractif ♂ - F0, maxF0, rangeF0, medianF0, intensity, durée, jitter, shimmer, HNR n.s. ♂ - Tout n.s. ♀
Jones et al. 2008	Anglais britannique	4 ♀	30 ♀ 30 ♂	Phrases naturelles ³	Choix forcé – 16 paires	- F0 élevée + attractif ♀
Leaderbrand et al. 2008	Anglais américain	1 ♀ 1 ♂	39 ♀ 9 ♂	Phrases manipulées	Échelle de Likert – 4 stimuli	- F0 bas + attractif ♂/♀

Vukovic et al. 2008	Anglais britannique	36 ♂	58 ♀ (contraceptifs) 65 ♀ (pas de contraceptifs)	Phrases manipulées	Choix forcé – 16 paires + échelle de Likert pour chaque voix préférée	- F0 bas + attractif ♂ et sans préférence différente avec ou sans contraceptifs
Saxton et al. 2009	Anglais britannique et écossais	6 ♀ 11-13 ans 6 ♀ 13-15 ans 6 ♂ 11-13 ans 6 ♂ 13-15ans	148 ♀ 177 ♂ Classe d'âge équivalente	Voyelles manipulées	Choix forcé – 6 paires + échelle de Likert pour chaque voix préférée	- F0 élevée + attractif ♀ mais seulement par le groupe 11-13 ans ♂ - F0 bas + attractif ♂ seulement par groupe de 13-15 ans ♀
Fraccaro et al. 2010	Anglais canadien	6 ♀	178 ♂	Voyelles manipulées	Choix forcé – 6 paires	- F0 élevée + attractif ♀ seulement en contexte de long- terme
Hodges-Simeon et al. 2010	Anglais américain	111 ♂	142 ♀	Parole spontanée	Échelle de Likert – 30 ou 31 stimuli	- F0 bas + attractif ♂ en contexte court/long-terme et fertile/non fertile - F0-SD bas + attractif ♂ en contexte long-terme fertile et court-terme non-fertile - intensité élevée + attractif en court/long-terme et fertile/non fertile - Df bas + attractif en court/long-terme et fertile ♂
Hughes 2010	Anglais américain	25 ♀ 20 ♂	27 ♀ 12 ♂	Appel téléphonique tronqué + manipulations naturelles	Choix forcé – Tous noté	- F0 bas + attractif ♂/♀
Jones et al. 2010	Anglais Britannique	6 ♀ 6 ♂	100 ♀ 100 ♂	Voyelles manipulées	Choix forcé – 6 paires	- F0 élevée + attractif ♀ - F0 bas + attractif ♂

Weiss & Burck 2010	Allemand	5 ♀ 5 ♂	7 ♀ 13 ♂	Phrases naturelles	Échelle de Likert – 9 voix	- F0 bas + attractif ♂ - Taux d'articulation élevée + attractif ♂/♀ - Centre de gravité spectral bas + attractif ♂/♀ - « <i>Spectral sd et skewness</i> » élevées + attractif ♀ - « <i>3rd central moment</i> » bas + attractif ♀
Borkowska & Pawlowski 2011	Polonais	58 ♀	144 ♂	Voyelles manipulées	Échelle de Likert – 13 voix	- F0 élevée + attractif ♀ (relation non-linéaire)
Pisanski & Rendall 2011	Anglais canadien	2 ♀ 6 ♂	30 ♀ 31 ♂	Série de mots naturels et manipulées	Échelle de Likert – 40 voix	- F0 et formants bas + attractif ♂ (même résultats sur stimuli naturel et manipulé) - F0 et formants bas - attractif ♀ (même résultats sur stimuli naturel et manipulé)
Puts, Welling & Buriss (2011)	Anglais américain	72 ♀	63 ♂	Lecture d'un texte ? manipulée	Échelle de Likert – 18 paires	- F0 élevé + attractif ♀ - Df élevé + attractif ♀
Liu & Xu 2011	Anglais britannique	1 ♀	10 ♂	Phrases modifiées	Échelle de Likert – Tous notés	- F0 élevé + attractif ♀ - Longueur tube vocale faible + attractif ♀
Simmons et al. 2011	Anglais australien	54 ♂	15 ♀	Voyelles naturelles	Échelle de Likert – Tous notés	- F0 bas + attractif ♂

Re et al. 2012	Anglais canadien	1 ♀ 1 ♂	9 ♀ 10 ♂	Voyelles manipulées	Choix forcé – 50 paires ♂/♀ + 6 paires supplémentaires ♂ + 42 paires supplémentaires ♀	- F0 élevée + attractif ♀ - F0 bas + attractif ♂
Fraccaro et al. 2013	Anglais canadien	4 ♀ 4 ♂	104 ♀ 110 ♂	Voyelles manipulées naturellement	Choix forcé – 16 paires	- F0 + attractif ♂ - F0 - attractif ♀
O'Connor et al. 2013	Anglais canadien	4 ♀ 4 ♂	128 ♂	Mots manipulés	Échelle de Likert – 40 stimuli	- F0 élevée + ♀ - F0 bas + attractif ♂
Xu et al. 2013	Anglais britannique (expérience 1)	1 ♀	10 ♂	Phrases naturellement modifiées puis manipulées	Échelle de Likert – Tous notés	- F0 bas + attractif ♂ - Formants n.s.
	Anglais britannique (expérience 2 – 5)	1 ♀ 1 ♂	16 ♀ 16 ♂	Phrase synthétisées	Échelle de Likert – 12 stimuli	- F0 élevée + attractif ♀ - F0 bas + attractif ♂ - Souffle élevée + attractif ♂/♀ - Formants bas + attractif ♂ mais n.s. pour ♀
Babel et al. 2014	Anglais américain	30 ♀ 30 ♂	15 ♀ 15 ♂	Mots naturels	Échelle de Likert – 15 stimuli de la même voix mais une seule note d'attractivité donnée à la fin	- Df élevée + ♀ - F0 n.s. ♀ et F0 + Df ♂
Hughes et al. 2014	Anglais américain	20 ♀ 20 ♂	20 ♀ 20 ♂	Nombres (1 à 10) Manipulés naturellement	Échelle de Likert – 40 voix	- F0 bas + attractif ♀ - jitter élevé + attractif ♀ - durée élevée + attractif ♀ - HNR/shimmer/intensité n.s ♀ - Tout n.s. ♂
Skrinda et al. 2014	Letton	60 ♂	29 ♂	Voyelles naturelles	Échelle de Likert – non spécifié	- F0 bas + attractif ♂ - F2 bas + attractif ♂ - Autres formants n.s. ♂

Tsantani et al. 2016	Anglais britannique	10 ♀ 9 ♂	183 ♀ 57 ♂	« Hello » manipulé	Choix forcé – 40 paires	- F0 bas + attractif ♂ - F0 n.s chez femmes
Sebesta et al. 2017	/	45 ♂ (Camerounais) 48 ♂(Namibien)	62 ♀ (Tchèque)	Phrase naturelle	Échelle de Likert – 45 ou 48 stimuli	- Cameroun : F0 bas + attractif ♂ - Namibie : HNR et Pf + bas attractif ♂
Shirazi et al. 2018	/	6 ♂ (Anglais américain)	63 ♀ allaitantes 65 ♀ nullipare (Philippines)	Phrases manipulées	Échelle de Likert – 12 stimuli	- F0 élevée + attractif ♂ (pas de différence entre les deux groupes)

Tableau 1. Chaque étude est caractérisée par la langue étudiée, le nombre de stimuli utilisés, le nombre de juges, le type de stimuli vocal utilisé, la méthodologie employée, le nombre de comparaison ou de stimuli entendu par un juge, les paramètres acoustiques étudiés et le sens des corrélations observés avec l'attractivité. Concernant le type de stimuli, l'adjectif « naturel » signifie que les stimuli n'ont pas été manipulés artificiellement pour changer certains aspects acoustiques des voix. Si l'adjectif « manipulée » est employé, tel est le cas et le nombre de stimuli est multiplié par le nombre de manipulation effectuée (par exemple, diminution ou augmentation du F0 pour un stimuli ; ces informations ne sont pas indiquées ici). Par « manipulation naturelle », nous entendons les études où les participants ont produit les stimuli vocaux après avoir modifié leurs voix volontairement ou non selon le protocole utilisé par les auteurs de l'étude. Enfin, la méthodologie peut être un choix forcé (i.e. choisir la voix la plus attractive entre deux voix) ou une évaluation via une échelle de Likert (i.e. noter une voix entre 1 à 7, 1 étant pas du tout attractif, 7 étant très attractif). Notons que les études impliquant des stimuli manipulés artificiellement ou naturellement impliquent souvent un choix forcé, afin de comparer pour un même locuteur deux versions différentes de sa voix (e.g. préférence entre deux stimuli où le F0 a été manipulé).

¹Tuomi & Fisher (1979) ont employé le terme « sexiness » dans leur étude en se référant plus globalement à l'attractivité vocale.

²Bruckert et al. (2006) ont employé le terme « pleasantness » dans leur étude, bien que celle-ci s'inscrit dans une perspective résolument évolutive dans le contexte d'un choix de partenaire.

i. Préférences pour la hauteur vocale (F0)

La plupart des études, tant corrélationnelles qu'expérimentales, ont montré une corrélation négative entre le F0 et l'attractivité. En d'autres termes, les femmes sont majoritairement attirées par des hommes aux voix relativement graves (Bruckert et al., 2006 ; Feinberg, Jones, Little, Burt, & Perrett, 2005 ; Hodges-Simeon, Gaulin, & Puts, 2010; Hughes, Farley, & Rhodes, 2010; Jones et al., 2010 ; Pisanski & Rendall, 2011; Vukovic, Feinberg, Jones, Debruine, Welling, Little, & Smith, 2008 ; Xu, Lee, Wu, Liu, & Birkholz, 2013). Notons toutefois quelques exceptions, dont les travaux de Babel, McGuire & King (2014) et de Hughes, Mogilski & Harrison (2014), qui ont rapporté une absence de corrélation et ceux de Shirazi, Puts & Escasa-Dorne (2018), qui ont observé une relation inverse.

Concernant les femmes, la plupart des études corrélationnelles et expérimentales ont montré que les hommes semblent être attirés par de hauts F0, c'est-à-dire des voix relativement plus aiguës (Borkowska & Pawlowski, 2011 ; Collins & Missing, 2003 ; Feinberg, DeBruine, Jones, & Perrett, 2008 ; Jones et al., 2010 ; Puts, Barndt, Welling, Dawood, & Burriss, 2011 ; Re, O'Connor, Bennett, & Feinberg, 2012). Toutefois, certaines études ont trouvé des relations inverses (Leaderbrand, Dekam, Morey, & Tuma, 2008 ; Oguchi & Kikuchi, 1997). Également, soulignons quelques autres résultats inattendus. Dans une première étude, Hughes et al. (2010) ont montré que lorsque des femmes s'adressent à des hommes qu'elles jugent attractifs, elles abaissent le registre de leurs voix dans les graves tandis qu'elles l'augmentent lorsqu'elles s'adressent à des hommes peu attractifs. De même, lorsque les stimuli sont évalués par des auditeurs masculins, ils jugent les voix plutôt graves de ces femmes comme étant plus attractives. Dans une seconde étude de Hughes et al. (2014), où il a été proposé à des sujets de sexe féminin d'adopter volontairement une voix attractive, il a été montré que, pour ce faire, les femmes ont tendance à user d'une voix plutôt grave. Et, en effet, l'expérience perceptive qui a suivi a confirmé que les voix attestant des F0 plutôt bas sont également perçues comme

plus attractives par des juges masculins. Enfin, Pisanski, Oleszkiewicz, Plachetka, Gmiterek & Reby (2018) ont également montré que les femmes ont tendance à baisser significativement leur F0 lorsqu'elles s'adressent à des hommes qu'elles jugent comme attractifs. Au regard de ces résultats contradictoires, de plus amples études sont nécessaires pour comprendre les préférences des hommes pour la hauteur vocale des femmes.

ii. Préférences pour le timbre (Formants)

Pour les préférences des femmes, les résultats issus de la littérature sont contrastés. Quelques études ont montré que les voix masculines présentant des résonances de basses fréquences sont perçues comme plus attractives (avec le Df, Hodges-Simeon et al., 2010 ; avec les formants sans calcul de leur espacement, Pisanski & Rendall, 2011 ; avec le Pf, Šebesta, Kleisner, Tureček, Kočnar, Akoko, Třebický, & Havlíček, 2017), mais deux autres études ont observé une absence de corrélation (Skrinda et al., 2014 ; Xu et al., 2013). De manière intéressante, deux autres études ont rapporté des résultats plus mitigés (Babel et al., 2014 ; Feinberg et al., 2005). La première a montré que seules les femmes de grande taille ont tendance à préférer des sons produits par des conduits vocaux relativement plus long (et plus le conduit est long, plus le signal résultant est grave), tandis que la deuxième étude a indiqué que des fréquences relativement basses seulement pour les deux premiers formants des voyelles /i/ et /u/ sont jugées plus attractives.

Concernant la préférence des hommes, des résonances perçues dans les hautes fréquences sont jugées comme plus attractives chez les femmes (Collins & Missing, 2003). De même, l'étude de Pisanski & Rendall (2011) a rapporté que lorsqu'ils sont situés dans les basses fréquences, les voix féminines sont perçues comme moins attractives. Toutefois, une étude a rapporté une absence de corrélation entre formants et attractivité (Xu et al., 2013). Dans ce contexte, les préférences pour le timbre dans les deux sexes ne semblent pas évidentes et les

différences de résultats pourraient être expliquées par la méthodologie employée pour le mesurer (i.e. utilisations diverses de ΔF , Fn, Df et Pf entre les études).

iii. Préférences pour le bruit (HNR et jitter)

Très peu d'études se sont intéressées aux liens entre le HNR (i.e. proxy du souffle vocal) et le jitter (i.e. proxy de la raucité vocale) avec l'attractivité vocale. Toutefois, les rares travaux ayant abordé la question ont suggéré que le souffle est un élément important de l'attractivité vocale pour les voix féminines (Babel et al., 2014 ; Van Borsel, Janssens, & De Bodt, 2009). Toutefois, des relations significatives ont été trouvées dans les deux sexes (Šebesta et al., 2017 ; Xu et al., 2013). Les voix les plus soufflées (i.e. présentant des valeurs moyennes de HNR relativement basses)¹ seraient ainsi plus attractives et ce, dans les deux sexes. Toutefois, deux autres études n'ont pas trouvé de corrélations significatives (Hughes et al., 2008/2014). Concernant la raucité, il n'existe à notre connaissance que ces deux mêmes études qui s'y sont intéressées : si la raucité n'est pas corrélée à l'attractivité chez les hommes, elle semble conférer aux voix féminines un certain degré d'attractivité (Hughes et al., 2008/2014). Dans ce contexte, de plus amples études sont nécessaires afin d'établir un réel consensus autour de ces deux paramètres.

iv. Préférences pour les variations de l'intonation (F0-SD)

Il existe également très peu d'études s'étant intéressées à l'effet de la micro-intonation sur les préférences vocales des hommes et des femmes. Par exemple, une étude a révélé que les voix présentant des valeurs de F0-SD relativement basses (i.e. voix monotones) sont plus attractives chez les hommes (Hodges-Simeon et al., 2010). Une autre étude suggère que ce résultat concerne à la fois les préférences vocales des hommes et des femmes pour le sexe opposé (Oguchi & Kikuchi, 1997). Toutefois, deux autres études parviennent à des résultats inverses : pour Bruckert et al. (2006) les voix monotones sont perçues comme étant moins

¹ Rappelons que les voix sont considérées comme pathologiques lorsqu'elles présentent des HNR inférieurs à 7 dB.

attractives dans les voix masculines. De même, selon Leongómez, Binter, Kubicová, Stolařová, Klapilová, Havlíček & Roberts (2014) ce sont les voix les plus mélodieuses (haut F0-SD) qui sont jugées comme étant les plus attractives tant pour les hommes que pour les femmes. Ces résultats contradictoires indiquent qu'il est difficile d'établir de façon définitive le rôle tenu par la micro-intonation dans les préférences des hommes et des femmes.

v. Préférences pour le débit

Il n'existe à notre connaissance aucune étude ayant directement abordé la question du débit et des liens qu'il entretiendrait avec l'attractivité. Toutefois, quelques études suggèrent qu'un débit relativement plus rapide serait perçu comme plus attractif dans les deux sexes.

En effet, dans les expériences utilisant des stimuli vocaux synthétisés et naturels, les hommes et les femmes s'exprimant avec un débit relativement rapide sont jugés comme étant plus compétents et bienveillants (Brown, Strong, & Rencher, 1973 ; Smith, Brown, Strong, & Rencher, 1975), plus crédibles et persuasifs (Miller, Maruyama, Beaber, & Valone, 1976), plus sympathiques (Weiss & Burkhardt, 2010), plus dynamiques et extravertis (Addington, 1968) et, *in fine*, socialement plus attractifs (Street, Brady, & Lee, 1984 ; Street, Brady, & Putman, 1983). Une autre étude avait également révélé que les individus perçus comme physiquement attractifs présentent des débits relativement plus rapides (Chaiken, 1979). À l'inverse, un débit relativement lent est associé à une impression de fausseté, à la passivité (Apple et al., 1979) et généralement à des locuteurs tristes voire dépressifs (Banse & Scherer, 1996 ; Siegman & Boyle, 1993 ; Sobin & Alpert, 1999).

IV.2. Sources de variation des préférences

i. Cycle menstruel

Il a été suggéré que la phase du cycle menstruel et les contextes d'accouplement – choisir un partenaire pour une relation romantique à long terme, ou juste un partenaire pour une relation de courte durée (i.e. relation sans lendemain) – pourrait influencer les préférences des

femmes pour certains attributs masculins, dont la qualité vocale (Feinberg, Jones, Law Smith, Moore, DeBruine, Corwell, Hillier, & Perrett, 2006 ; Pisanski, Hahn, Fisher, DeBruine, Feinberg, & Jones, 2014 ; Puts, 2005). L'hypothèse du « *good genes ovulatory shift* » suppose en effet que les femmes en phase d'ovulation expriment des préférences en faveur de partenaires plus masculins (la masculinité perçue étant supposément liée à une plus haute qualité génétique), plus particulièrement dans le contexte d'une relation à court terme (Jünger, Motta-Mena, Cardenas, Bailey, Rosenfield, Schild, Penke, & Puts, 2018). À l'inverse, les femmes en phase non-fertile (phases folliculaire et/ou lutéale) ont tendance à préférer des partenaires moins masculins (i.e. indiquant un investissement parental important), vers lesquels elles pourraient être particulièrement attirées pour une relation à long terme. Cette variabilité au niveau des préférences constituerait une stratégie adaptative permettant aux femmes d'optimiser leur fitness selon la période de leur cycle menstruel.

Concernant les préférences vocales, Puts (2005) a constaté que pour une même voix, les femmes en phase ovulatoire préfèrent les voix graves (F0 bas) dans un contexte de recherche de partenaire sur du court-terme. De même, Feinberg et al. (2006) ont constaté que les préférences des femmes en matière de masculinité pour les voix graves sont plus fortes pendant la phase fertile qu'en dehors de cette période. Bien qu'elle ne soit pas significative, l'étude de Pisanski et al. (2014) a rapporté la même tendance. Enfin, une étude a indiqué que les femmes en phase fertile préfèrent des voix présentant des timbres situés dans les basses fréquences (bas Df), que cela soit dans le contexte d'une relation à court et long-terme (Hodges-Simeon et al., 2010). Les auteurs de cette étude ont également constaté que la préférence pour les voix graves (bas F0) sont plus marquées en phase fertile et dans le contexte d'une relation à court-terme, tandis que les voix les plus monotones (valeurs basses de F0-SD) sont plus attractives dans les phases non fertiles pour le court-terme et la phase fertile pour le long-terme.

Cependant, le paradigme de la variation des préférences des femmes en fonction du cycle menstruel a récemment été remis en question. En effet, des études récentes suggèrent que le cycle menstruel des femmes n'influence pas leurs préférences pour des corps et des visages plus ou moins masculins (Jones, Hahn, Fisher, Wang, Kandrik, Han, Fasolt, Morrison, Lee, Holzleitner, O'Shea, Roberts, Little, & DeBruine, 2018 ; Marcinkowska, Galbarczyk, & Jasienska, 2018). En utilisant un échantillon de plus grande taille et une méthodologie plus rigoureuse, Jünger et al. (2018) n'ont effectivement retrouvé aucun effet de la phase du cycle, du risque de conception et des niveaux d'hormones stéroïdes pouvant expliquer les préférences des femmes pour la voix des hommes.

Enfin, notons qu'il a également été suggéré que l'attractivité vocale des femmes pouvait changer au cours du cycle menstruel : les voix des femmes n'étant pas en phase d'ovulation sont perçues comme moins attractives par les hommes, alors qu'elles sont jugées comme étant attractives lorsque les locutrices approchent de cette phase (Pipitone & Gallup, 2008). En outre, il a été montré que plus elles approchent de la phase fertile, plus la voix des femmes monte dans les aigus (augmentation du F0), tandis qu'elle descend dans les graves en dehors de cette phase (Bryant & Haselton, 2009 ; Fischer, Semple, Fickenscher, Jürgens, Kruse, Heistermann, Amir, & Halsey, 2011). En effet, comme les plis vocaux possèdent des récepteurs aux hormones sexuelles au niveau des cellules épithéliales, la variation des œstrogènes et de la progestérone au cours du cycle menstruel modifieraient leur masse, leur viscosité et leur tension, et donc leurs propriétés oscillatoires. Il a été avancé que cette variation cyclique pourrait être un signal adaptatif permettant d'augmenter son attractivité pour attirer des partenaires sexuels lorsque le risque de conception est important, ceci dans le but de favoriser la conception d'un enfant (Fischer et al., 2011; Pipitone & Gallup, 2008 ; Puts et al., 2013).

ii. Environnement socio-culturel

Bien qu'elles soient relativement rares, certaines études suggèrent que les préférences vocales ne sont pas universelles mais varient d'une langue/culture à l'autre.

L'étude pionnière de van Bezooijen (1995) a cherché à montrer l'effet de l'environnement socio-culturel sur les préférences vocales. L'auteur a souligné que le F0 moyen des Japonaises se situe parmi les plus élevés dans l'ensemble des sociétés humaines étudiées, avec une moyenne située autour de 232 Hz, comparativement à celles des Américaines, en moyenne à 214 Hz, et des Suédoises, en moyenne à 196 Hz. A titre comparatif, les Françaises se situent autour des 190 Hz (Vaissière, 2015). Les conclusions générales de l'étude suggèrent que ces différences de F0 s'expliquent par le fait que les hommes et les femmes doivent se conformer à des attentes sociales particulières lesquelles peuvent varier d'une culture à l'autre suivant le rôle social attribué à chaque sexe. Ainsi, au Japon, les traits de caractère traditionnellement associés à la féminité comprennent la modestie, l'innocence, la dépendance et plus généralement l'impuissance physique et la soumission psychologique : les Japonaises les signalent vocalement afin de répondre aux préférences des hommes partageant la même culture. A l'inverse, dans une culture relativement plus égalitaire comme les Pays-Bas, les Néerlandais tendent à valoriser des traits relatifs à l'indépendance féminine, ce qui se traduit par une préférence tournée vers des voix plus masculines et donc plus graves. En conclusion, il semblerait donc que les hommes japonais et/ou néerlandais tournent leurs préférences vers des voix culturellement congruentes par rapport aux valeurs attribuées à la féminité (i.e. soumission vs. indépendance, van Bezooijen, 1995).

Enfin, deux autres études montrent d'une part que dans une population namibienne, l'attractivité vocale des hommes n'est pas prédite par la hauteur de leur voix (F0), mais par le caractère plus ou moins soufflé de celle-ci (Šebesta et al., 2017), et d'autre part que les

Philippines préfèrent les voix masculines présentant des voix relativement aiguës (haut F0, Shirazi et al., 2018).

Notons pour finir qu'il n'existe pas à notre connaissance d'étude ayant mesuré systématiquement les préférences vocales d'une culture à l'autre en tentant d'identifier les prédicteurs liés à une possible variation. Pourtant, des variations socio-culturelles pour les préférences au niveau vocal ne semblent pas improbables, comme cela a été observé pour d'autres composantes de l'attractivité tels que le ratio taille-hanche, l'indice de masse corporelle ou la stature, lesquelles dépendent en partie de l'écologie (discuté dans Pisanski & Feinberg, 2013). L'influence de l'environnement culturel constitue ainsi une piste de recherche intéressante à explorer dans l'avenir.

IV.3. Limites des préférences

Il existe peu d'études s'étant intéressées aux limites inférieures ou supérieures des préférences vocales. Néanmoins, une étude a constaté qu'en deçà du seuil de 96 Hz, les préférences des femmes pour les voix masculines ne varient plus (Re, O'Connor, Bennett, Feinberg, & Reby, 2012). En revanche, quand elles ont à choisir entre deux stimuli dont le F0 est supérieur à 96 Hz, elles préfèrent systématiquement le stimulus le plus grave. Concernant les préférences des hommes, deux études ont signalé l'absence d'une limite supérieure concernant la hauteur vocale des femmes, avec une préférence systématique pour des F0 toujours plus élevés, du moins dans la plage de valeur testée (i.e. 160-300 Hz, Feinberg, DeBruine, Jones, & Perrett, 2008 ; Re et al., 2012). En revanche, l'étude de Borkowska & Pawlowski (2011) ont rapporté une relation non linéaire entre la hauteur vocale et l'attractivité, cette dernière commençant à diminuer dès lors que la valeur du F0 avoisine les 260 Hz. Dans le cas des voix se situant en dessous ou au-dessus des valeurs limites, ces observations peuvent s'expliquer par le fait que ces voix-ci présentent une pathologie vocale, comme c'est le cas pour les voix présentant un degré élevé de raucité et/ou de souffle vocal. De même, il semblerait que

la discrimination des voyelles soit significativement diminuée lorsque le F0 est trop bas (Smith, Patterson, Turner, Kawahara, & Irino, 2008). Dans le cas d'un F0 relativement élevé, les voix féminines pourraient être associées à des individus identifiés comme sexuellement non mature (Borkowska & Pawlowski, 2011). De manière générale, dès lors que les voix présentent des valeurs acoustiques distribuées en dehors de celles observées dans la population générale, elles sont associées à des voix perçues comme pathologiques (Barkat-Defradas, Busseuil, Chauvy, Hirsch, Fauth, Révis, & Bretèque, 2012) et renvoient à des traits de personnalité jugés négatifs (Révis, 2017).

IV.4. Autres composantes de l'attractivité vocale

Quelques autres facteurs liés à l'attractivité vocale méritent également d'être mentionnés. Un de ceux-ci correspond à la notion de « *voix moyenne* ». Dans ce cadre, Bruckert, Bestelmeyer, Latinus, Rouger, Charest, Rousselet, Kawahara & Belin (2010) ont montré que l'attractivité vocale peut être influencée par le phénomène appelé « *averaging* ». En effet, plus nous augmentons le nombre de voix pour construire des voix composites – dites « *moyennées* » (ou « *averaged voices* ») – plus l'attractivité vocale perçue augmente, et ce indépendamment du sexe. Les auteurs ont expliqué ce phénomène par deux processus indépendants. Premièrement, par ce qui peut être défini comme une opération de lissage spectro-temporel diminuant les apériodicités du signal (i.e. bruit), ayant pour effet d'éliminer les « *imperfections* » perceptibles d'une voix. Deuxièmement, parce que le « *moyennage* » a tendance à réduire la distance à la moyenne observée dans la population, dans un contexte où chaque individu évalue les stimuli vocaux par rapport à une moyenne, laquelle est construite à partir d'informations provenant de contextes sociaux plus généraux. Ce phénomène a également été observé pour les visages (Baudouin & Tiberghien, 2004 ; Halberstadt & Rhodes, 2000 ; Rhodes, Yoshikawa, Clark, Lee, McKay, & Akamatsu, 2001). Toutefois, notons qu'une certaine proportion de voix dans la population est tout de même considérée comme plus

attractive que ces voix composites, suggérant que cette hypothèse n'explique qu'une partie de l'attractivité vocale (Bruckert et al., 2010), tout comme cela a été observé pour les visages (DeBruine, Jones, Unger, Little, & Feinberg, 2007).

Une autre composante liée à l'attractivité vocale et pouvant expliquer les préférences serait l'étude de la « *sparseness* » ou « *parcimonie* » d'une voix. Il a été montré que les systèmes perceptuels (i.e. visuel, olfactif et auditif) ont été sélectionnés pour coder de l'information de manière efficace, c'est-à-dire rapidement et en fournissant le moins d'énergie possible (i.e. loi générale du moindre effort), dans les environnements dans lesquels nos ancêtres ont vécu (Renoult, Bovet, & Raymond, 2016). Ce codage efficace peut s'effectuer par l'élimination des composantes redondantes d'un signal. En conséquence, les signaux sont plus rapidement et plus précisément traités, ce qui facilite leur stockage et leur récupération. Comme cela requiert un nombre limité de neurones, un codage parcimonieux limiterait les coûts métaboliques et permettrait de sauvegarder de l'énergie pour d'autres fonctions. Il a ainsi été montré que la parcimonie d'un visage peut expliquer une proportion significative de son attractivité (Holzleitner, Lee, Hahn, Kandrik, Bovet, Renoult, Simmons, Garrod, DeBruine, & Jones, 2018 ; Renoult et al., 2016). Bien que considéré comme un biais perceptif, et au même titre que des biais cognitifs et sensoriels, il a été suggéré que le phénomène de parcimonie a pu être une force importante dans l'évolution et la diversification des ornements sexuels secondaires (consulter Renoult & Mendelson, 2019 pour une discussion de ces aspects). Bien que les mécanismes neurophysiologiques sous-jacents liés à l'efficacité du codage d'un signal acoustique au regard de l'attractivité vocale soit encore inconnus, et bien que la redondance d'un signal sonore soit un concept relativement flou, il n'existe à notre connaissance aucune étude s'étant intéressé à ces aspects pour la voix.

IV.4. Conclusion

A l'issue de ce chapitre, nous avons montré qu'il existait un relatif désaccord concernant certaines préférences, notamment pour la hauteur vocale féminine ainsi que pour le timbre et la micro-intonation dans les deux sexes. De plus, nous avons aussi indiqué que d'autres paramètres acoustiques, tel que le jitter, le HNR et le débit, sont relativement peu étudiés alors même qu'ils pourraient fortement participer à l'attractivité vocale.

Généralement, plusieurs facteurs peuvent expliquer ces résultats hétérogènes, citons à titre d'exemples : les différences de tailles d'échantillon (pour les stimuli comme pour les juges) ou encore la nature des stimuli vocaux utilisés dans les expériences psychoacoustiques visant à étudier les préférences vocales (e.g. parole lue vs. spontanée). Enfin, une limite importante réside dans le fait que la très grande majorité des études mentionnées précédemment ont été conduites en anglais auprès de juges eux-mêmes anglophones, ce qui d'une part, occulte la dimension transculturelle et, d'autre part, rend toute tentative de comparaison difficile.

Dans la perspective de comprendre l'évolution du dimorphisme vocal, nous nous attacherons à discuter dans le chapitre suivant les interprétations évolutives expliquant l'ensemble des préférences jusque-là décrites.

Article 1

Male vocal quality and its relation to females' preferences

Alexandre Suire, Michel Raymond & Melissa Barkat-Defradas

Article accepté dans *Evolutionary Psychology* (sous presse)

Résumé

Les préférences vocales des femmes semblent relativement consensuelles concernant la hauteur vocale (i.e. F0) des hommes. En effet, quasi-toutes les études corrélationnelles et expérimentales montrent que les femmes préfèrent des voix relativement graves. Toutefois, le peu de résultats connus à ce jour sont plus contrastés concernant le timbre et la micro-intonation, de même qu'il existe peu d'études s'étant intéressées au jitter et au HNR. De plus, la plupart des études ont été conduites dans des populations anglo-saxonnes, mais un petit nombre de travaux suggèrent que les préférences usuellement décrites ne sont pas universelles. Afin d'éclaircir ces différents aspects et contribuer plus généralement à l'étude des préférences vocales, nous avons mené une étude sur les préférences des femmes françaises pour la voix des hommes français.

Dans ce contexte, nous avons enregistré 58 participants prononçant une même phrase imposée par l'expérimentateur. Pour chacun de ces stimuli, nous avons analysé les composantes acoustiques suivantes : le F0, le F0-SD, le Df, le Pf, le jitter, le HNR et l'intensité (notons que le Pf a été déterminée avec les voix de 68 femmes prononçant la même phrase). Via une plateforme internet, nous avons présenté à 224 juges femmes (dont 165 éligibles) 11 paires de voix, chacune contenant une voix aléatoirement choisie parmi le pool entier de stimuli. A chaque paire, la juge devait choisir la voix qu'elle considérait comme la plus attractive. Notons que la 2^{ème} et la 3^{ème} paires étaient identiques à la 10^{ème} et la 11^{ème} afin d'évaluer la fiabilité du juge (e.g. détecter des clics aléatoires).

Nos résultats montrent que les femmes françaises exhibent une préférence significative pour les voix relativement plus graves (bas F0) et/ou celles présentant une intonation relativement marquée (haut F0-SD). En revanche, elles ne montrent aucune préférence particulière pour tous les autres paramètres acoustiques. Le paradigme de choix forcé nous permet de mettre en évidence qu'une voix ayant un F0 environ 11 à 12 Hz plus bas et/ou un F0-

SD de 5 Hz plus élevée qu'une autre voix a 65% de chance d'être considérée comme la plus attractive. Globalement, nos résultats suggèrent que la préférence des femmes a évolué pour être attentif à la hauteur vocale et à l'intonation plutôt que le timbre, la raucité ou le souffle. Dans ce contexte, le F0 et le F0-SD pourraient refléter des traits particulièrement pertinents dans le choix d'un partenaire masculin.

Male vocal quality and its relation to females' preferences

Authors: Alexandre Suire¹, Michel Raymond¹, Melissa Barkat-Defradas¹

Corresponding author: Alexandre Suire

E-mail: alexandre.suire@umontpellier.fr

Fax: +33 4 67 14 36 22

Tel: +33 4 67 14 49 66

¹ Institut des Sciences de l'Évolution de Montpellier, Université de Montpellier, Centre National de la Recherche Scientifique, Institut pour la Recherche et le Développement, École Pratique des Hautes Études – Place Eugène Bataillon 34095 Montpellier, France.

Abstract

In both correlational and experimental settings, studies on women's vocal preferences have reported negative relationships between perceived attractiveness and men's vocal pitch, emphasizing the idea of an adaptive preference. However, such consensus on vocal attractiveness has been mostly conducted with native English speakers, but a few evidence suggest that it may be culture-dependent. Moreover, other overlooked acoustic components of vocal quality, such as intonation, perceived breathiness and roughness may influence vocal attractiveness. In this context, the present study aims to contribute to the literature by investigating vocal attractiveness in an underrepresented language (i.e. French) as well as shedding light on its relationship with understudied acoustic components of vocal quality. More specifically, we investigated the relationships between attractiveness ratings as assessed by female raters and male voice pitch, its variation, the formants' dispersion and position, and the harmonics-to-noise and jitter ratios. Results show that women were significantly more attracted to lower vocal pitch and higher intonation patterns. However, they did not show any directional preferences for all the other acoustic features. We discuss our results in light of the adaptive functions of vocal preferences in a mate choice context.

Keywords: attractiveness; fundamental frequency; formants; intonation; breathiness; roughness; mate choice.

Introduction

Voice is one of the fundamental aspects of human communication. Indeed, research has reported that acoustic signals provide listeners with information on the quality or condition of the speaker such as sex (Bachorowski and Owren, 1999; Gelfer and Bennett, 2013; Gelfer and Mikos, 2005; Hillenbrand and Clark, 2009), age (Linville and Fisher, 1985; Ptacek, and Sander, 1966; Shipp, Qi, Huntley, and Hollien, 1992), sexual orientation (Lyons, Lynch, Brewer, and Bruno, 2014; Munson, McDonald, DeBoe, and White, 2006), physical strength (Sell, Bryant, Cosmides, Tooby, Sznycer, von Rueden, Krauss and Gurven, 2010), sexual behavior and body configuration (Hughes, Dispenza, and Gallup, 2004). In this context, numerous studies have explored the relationships between acoustic features of speech and several auditory impressions, among which, attractiveness as assessed by opposite-sex members. Focus has especially been given to sexually dimorphic acoustic traits such as the fundamental frequency (i.e., F_0 , the acoustic correlate of voice pitch) and the formant frequencies (i.e., the resonances of the vocal tract, the acoustic correlate of perceived timbre) (Titze, 1989).

In both correlational and experimental settings, most studies have reported a consistent negative relationship between men's F_0 and attractiveness, that is, women are attracted to relatively low-pitched voices (Bruckert, Lienard, Lacroix, Kreutzer, and Leboucher, 2006; Collins, 2000; Feinberg, Jones, Little, Burt, and Perrett, 2005; Hodges-Simeon, Gaulin, and Puts, 2010; Hughes, Farley, and Rhodes, 2010; Jones, Feinberg, DeBruine, Little, and Vukovic, 2010; Pisanski and Rendall, 2011; Vukovic, Feinberg, Jones, DeBruine, Welling, Little and Smith, 2008; Xu, Lee, Wu, Liu, and Birkholz, 2013). Relatively lower formants' dispersion (i.e., D_f , the relative distance between two consecutive formants, which is correlated to the vocal tract length), were also found to be more attractive in male voices (Hodges-Simeon et al., 2010; Pisanski and Rendall, 2011). Although two studies have found non-significant relationships (Babel, McGuire, and King, 2014; Feinberg et al., 2005), the former reported that

larger females tended to prefer increased apparent vocal tract size (which positively correlates with a larger body size) while the latter reported that lower first formants' frequencies for the vowels /i/ and /u/ were judged as more attractive; still, both studies suggested that apparent vocal tract size influences vocal attractiveness. Additionally, although it has received little attention compared to the F0 and Df, one study has reported that lower F0-SD (i.e., the evolution of F0 through time, which acoustically correlates to micro variations of intonation patterns in continuous speech) was more attractive in men (Hodges-Simeon et al., 2010), although two other studies have reported the opposite relationship (Bruckert et al., 2006; Leongómez, Binter, Kubicová, Stolařová, Klapilová, Havlíček, and Roberts, 2014).

Under the scope of human sexual selection, three ultimate accounts can be invoked to explain the relationships between females' preferences and men's voices. Firstly, there is intersexual selection, which corresponds to the selection exerted by one sex over another. For instance, lower F0s were found to be positively associated to higher circulating testosterone levels in men (Dabbs and Mallinger, 1999; Evans, Neave, Wakelin, and Hamilton, 2008; Hodges-Simeon, Gurven, and Gaulin, 2015; Jost, Fuchs, Loeffler, Thiery, Kratzsch, Berger, and Engel, 2018; although see Arnocky, Hodges-Simeon, Ouellette, and Albert, 2018; Bruckert et al., 2006; Puts, Apicella, and Cardenas, 2012), which is known to act as an immunosuppressant (Foo, Nakagawa, Rhodes, and Simmons, 2017). As men possessing high testosterone levels should have a better immune system to bear its costs, lower F0s may thus signal health status as a result of possessing 'good genes' (Folstad and Karter, 1992). If so, females may then be attracted to such men as they represent higher genetic quality mates (Arnocky et al., 2018; Hodges-Simeon et al., 2015). Secondly, there is intrasexual selection, which corresponds to competition among same-sex individuals. For instance, it has been regularly shown that lower F0s and Dfs were perceptually associated to larger, stronger, more masculine and more socially and physically dominant men (Hodges-Simeon et al., 2010;

Pisanski, Fraccaro, Tigue, O'Connor, and Feinberg, 2014a; Puts, Gaulin, and Verdolini, 2006; Puts, Hodges, Cárdenas, and Gaulin, 2007; Rendall, Vokey, and Nemeth, 2007; Sell et al., 2010), with F0 being recently argued to signal formidability (Puts and Aung, 2019; although see Feinberg, Jones, and Armstrong, 2019). Additionally, lower F0-SD (i.e., monotonous voices) has been hypothesized to be a marker of self-confidence and experience and is also associated to perceived dominance in men (Hodges-Simeon et al., 2010). In this context, if women are attracted to more dominant and formidable men, then the formers might display a preference for lower F0s and Dfs. Lastly, a sensory bias may explain vocal attractiveness relationships. Humans possess a cognitive bias to associate deeper vocal frequencies to perceptually larger individuals (Pisanski and Rendall, 2011; Rendall, Vokey, and Nemeth, 2007; Xu et al., 2013), although the relationships between vocal pitch and resonant frequencies with height and weight are relatively weak (Pisanski, Fraccaro, Tigue, O'Connor, Röder, Andrews, Fink, DeBruine, Jones, and Feinberg, 2014b). Nonetheless, if women actually prefer larger men as mates, then they might also prefer men with perceptually deeper vocal features.

According to the source-filter theory of speech production (Taylor and Reby, 2010), the underlying mechanisms of phonation in humans rests on the larynx (the source) and the subsequent filtering of vocal signals by the supralaryngeal vocal tract (the filter). The airflow expelled from the lungs and forced out through the glottis causes mechanical oscillations of the vocal folds within the larynx (i.e., Bernoulli's principle). The tension, length and thickness of vocal folds determine the vocal height, which acoustically correlates to the fundamental frequency (i.e., F0). Namely, the sound waves produced by the vocal folds' oscillations travel through the pharyngeal, the oral and (possibly) the nasal cavities before being expelled. During this process, the vocal tract configuration filters the laryngeal flow generated at the glottis by amplifying some frequencies to the detriment of others and, thereby, producing the formant frequencies that lead to the perception of vocal timbre. Moreover, the movements of the

articulatory organs involved in speech production such as the tongue, the lips and the palate modify the shape of the vocal tract, which determine the frequencies associated to the different speech sounds. In humans, both pitch and resonant frequencies display salient sex differences. Indeed, at puberty, males experience a significant influence of androgens, especially testosterone, which entails important consequences on larynx size and vocal folds thickness and length, which acoustically lower the voice pitch, deepen the resonant frequencies and reduce their spacing. This proximate mechanism explains why before puberty, boys and girls exhibit similar vocal frequencies (Fitch, 1999), until the former practically do not overlap with those of adult females (Titze, 1989). Additionally, in the adult life, inter-individual variations in vocal features are influenced by age (Linville and Fisher, 1985; Shipp et al., 1992), circulating androgens level (Abitbol, Abitbol, and Abitbol, 1999; Akcam, Bolu, Merati, Durmus, Gerek, and Ozkaptan, 2004; Dabbs and Mallinger, 1999) and, possibly, to the exposure of testosterone in-utero (Fouquet, Pisanski, Mathevon, and Reby, 2016).

Fundamental and formant frequencies aside, a few understudied vocal features also seem to contribute to vocal quality, such as vocal breathiness and vocal roughness. Firstly, vocal breathiness can be captured by the harmonics-to-noise ratio (HNR), which corresponds to a ratio between periodic components (i.e., the harmonics, which are multiple integer of the F_0) and a non-periodic component (i.e., noise) comprising a segment of voiced speech (Teixeira, Oliveira, and Lopes, 2013). More specifically, this ratio reflects the efficiency of speech production. The greater the airflow expelled from the lungs into energy of vibration of the vocal folds, the higher the HNR, which is perceptually associated with a more sonorant and harmonic voice. Conversely, a lower HNR is generally associated with a perceptually asthenic, dysphonic and breathier voice. Secondly, vocal roughness can be captured by the jitter, a measure of the F_0 disturbance, which is defined as the parameter capturing the frequency variation at the glottis from cycle to cycle in the sound wave (Hillenbrand, 1988; Rabinov, Kreiman, Gerratt, and

Bielamowicz, 1995; Wendahl, 1966). More specifically, the jitter measures the regularity of the vocal folds during successive periods of oscillations. The higher the jitter, the “rougher” sounds the voice. Although little is known about their physiological mechanisms, it has been suggested that both acoustic components may be sensitive to hormonal influx as they both relate to the oscillations of the vocal folds, which possess receptors to circulating androgens (Pisanski, Jones, Fink, O'Connor, DeBruine, Röder, and Feinberg, 2016).

Vocal breathiness has been suggested to be an important component of vocal attractiveness in female voices (Babel et al., 2014; Van Borsel, Janssens, and De Bodt, 2009), but significant relationships have been reported in both sexes (Šebesta, Kleisner, Tureček, Kočnar, Akoko, Třebický, and Havlíček, 2017; Xu et al., 2013). Thus, lower HNR profiles (i.e., breathy voices) have been suggested to be more attractive. Additionally, it has been suggested to soften the aggressiveness of males with larger body size (Xu et al., 2013), which in turn could increase their overall attractiveness towards females. On the other hand, little evidence is actually known on whether vocal roughness (as measured with the jitter) significantly contributes to perceived vocal attractiveness as studies that have directly tackled the topic have led to mixed results (Babel et al., 2014; Hughes, Mogilski, & Harrison, 2014; Hughes, Pastizzo, & Gallup, 2008).

Interestingly, experimental consensus regarding the F0 strongly suggests that women’s vocal preferences are consistent independently of the culture under study. Negative relationships have been mostly reported in English-speaking populations such as Americans (Hodges-Simeon et al., 2010), Canadians (Feinberg et al., 2005; Pisanski and Rendall, 2011), British (Jones et al., 2010; Vukovic et al., 2008), Scottish (Saxton, DeBruine, Jones, Little, and Roberts, 2009), and Australians (Simmons, Peters, and Rhodes, 2011), but also in Dutch (Collins, 2000), German (Weiss and Burkhardt, 2010), Czech (Valentová, Roberts, and Havlíček, 2013), Latvians (Skrinda, Krama, Kecko, Moore, Kaasik, Meija, Lietuviētis, Rantala,

and Krams, 2014) and in a small sample of French speakers (Bruckert et al., 2006). Although evidence is scarce, a few findings challenge this view, suggesting that vocal attractiveness may rest on different acoustic cues depending on the culture under study. For instance, one study reported that in a Filipino-speaking group sample, both nulliparous and breastfeeding women showed a preference for feminized (i.e., higher F0) rather than masculinized voice pitch (i.e., lower F0) (Shirazi, Puts, and Escasa-Dorne, 2018). In the Hadzas, it has also been reported that women who are breastfeeding prefer men with higher pitch voices as mates, those who are not breastfeeding preferring lower pitch male voices (Apicella and Feinberg, 2009). Interestingly, another study found that Namibian men's vocal attractiveness could be predicted by their degree of vocal breathiness (measured through the HNR) and not by their voice pitch (Šebesta et al., 2017).

In this context, the aim of this replication study is to investigate culture-dependency for vocal attractiveness in an underrepresented language (i.e., French) as well as investigating attractiveness relationships with understudied acoustic features of vocal quality.

Material and Methods

This study was conducted in Montpellier, France. The French National Commission of Informatics and Liberty approved the experimental designs of the present study (CNIL number 2-17029). Prior to the study, all participants provided the investigator with their written consent.

a. Stimuli

An aggregate of 58 male participants (mean age = 23; SD = 3.36), native speakers of French, produced the vocal stimuli. These participants were drawn from another study (Suire, Raymond, and Barkat-Defradas, 2018; two of which were not included in that study). They were seated in a quiet, anechoic, soundproof room equipped with a Sennheiser™ BF 515 microphone connected to a PC located in another room. Vocal samples consisted in the recording of a short utterance '*Dans la vie, je pense toujours prendre les bonnes décisions et*

c'est pour cela que je vais gagner' (i.e., 'In life, I always think I'll make the right decision and that is why I will win'). To control for intensity, participants were asked to speak at a constant distance of 15 cm from the microphone. All recordings were encoded using the Adobe© Audition CS6 at a sampling rate of 44 kHz – 32 bit – mono then saved as .wav files.

b. Acoustic analyses

All recordings were analyzed using the Praat© voice analysis software (version 6.0.31, Boersma and Weenink, 2018). The mean fundamental frequency (F0) and its variation (F0-SD) were measured using the autocorrelation method with a pitch floor of 75 Hz and a ceiling of 300 Hz (Praat's recommendation), with other settings kept as default. The harmonics-to-noise ratio (HNR, in dB) and the local jitter (%), which corresponds to the average absolute difference between consecutive periods, divided by the average period, and calculated in percentage, were measured across the entire utterance using the same settings as the F0. The local jitter corresponds to the jitter ratio, which is commonly used to describe vocal perturbations (Jones, Trabold, Plante, Cheetham, and Earis, 2001). Additionally, intensity (dB) was retrieved using Praat's default settings. Formant frequencies (F1 to F4) were measured at each glottal pulse, targeting voiced speech only, using a formant ceiling of 5000 Hz (Praat's recommendation), then averaged across the entire utterance. Then, the formants' dispersion (Df) was calculated using the following formula (Fitch, 1997):

$$Df = \frac{\sum_{i=1}^{N-1} F_{i+1} - F_i}{N - 1},$$

where Df is the formant dispersion (in Hz), N is the total number of formants measured, and F_i is the frequency (in Hz) of formant i . Lastly, we computed the formants' position (Pf) using the method described in Puts et al. (2012), which has been argued to be sexually more dimorphic than Df. To compute the formants' position, we used female vocal stimuli that were drawn from the same study of the male vocal stimuli ($n_{\text{female}} = 68$, Suire et al. 2018).

Descriptive statistics of the male vocal stimuli for each acoustic feature are reported in Table 1 and their zero-order correlations in Table 2. Mean F0 was positively correlated with F0-SD ($r = 0.56$, $p < 0.001$). Df was positively associated to Pf ($r = 0.31$, $p = 0.019$) and HNR ($r = 0.35$, $p = 0.008$). Lastly, HNR was negatively correlated with jitter ($r = -0.57$, $p < 0.001$). All these correlations are consistent with those reported in the literature (for F0 and F0-SD, see Hodges-Simeon et al., 2010; for Df and Pf see the open data of Han, Wang, Fasolt, Hahn, Holzleitner, Lao, DeBruine, Feinberg and Jones, 2018; for jitter and HNR, see de Krom, 1993), except the correlation between Df and HNR, which to our knowledge was not reported elsewhere.

n = 58	Mean	SD	Ranges
Mean F0 (Hz)	114.47	11.84	85.44 – 140.07
F0-SD (Hz)	15.16	5.06	6.97 – 28.31
Df (Hz)	1086.78	36.60	1005 – 1181
Pf (Hz)	-1.61	0.47	-2.47 – -0.65
HNR (dB)	11.32	1.37	7.93 – 14.94
Jitter (%)	2.68	0.47	1.83 – 4.41
Intensity (dB)	64.73	3.61	53.96 – 76.93

Table 1. Descriptive statistics of the acoustic characteristics of the vocal stimuli.

	Mean F0 (Hz)	F0-SD (Hz)	Df (Hz)	Pf (Hz)	HNR (dB)	Jitter (%)	Intensity (dB)
Mean F0 (Hz)	1						
F0-SD (Hz)	0.56***	1					
Df (Hz)	-0.16	-0.13	1				
Pf (Hz)	0.16	0.10	0.31*	1			
HNR (dB)	0.13	-0.24	0.35**	-0.06	1		
Jitter (%)	-0.15	0.20	0.13	-0.14	-0.57***	1	
Intensity (dB)	0.13	0.02	0.24	0.05	0.22	-0.08	1

Table 2. Zero-order correlations between each acoustic feature for the vocal stimuli.

Significance code: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

c. Experimental procedure

The experimental procedure was automated on an online computer-interfaced program. 224 French female raters participated in a perceptual study after they self-reported in a questionnaire

their age, origins of parents and grandparents (to control for potential cultural preferences), sexual orientation (to control for sexual preferences) and whether they suffered from a hearing impairment (note that other information were reported but are not used in the present study). After filling out the questionnaire, female raters were presented with a series of 11 choices each including a pair of voices. For each pair, two stimuli were randomly selected from the whole pool of vocal stimuli. The two vocal stimuli were randomized in their position presented in each pair (left or right position) on the computer screen. Judges were asked to choose the most attractive vocal stimulus by clicking on it. Participants were allowed to listen to the stimuli as much as they wanted. However, when the female judge made her choice, she could not go back to the previous one anymore. To measure intra-rater reliability, the second and third pairs were the same as the tenth and eleventh pairs.

Although a forced choice paradigm is usually implemented with experimentally manipulated vocal stimuli (e.g. Jones et al., 2010; Re, O'Connor, Bennett, and Feinberg, 2012), there is fundamentally no advantage or disadvantage between a forced-choice paradigm and a correlational rating study for either manipulated or non-manipulated stimuli. Crucially, it does not yield different results (e.g. for women's preferences of men's F0, for experimental designs see: Vukovic et al. 2008; Jones et al. 2010; Re et al. 2012; and for correlational designs see: Feinberg et al. 2005; Hodges-Simeon et al. 2010; Pisanski and Rendall 2011).

We stopped collecting data when each voice of the 58 voices was heard at least 40 times in order to obtain statistically relevant data. In the end, the mean number of times a voice has been heard is $M \pm SD = 54.14 \pm 6.55$, with 72 and 42 times respectively for the most and least heard voices.

Out of the 225 female participants who completed the questionnaire, 137 participants completed all 11 decisions, 28 participants skipped some of the decisions (mean number of skipped decisions = 8.75), for a total of 1570 decisions in our analyses. Description of the

judges' characteristics that completed at least one pair ($n = 165$, $M \pm SD = 28.95 \pm 14.16$) are given in Table 3.

		n
Completed the full test		
	<i>No</i>	28
	<i>Yes</i>	137
Ancestry		
	<i>European</i>	135
	<i>Non-European</i>	30
Sexual orientation		
	<i>Heterosexual</i>	142
	<i>Homosexual</i>	4
	<i>Bisexual</i>	11
	<i>Not reported</i>	8
Hearing impairment		
	<i>No</i>	161
	<i>Yes</i>	3
	<i>Not reported</i>	1

Table 3. Number of judges for each of the following categories: those who completed the full test (i.e., heard all the pairs), grandparents' ancestry, sexual orientation and hearing impairments.

d. Data analysis

To analyze women's preferences for men's voices, a generalized linear mixed model (GLMM) was used with the response variable being if the female judge chose or not the voice presented to her on the left position. The GLMM was fitted with a binomial error structure since the response variable consisted in a discrete probability distribution of the number of successes in a sequence of several independent trials. In order to explore acoustics' preferences, seven predictor variables were computed and corresponded to the differences observed in mean F0, F0-SD, Df, Pf, HNR, jitter and intensity between the two vocal stimuli (numerical variables that were standardized). Judges' age (standardized variable), ancestry (i.e., European or non-European grandparents') and sexual orientation (i.e., heterosexual and non-heterosexual) were added as control variables and put in interaction with the differences in acoustics characteristics

to assess their influence on voice preferences. Judges' identities and the vocal stimuli were added as random effects as intercepts only. A symbolic representation of the GLMM is given in the supplementary material.

GLMMs with and without the control variables were performed to explore any statistical differences. Moreover, we performed two additional GLMMs, one without individuals with hearing impairment and one without individuals who did not report sexual orientation (these individuals were treated as non-heterosexual in the main GLMM). The significance of each predictor in all GLMMs was assessed from the comparison of the model excluding the predictor with the model including all the other predictors (i.e., likelihood-ratio chi-square tests, ANOVA type III). Additionally, since some acoustic variables are highly correlated (see Table 2), we conducted multicollinearity checks on the GLMMs using the variation inflation factors (VIFs).

All statistical analyses were performed under the R software (version 3.4.0), using the following packages: 'lme4' to build the generalized linear models with random effects (Bates, Mächler, Bolker, and Walker, 2014), 'car' to compute the statistical significance of each predictor and check potential multicollinearity problems for the GLMMs (Fox, Weisberg, and Fox, 2011) and 'MuMIn' to compute the pseudo-R₂ (Bartoń, 2018). In order to illustrate the results with figures, we used 'boot' to transform the coefficients of the GLMMs back into probabilities (Canty and Ripley, 2012), 'dplyr' to compute the predictions of the model (Wickham, François, Henry, and Müller, 2018) and 'ggplot2' for the resulting figures (Wickham, 2009).

Results

Descriptive statistics of the mean difference in acoustic features are reported in Table 4.

	Mean	SD	Ranges
Difference in mean F0	-0.38	16.70	-53.28 – 49.84
Difference in F0-SD	-0.066	6.89	-20.79 – 20.43
Difference in Df	1.25	51.73	-176.66 – 176.66
Difference in Pf	0.003	0.66	-1.81 – 1.81
Difference in HNR	-0.0086	1.91	-5.73 – 5.58
Difference in jitter	0.013	0.64	-2.58 – 2.58
Difference in intensity	0.065	5.06	-20.63 – 22.97

Table 4. Descriptive statistics for the unstandardized mean difference for each acoustic feature summarized over the total number of observations ($n = 1570$).

We computed intra-rater reliability scores by calculating the proportion of identical chosen vocal stimuli between the second and third first pairs with the tenth and eleventh pairs. Intra-rater reliability was high: $M \pm SD = 0.791 \pm 0.257$, i.e., judges considered on average more than 2/3 the same voices as attractive.

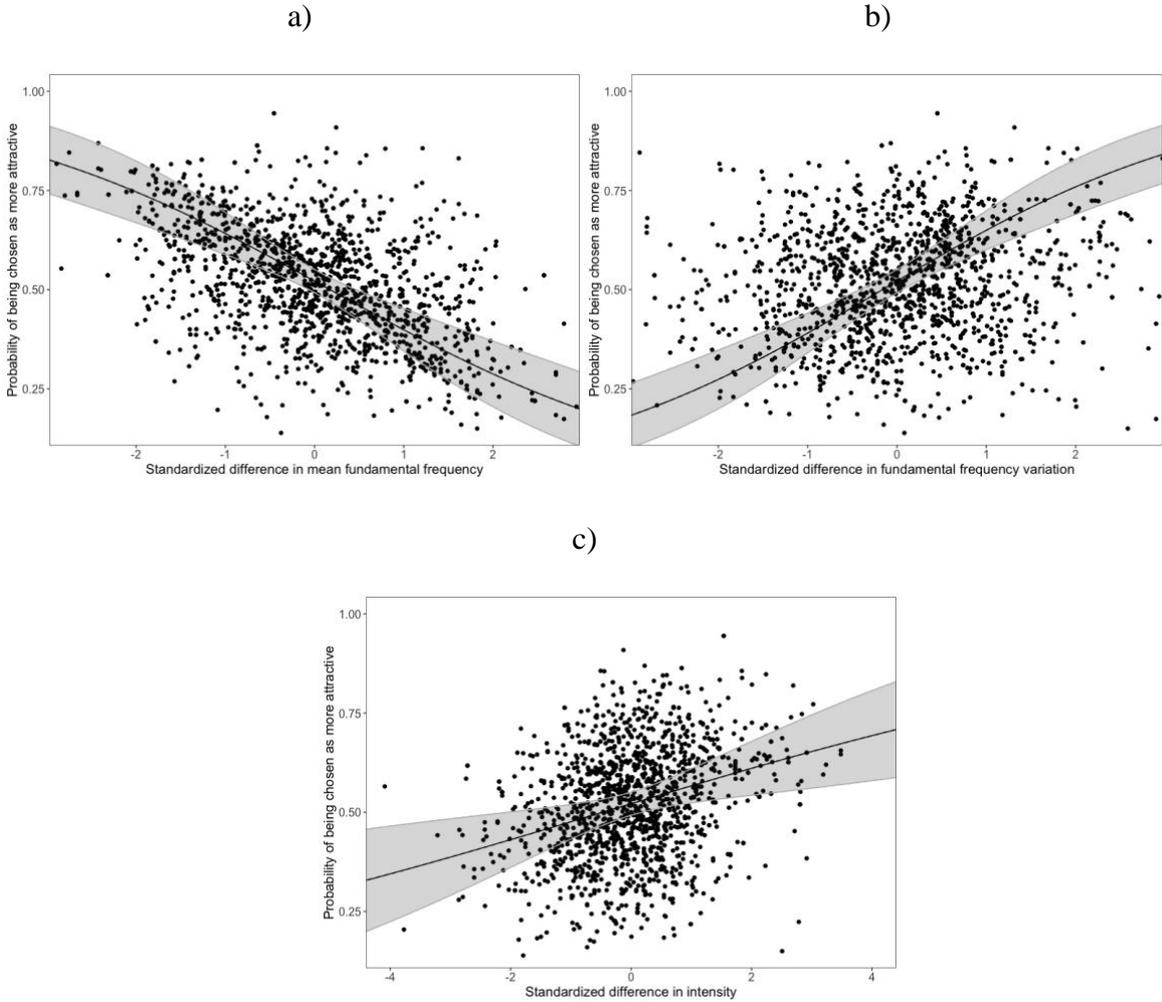
Results of the main GLMM are reported in Table 5. VIFs were all inferior to 4, indicating no problems of multicollinearity. When presented with two voices, women preferred lower F0 ($\chi^2_1 = 24.89$, $p < 0.001$), higher F0-SD profiles ($\chi^2_1 = 34.00$, $p < 0.001$) and louder stimuli ($\chi^2_1 = 7.52$, $p = 0.006$).

	Estimate	SE	χ^2	<i>p</i> value
Intercept	0.09	0.06	/	/
Difference in mean F0	-0.49	0.10	24.89	<0.001
Difference in F0-SD	0.53	0.09	34.00	<0.001
Difference in Df	0.18	0.10	3.26	0.070
Difference in Pf	-0.06	0.08	0.56	0.452
Difference in HNR	-0.12	0.10	1.23	0.266
Difference in jitter	-0.04	0.09	0.27	0.602
Difference in intensity	0.18	0.06	7.52	0.006
Interactions with age				
Difference in F0	0.16	0.09	2.86	0.090
Difference in F0-SD	0.04	0.09	0.25	0.616
Difference in Df	0.13	0.09	2.06	0.151
Difference in Pf	-0.06	0.07	0.70	0.399
Difference in HNR	-0.11	0.09	1.31	0.251
Difference in jitter	0.10	0.08	1.61	0.204
Difference in intensity	0.15	0.06	5.65	0.017
Interactions with ancestry				
Difference in F0	-0.008	0.22	0.001	0.968
Difference in F0-SD	-0.41	0.20	3.97	0.046
Difference in Df	0.04	0.23	0.03	0.863
Difference in Pf	-0.17	0.18	0.82	0.364
Difference in HNR	-0.01	0.25	0.003	0.953
Difference in jitter	0.06	0.21	0.09	0.752
Difference in intensity	-0.10	0.17	0.37	0.539
Interactions with sexual orientation				
Difference in F0	0.15	0.24	0.38	0.534
Difference in F0-SD	-0.54	0.23	5.49	0.019
Difference in Df	-0.14	0.23	0.36	0.544
Difference in Pf	-0.10	0.18	0.28	0.593
Difference in HNR	-0.11	0.28	0.15	0.691
Difference in jitter	0.18	0.24	0.60	0.436
Difference in intensity	0.27	0.18	2.29	0.130

Table 5. Results of the GLMM predicting women’s preferences for men’s voices, ($n_{\text{stimuli}} = 58$, $n_{\text{judges}} = 165$, $n_{\text{observations}} = 1570$). For each variable, the χ^2 and the *p* values associated from the likelihood-ratio chi-square test of the comparison between the full model and the model without the predictors and the control variables are given (ANOVA type III). For the categorical variables’ ‘ancestry’ and ‘sexual orientation’, the estimates are given compared to the reference category (1 = European ancestry and 1 = heterosexual). *P* values are considered significant at the 0.05 threshold (in bold). The degrees of freedom is 1 for every test.

For easier understanding of the model's output, the predicted probabilities of considering a voice more attractive than the other within the same pair were plotted against the range of differences in mean F0, F0-SD and intensity between the two voices (Figure 1).

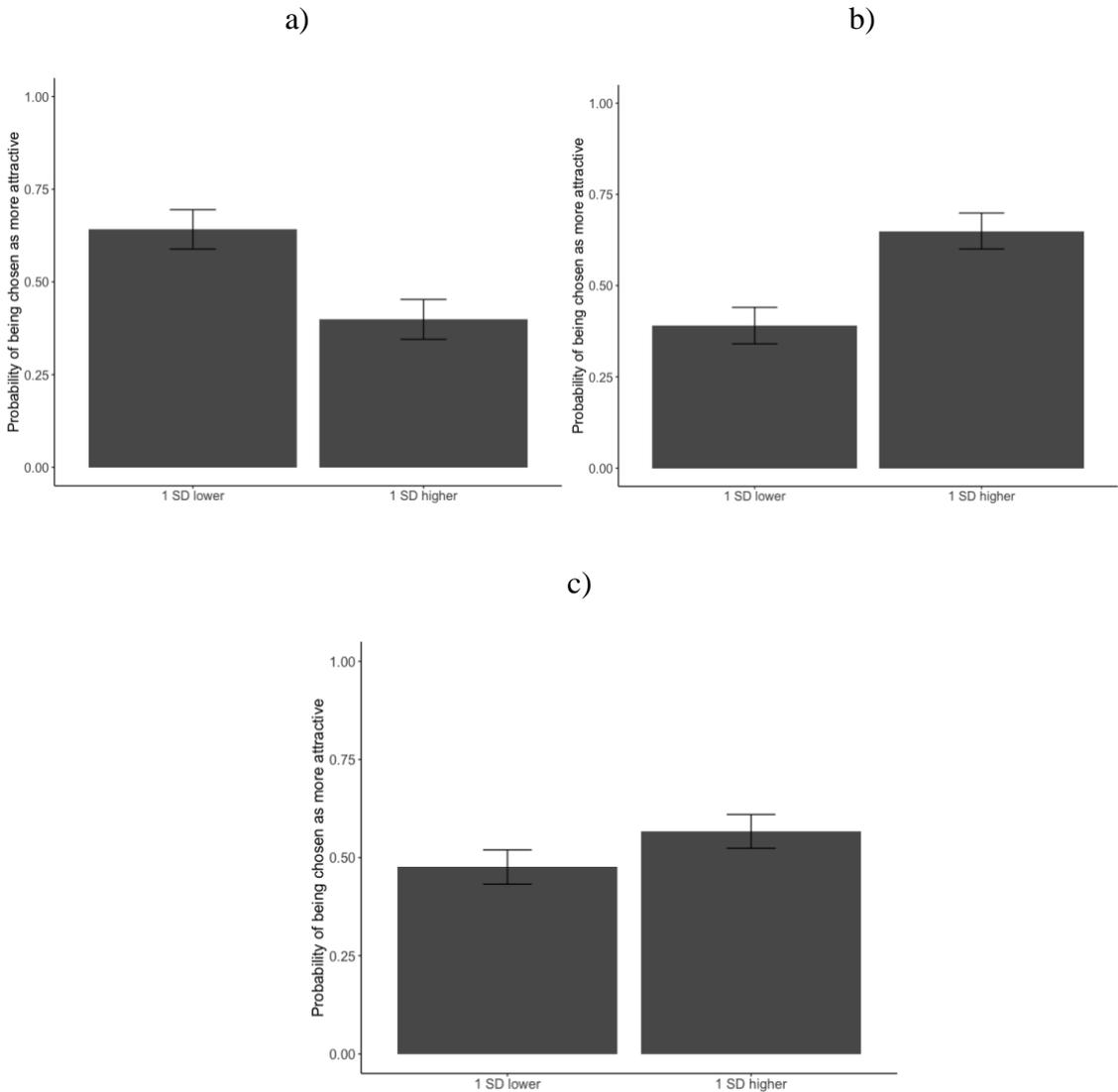
Figure 1. Probabilities of being picked as more attractive plotted against the standardized differences between the two voices heard in a) mean F0, b) F0-SD and c) intensity. The black curves represent the model's predictions associated with 95% confidence intervals (in grey).



We also computed the predicted probability that a voice would be considered more attractive when it is 1 standard deviation lower and 1 standard deviation higher than the opposite one on the basis of their F0, F0-SD and intensity (Figure 2). A voice with a mean F0 that is one standard deviation lower than the other in the same pair has a probability of being picked as

more attractive up to ~65%, likewise, a voice with a F0-SD which is 1 standard deviation higher has a probability of being picked as more attractive up to ~65%.

Figure 2. Barplots of the predicted probabilities that a voice would be considered more attractive when it is 1 standard deviation lower and 1 standard deviation higher than the other voice, as a function of its a) mean F0, b) F0-SD and c) intensity. Bars are associated with 95% confidence intervals.



Additionally, female judges did not show directional preferences for Df, Pf, HNR or jitter (all p values > 0.05). Judges' age had a significant influence on their preferences for intensity ($\chi^2_1 = 7.52, p = 0.006$), i.e., relatively older women preferred louder vocal profiles.

Women with non-European ancestry and non-heterosexual women showed a preference for lower F0-SD profiles (respectively $\chi^2_1 = 3.97$, $p = 0.046$; $\chi^2_1 = 5.49$, $p = 0.019$). The model explained 12% of the variance in vocal preferences, including fixed and random effects. Lastly, the variance of the random intercept for judges was higher than the vocal stimuli ($\sigma_{\text{judges}} = 0.07$; $\sigma_{\text{stimuli}} = 0.01$).

The model without ancestry and the one without sexual orientation were not statistically different from the full model (respectively $\chi^2_7 = 10.42$, $p = 0.165$; $\chi^2_7 = 9.96$, $p = 0.190$). Removing age from the model was statistically different from the full model ($\chi^2_7 = 18.74$, $p = 0.009$). The models without judges with hearing impairment and without judges who did not report sexual orientation did not qualitatively change the results. In all models, the main results remained the same: female judges still considered voices with lower F0, higher F0-SD and higher intensity as more attractive. All models without the control variables are given in the supplementary material.

Discussion

Women significantly preferred lower vocal pitch in men. This result is consistent with previous findings in English-speaking populations (Feinberg et al., 2005; Hodges-Simeon et al., 2010; Hughes et al., 2010; Jones et al., 2010; Pisanski and Rendall, 2011; Vukovic et al., 2008) and several other languages (Bruckert et al., 2006; Skrinda et al., 2014; Valentová et al., 2013; Weiss and Burkhardt, 2010). Moreover, this finding has been replicated with a similar or higher number of stimuli and judges than most of these studies (see Hodges-Simeon et al., 2010 for an example of a study with a higher number of stimuli). As vocal height correlates to several biological and social information about men, such as testosterone levels (Dabbs and Mallinger, 1999; Evans et al., 2008; Hodges-Simeon et al., 2015), sexually related behaviors (Hughes et al., 2004), body size assessments (Pisanski et al., 2014a), as well as signaling social dominance (Puts et al., 2007) and social rankings (Cheng, Tracy, Ho, and Henrich, 2016), women may rely

on this salient acoustic cue as an assessment of sexual partner quality. Several studies have reported that men exhibiting relatively low-pitched voices reported a higher mating success in industrialized societies (Hodges-Simeon, Gaulin, and Puts, 2011; Puts, 2005; Puts et al., 2006; although see Suire et al., 2018) and a higher reproductive success in a hunter-gatherer society (Apicella, Feinberg, and Marlowe, 2007; although see Smith, Olkhov, Puts, and Apicella, 2017).

Moreover, French women also significantly preferred higher F0-SD profiles in men, that is, more expressive (or less monotonous) voices. Although our study had a higher number of judges and stimuli than the two others that reported the same relationship (Bruckert et al., 2006; Leongómez et al., 2014), another study had a higher number of stimuli but less judges (Hodges-Simeon et al., 2010). Nonetheless, while self-confidence and experience can be expressed through monotonous voices, to which some women may be more attracted to (Hodges-Simeon et al., 2010), our results do not follow the same tendency. A possible explanation may be that more marked intonation patterns might be perceived as more attractive as it is a marker of perceived state-dependent qualities such as positive emotions (e.g. joy and happiness) (Banse and Scherer, 1996), conversational interest as well as emotional activation (i.e., arousal) and intensity (Laukka, Juslin, and Bresin, 2005). Ultimately, expressive voices could reflect the speaker's current mental-health state since it has been previously reported that clinically depressed patients show typically reduced F0-SD values (Ellgring and Scherer, 1996). Thus, higher F0 variability may be associated to more enthusiastic and extroverted individuals, to which women may be more attracted. In this sense, our result is consistent with previous findings in both men and women (Bruckert et al., 2006; Leongómez et al., 2014). Although it has been suggested to be a cue of femininity, as women display twice as much F0 variation, we suggest that irrespective of sex, higher F0-SD profiles should be perceived as more attractive.

No directional preferences were observed for the formants' dispersion and position, which corroborates some previous findings (Babel et al., 2014; Feinberg et al., 2005), using a higher or similar number of stimuli and a higher number of judges. Several studies have suggested that Df may be a more important vocal cue to assess in human competitive settings. Indeed, it has been reported that lower Df patterns were associated to perceived dominance in men (Puts et al., 2007; Wolff and Puts, 2010). This can be explained by the fact that lower Df patterns are associated to larger body size (Pisanski et al., 2016) and to perceived larger individuals (Bruckert et al., 2006; Collins, 2000; Rendall et al., 2007). Interestingly, females were also found to be more sensitive to this vocal cue than men after hearing women's voices (Puts, Barndt, Welling, Dawood, and Burriss, 2011). Such results emphasize the idea that same-sex individuals may use Df to track competitor's masculinity and/or femininity. Similarly, some research suggest that the formants' position may signal threat potential among men (Puts et al., 2012), although a recent study found no correlations to physical strength (Han et al., 2018).

Our results also indicated that vocal breathiness and roughness (assessed respectively through the HNR and the jitter ratio) did not significantly contribute to men's vocal attractiveness, using a higher number of stimuli and judges than previous studies (Babel et al., 2014; Hughes et al., 2014, 2008). Although one study reported that breathier voices were found to be more attractive in Namibian men, ours did not (Šebesta et al., 2017). Another study found that perceived 'breathy' voices were significantly more attractive in both sexes (Xu et al., 2013), although the underlying acoustic component was not clearly identified in this study. Lack of significant findings for breathiness suggests that it is more associated with feminine vocal quality, as previously suggested (Henton and Bladon, 1985; Van Borsel et al., 2009). It is also possible that when assessing attractiveness, women may be particularly attuned to the vocal features that are indicative of one's heritable mate quality, such as the F0. In this context, breathiness and roughness may not reliably indicate mate or competitor quality for listeners, at

least in men. Although they are correlated to other body features (see Pisanski et al., 2016 for an extensive study on that matter), further studies are needed to understand whether these two acoustic components of the human voice are perceptually salient in influencing vocal attractiveness. Otherwise, it has been suggested that HNR and jitter may be indicative of current hormonal profiles as both parameters relate to the oscillations of the vocal folds, which possess many cellular receptors to androgens (Pisanski et al., 2016).

An important limitation to the current study is that we did not investigate the effects of women's menstrual cycle upon perceived vocal attractiveness. Indeed, there was more variations between females judges than between vocal stimuli ($\sigma_{\text{judges}} = 0.07$; $\sigma_{\text{stimuli}} = 0.01$), suggesting, for example, that the timing of the ovulatory cycle may play a role. In fact, it has been long suggested that menstrual phase and mating contexts may influence women's preferences for masculine vocal attributes (Feinberg et al., 2006; Pisanski, Hahn, Fisher, DeBruine, Feinberg, and Jones, 2014c; Puts, 2005). Under the 'good genes ovulatory shift hypothesis', women in their fertile phase are predicted to shift their preferences towards mates indicating high genetic quality (i.e., more masculine men, to which women may be particularly attracted to for a short-term relationship, such as a one-night stand), as opposed to mates indicating high parental investment in their non-fertile phase (i.e., less masculine men, to which women may be particularly attracted to for a long-term, committed and romantic relationship) (Jünger, Kordsmeyer, Gerlach, and Penke, 2018). These shifting preferences have been suggested to be an adaptive strategy in order to maximize fitness benefits for women.

For instance, Puts (2005) found that females judged lowered pitch voices more attractive than the same voices raised in pitch in their fertile phase of their ovulatory cycle with respect to a short-term context. Similarly, Feinberg et al. (2006) found that women's masculinity preferences for low-pitched voices were stronger during the fertile phase. Although the effect was not significant, Pisanski et al. (2014c) also reported stronger preferences for

masculinized voice pitch. Lastly, one study has reported that women in their fertile phase significantly preferred lowered Df when questioned for both short- and long-term relationships (Hodges-Simeon et al., 2010). The authors also found that mean $F0$ and attractiveness was strongest for fertile-phase women rating short-term attractiveness, while $F0-SD$ was more attractive for non-fertile phase female rating short-term attractiveness and fertile females rating long-term attractiveness. However, recent evidence have suggested that women menstrual cycle does not influence their preferences for masculinized bodies and faces (Jones, Hahn, Fisher, Wang, Kandrik, Han, Fasolt, Morrison, Holzleitner, O'Shea, Roberts, Little, and DeBruine, 2017; Marcinkowska, Galbarczyk, and Jasienska, 2018). Using a large sample size and a more methodologically grounded procedure, Jünger et al. (2018) found no effect of the cycle phase, conception risk and steroid hormone levels on women's auditory preferences for men's voices. Further research is thus needed to reliably investigate if the menstrual cycle has a significant effect over shifted preferences. In any case, not controlling for this factor will only provide conservative results, under the hypothesis that the time of the menstrual cycle is randomly distributed among the participating women.

Other limitations include the difference in age between men who provided the vocal stimuli and the female judges. However, in our sample both the youngest individual who provided the vocal stimuli and the youngest female judge were aged 18, which is largely above the age where mate preferences develop and become relevant (age 13-15, Saxton, Caryl, and Craig Roberts, 2006; Saxton, DeBruine, Jones, Little and Roberts, 2009). Moreover, an interesting perspective for future research would be to investigate possible non-linear effects of preferences as a function of vocal parameters. Indeed, extreme values for a particular vocal parameter may be perceived as pathological (as it is the case for high values of jitter and low values of HNR, Teixeira et al., 2013) or perceived as immature and/or too feminine (e.g. high $F0$). To our knowledge, only one study has tackled this topic in women's preferences for men's

F0, and it was found that women did not prefer vocal pitches below the ~96 Hz threshold. This suggests that preferences may contribute to stabilizing selection pressure for low pitch in men's voices (Re et al., 2012). Interestingly, in men's preferences for the F0 of women, one study reported a non-linear relationship with attractiveness ratings starting to decrease when the F0 is higher than ~260 Hz (Borkowski and Pawlowski 2011), although two studies have reported that there was no upper limit (Feinberg, DeBruine, Jones, and Perrett, 2008; Re et al., 2012).

Conclusions

The current study adds to the body of literature on vocal attractiveness in an underrepresented language (i.e., French). Although voice pitch findings were replicated, confirming women's preferences for low-pitched masculine voices, most of the other acoustic features investigated in this study did not yield to significant results, leading us to conclude that variations in resonant frequencies' spacing, breathiness and roughness do not seem to be important contributors of men's vocal attractiveness, at least in a French-speaking sample. Further studies should explore these relationships in other cultures so as to reaffirm these findings.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of conflicting interests

The authors declare that there is no conflict of interest.

Data availability

The data and the R code from this study can be found at:

<https://figshare.com/s/cab62d1e411503982c91>

References

- Abitbol, J., Abitbol, P., & Abitbol, B. (1999). Sex hormones and the female voice. *Journal of Voice*, 13(3), 424-446. [https://doi.org/10.1016/S0892-1997\(99\)80048-4](https://doi.org/10.1016/S0892-1997(99)80048-4)
- Akcam, T., Bolu, E., Merati, A. L., Durmus, C., Gerek, M., & Ozkaptan, Y. (2004). Voice Changes after Androgen Therapy for Hypogonadotrophic Hypogonadism: *The Laryngoscope*, 114(9), 1587-1591. <https://doi.org/10.1097/00005537-200409000-00016>
- Apicella, C. L., & Feinberg, D. R. (2009). Voice pitch alters mate-choice-relevant perception in hunter-gatherers. *Proceedings of the Royal Society B: Biological Sciences*, 276(1659), 1077-1082. <https://doi.org/10.1098/rspb.2008.1542>
- Apicella, C. L., Feinberg, D. R., & Marlowe, F. W. (2007). Voice pitch predicts reproductive success in male hunter-gatherers. *Biology Letters*, 3(6), 682-684. <https://doi.org/10.1098/rsbl.2007.0410>
- Arnocky, S., Hodges-Simeon, C. R., Ouellette, D., & Albert, G. (2018). Do men with more masculine voices have better immunocompetence? *Evolution and Human Behavior*. <https://doi.org/10.1016/j.evolhumbehav.2018.06.003>
- Babel, M., McGuire, G., & King, J. (2014). Towards a More Nuanced View of Vocal Attractiveness. *PLoS ONE*, 9(2), e88616. <https://doi.org/10.1371/journal.pone.0088616>
- Bachorowski, J. A., & Owren, M. J. (1999). Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. *The Journal of the Acoustical Society of America*, 106(2), 1054-1063.
- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of personality and social psychology*, 70(3), 614.
- Bartoń, K. (2018). *MuMIn: Multi-Model Inference*. R package version 1.42.1. Consulté à l'adresse <https://CRAN.R-project.org/package=MuMIn>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4. *ArXiv:1406.5823 [Stat]*. Consulté à l'adresse <http://arxiv.org/abs/1406.5823>

- Bruckert, L., Lienard, J.-S., Lacroix, A., Kreutzer, M., & Leboucher, G. (2006). Women use voice parameters to assess men's characteristics. *Proceedings of the Royal Society B: Biological Sciences*, 273(1582), 83-89. <https://doi.org/10.1098/rspb.2005.3265>
- Canty, A., & Ripley, B. (2012). boot: Bootstrap R (S-Plus) functions. (Version R package version, 1(7)).
- Cheng, J. T., Tracy, J. L., Ho, S., & Henrich, J. (2016). Listen, follow me: Dynamic vocal signals of dominance predict emergent social rank in humans. *Journal of Experimental Psychology: General*, 145(5), 536-547. <https://doi.org/10.1037/xge0000166>
- Collins. (2000). Men's voices and women's choices. *Animal Behaviour*, 60(6), 773-780.
- Dabbs, J. M., & Mallinger, A. (1999). High testosterone levels predict low voice pitch among men. *Personality and Individual Differences*, 27(4), 801-804. [https://doi.org/10.1016/S0191-8869\(98\)00272-4](https://doi.org/10.1016/S0191-8869(98)00272-4)
- de Krom, G. D. (1993). A cepstrum-based technique for determining a harmonics-to-noise ratio in speech signals. *Journal of Speech, Language, and Hearing Research*, 36(2), 254-266.
- Ellgring, H., & Scherer, K. R. (1996). Vocal indicators of mood change in depression. *Journal of Nonverbal Behavior*, 20(2), 83-110. <https://doi.org/10.1007/BF02253071>
- Evans, S., Neave, N., Wakelin, D., & Hamilton, C. (2008). The relationship between testosterone and vocal frequencies in human males. *Physiology & Behavior*, 93(4-5), 783-788. <https://doi.org/10.1016/j.physbeh.2007.11.033>
- Feinberg, David R, DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2008). The Role of Femininity and Averageness of Voice Pitch in Aesthetic Judgments of Women's Voices. *Perception*, 37(4), 615-623. <https://doi.org/10.1068/p5514>
- Feinberg, D.R., Jones, B. C., & Armstrong, M. M. (2019). No Evidence That Men's Voice Pitch Signals Formidability. *Trends in Ecology & Evolution*, 34(3), 190-192. <https://doi.org/10.1016/j.tree.2018.12.014>

- Feinberg, D.R., Jones, B. C., Law Smith, M. J., Moore, F. R., DeBruine, L. M., Cornwell, R. E., ... Perrett, D. I. (2006). Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. *Hormones and Behavior*, 49(2), 215-222. <https://doi.org/10.1016/j.yhbeh.2005.07.004>
- Feinberg, D.R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, 69(3), 561-568. <https://doi.org/10.1016/j.anbehav.2004.06.012>
- Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *The Journal of the Acoustical Society of America*, 102(2), 1213-1222.
- Foo, Y. Z., Nakagawa, S., Rhodes, G., & Simmons, L. W. (2017). The effects of sex hormones on immune function : A meta-analysis: Sex hormones and immune function. *Biological Reviews*, 92(1), 551-571. <https://doi.org/10.1111/brv.12243>
- Fouquet, M., Pisanski, K., Mathevon, N., & Reby, D. (2016). Seven and up : Individual differences in male voice fundamental frequency emerge before puberty and remain stable throughout adulthood. *Royal Society Open Science*, 3(10), 160395. <https://doi.org/10.1098/rsos.160395>
- Fox, J., Weisberg, S., & Fox, J. (2011). *An R companion to applied regression* (2nd ed). Thousand Oaks, Calif: SAGE Publications.
- Gelfer, M. P., & Bennett, Q. E. (2013). Speaking Fundamental Frequency and Vowel Formant Frequencies : Effects on Perception of Gender. *Journal of Voice*, 27(5), 556-566. <https://doi.org/10.1016/j.jvoice.2012.11.008>
- Gelfer, M. P., & Mikos, V. A. (2005). The Relative Contributions of Speaking Fundamental Frequency and Formant Frequencies to Gender Identification Based on Isolated Vowels. *Journal of Voice*, 19(4), 544-554. <https://doi.org/10.1016/j.jvoice.2004.10.006>

- Han, C., Wang, H., Fasolt, V., Hahn, A., Holzleitner, I. J., Lao, J., ... Jones, B. (2018). No clear evidence for correlations between handgrip strength and sexually dimorphic acoustic properties of voices. *BioRxiv*. <https://doi.org/10.1101/227165>
- Henton, C. G., & Bladon, R. A. (1985). Breathiness in normal female speech : Inefficiency versus desirability. *Language & Communication*.
- Hillenbrand, J. (1988). Perception of aperiodicities in synthetically generated voices. *The Journal of the Acoustical Society of America*, 83(6), 2361-2371. <https://doi.org/10.1121/1.396367>
- Hillenbrand, J. M., & Clark, M. J. (2009). The role of f₀ and formant frequencies in distinguishing the voices of men and women. *Attention, Perception, & Psychophysics*, 71(5), 1150-1166. <https://doi.org/10.3758/APP.71.5.1150>
- Hodges-Simeon, C. R., Gaulin, S. J. C., & Puts, D. A. (2010). Different Vocal Parameters Predict Perceptions of Dominance and Attractiveness. *Human Nature*, 21(4), 406-427. <https://doi.org/10.1007/s12110-010-9101-5>
- Hodges-Simeon, C. R., Gaulin, S. J. C., & Puts, D. A. (2011). Voice Correlates of Mating Success in Men : Examining “Contests” Versus “Mate Choice” Modes of Sexual Selection. *Archives of Sexual Behavior*, 40(3), 551-557. <https://doi.org/10.1007/s10508-010-9625-0>
- Hodges-Simeon, C. R., Gurven, M., & Gaulin, S. J. C. (2015). The low male voice is a costly signal of phenotypic quality among Bolivian adolescents. *Evolution and Human Behavior*, 36(4), 294-302. <https://doi.org/10.1016/j.evolhumbehav.2015.01.002>
- Hughes, S. M., Dispenza, F., & Gallup, G. G. (2004). Ratings of voice attractiveness predict sexual behavior and body configuration. *Evolution and Human Behavior*, 25(5), 295-304. <https://doi.org/10.1016/j.evolhumbehav.2004.06.001>

- Hughes, S. M., Farley, S. D., & Rhodes, B. C. (2010). Vocal and Physiological Changes in Response to the Physical Attractiveness of Conversational Partners. *Journal of Nonverbal Behavior*, 34(3), 155-167. <https://doi.org/10.1007/s10919-010-0087-9>
- Hughes, S. M., Mogilski, J. K., & Harrison, M. A. (2014). The Perception and Parameters of Intentional Voice Manipulation. *Journal of Nonverbal Behavior*, 38(1), 107-127. <https://doi.org/10.1007/s10919-013-0163-z>
- Hughes, S. M., Pastizzo, M. J., & Gallup, G. G. (2008). The Sound of Symmetry Revisited : Subjective and Objective Analyses of Voice. *Journal of Nonverbal Behavior*, 32(2), 93-108. <https://doi.org/10.1007/s10919-007-0042-6>
- I. Folstad, & Karter A.J. (1992). Parasites, bright males, and the immunocompetence handicap. *The American Naturalist*, 139(3), 603-622.
- Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., & Vukovic, J. (2010). A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. *Animal Behaviour*, 79(1), 57-62. <https://doi.org/10.1016/j.anbehav.2009.10.003>
- Jones, B. C., Hahn, A. C., Fisher, C. I., Wang, H., Kandrik, M., Han, C., ... DeBruine, L. M. (2017). No compelling evidence that preferences for facial masculinity track changes in women's hormonal status. *BioRxiv*. <https://doi.org/10.1101/136549>
- Jones, T. M., Trabold, M., Plante, F., Cheetham, B. M. G., & Earis, J. E. (2001). Objective assessment of hoarseness by measuring jitter. *Clinical Otolaryngology and Allied Sciences*, 26(1), 29-32. <https://doi.org/10.1046/j.1365-2273.2001.00413.x>
- Jost, L., Fuchs, M., Loeffler, M., Thiery, J., Kratzsch, J., Berger, T., & Engel, C. (2018). Associations of Sex Hormones and Anthropometry with the Speaking Voice Profile in the Adult General Population. *Journal of Voice*, 32(3), 261-272. <https://doi.org/10.1016/j.jvoice.2017.06.011>

- Jünger, J., Kordsmeyer, T. L., Gerlach, T. M., & Penke, L. (2018). Fertile women evaluate male bodies as more attractive, regardless of masculinity. *Evolution and Human Behavior*, 39(4), 412-423. <https://doi.org/10.1016/j.evolhumbehav.2018.03.007>
- Jünger, J., Motta-Mena, N. V., Cardenas, R., Bailey, D., Rosenfield, K. A., Schild, C., ... Puts, D. A. (2018). Do women's preferences for masculine voices shift across the ovulatory cycle? *Hormones and Behavior*, 106, 122-134. <https://doi.org/10.1016/j.yhbeh.2018.10.008>
- Laukka, P., Juslin, P., & Bresin, R. (2005). A dimensional approach to vocal expression of emotion. *Cognition & Emotion*, 19(5), 633-653. <https://doi.org/10.1080/02699930441000445>
- Leongómez, J. D., Binter, J., Kubicová, L., Stolařová, P., Klapilová, K., Havlíček, J., & Roberts, S. C. (2014). Vocal modulation during courtship increases perceptivity even in naive listeners. *Evolution and Human Behavior*, 35(6), 489-496. <https://doi.org/10.1016/j.evolhumbehav.2014.06.008>
- Linville, S. E., & Fisher, H. B. (1985). Acoustic characteristics of perceived versus actual vocal age in controlled phonation by adult females. *The Journal of the Acoustical Society of America*, 78(1), 40-48. <https://doi.org/10.1121/1.392452>
- Lyons, M., Lynch, A., Brewer, G., & Bruno, D. (2014). Detection of Sexual Orientation (“Gaydar”) by Homosexual and Heterosexual Women. *Archives of Sexual Behavior*, 43(2), 345-352. <https://doi.org/10.1007/s10508-013-0144-7>
- Marcinkowska, U. M., Galbarczyk, A., & Jasienska, G. (2018). La donna è mobile? Lack of cyclical shifts in facial symmetry, and facial and body masculinity preferences—A hormone based study. *Psychoneuroendocrinology*, 88, 47-53.
- Munson, B., McDonald, E. C., DeBoe, N. L., & White, A. R. (2006). The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics*, 34(2), 202-240. <https://doi.org/10.1016/j.wocn.2005.05.003>

- Pisanski, K., Fraccaro, P. J., Tigue, C. C., O'Connor, J. J. M., & Feinberg, D. R. (2014a). Return to Oz: Voice pitch facilitates assessments of men's body size. *Journal of Experimental Psychology: Human Perception and Performance*, 40(4), 1316-1331. <https://doi.org/10.1037/a0036956>
- Pisanski, K., Fraccaro, P. J., Tigue, C. C., O'Connor, J. J. M., & Feinberg, D. R. (2014b). Return to Oz: Voice pitch facilitates assessments of men's body size. *Journal of Experimental Psychology: Human Perception and Performance*, 40(4), 1316-1331. <https://doi.org/10.1037/a0036956>
- Pisanski, K., Fraccaro, P. J., Tigue, C. C., O'Connor, J. J. M., Röder, S., Andrews, P. W., ... Feinberg, D. R. (2014). Vocal indicators of body size in men and women: A meta-analysis. *Animal Behaviour*, 95, 89-99. <https://doi.org/10.1016/j.anbehav.2014.06.011>
- Pisanski, K., Hahn, A. C., Fisher, C. I., DeBruine, L. M., Feinberg, D. R., & Jones, B. C. (2014). Changes in salivary estradiol predict changes in women's preferences for vocal masculinity. *Hormones and Behavior*, 66(3), 493-497.
- Pisanski, K., Jones, B. C., Fink, B., O'Connor, J. J. M., DeBruine, L. M., Röder, S., & Feinberg, D. R. (2016). Voice parameters predict sex-specific body morphology in men and women. *Animal Behaviour*, 112, 13-22. <https://doi.org/10.1016/j.anbehav.2015.11.008>
- Pisanski, K., & Rendall, D. (2011). The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. *The Journal of the Acoustical Society of America*, 129(4), 2201-2212. <https://doi.org/10.1121/1.3552866>
- Ptacek, P. H., & Sander, E. K. (1966). Age recognition from voice. *Journal of speech and hearing Research*, 9(2), 273-277.
- Puts, D. A., Apicella, C. L., & Cardenas, R. A. (2012). Masculine voices signal men's threat potential in forager and industrial societies. *Proceedings of the Royal Society B: Biological Sciences*, 279(1728), 601-609. <https://doi.org/10.1098/rspb.2011.0829>

- Puts, David A., & Aung, T. (2019). Does Men's Voice Pitch Signal Formidability? A Reply to Feinberg et al. *Trends in Ecology & Evolution*, 34(3), 189-190. <https://doi.org/10.1016/j.tree.2018.12.004>
- Puts, David A., Barndt, J. L., Welling, L. L. M., Dawood, K., & Burriss, R. P. (2011). Intrasexual competition among women : Vocal femininity affects perceptions of attractiveness and flirtatiousness. *Personality and Individual Differences*, 50(1), 111-115. <https://doi.org/10.1016/j.paid.2010.09.011>
- Puts, David Andrew. (2005). Mating context and menstrual phase affect women's preferences for male voice pitch. *Evolution and Human Behavior*, 26(5), 388-397. <https://doi.org/10.1016/j.evolhumbehav.2005.03.001>
- Puts, David Andrew, Gaulin, S. J. C., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior*, 27(4), 283-296. <https://doi.org/10.1016/j.evolhumbehav.2005.11.003>
- Puts, David Andrew, Hodges, C. R., Cárdenas, R. A., & Gaulin, S. J. C. (2007). Men's voices as dominance signals : Vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior*, 28(5), 340-344. <https://doi.org/10.1016/j.evolhumbehav.2007.05.002>
- Rabinov, C. R., Kreiman, J., Gerratt, B. R., & Bielałowicz, S. (1995). Comparing Reliability of Perceptual Ratings of Roughness and Acoustic Measures of Jitter. *Journal of Speech Language and Hearing Research*, 38(1), 26. <https://doi.org/10.1044/jshr.3801.26>
- Re, D. E., O'Connor, J. J. M., Bennett, P. J., & Feinberg, D. R. (2012). Preferences for Very Low and Very High Voice Pitch in Humans. *PLoS ONE*, 7(3), e32719. <https://doi.org/10.1371/journal.pone.0032719>

- Rendall, D., Vokey, J. R., & Nemeth, C. (2007). Lifting the curtain on the Wizard of Oz : Biased voice-based impressions of speaker size. *Journal of Experimental Psychology: Human Perception and Performance*, 33(5), 1208-1219. <https://doi.org/10.1037/0096-1523.33.5.1208>
- Saxton, T. K., Caryl, P. G., & Craig Roberts, S. (2006). Vocal and Facial Attractiveness Judgments of Children, Adolescents and Adults : The Ontogeny of Mate Choice. *Ethology*, 112(12), 1179-1185. <https://doi.org/10.1111/j.1439-0310.2006.01278.x>
- Saxton, T. K., Debruine, L. M., Jones, B. C., Little, A. C., & Roberts, S. C. (2009). Face and voice attractiveness judgments change during adolescence. *Evolution and Human Behavior*, 30(6), 398-408. <https://doi.org/10.1016/j.evolhumbehav.2009.06.004>
- Šebesta, P., Kleisner, K., Tureček, P., Kočnar, T., Akoko, R. M., Třebický, V., & Havlíček, J. (2017). Voices of Africa : Acoustic predictors of human male vocal attractiveness. *Animal Behaviour*, 127, 205-211. <https://doi.org/10.1016/j.anbehav.2017.03.014>
- Sell, A., Bryant, G. A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., ... Gurven, M. (2010). Adaptations in humans for assessing physical strength from the voice. *Proceedings of the Royal Society B: Biological Sciences*, 277(1699), 3509-3518. <https://doi.org/10.1098/rspb.2010.0769>
- Shipp, T., Qi, Y., Huntley, R., & Hollien, H. (1992). Acoustic and temporal correlates of perceived age. *Journal of Voice*, 6(3), 211-216. [https://doi.org/10.1016/S0892-1997\(05\)80145-6](https://doi.org/10.1016/S0892-1997(05)80145-6)
- Shirazi, T. N., Puts, D. A., & Escasa-Dorne, M. J. (2018). Filipino Women's Preferences for Male Voice Pitch : Intra-Individual, Life History, and Hormonal Predictors. *Adaptive Human Behavior and Physiology*, 4(2), 188-206. <https://doi.org/10.1007/s40750-018-0087-2>
- Simmons, L. W., Peters, M., & Rhodes, G. (2011). Low Pitched Voices Are Perceived as Masculine and Attractive but Do They Predict Semen Quality in Men? *PLoS ONE*, 6(12), e29271. <https://doi.org/10.1371/journal.pone.0029271>

- Skrinda, I., Krama, T., Kecko, S., Moore, F. R., Kaasik, A., Meija, L., ... Krams, I. (2014). Body height, immunity, facial and vocal attractiveness in young men. *Naturwissenschaften*, *101*(12), 1017-1025. <https://doi.org/10.1007/s00114-014-1241-8>
- Smith, K. M., Olkhov, Y. M., Puts, D. A., & Apicella, C. L. (2017). Hadza Men With Lower Voice Pitch Have a Better Hunting Reputation. *Evolutionary Psychology*, *15*(4), 147470491774046. <https://doi.org/10.1177/1474704917740466>
- Suire, A., Raymond, M., & Barkat-Defradas, M. (2018). Human vocal behavior within competitive and courtship contexts and its relation to mating success. *Evolution and Human Behavior*, *39*(6), 684-691. <https://doi.org/10.1016/j.evolhumbehav.2018.07.001>
- Taylor, A. M., & Reby, D. (2010). The contribution of source-filter theory to mammal vocal communication research : Advances in vocal communication research. *Journal of Zoology*, *280*(3), 221-236. <https://doi.org/10.1111/j.1469-7998.2009.00661.x>
- Teixeira, J. P., Oliveira, C., & Lopes, C. (2013). Vocal Acoustic Analysis – Jitter, Shimmer and HNR Parameters. *Procedia Technology*, *9*, 1112-1122. <https://doi.org/10.1016/j.protcy.2013.12.124>
- Titze, I. R. (1989). Physiologic and acoustic differences between male and female voices. *The Journal of the Acoustical Society of America*, *85*(4), 1699-1707. <https://doi.org/10.1121/1.397959>
- Valentová, J., Roberts, S. C., & Havlíček, J. (2013). Preferences for Facial and Vocal Masculinity in Homosexual Men : The Role of Relationship Status, Sexual Restrictiveness, and Self-Perceived Masculinity. *Perception*, *42*(2), 187-197. <https://doi.org/10.1068/p6909>
- Van Borsel, J., Janssens, J., & De Bodt, M. (2009). Breathiness as a Feminine Voice Characteristic : A Perceptual Approach. *Journal of Voice*, *23*(3), 291-294. <https://doi.org/10.1016/j.jvoice.2007.08.002>

- Vukovic, J., Feinberg, D. R., Jones, B. C., DeBruine, L. M., Welling, L. L. M., Little, A. C., & Smith, F. G. (2008). Self-rated attractiveness predicts individual differences in women's preferences for masculine men's voices. *Personality and Individual Differences*, 45(6), 451-456. <https://doi.org/10.1016/j.paid.2008.05.013>
- Weiss, B., & Burkhardt, F. (2010). Voice Attributes Affecting Likability Perception. *Eleventh Annual Conference of the International Speech Communication Association*, 4.
- Wendahl, R. W. (1966). Some Parameters of Auditory Roughness. *Folia Phoniatica et Logopaedica*, 18(1), 26-32. <https://doi.org/10.1159/000263081>
- Wickham, H. (2009). *Ggplot2 : Elegant graphics for data analysis*. New York: Springer.
- Wickham, H., François, R., Henry, L., & Müller, K. (2018). *Dplyr : A Grammar of Data Manipulation*. R package version 0.7.8. Consulté à l'adresse <https://CRAN.R-project.org/package=dplyr>
- Wolff, S. E., & Puts, D. A. (2010). Vocal masculinity is a robust dominance signal in men. *Behavioral Ecology and Sociobiology*, 64(10), 1673-1683. <https://doi.org/10.1007/s00265-010-0981-5>
- Xu, Y., Lee, A., Wu, W.-L., Liu, X., & Birkholz, P. (2013). Human Vocal Attractiveness as Signaled by Body Size Projection. *PLoS ONE*, 8(4), e62397. <https://doi.org/10.1371/journal.pone.0062397>

Supplementary material S1.

Symbolic representation of the GLMM.

Left / Right ~ differences in vocal parameters +
differences in vocal parameters*age +
differences in vocal parameters*ancestry +
differences in vocal parameters*sexual orientation +
Random effect of the judge +
Random effect of the vocal stimuli

The response variable ('Left / Right') corresponds to whether the judge chose the voice on the left or the one on the right.

The differences in vocal parameters correspond to the difference between F0, F0-SD, Df, Pf, HNR, jitter and intensity between the voice on the left and the one on the right. Thus, if one of those differences is negative for a given pair of voices, it means that the voice on the left has lower values than the one on the right. If it is positive, the voice on the left has higher values.

Now for each judge, she chose one of these two voices for each of her choices, and we know whether she clicked on the left or the right. Thus, we know if she preferred for each vocal parameter, the voice with the lower F0 (negative difference, clicked left) or the highest F0 (positive difference, clicked right) for each single pair of voices. Overall, if the estimate of the GLMM for the difference in mean F0 is negative, judges chose on average voices with the lower F0; if the estimate is positive, it means that the judges chose on average voices with the higher F0.

Lastly, implementing control variables as main factors would not make any sense as it would mean that age, ancestry and sexual orientation have an influence on whether a judge just preferred to click on the left or the right. Thus, they need to be put in interaction with the differences in vocal parameters to know whether it influenced the judge's choice.

Supplementary material S2.

Below are presented all the models without the control variables (sexual orientation and ancestry) and the model without female judges with hearing impairment.

1. GLMM without ‘sexual orientation’

	Estimate	SE	χ^2	<i>p value</i>
Intercept	0.08	0.06	/	/
Difference in mean F0	-0.47	0.09	25.92	<0.001
Difference in F0-SD	0.46	0.08	25.29	<0.001
Difference in Df	0.16	0.09	2.99	0.083
Difference in Pf	-0.08	0.07	1.37	0.241
Difference in HNR	-0.12	0.10	1.43	0.231
Difference in jitter	-0.02	0.08	0.09	0.761
Difference in intensity	0.22	0.06	12.40	<0.001
Interactions with age				
Difference in F0	0.14	0.09	2.44	0.117
Difference in F0-SD	0.07	0.09	0.52	0.468
Difference in Df	0.11	0.08	1.75	0.185
Difference in Pf	-0.05	0.07	0.52	0.469
Difference in HNR	-0.10	0.10	1.07	0.300
Difference in jitter	0.10	0.08	1.42	0.232
Difference in intensity	0.15	0.06	5.77	0.016
Interactions with ancestry				
Difference in F0	-0.005	0.22	0.00	0.981
Difference in F0-SD	-0.38	0.20	3.58	0.058
Difference in Df	0.05	0.23	0.05	0.815
Difference in Pf	-0.17	0.18	0.87	0.349
Difference in HNR	0.01	0.25	0.002	0.961
Difference in jitter	0.04	0.21	0.04	0.829
Difference in intensity	-0.14	0.17	0.66	0.414

2. GLMM without ‘ancestry’

	Estimate	SE	χ^2	<i>p value</i>
Intercept	0.09	0.06	/	/
Difference in mean F0	-0.50	0.09	30.88	<0.001
Difference in F0-SD	0.46	0.08	31.33	<0.001
Difference in Df	0.16	0.09	3.03	0.081
Difference in Pf	-0.06	0.07	0.74	0.388
Difference in HNR	-0.09	0.10	0.86	0.352
Difference in jitter	-0.01	0.08	0.01	0.900
Difference in intensity	0.17	0.06	7.87	0.004
Interactions with age				
Difference in F0	0.15	0.09	2.78	0.097
Difference in F0-SD	0.03	0.09	0.11	0.736
Difference in Df	0.12	0.09	2.08	0.148
Difference in Pf	-0.07	0.07	0.81	0.368
Difference in HNR	-0.12	0.10	1.56	0.210
Difference in jitter	0.10	0.08	1.61	0.204
Difference in intensity	0.15	0.06	5/77	0.016
Interactions with sexual orientation				
Difference in F0	0.16	0.24	0.46	0.497
Difference in F0-SD	-0.54	0.22	5.71	0.016
Difference in Df	-0.10	0.23	0.20	0.652
Difference in Pf	-0.12	0.18	0.43	0.511
Difference in HNR	-0.15	0.28	0.30	0.581
Difference in jitter	0.14	0.24	0.35	0.553
Difference in intensity	0.25	0.17	2.04	0.153

3. GLMM without ‘age’

	Estimate	SE	χ^2	<i>p value</i>
Intercept	0.08	0.06	/	/
Difference in mean F0	-0.50	0.10	26.54	<0.001
Difference in F0-SD	0.50	0.10	32.47	<0.001
Difference in Df	0.15	0.10	2.39	0.121
Difference in Pf	-0.04	0.07	0.23	0.629
Difference in HNR	-0.11	0.10	1.12	0.289
Difference in jitter	-0.04	0.09	0.18	0.667
Difference in intensity	0.17	0.06	6.90	0.008
Interactions with ancestry				
Difference in F0	0.10	0.24	0.17	0.67
Difference in F0-SD	-0.50	0.22	4.98	0.025
Difference in Df	-0.11	0.24	0.24	0.625
Difference in Pf	-0.13	0.18	0.53	0.465
Difference in HNR	-0.09	0.28	0.11	0.738
Difference in jitter	0.14	0.24	0.36	0.548
Difference in intensity	-0.22	0.17	1.61	0.20
Interactions with sexual orientation				
Difference in F0	0.06	0.21	0.09	0.765
Difference in F0-SD	-0.40	0.20	4.02	0.044
Difference in Df	0.03	0.23	0.02	0.886
Difference in Pf	-0.16	0.18	0.79	0.372
Difference in HNR	0.07	0.25	0.08	0.778
Difference in jitter	0.08	0.21	0.16	0.687
Difference in intensity	-0.12	0.17	0.575	0.448

4. GLMM without judges with hearing impairment (n = 4)

	Estimate	SE	χ^2	<i>p</i> value
Intercept	0.10	0.06	/	/
Difference in mean F0	-0.49	0.10	23.66	<0.001
Difference in F0-SD	0.54	0.09	34.44	<0.001
Difference in Df	0.18	0.10	3.19	0.074
Difference in Pf	-0.06	0.08	0.54	0.461
Difference in HNR	-0.12	0.10	1.19	0.274
Difference in jitter	-0.04	0.09	0.21	0.643
Difference in intensity	0.17	0.07	6.94	0.008
Interactions with age				
Difference in F0	0.23	0.11	4.35	0.036
Difference in F0-SD	0.02	0.10	0.02	0.871
Difference in Df	0.19	0.12	2.70	0.100
Difference in Pf	-0.05	0.09	0.35	0.562
Difference in HNR	-0.15	0.12	1.73	0.188
Difference in jitter	0.10	0.10	1.04	0.308
Difference in intensity	0.15	0.07	4.32	0.037
Interactions with ancestry				
Difference in F0	0.04	0.22	0.03	0.861
Difference in F0-SD	-0.45	0.20	4.79	0.028
Difference in Df	0.09	0.24	0.14	0.708
Difference in Pf	-0.18	0.19	0.87	0.349
Difference in HNR	0.03	0.21	0.02	0.875
Difference in jitter	0.01	0.20	0.03	0.869
Difference in intensity	-0.07	0.17	0.17	0.677
Interactions with sexual orientation				
Difference in F0	0.22	0.24	0.82	0.363
Difference in F0-SD	-0.63	0.23	7.18	0.007
Difference in Df	-0.09	0.25	0.12	0.720
Difference in Pf	-0.07	0.19	0.15	0.692
Difference in HNR	-0.14	0.29	0.25	0.617
Difference in jitter	0.17	0.25	0.46	0.494
Difference in intensity	0.26	0.18	2.00	0.156

5. GLMM without judges who did not report sexual orientation (n = 8)

	Estimate	SE	χ^2	<i>p</i> value
Intercept	0.10	0.06	/	/
Difference in mean F0	-0.49	0.10	24.17	<0.001
Difference in F0-SD	0.52	0.09	32.65	<0.001
Difference in Df	0.19	0.10	3.39	0.065
Difference in Pf	-0.05	0.08	0.45	0.503
Difference in HNR	-0.12	0.10	1.39	0.236
Difference in jitter	-0.06	0.09	0.39	0.532
Difference in intensity	0.17	0.07	6.93	0.008
Interactions with age				
Difference in F0	0.24	0.10	5.55	0.018
Difference in F0-SD	-0.05	0.10	0.24	0.624
Difference in Df	0.20	0.11	3.38	0.065
Difference in Pf	-0.05	0.08	0.31	0.577
Difference in HNR	-0.17	0.11	2.34	0.125
Difference in jitter	0.07	0.09	0.55	0.457
Difference in intensity	0.15	0.07	5.62	0.017
Interactions with ancestry				
Difference in F0	0.05	0.22	0.05	0.817
Difference in F0-SD	-0.42	0.21	4.13	0.042
Difference in Df	0.02	0.24	0.009	0.922
Difference in Pf	-0.15	0.19	0.63	0.427
Difference in HNR	0.04	0.26	0.02	0.875
Difference in jitter	0.11	0.22	0.25	0.618
Difference in intensity	-0.05	0.17	0.07	0.792
Interactions with sexual orientation				
Difference in F0	0.06	0.31	0.04	0.836
Difference in F0-SD	-0.72	0.29	6.08	0.013
Difference in Df	0.12	0.29	0.18	0.671
Difference in Pf	0.02	0.23	0.009	0.924
Difference in HNR	-0.65	0.38	2.80	0.094
Difference in jitter	-0.35	0.31	1.28	0.257
Difference in intensity	0.45	0.24	3.31	0.068

Article 2

**Men's preferences for women's voice pitch, its variation, timbre,
breathiness, roughness and tempo.**

Alexandre Suire, Michel Raymond & Melissa Barkat-Defradas

Article en préparation

Résumé

Les études sur les préférences vocales des hommes semblent relativement conflictuelles concernant la hauteur vocale (i.e. F0) des femmes. En effet, si la majorité des études a montré que les hommes préfèrent des voix relativement aiguës, un certain nombre d'études suggère des relations inverses. Également, si des timbres plus élevés (mesurée via le Df) sont considérés comme plus attractifs, il existe toutefois peu d'études s'étant intéressé au F0-SD, au jitter et au HNR des femmes. De plus, il n'existe à notre connaissance aucune étude s'étant intéressé à l'attractivité du débit syllabique. Enfin, comme pour les préférences des femmes, la plupart des études ont été conduites dans des populations anglo-saxonnes, mais un petit nombre d'études suggère que les préférences ne sont pas universelles. Afin d'éclaircir ces différents aspects et contribuer plus généralement à l'étude des préférences vocales, nous avons mené une seconde étude sur les préférences des hommes français pour la voix des femmes françaises.

Dans ce contexte, nous avons enregistré 13 participantes prononçant plusieurs phrases imposées par l'expérimentateur, dont les 5 plus longues ont été retenues. Chacune de ces phrases a été ensuite manipulée en débit afin de produire une version accélérée ou diminuée par rapport à sa version naturelle ($\pm 10\%$). Nous avons ensuite analysé pour chaque voix et chaque phrase les composantes acoustiques suivantes : le F0, le F0-SD, le Df, le jitter, le HNR et l'intensité. Puis, nous avons présenté à 152 juges hommes (dont 135 éligibles) 13 paires de voix, chacune contenant une voix accélérée et une diminuée ne provenant pas de la même locutrice et aléatoirement choisie parmi le pool entier de stimuli. Ceci nous a permis d'étudier à la fois les préférences pour le débit ainsi que celles pour les paramètres acoustiques. De plus, nous avons tenté de voir si les préférences des hommes pourraient être influencées selon qu'ils recherchent un partenaire sur le court ou sur le long terme. Ainsi, à chaque paire, le juge devait choisir la voix qu'il considérait comme la plus attractive selon la même question contextuelle suivante : « *Quelle femme préférez-vous pour une relation sans lendemain (i.e. relation court-*

terme) ? » ou « *Quelle voix préférez-vous pour fonder une famille (i.e. relation long-terme) ?* ».

Nos résultats montrent que les hommes français exhibent une préférence significative pour les voix relativement plus rapides (haut débit), relativement plus graves (bas F0), des timbres relativement plus aiguës (haut Df), une intonation relativement marquée (haut F0-SD), plus de raucité (haut jitter) et moins de souffle (haut HNR). Enfin, la question contextuelle n'a pas influencé le choix des juges. Nous pouvons noter quelques remarques concernant ces résultats. D'une part, des débits plus rapides refléteraient de meilleurs processus de planification mentale : les individus affichant un rythme de parole plus rapide nécessitent moins d'efforts cognitifs et de temps supplémentaire pour produire un discours naturel, compréhensible, cohérent et fluide. Ainsi, des débits relativement plus rapides chez les hommes devraient être aussi perçus comme plus attractifs par les femmes. D'autre part, les préférences des hommes français ne semblent pas suivre la même tendance que les locuteurs anglophones, plus particulièrement pour le F0. Ceci suggère que des facteurs sociaux et/ou environnementaux jouent possiblement un rôle important dans le façonnement des préférences.

**Men's preferences for women's voice pitch, its variation, timbre, breathiness, roughness
and tempo.**

Authors: Alexandre Suire¹, Michel Raymond¹, Melissa Barkat-Defradas¹

Corresponding author: Alexandre Suire

E-mail: alexandre.suire@umontpellier.fr

Fax: +33 4 67 14 36 22

Tel: +33 4 67 14 49 66

¹ Institut des Sciences de l'Évolution de Montpellier, Université de Montpellier, Centre National de la Recherche Scientifique, Institut pour la Recherche et le Développement, École Pratique des Hautes Études – Place Eugène Bataillon 34095 Montpellier, France.

Abstract

Men's preferences for women's speech acoustic features have been extensively studied. However, some acoustic preferences have remained unexplored, such as micro-intonation patterns, vocal breathiness and speech tempo, while others need to be further investigated, such as voice pitch and timbre. Moreover, most studies have assessed such relationships in English speaking populations, but some evidence suggest that it may be partly influenced by the sociocultural environment. Thus, men's preferences for women's voices may vary from one population to another. In this study, we assessed men's preferences for women's voice pitch (i.e., fundamental frequency, F0), its variation (i.e., F0-SD), timbre (i.e., formant dispersion, Df), vocal breathiness (i.e., harmonics-to-noise ratio), vocal roughness (i.e., local jitter) and tempo (i.e., speaking rate) in a French-speaking population. Using a pairwise comparison paradigm, French men significantly preferred lower F0, higher F0-SD, Df, HNR and jitter profiles as well as faster speaking rates. Additionally, the type of mating relationship (short, uncommitted vs. a long, romantic relationship) did not influence these preferences. We discuss our results in light of voice attractiveness and its adaptive function in a mate choice context.

Keywords: voice; attractiveness; fundamental frequency; formants; tempo; voice quality.

Introduction

Humans vary in how they produce speech, and those differences can depend on a number of biological and sociocultural factors. The fact that beyond the linguistic message, vocal properties of speech seem to convey substantial bio-psychological information concerning the speaker has raised much interest. So far, studies have suggested that acoustic signals are likely to provide humans with information on the quality or condition of the speaker regarding sex (Bachorowski and Owren, 1999; Gelfer and Bennett, 2013; Gelfer and Mikos, 2005; Hillenbrand and Clark, 2009), age (Linville and Fisher, 1985; Ptacek and Sander, 1966; Shipp et al., 1992), sexual orientation (Lyons et al., 2014; Munson et al., 2006), physical strength (Sell et al., 2010), sexual behavior and body configuration (Hughes et al., 2004). The most salient perceptual features of speech that seem to carry such information are the fundamental frequency (i.e., F0, the acoustic correlate of voice pitch) and the formants (i.e., the resonances of the vocal tract), both being sexually dimorphic in humans (Titze, 1989).

Such studies have also prompted a whole body of research that studies the relationships between these vocal acoustic features with several auditory impressions, among which, attractiveness. Regarding men's preferences for women's voices, most experimental and correlational studies have found a consistent positive relationship between attractiveness and the F0, that is, men are attracted in average to relatively higher voice pitch (Borkowska and Pawlowski, 2011; Collins and Missing, 2003; Feinberg et al., 2008; Jones et al., 2010; Puts et al., 2011; Re et al., 2012; however see Hughes et al., 2010, 2014). Relatively higher formant dispersions (i.e., D_f , the relative distance between two consecutive formants, which correlates to the vocal tract length and perceived timbre) are also perceived as more attractive by men (Babel et al., 2014; Puts et al., 2011). Additionally, the variation of the F0 (i.e., F0-SD, the acoustic correlate of speech micro-intonation patterns) has also been hypothesized to play upon the perception of indexical cues relevant to human competing and mating contexts (Hodges-

Simeon et al., 2010, 2011; Leongómez et al., 2014). Although sexually dimorphic, it has only been tested for women's preferences in men's voices (Bruckert et al., 2006; Hodges-Simeon et al., 2010), but one study suggests that men may be attracted to higher F0-SD profiles in women as it may be a cue of femininity (Leongómez et al., 2014).

Fundamental and formant frequencies aside, a few understudied acoustic features that are known to participate to vocal quality may also influence vocal attractiveness. For instance, vocal breathiness could be positively related to perceived vocal attractiveness in both sexes (Šebesta et al., 2017; Xu et al., 2013). As it is sexually dimorphic, it has been argued to be a more feminine trait, which is related to desirability in women (Babel et al., 2014; Van Borsel et al., 2009). Vocal breathiness can be captured by the harmonics-to-noise ratio (HNR), which corresponds to a ratio between periodic components and non-periodic component comprising a segment of voiced speech. HNR represents an evaluation of the vibration of the vocal folds against the glottal noise and reflects the efficiency of speech (Teixeira et al., 2013). The greater the flow of air expelled from the lungs into energy of vibration of vocal cords, the greater is the HNR, which acoustically correlates to lower vocal breathiness and, reversely, the more important the air leakage at the glottis the higher the auditory sensation of breathiness. Additionally, another acoustic feature of interest is the jitter, a feature linked to perceived vocal roughness (Hillenbrand, 1988; Rabinov et al., 1995; Wendahl, 1966). Jitter is a measure of the F0 disturbance, and is defined as the parameter capturing the frequency variation from cycle to cycle in the sound wave. The higher the jitter, the "rougher" a voice sounds. Whether jitter contributes significantly to perceived vocal attractiveness has yet to be fully explored, as to our knowledge, only two studies have tackled the issue, with mixed results (Babel et al., 2014; Hughes et al., 2008).

Interestingly, men's preferences have also been recorded for distinctive relationship contexts, as they can be modulated on the type of relationship sought: a short, uncommitted vs.

a romantic, committed, long relationship. Within this paradigm, men may be attracted to different vocal cues depending on the mating context. So far, two studies have reported that men are more attracted to higher feminine voices for a short-term than a long-term relationship (Little et al., 2011; Puts et al., 2011), with formant dispersion having a larger overall effect than the F0 (Puts et al., 2011). Theoretically speaking, the explanation for such correlations between acoustic features of speech and perceived vocal attractiveness have been suggested to be linked to the importance of perceived femininity in women, a feature tightly linked to higher frequencies in women's voices and important in a mate choice context as it has been suggested to be a cue of fertility (Bryant and Haselton, 2009; Collins and Missing, 2003; Fraccaro et al., 2010; O'Connor et al., 2013; Röder et al., 2013).

However, most studies on acoustics' preferences have been conducted with English native speakers, but few evidences suggest that relationships between attractiveness and acoustic properties of women's voices may not be consistent across populations of different cultures. Indeed, physiological and anatomical differences do not explain the full variation in mean F0 between men and women, as individuals of both sexes exhibit considerable variation from one language to another (Andreeva et al., 2014; Keating and Kuo, 2012; Pépiot, 2014; Rose, 1991; Traunmüller and Eriksson, 1995; Yamazawa and Hollien, 1992). For instance, even under the same speaking conditions, a pool of American women balanced in age, exhibit a lower F0 than Japanese women (mean F0: 211 vs. 224 Hz, Yamazawa and Hollien, 1992), while Bulgarian and Polish women exhibit a higher F0 than German and English women (mean F0: 272 and 266 Hz vs. 210 and 217 Hz, Andreeva et al., 2014). As men and women vary in mean F0 across various languages, this strongly suggests that some of the differences must be accounted for by learned behavior or specific sociocultural practices (Simpson, 2009). For instance, it has been suggested that Dutch women display a lower F0 than Japanese women, and interestingly, Dutch and Japanese men tend to prefer female voices that exhibit culturally

congruent vocal heights that is: low female voices vs. high female voices for Dutch vs. Japanese men respectively (Van Bezooijen, 1995). Although such differences can be accounted by the inherent properties of the spoken language, perceived vocal femininity may be influenced by the culture under study.

Furthermore, in comparison to the above-mentioned speech acoustic features that have been investigated so far in the framework of vocal attractiveness, speech tempo – that is the articulation rate or the number of syllables produced per second – has also received little attention. Yet, some researchers have long recognized that variation in speech rate is a way among others to inform listeners about various biological and sociocultural features of the speaker. Indeed, it has been shown for instance that speech tempo is greatly affected by age (Harnsberger et al., 2008; Ramig, 1983; Shipp et al., 1992; Smith et al., 1987), geographical origin (Byrd, 1994; Jacewicz et al., 2009; Verhoeven et al., 2004) and emotional states (Fairbanks and Hoaglin, 1941; Scherer et al., 2015; Williams and Stevens, 1972). Although substantial variation exists in speech rate between languages, in regard to sex, most available evidence suggests that men actually speak relatively faster than women do, with a significant difference amounting from 5 to 7% (Block and Killen, 1996; Byrd, 1994; Lutz and Mallard, 1986; Robb et al., 2004; Whiteside, 1996; Yuan et al., 2006). Although the context of interaction can affect tempo (e.g. child-directed speech or conversational convergence), individuals do exhibit variation in their habitual speaking rate styles. In this context, as hypothesized and shown for the acoustic properties of speech, the variation in speech tempo may lead to different perceived impressions about a speaker.

Indeed, in synthetically manipulated speaking rate experiments, both men and women with faster speech tempo are judged as more competent and benevolent (Brown et al., 1973; Smith et al., 1975) as well as more credible and persuasive (Miller et al., 1976), while decreased speech tempo is associated to less truthful, less fluent and more “passive” individuals (Apple et

al., 1979). In several non-manipulated speech tempo experiments, it has also been shown for both sexes that speech rate was also positively linked to competence, social attractiveness (Street et al., 1984, 1983), likeability (Weiss and Burkhardt, 2010) and were associated to more animated and extroverted individuals (Addington, 1968). Conversely, speakers with slower speech tempo are generally associated with sadness and depression (Banse and Scherer, 1996; Siegman and Boyle, 1993; Sobin and Alpert, 1999). Interestingly, one study has also reported that physically perceived attractive individuals exhibited relatively faster speaking rates (Chaiken, 1979).

A possible explanation for such correlations is that speech rate may signal a higher capacity in mental planning processes: individuals displaying slower speech tempo require additional effort and time to produce natural, understandable, coherent and fluent linguistic output. Thus, these studies draw the hypothesis that faster speaking rates could be perceived as more attractive in both men and women since it may be a proxy of cognitive flexibility. However, it still remains to formally address men's preferences for women's speech tempo.

Here, men's preferences for women's speech tempo and five other acoustic vocal features were evaluated, namely F0, F0-SD, Df, HNR and jitter. Additionally, we aim to study such relationships by taking into consideration possibly context-dependent men's mating preferences (i.e., short-term vs. long-term) and in a French speaking population to add insight into the relative importance of cultural factors (vs. biological factors) regarding men's preferences for vocal attractiveness.

1. Material and Methods

All protocols used in this study have been approved by the French National Commission of Informatics and Liberty (CNIL number 1261003). All participants provided their written consent prior to the study.

a. Stimuli manipulations

13 French female native speakers (mean age = 23.2 y.o., standard-deviation = 2.3), undergraduates from the University of Montpellier (France), were recruited to produce the vocal stimuli. Each were audio recorded while reading the same set of 24 sentences at conversational speech level in an anechoic, soundproof recording room. We chose to record sentences as we believe it is more natural and ecologically valid than using sustained vowels or counted numbers (e.g. Collins and Missing, 2003; Feinberg et al., 2008). The microphone was placed at 15 cm of the speaker's mouth so as to control for variation in intensity. Acoustic data were acquired using a Sennheiser BF 515 microphone, connected to a PC located in an adjacent room. All recordings were encoded using the Adobe© Audition CS6 at a sampling rate of 44 kHz – 32 bit – mono and stimuli were saved as .wav files.

The five longest utterances were then selected for further manipulations: speech tempo was modified using Audacity© freeware at both +10% or –10% of the original tempo variation in speech rate. The five longest utterances were kept so as to be sure that, in the experimental procedure (see next section), raters would have enough auditory material to evaluate the vocal samples in terms of attractiveness and enough acoustic variation for each female voice. This resulted in a sample of 130 vocal stimuli (13 speakers x 5 utterances x 2 speech tempo manipulations) that were used as auditory stimuli in a psychoacoustic experiment (see Supplementary Material for further information on the auditory stimuli).

The new speech tempo versions showed that they significantly differ from each other and from the unmanipulated stimuli in their mean number of syllables per second: accelerated vs. natural ($t = -4.41, p < 0.001$), slowed vs. natural ($t = 4.52, p < 0.001$), accelerated vs. slowed ($t = 8.94, p < 0.001$).

b. Acoustic analyses

Six acoustic parameters were analyzed for each version (i.e., natural, accelerated and slowed tempo) using the Praat© voice analysis software (Boersma, P., and Weenink, D., 2017). The fundamental frequency (F0) and its variation across time (F0-SD) were measured using the autocorrelation method with a pitch floor of 100 Hz and a ceiling of 500 Hz (Praat's recommendation), with other settings kept as default. The harmonics-to-noise ratio (HNR, in dB) and the local jitter, which corresponds to the average absolute difference between consecutive periods, divided by the average period, and calculated in percentage, were measured across the entire utterance using the same settings as the F0. Additionally, intensity (dB) was retrieved using Praat's default settings. Formant frequencies (F1 to F4) were measured at each glottal pulse, targeting voiced speech only, using a formant ceiling of 5500 Hz (Praat's recommendation), then averaged across the entire utterance. Then, we calculated the formants' dispersion (Df) using the following formula (Fitch, 1997):

$$Df = \frac{\sum_{i=1}^{N-1} F_{i+1} - F_i}{N - 1},$$

where Df is the formant dispersion (in Hz), N is the total number of formants measured, and F_i is the frequency (in Hz) of formant i .

Comparisons of all acoustic characteristics between the three versions showed no significant differences in F0 (accelerated vs. natural, $t = 0.038$, $p = 0.99$; slowed vs. natural, $t = 0.37$, $p = 0.92$; accelerated vs. slowed, $t = 0.33$, $p = 0.94$), F0-SD (accelerated vs. natural, $t = 0.12$, $p = 0.99$; slowed vs. natural, $t = -0.65$, $p = 0.79$; accelerated vs. slowed, $t = -0.77$, $p = 0.71$), Df (accelerated vs. natural, $t = -0.20$, $p = 0.97$; slowed vs. natural, $t = -0.37$, $p = 0.92$; accelerated vs. slowed, $t = -0.17$, $p = 0.098$) and jitter (accelerated vs. natural, $t = -0.97$, $p = 0.59$; slowed vs. natural, $t = 0.21$, $p = 0.97$; accelerated vs. slowed, $t = 1.18$, $p = 0.46$). However, HNR values differed between the accelerated and slowed tempo version ($t = -3.14$, $p < 0.05$) but not between the unmanipulated and manipulated versions (accelerated vs. natural, $t = 1.24$,

$p = 0.42$; slowed vs. natural, $t = -1,90$, $p = 0.14$). Descriptive statistics of the speech tempo and acoustic parameters for the unmanipulated and manipulated versions are given in Table 1.

	Unmanipulated	Accelerated	Slowed
Mean tempo (syl/sec)	5.23 ± 0.71 (3.67 – 6.55)	5.82 ± 0.87 (4.06 – 9.11)	4.73 ± 0.62 (3.31 – 5.88)
F0 (Hz)	229.57 ± 25.68 (168.6 – 304)	229.58 ± 25.57 (168.9 – 304)	228.77 ± 25.50 (168.6 – 302.1)
F0-SD (Hz)	33.87 ± 7.48 (17.72 – 50.20)	33.74 ± 7.50 (18.81 – 49.98)	34.73 ± 7.55 (18.54 – 50.93)
Df (Hz)	1152.72 ± 29.79 (1098 – 1229)	1153.44 ± 30.06 (1096 – 1227)	1154.15 ± 29.75 (1095 – 1227)
HNR (dB)	14.78 ± 1.20 (12.50 – 16.97)	14.52 ± 1.20 (12.17 – 16.59)	15.18 ± 1.18 (12.86 – 17.43)
Jitter (%)	1.88 ± 0.28 (1.27 – 2.77)	1.93 ± 0.29 (1.23 – 2.93)	1.87 ± 0.27 (1.22 – 2.67)
Intensity (dB)	76.22 ± 2.08 (71.87 – 80.84)	76.27 ± 2.06 (72.21 – 81.25)	76.19 ± 2.08 (71.78 – 80.95)

Table 1. Descriptive statistics of the mean speech tempo and acoustic parameters for the unmanipulated and manipulated vocal stimuli (mean \pm standard-deviation, ranges are given in parentheses). The unmanipulated versions were not used in the experimental procedure.

c. Pre-test

A pre-test was conducted to verify whether speech tempo modifications (accelerated vs. slowed) were perceived as natural speech by naive judges. Random pairs of utterances from the same female speaker were presented to 40 independent judges (graduate students from the University of Montpellier, France), whose task was to discriminate the modified version (i.e., either accelerated or slowed tempo) from the original stimuli using a pairwise comparison paradigm. Each judge was presented 14 pairs of voices. Results revealed that accelerated and slowed down speech tempo were not perceived as unnatural from the unmanipulated stimuli (exact binomial test, $p = 0.42$). This result is consistent with previous literature suggesting that « the average just noticeable difference for speech tempo is about 5% variation » (Quene, 2007).

d. Experimental procedure

The experimental procedure was automated on a computer-interfaced program. 152 French male raters participated in a perceptual study after they filled a questionnaire assessing several personal and sociological information. After filling out the questionnaire, raters were presented with a series of 13 choices each including a pair of voices. For each pair, two stimuli with a

+10% and -10% tempo variation were presented, both randomly taken from the set of accelerated and slowed speech tempo versions. Pairs did not include two stimuli from the same female speaker in order to test for voice preferences as well. In addition, the two manipulated vocal stimuli were randomized in their position presented in each pair (left or right position) on the computer screen. Judges were then asked to choose the most attractive vocal stimulus by clicking on it. Participants were allowed to listen to the stimuli as much as they wanted. However, when the participant made his choice, he could not go back to the previous one. Preferences were recorded for one of the two following contextual questions selected on a random basis for each judge: “which one of these two women would you prefer for a one-night stand (i.e., short-term relationship)?” or “which one of these two women would you prefer to found a family (i.e., long-term relationship)?”

e. Statistical analyses

Judges who declared a hearing impairment, a homosexual orientation and non-European ascendants (to control for potential cultural differences in preferences) were not considered further, resulting in a final sample of 135 male raters (mean age = 38.70 y.o., standard-deviation = 12.18).

To analyze men’s preferences for speech tempo and acoustics characteristics, a generalized linear mixed model (GLMM) was used with the response variable being if the judge chose or not the voice presented to him on the left position. The GLMM was fitted with a binomial error structure since the response variable consisted in a discrete probability distribution of the number of successes in a sequence of several independent trials. To test preferences for speech tempo, a variable called ‘Faster speaking rate’ (FSR) was created, with two modalities being either left or right. This variable recorded if the voice presented on the left corresponded to the accelerated speech tempo version. To test men’s acoustics preferences, six explanatory variables were created and corresponded to the difference in F0, F0-SD, Df,

HNR and jitter and intensity between the two female vocal stimuli. Context-dependent questions (categorical variable), with two modalities being short-term and long-term relationships were added as another explanatory variable. This variable was put in interaction with FSR and the differences in acoustics characteristics to assess its influence on men's preferences. In a similar fashion, men's age was added as a control variable and put in interaction with FSR and the differences in acoustics characteristics to assess its influence on men's preferences. As judges could hear two vocal stimuli that differed in length (in seconds), the difference in stimuli length was added as a control variable and put in interaction with the context-dependent questions and age. Finally, judges' identities and the vocal stimuli were added as random effects.

The significance of each variable was assessed from the comparison of the model excluding the variable with the model including all the other variables. All statistical analyses were performed using the R software (version 3.4.0).

2. Results

Results of the GLMM are reported in Table 2.

Men's preferences for women's voices					
	Estimate	SE	χ^2	df	<i>p value</i>
Intercept	-0.74	0.18	/	/	/
Faster speaking rate (FSR)	1.65	0.32	46.97	1	<0.001
Left / Right					
Difference in F0	-0.55	0.19	28.78	1	<0.001
Difference in F0-SD	0.44	0.13	15.73	1	<0.001
Difference in Df	0.23	0.10	12.35	1	<0.001
Difference in HNR	0.28	0.24	10.88	1	<0.05
Difference in jitter	0.13	0.17	7.03	1	<0.05
Difference in intensity	0.12	0.13	0.76	1	0.38
Difference in stimuli length	0.096	0.22	0.013	1	0.90
FSR * Question			1.03	2	0.59
FSR-Right * Question	0.22	0.24			
Long-term / Short-term					
FSR-Left * Question	-0.040	0.26			
Long-term / Short-term					
Difference in F0 * Question	-0.21	0.25	0.72	1	0.39
Difference in F0-SD * Question	-0.15	0.18	0.75	1	0.38
Difference in Df * Question	0.037	0.14	0.067	1	0.79
Difference in HNR * Question	0.52	0.33	2.46	1	0.11
Difference in jitter * Question	0.38	0.24	2.51	1	0.11
Difference in intensity * Question	-0.31	0.17	3.26	1	0.070
Difference in stimuli length * Question	-0.17	0.27	0.41	1	0.52
FSR * Age			0.74	2	0.68
FSR-Right * Age	-0.099	0.12			
FSR-Left * Age	0.0061	0.13			
Difference in F0 * Age	-0.096	0.12	0.60	1	0.43
Difference in F0-SD * Age	0.041	0.092	0.20	1	0.65
Difference in Df * Age	-0.022	0.069	0.10	1	0.74
Difference in HNR * Age	-0.16	0.15	1.03	1	0.30
Difference in jitter * Age	-0.18	0.12	2.43	1	0.11
Difference in intensity * Age	0.28	0.083	11.44	1	<0.001
Difference in stimuli length * Age	0.094	0.13	0.47	1	0.48

Table 2. Results of the generalized linear mixed model predicting judges' preferences (n = 135 male raters). For each variable, the χ^2 , the degrees of freedom (df) and the p values associated from the chi-square test of the comparison between the full model and the model without the

variable are given. For the categorical variables “FSR” and “Question” the estimates are given for one category compared to the reference category (bold underlined term). P values are considered significant at the 0.05 threshold (in bold).

When presented with two voices, men significantly preferred faster speech tempo ($\chi^2_2 = 46.97$, $p < 0.001$). Preferences for the accelerated speech tempo was up to 62.8% compared to slowed speech tempo and is significantly above chance-level (Figure 1). Additionally, men significantly preferred voices with relatively lower F0 ($\chi^2_1 = 28.78$, $p < 0.001$), higher F0-SD ($\chi^2_1 = 15.73$, $p < 0.001$), higher Df ($\chi^2_1 = 12.35$, $p < 0.001$), higher HNR ($\chi^2_1 = 10.88$, $p < 0.05$) and higher jitter vocal profiles ($\chi^2_1 = 7.03$, $p < 0.05$). Stimuli length did not influence men’s preferences ($\chi^2_1 = 0.013$, $p = 0.90$). Context-dependent questions did not influence men’s preferences for speech tempo and any acoustic characteristics (all $p > 0.05$), i.e., men did not show specific preferences for the short- or long-term relationship. Similarly, men’s age did not influence their preferences for tempo and acoustics characteristics ($p > 0.05$), except for intensity ($\chi^2_1 = 11.44$, $p < 0.001$), i.e., older men preferred louder vocal stimuli.

3. Discussion

The present study highlights some original findings concerning men’s preferences for women’s voices.

Within a forced-choice paradigm, men significantly preferred a lower vocal pitch in French-speaking women. Such results contradict most previous findings in English-speaking populations (Babel et al., 2014; Collins and Missing, 2003; Feinberg et al., 2008; Re et al., 2012). At a first glance, such differences can be explained by the usage of distinct acoustic stimuli as most studies used isolated vowels or counted numbers instead of short utterances. Longer acoustic stimuli should give a more representative impression of one’s habitual speaking style, which is more ecologically valid. However, sustained vowels or count numbers give a more precise account of one’s speech acoustic features. To our knowledge, only one

study has studied the effects of stimulus type (words vs. isolated vowel vs. three vowels sequence) upon vocal attractiveness (Ferdenzi et al., 2013). In this study, there were no main effect of the stimulus type on attractiveness ratings, but the more ecologically valid stimuli (i.e., words) were found to be significantly more attractive than the isolated vowels condition. Moreover, although our stimuli sample included only a small amount of different female voices ($n = 13$), this was compensated by a large sample of male judges ($n = 135$).

Methodological variations aside, our findings suggest that cultural factors may have influenced men's preferences for females' low-pitched voices, at least in a French speaking population. Interestingly, in two American samples, it has been reported that women who lowered their vocal pitch when addressing attractive conversational partners were perceived as more attractive (Hughes et al., 2010, 2014). Additionally, two other studies also found that men preferred lower vocal pitch in an American and Japanese speaking populations (Leaderbrand et al., 2008; Oguchi and Kikuchi, 1997). A possible explanation for such results is that representations of women through mass media, TV, films and radio broadcastings, and more generally contemporary cultural pressures, may have influenced such preferences, which may well be correlated to an increasingly more implicated occupational roles for women in traditionally male-oriented societies (Hughes et al., 2010). Gender representations in modern times may also have projected a stereotype of a "sexy voice" characterized by a lower vocal pitch in women. For instance, one study reported that women's voice pitch has significantly deepened over time (Pemberton et al., 1998). In this context, relatively lower F0 may lead to the perception of a more dominant (and therefore a more masculine) woman, as well as being perceived as more confident and trustworthy (Anderson and Klofstad, 2012; O'Connor et al., 2011). Thus, women exhibiting a relatively lower vocal pitch may be favored in gaining higher social statuses. Nonetheless, perceived femininity could also be mediated by other acoustic features, such as the F0-SD and the formant frequencies.

Indeed, men also significantly preferred more dynamic intonation patterns (i.e., higher F0-SD) over monotonous voices (i.e., lower F0-SD). F0-SD is a sexually dimorphic trait as men display twice as less variations, but less is known onto what function it yields. From the few available evidence, one study has shown that men increased their F0 variability when speaking to attractive opposite-sex individuals (as opposed to non-attractive individuals), while women did so but only after viewing an attractive same-sex competitor (Leongómez et al., 2014). This result could be interpreted as a potential marker of social interest or a simple physiological arousal in both sexes. But ultimately, two factors can be invoked. Firstly, increased F0-SD in women may be a mean to attract mates' attention by displaying higher perceived vocal femininity in a courtship context. Indeed, recordings of both men and women speaking to attractive individuals were preferred by naive listeners over recordings directed towards non-attractive individuals (Leongómez et al., 2014). Secondly, irrespective of sex, F0-SD may be a cue to current health and mental conditions. For instance, it has been previously reported that clinically depressed men and women show typically more monotonous voices as it is a marker of positive emotions (Bänziger and Scherer, 2005; Ellgring and Scherer, 1996). Thus, such results suggest that in both sexes higher F0-SD may be more attractive, which is corroborated by our results and previous findings for men (Bruckert et al., 2006; however see Hodges-Simeon et al., 2010).

Men also showed a significant preference for higher Df patterns. This result is consistent with those observed in an American-English population (Puts et al., 2011) and has been hypothesized to indicate perceived smaller body size characteristics in women, to which men seem to be more attracted to (Beck et al., 1976). Moreover, it has been shown that higher Df patterns are perceived as more flirtatious by other women, which may be a mean to track competitors' femininity. Interestingly, it has also been suggested to be an important vocal tracker of men's physical and social dominance (Puts et al., 2007). Another study found that

higher formant position, i.e., the average standardized formant value for the first n formants where formants are standardized using between-sex means and standard deviation, was perceived as more attractive by men (Babel et al., 2014).

Although the size effects were somewhat weaker than F0 and F0-SD, men significantly preferred voices richer in harmonics with relatively lower glottal noise (i.e., higher HNR profiles) as well as voices presenting higher F0 disturbance (i.e., relatively higher jitter profiles). It has been suggested that breathier voices are more characteristic of female voices than male voices (Van Borsel et al., 2009), which increase women's vocal attractiveness (Xu et al., 2013). HNR and jitter have been suggested to indicate body characteristics in both sexes (González, 2006; Pisanski et al., 2016). The study with the largest sample size reported that HNR and jitter were respectively negatively and positively associated to women's waist-to-hip ratios (Pisanski et al., 2016). Thus, women with less noise perturbations in their voices (i.e., unbreathy) were associated to physically more masculine women. Such findings are surprising given that men in our study preferred both higher HNR and jitter vocal profiles. Nonetheless, it has also been suggested that relationships between jitter, HNR and body characteristics could be mediated through hormonal profiles (Pisanski et al., 2016). As vocal folds are sensitive to circulating androgens levels, higher testosterone or low estrogens levels should lead to an increase of vocal folds' mass and size. Following, more massive and thicker vocal folds oscillate less regularly, resulting in a higher jitter, which could be associated to more masculine voices in women. As research produced mixed results (including ours), further studies are needed to understand the influence of HNR and jitter on perceived vocal attractiveness and their relationships to body characteristics and hormonal profiles.

Finally, men significantly preferred faster speech tempo in women. Speech tempo may be a signal of the speaker's current physical conditions. Indeed, on a biological level, the available evidence shows that speech rate can be ultimately traced to the individual's cognitive

abilities. For instance, it has been reported that individuals in poor health conditions suffering from cognitive disorders show typically reduced speech tempo, from depression (France et al., 2000; Teasdale et al., 1980) to neurodegenerative brain diseases such as amyotrophic lateral sclerosis (Turner et al., 1995), Parkinsonian dysarthria (Caligiuri, 1989), as well as speech stuttering (Andrade et al., 2003) and Alzheimer's disease (Gayraud et al., 2011). This can be explained by the fact that tempo represents an individual's capacity in speech mental planning processes, as individuals suffering from such diseases require additional effort and time to phonate naturally and fluently. Thus, faster speech tempo may indicate a more forceful and, possibly, a more skillful use of the vocal organs generally. Although significant differences exist between men and women across population of different languages, such preferences may be adaptive, but further studies are needed to understand its significance.

Eventually, no difference in men's preferences for tempo and acoustics features was observed between the two contextual questions as previously reported in the literature (Little et al., 2011; Puts et al., 2011). This does not contradict previous finding, but rather suggests that men are attracted to the same vocal features independently of the context of the relationship.

One possible limitation is that our study did not explore the direction (i.e., linearity or non-linearity) of the relationships between attractiveness and acoustics characteristics. For instance, a few studies have reported a linear positive relationship between the F0 and attractiveness (e.g. Collins and Missing, 2003; Feinberg et al., 2008; Re et al., 2012) but a few evidences suggests that it may show upper and lower thresholds. On one hand, voices with too high F0 may be perceived as 'babyish' (Burnham et al., 2002; Trehub et al., 1997), a feature linked to perceived immaturity (Berry, 1990, 1992). On the other hand, voices with very low F0 may be associated to masculine women, while also associated to dysphonia (Wolfe and Martin, 1997). Thus, voices with relatively too high or too low F0s may not be necessarily preferred. To our knowledge, only one study has reported a non-linear relationship in a Polish-

speaking population, with attractiveness ratings starting to decrease when the F0 is lower to 260 Hz (Borkowska and Pawlowski, 2011). Although tempo modifications only affected HNR values, another limitation lies in possible interactions between tempo and acoustic features. Furthermore, one individual might also prioritize one acoustic component over another when judging vocal attractiveness as it has been previously suggested (Puts et al., 2011).

Overall, the present study adds to the body of literature on acoustics preferences in women's voices. Although some results on preferences complement the same findings in regard to previous studies, further research is needed to assess vocal preferences across populations of different languages, as well as understanding the role of understudied acoustic features and their relative importance to vocal attractiveness within mate choice.

Data availability

The data associated with this research article are available at [link].

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- Addington, D. W. (1968). The relationship of selected vocal characteristics to personality perception. *Speech Monographs*, 35(4), 492–503. <https://doi.org/10.1080/03637756809375599>
- Anderson, R. C., and Klobstad, C. A. (2012). Preference for Leaders with Masculine Voices Holds in the Case of Feminine Leadership Roles. *PLoS ONE*, 7(12), e51216. <https://doi.org/10.1371/journal.pone.0051216>
- Andrade, C. R. F. de, Cervone, L. M., and Sassi, F. C. (2003). Relationship between the stuttering severity index and speech rate. *Sao Paulo Medical Journal*, 121(2), 81–84. <https://doi.org/10.1590/S1516-31802003000200010>

- Andreeva, B., Demenko, G., Mobius, B., Zimmerer, F., Jugler, J., and Oleskowicz-Popiel, M. (2014). Differences of Pitch Profiles in Germanic and Slavic Languages. *Fifteenth Annual Conference of the International Speech Communication Association*.
- Apple, W., Streeter, L. A., and Krauss, R. M. (1979). Effects of Pitch and Speech Rate on Personal Attributions. *Journal of Personality and Social Psychology*, 37(5), 715.
- Babel, M., McGuire, G., and King, J. (2014). Towards a More Nuanced View of Vocal Attractiveness. *PLoS ONE*, 9(2), e88616. <https://doi.org/10.1371/journal.pone.0088616>
- Bachorowski, J. A., and Owren, M. J. (1999). Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. *The Journal of the Acoustical Society of America*, 106(2), 1054–1063.
- Banse, R., and Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70(3), 614.
- Bänziger, T., and Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech Communication*, 46(3–4), 252–267. <https://doi.org/10.1016/j.specom.2005.02.016>
- Beck, S. B., Ward-Hull, C. I., and McLear, P. M. (1976). Variables Related to Women's Somatic Preferences of the Male and Female Body. *Journal of Personality and Social Psychology*, 34(6), 1200.
- Berry, D. S. (1990). Vocal attractiveness and vocal babyishness: Effects on stranger, self, and friend impressions. *Journal of Nonverbal Behavior*, 14(3), 141–153. <https://doi.org/10.1007/BF00996223>
- Berry, D. S. (1992). Vocal types and stereotypes: Joint effects of vocal attractiveness and vocal maturity on person perception. *Journal of Nonverbal Behavior*, 16(1), 41–54. <https://doi.org/10.1007/BF00986878>
- Block, S., and Killen, D. (1996). Speech rates of Australian English-speaking children and adults. *Australian Journal of Human Communication Disorders*, 24(1), 39–44.

- Boersma, P., and Weenink, D. (2017). Praat: Doing phonetics by computer [Computer software] (Version 6.0.33).
- Borkowska, B., and Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Animal Behaviour*, 82(1), 55–59. <https://doi.org/10.1016/j.anbehav.2011.03.024>
- Brown, B. L., Strong, W. J., and Rencher, A. C. (1973). Perceptions of personality from speech: effects of manipulations of acoustical parameters. *The Journal of the Acoustical Society of America*, 54(1), 29–35. <https://doi.org/10.1121/1.1913571>
- Bruckert, L., Lienard, J.-S., Lacroix, A., Kreutzer, M., and Leboucher, G. (2006). Women use voice parameters to assess men's characteristics. *Proceedings of the Royal Society B: Biological Sciences*, 273(1582), 83–89. <https://doi.org/10.1098/rspb.2005.3265>
- Bryant, G. A., and Haselton, M. G. (2009). Vocal cues of ovulation in human females. *Biology Letters*, 5(1), 12–15. <https://doi.org/10.1098/rsbl.2008.0507>
- Burnham, D., Kitamura, C., and Vollmer-Conna, U. (2002). What's New, Pussycat? On Talking to Babies and Animals. *Science*, 296(5572), 1435–1435.
- Byrd, D. (1994). Relations of sex and dialect to reduction. *Speech Communication*, 15(1–2), 39–54.
- Caligiuri, M. (1989). The influence of speaking rate on articulatory hypokinesia in parkinsonian dysarthria*1. *Brain and Language*, 36(3), 493–502. [https://doi.org/10.1016/0093-934X\(89\)90080-1](https://doi.org/10.1016/0093-934X(89)90080-1)
- Chaiken, S. (1979). Communicator Physical Attractiveness and Persuasion. *Journal of Personality and Social Psychology*, 37(8), 1387.
- Collins, S. A., and Missing, C. (2003). Vocal and visual attractiveness are related in women. *Animal Behaviour*, 65(5), 997–1004. <https://doi.org/10.1006/anbe.2003.2123>

- Ellgring, H., and Scherer, K. R. (1996). Vocal indicators of mood change in depression. *Journal of Nonverbal Behavior*, 20(2), 83–110. <https://doi.org/10.1007/BF02253071>
- Fairbanks, G., and Hoaglin, L. W. (1941). An experimental study of the durational characteristics of the voice during the expression of emotion. *Speech Monographs*, 8(1), 85–90. <https://doi.org/10.1080/03637754109374888>
- Feinberg, D. R., DeBruine, L. M., Jones, B. C., and Perrett, D. I. (2008). The Role of Femininity and Averageness of Voice Pitch in Aesthetic Judgments of Women’s Voices. *Perception*, 37(4), 615–623. <https://doi.org/10.1068/p5514>
- Ferdenzi, C., Patel, S., Mehu-Blantar, I., Khidasheli, M., Sander, D., and Delplanque, S. (2013). Voice attractiveness: Influence of stimulus duration and type. *Behavior Research Methods*, 45(2), 405–413. <https://doi.org/10.3758/s13428-012-0275-0>
- Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *The Journal of the Acoustical Society of America*, 102(2), 1213–1222.
- Fracarro, P. J., Feinberg, D. R., DeBruine, L. M., Little, A. C., Watkins, C. D., and Jones, B. C. (2010). Correlated Male Preferences for Femininity in Female Faces and Voices. *Evolutionary Psychology*, 8(3), 147470491000800. <https://doi.org/10.1177/147470491000800311>
- France, D. J., Shiavi, R. G., Silverman, S., Silverman, M., and Wilkes, M. (2000). Acoustical properties of speech as indicators of depression and suicidal risk. *IEEE Transactions on Biomedical Engineering*, 47(7), 829–837. <https://doi.org/10.1109/10.846676>
- Gayraud, F., Lee, H.-R., and Barkat-Defradas, M. (2011). Syntactic and lexical context of pauses and hesitations in the discourse of Alzheimer patients and healthy elderly subjects. *Clinical Linguistics and Phonetics*, 25(3), 198–209. <https://doi.org/10.3109/02699206.2010.521612>

- Gelfer, M. P., and Bennett, Q. E. (2013). Speaking Fundamental Frequency and Vowel Formant Frequencies: Effects on Perception of Gender. *Journal of Voice*, 27(5), 556–566. <https://doi.org/10.1016/j.jvoice.2012.11.008>
- Gelfer, M. P., and Mikos, V. A. (2005). The Relative Contributions of Speaking Fundamental Frequency and Formant Frequencies to Gender Identification Based on Isolated Vowels. *Journal of Voice*, 19(4), 544–554. <https://doi.org/10.1016/j.jvoice.2004.10.006>
- González, J. (2006). Research in acoustics of human speech sounds: Correlates and perception of speaker body size. *Recent Research Developments in Applied Physics*, 9, 1–15.
- Harnsberger, J. D., Shrivastav, R., Brown, W. S., Rothman, H., and Hollien, H. (2008). Speaking Rate and Fundamental Frequency as Speech Cues to Perceived Age. *Journal of Voice*, 22(1), 58–69. <https://doi.org/10.1016/j.jvoice.2006.07.004>
- Hillenbrand, J. (1988). Perception of aperiodicities in synthetically generated voices. *The Journal of the Acoustical Society of America*, 83(6), 2361–2371. <https://doi.org/10.1121/1.396367>
- Hillenbrand, J. M., and Clark, M. J. (2009). The role of f_0 and formant frequencies in distinguishing the voices of men and women. *Attention, Perception, and Psychophysics*, 71(5), 1150–1166. <https://doi.org/10.3758/APP.71.5.1150>
- Hodges-Simeon, C. R., Gaulin, S. J. C., and Puts, D. A. (2010). Different Vocal Parameters Predict Perceptions of Dominance and Attractiveness. *Human Nature*, 21(4), 406–427. <https://doi.org/10.1007/s12110-010-9101-5>
- Hodges-Simeon, C. R., Gaulin, S. J. C., and Puts, D. A. (2011). Voice Correlates of Mating Success in Men: Examining “Contests” Versus “Mate Choice” Modes of Sexual Selection. *Archives of Sexual Behavior*, 40(3), 551–557. <https://doi.org/10.1007/s10508-010-9625-0>

- Hughes, S. M., Dispenza, F., and Gallup, G. G. (2004). Ratings of voice attractiveness predict sexual behavior and body configuration. *Evolution and Human Behavior*, 25(5), 295–304. <https://doi.org/10.1016/j.evolhumbehav.2004.06.001>
- Hughes, S. M., Farley, S. D., and Rhodes, B. C. (2010). Vocal and Physiological Changes in Response to the Physical Attractiveness of Conversational Partners. *Journal of Nonverbal Behavior*, 34(3), 155–167. <https://doi.org/10.1007/s10919-010-0087-9>
- Hughes, S. M., Mogilski, J. K., and Harrison, M. A. (2014). The Perception and Parameters of Intentional Voice Manipulation. *Journal of Nonverbal Behavior*, 38(1), 107–127. <https://doi.org/10.1007/s10919-013-0163-z>
- Hughes, S. M., Pastizzo, M. J., and Gallup, G. G. (2008). The Sound of Symmetry Revisited: Subjective and Objective Analyses of Voice. *Journal of Nonverbal Behavior*, 32(2), 93–108. <https://doi.org/10.1007/s10919-007-0042-6>
- Jacewicz, E., Fox, R. A., O'Neill, C., and Salmons, J. (2009). Articulation rate across dialect, age, and gender. *Language Variation and Change*, 21(02), 233. <https://doi.org/10.1017/S0954394509990093>
- Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., and Vukovic, J. (2010). A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. *Animal Behaviour*, 79(1), 57–62. <https://doi.org/10.1016/j.anbehav.2009.10.003>
- Keating, P., and Kuo, G. (2012). Comparison of speaking fundamental frequency in English and Mandarin. *The Journal of the Acoustical Society of America*, 132(2), 1050–1060. <https://doi.org/10.1121/1.4730893>
- Leaderbrand, K., Dekam, J., Morey, A., and Tuma, L. (2008). The Effects of Voice Pitch on Perceptions of Attractiveness: Do You Sound Hot or Not? *Winona State University Psychology Student Journal*, 6.

- Leongómez, J. D., Binter, J., Kubicová, L., Stolařová, P., Klapilová, K., Havlíček, J., and Roberts, S. C. (2014). Vocal modulation during courtship increases perceptivity even in naive listeners. *Evolution and Human Behavior*, 35(6), 489–496. <https://doi.org/10.1016/j.evolhumbehav.2014.06.008>
- Linville, S. E., and Fisher, H. B. (1985). Acoustic characteristics of perceived versus actual vocal age in controlled phonation by adult females. *The Journal of the Acoustical Society of America*, 78(1), 40–48. <https://doi.org/10.1121/1.392452>
- Little, A. C., Connely, J., Feinberg, D. R., Jones, B. C., and Roberts, S. C. (2011). Human preference for masculinity differs according to context in faces, bodies, voices, and smell. *Behavioral Ecology*, 22(4), 862–868. <https://doi.org/10.1093/beheco/arr061>
- Lutz, K. C., and Mallard, A. R. (1986). Disfluencies and rate of speech in young adult nonstutterers. *Journal of Fluency Disorders*, 11(4), 307–316.
- Lyons, M., Lynch, A., Brewer, G., and Bruno, D. (2014). Detection of Sexual Orientation (“Gaydar”) by Homosexual and Heterosexual Women. *Archives of Sexual Behavior*, 43(2), 345–352. <https://doi.org/10.1007/s10508-013-0144-7>
- Miller, N., Maruyama, G., Beaber, R. J., and Valone, K. (1976). Speed of Speech and Persuasion. *Journal of Personality and Social Psychology*, 34(4), 615.
- Munson, B., McDonald, E. C., DeBoe, N. L., and White, A. R. (2006). The acoustic and perceptual bases of judgments of women and men’s sexual orientation from read speech. *Journal of Phonetics*, 34(2), 202–240. <https://doi.org/10.1016/j.wocn.2005.05.003>
- O’Connor, J. J. M., Fraccaro, P. J., Pisanski, K., Tigue, C. C., and Feinberg, D. R. (2013). Men’s Preferences for Women’s Femininity in Dynamic Cross-Modal Stimuli. *PLoS ONE*, 8(7), e69531. <https://doi.org/10.1371/journal.pone.0069531>

- O'Connor, J. J. M., Re, D. E., and Feinberg, D. R. (2011). Voice Pitch Influences Perceptions of Sexual Infidelity. *Evolutionary Psychology*, 9(1), 147470491100900. <https://doi.org/10.1177/147470491100900109>
- Oguchi, T., and Kikuchi, H. (1997). Voice and interpersonal attraction. *Japanese Psychological Research*, 39(1), 56–61. <https://doi.org/10.1111/1468-5884.00037>
- Pemberton, C., McCormack, P., and Russell, A. (1998). Have women's voices lowered across time? A cross sectional study of Australian women's voices. *Journal of Voice*, 12(2), 208–213. [https://doi.org/10.1016/S0892-1997\(98\)80040-4](https://doi.org/10.1016/S0892-1997(98)80040-4)
- Pépiot, E. (2014). 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, 305–309.
- Pisanski, K., Jones, B. C., Fink, B., O'Connor, J. J. M., DeBruine, L. M., Röder, S., and Feinberg, D. R. (2016). Voice parameters predict sex-specific body morphology in men and women. *Animal Behaviour*, 112, 13–22. <https://doi.org/10.1016/j.anbehav.2015.11.008>
- Ptacek, P. H., and Sander, E. K. (1966). Age recognition from voice. *Journal of Speech and Hearing Research*, 9(2), 273–277.
- Puts, David A., Barndt, J. L., Welling, L. L. M., Dawood, K., and Burriss, R. P. (2011). Intrasexual competition among women: Vocal femininity affects perceptions of attractiveness and flirtatiousness. *Personality and Individual Differences*, 50(1), 111–115. <https://doi.org/10.1016/j.paid.2010.09.011>
- Puts, David Andrew, Hodges, C. R., Cárdenas, R. A., and Gaulin, S. J. C. (2007). Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior*, 28(5), 340–344. <https://doi.org/10.1016/j.evolhumbehav.2007.05.002>
- Quene, H. (2007). On the just noticeable difference for tempo in speech. *Journal of Phonetics*, 35(3), 353–362.

- Rabinov, C. R., Kreiman, J., Gerratt, B. R., and Bielałowicz, S. (1995). Comparing Reliability of Perceptual Ratings of Roughness and Acoustic Measures of Jitter. *Journal of Speech Language and Hearing Research*, 38(1), 26. <https://doi.org/10.1044/jshr.3801.26>
- Ramig, L. A. (1983). Effects of physiological aging on speaking and reading rates. *Journal of Communication Disorders*, 16(3), 217–226. [https://doi.org/10.1016/0021-9924\(83\)90035-7](https://doi.org/10.1016/0021-9924(83)90035-7)
- Re, D. E., O'Connor, J. J. M., Bennett, P. J., and Feinberg, D. R. (2012a). Preferences for Very Low and Very High Voice Pitch in Humans. *PLoS ONE*, 7(3), e32719. <https://doi.org/10.1371/journal.pone.0032719>
- Re, D. E., O'Connor, J. J. M., Bennett, P. J., and Feinberg, D. R. (2012b). Preferences for Very Low and Very High Voice Pitch in Humans. *PLoS ONE*, 7(3), e32719. <https://doi.org/10.1371/journal.pone.0032719>
- Robb, M. P., Maclagan, M. A., and Chen, Y. (2004). Speaking rates of American and New Zealand varieties of English. *Clinical Linguistics and Phonetics*, 18(1), 1–15.
- Röder, S., Fink, B., and Jones, B. C. (2013). Facial, Olfactory, and Vocal Cues to Female Reproductive Value. *Evolutionary Psychology*, 11(2), 147470491301100. <https://doi.org/10.1177/147470491301100209>
- Rose, P. (1991). How effective are long term mean and standard deviation as normalisation parameters for tonal fundamental frequency? *Speech Communication*, 10(3), 229–247. [https://doi.org/10.1016/0167-6393\(91\)90014-K](https://doi.org/10.1016/0167-6393(91)90014-K)
- Scherer, K. R., Sundberg, J., Tamarit, L., and Salomão, G. L. (2015). Comparing the acoustic expression of emotion in the speaking and the singing voice. *Computer Speech and Language*, 29(1), 218–235. <https://doi.org/10.1016/j.csl.2013.10.002>
- Šebesta, P., Kleisner, K., Tureček, P., Kočnar, T., Akoko, R. M., Třebický, V., and Havlíček, J. (2017). Voices of Africa: acoustic predictors of human male vocal attractiveness. *Animal Behaviour*, 127, 205–211. <https://doi.org/10.1016/j.anbehav.2017.03.014>

- Sell, A., Bryant, G. A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., ... Gurven, M. (2010). Adaptations in humans for assessing physical strength from the voice. *Proceedings of the Royal Society B: Biological Sciences*, 277(1699), 3509–3518. <https://doi.org/10.1098/rspb.2010.0769>
- Shipp, T., Qi, Y., Huntley, R., and Hollien, H. (1992). Acoustic and temporal correlates of perceived age. *Journal of Voice*, 6(3), 211–216. [https://doi.org/10.1016/S0892-1997\(05\)80145-6](https://doi.org/10.1016/S0892-1997(05)80145-6)
- Siegmán, A. W., and Boyle, S. (1993). Voices of Fear and Anxiety and Sadness and Depression: The Effects of Speech Rate and Loudness on Fear and Anxiety and Sadness and Depression. *Journal of Abnormal Psychology*, 102(3), 430.
- Simpson, A. P. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass*, 3(2), 621–640. <https://doi.org/10.1111/j.1749-818X.2009.00125.x>
- Smith, B. L., Brown, B. L., Strong, W. J., and Rencher, A. C. (1975). Effects of Speech Rate on Personality Perception. *Language and Speech*, 18(2), 145–152. <https://doi.org/10.1177/002383097501800203>
- Smith, B. L., Wasowicz, J., and Preston, J. (1987). Temporal characteristics of the speech of normal elderly adults. *Journal of Speech, Language, and Hearing Research*, 30(4), 522–529.
- Sobin, C., and Alpert, M. (1999). Emotion in Speech: The Acoustic Attributes of Fear, Anger, Sadness, and Joy. *Journal of Psycholinguistic Research*, 28(4), 347–365.
- Street, R. L., Brady, R. M., and Lee, R. (1984). Evaluative responses to communicators: The effects of speech rate, sex, and interaction context. *Western Journal of Speech Communication*, 48(1), 14–27. <https://doi.org/10.1080/10570318409374138>
- Street, R. L., Brady, R. M., and Putman, W. B. (1983). The Influence of Speech Rate Stereotypes and Rate Similarity on Listeners' Evaluations of Speakers. *Journal of Language and Social Psychology*, 2(1), 37–56. <https://doi.org/10.1177/0261927X8300200103>

- Teasdale, J. D., Fogarty, S. J., and Williams, J. M. G. (1980). Speech rate as a measure of short-term variation in depression. *British Journal of Social and Clinical Psychology*, 19(3), 271–278. <https://doi.org/10.1111/j.2044-8260.1980.tb00353.x>
- Teixeira, J. P., Oliveira, C., and Lopes, C. (2013). Vocal Acoustic Analysis – Jitter, Shimmer and HNR Parameters. *Procedia Technology*, 9, 1112–1122. <https://doi.org/10.1016/j.protcy.2013.12.124>
- Titze, I. R. (1989). Physiologic and acoustic differences between male and female voices. *The Journal of the Acoustical Society of America*, 85(4), 1699–1707. <https://doi.org/10.1121/1.397959>
- Traunmüller, H., and Eriksson, A. (1995). The frequency range of the voice fundamental in the speech of male and female adults. *Unpublished Manuscript*, 11.
- Trehub, S. E., Unyk, A. M., Kamenetsky, S. B., Hill, D. S., Trainor, L. J., Henderson, J. L., and Saraza, M. (1997). Mothers' and fathers' singing to infants. *Developmental Psychology*, 33(3), 500.
- Turner, G. S., Tjaden, K., and Weismer, G. (1995). The Influence of Speaking Rate on Vowel Space and Speech Intelligibility for Individuals With Amyotrophic Lateral Sclerosis. *Journal of Speech Language and Hearing Research*, 38(5), 1001. <https://doi.org/10.1044/jshr.3805.1001>
- Van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*, 38(3), 253–265.
- Van Borsel, J., Janssens, J., and De Bodt, M. (2009). Breathiness as a Feminine Voice Characteristic: A Perceptual Approach. *Journal of Voice*, 23(3), 291–294. <https://doi.org/10.1016/j.jvoice.2007.08.002>

- Verhoeven, J., De Pauw, G., and Kloots, H. (2004). Speech Rate in a Pluricentric Language: A Comparison Between Dutch in Belgium and the Netherlands. *Language and Speech*, 47(3), 297–308. <https://doi.org/10.1177/00238309040470030401>
- Weiss, B., and Burkhardt, F. (2010). Voice Attributes Affecting Likability Perception. *Eleventh Annual Conference of the International Speech Communication Association*, 4.
- Wendahl, R. W. (1966). Some Parameters of Auditory Roughness. *Folia Phoniatica et Logopaedica*, 18(1), 26–32. <https://doi.org/10.1159/000263081>
- Whiteside SP. (1996). Temporal-based acoustic-phonetic patterns in read speech: Some evidence for speaker sex differences. *Journal of the International Phonetic Association*, 26(1), 23–40.
- Williams, C. E., and Stevens, K. N. (1972). Emotions and Speech: Some Acoustical Correlates. *The Journal of the Acoustical Society of America*, 52(4B), 1238–1250. <https://doi.org/10.1121/1.1913238>
- Wolfe, V., and Martin, D. (1997). Acoustic correlates of dysphonia: type and severity. *Journal of Communication Disorders*, 30(5), 403–416. [https://doi.org/10.1016/S0021-9924\(96\)00112-8](https://doi.org/10.1016/S0021-9924(96)00112-8)
- Xu, Y., Lee, A., Wu, W.-L., Liu, X., and Birkholz, P. (2013). Human Vocal Attractiveness as Signaled by Body Size Projection. *PLoS ONE*, 8(4), e62397. <https://doi.org/10.1371/journal.pone.0062397>
- Yamazawa, H., and Hollien, H. (1992). Speaking fundamental frequency patterns of Japanese women. *Phonetica*, 49(2), 128–140.
- Yuan, J., Liberman, M., and Cieri, C. (2006). Towards an Integrated Understanding of Speaking Rate in Conversation. *Ninth International Conference on Spoken Language Processing*, 4.

V. Évolution du dimorphisme vocal

V.1. Poids relatif des pressions de sélection sexuelle dans les deux sexes

Après avoir mis en évidence les préférences vocales des deux sexes, il convient de discuter de leurs interprétations évolutives. Tout d'abord, il est nécessaire de comprendre que le poids relatif des deux pressions de sélection sexuelle que sont la compétition intrasexuelle et la compétition intersexuelle ne sont pas équivalentes pour les deux sexes. Bien que les deux jouent évidemment un rôle, la compétition intrasexuelle semble avoir été plus importante chez les hommes quand le choix de partenaire a, semble-t-il, été d'une plus grande importance chez les femmes. Cette dichotomie peut encore s'expliquer par la théorie de l'investissement parental (Trivers, 1972). Rappelons ainsi que l'investissement parental – c'est-à-dire le temps et l'effort investi dans la production des gamètes, les soins, la protection et l'éducation des enfants – varie considérablement entre les sexes et les coûts physiologiques sont plus importants chez les femmes que chez les hommes. De plus, dans toutes les sociétés actuelles (et possiblement ancestrales), les hommes investissent systématiquement moins dans l'effort parental que ne le font les femmes.

De manière générale, cette asymétrie a pour conséquence une pression de sélection intrasexuelle plus importante chez les mâles d'une espèce, humains compris. En effet, de nombreuses études ont montré que les hommes ont une force corporelle et une propension à l'agression compétitive plus élevée, des histoires de vies plus risquées, une maturation sexuelle plus tardive, un taux de mortalité plus précoce et sont beaucoup moins regardants sur la qualité phénotypique de leur partenaire, particulièrement pour les relations à court terme (voir Puts et al., 2015 et Sugiyama, 2015 pour une discussion complète de ces aspects). De plus, les hommes présentent une variance plus large dans leur succès reproducteur (Wilson, Miller, & Crouse, 2017). Enfin, il a été montré que la compétition intrasexuelle a joué un rôle plus important chez les hommes dans l'évolution du dimorphisme sexuel aux plans facial, corporel, comportemental

et vocal (Hill, Hunt, Welling, Cárdenas, Rotella, Wheatley, Dawood, Shriver, & Puts, 2013 ; Kordsmeyer, Hunt, Puts, Ostner, & Penke, 2018 ; Puts, 2010 ; Saxton, Mackey, McCarty, & Neave, 2016). Dans ce contexte, il a été suggéré que la voix des hommes a essentiellement évolué sous la pression de la compétition intrasexuelle afin de signaler la masculinité, la dominance et l'agression, tandis que celle des femmes a plutôt évolué sous la pression de la compétition intersexuelle afin de signaler la féminité et la fertilité (Puts et al., 2014).

V.2. La voix : un signal « *honnête* » ?

L'idée selon laquelle les propriétés vocales d'un individu constituent un signal « *honnête* » est l'hypothèse évolutive la plus citée pour comprendre l'évolution du dimorphisme vocal dans l'espèce humaine (e.g. Collins & Missing, 2003 ; Feinberg et al., 2005 ; Hodges-Simeon et al., 2015 ; Jones et al., 2010; O'Connor et al., 2013 ; Pipitone & Gallup, 2008 ; Puts & Aung, 2019). On définit un signal « *honnête* » comme un trait — physique ou comportemental — dont la mise en place et le maintien nécessitent une importante allocation d'énergie. Dans ce contexte-là, nous parlons également de « *handicap* » (Zahavi, 1975). Toutefois, nous pouvons parler d'« *indice* » lorsque le trait est contraint par d'autres aspects du phénotype (Fitch & Giedd, 1999 ; Fitch, 1997). Dans notre contexte, la voix humaine constituerait un signal « *honnête* » de la qualité d'un individu en tant que partenaire et/ou compétiteur, et refléterait un ou plusieurs autres traits particulièrement pertinents dans le cadre de la compétition et/ou du choix d'un partenaire.

Dans le cas des « *handicaps* », seuls les individus en bonne condition physique peuvent supporter le coût de production de tels signaux. Ainsi, un signal « *honnête* » indique d'une façon fiable la qualité phénotypique (et potentiellement génotypique) d'un individu, celui-ci ne pouvant pas « *tricher* » sur ce trait. Dès lors qu'on l'intègre au processus de sélection sexuelle, la théorie du handicap permet d'expliquer pourquoi certains traits extravagants (e.g. ornements plumeux très colorés, chants complexes et parades nuptiales chez les oiseaux), pouvant

contribuer à première vue à diminuer la fitness des individus, ont été sélectionnés au cours de l'évolution.

Pour la voix, nous avons vu que les femmes expriment une nette préférence pour des voix masculines relativement graves (bas F0). Pour expliquer ces corrélations perceptives, l'hypothèse de l'« *immunohandicap* » a été avancée (Folstad & Karter, 1992). Cette hypothèse propose que, si les androgènes sont immunosuppresseurs (Foo, Nakagawa, Rhodes, & Simmons, 2017), alors la baisse du système immunitaire par des taux élevés d'androgènes n'est réalisable que pour les individus en bonne santé. En conséquence, les hommes dotés de voix graves, c'est-à-dire d'attributs vocaux coûteux, sont préférés des femmes du fait qu'ils reflètent honnêtement leur état de santé. Dans ce contexte, les études se sont donc surtout attachées à décrire des corrélations entre paramètres vocaux et testostérone. Toutefois, comme précédemment discuté (cf section II.2.), il apparaît à la lecture de ces travaux qu'aucun consensus ne puisse être dégagé à ce sujet (e.g. Dabbs & Mallinger, 1998 ; Puts et al., 2016 vs. Arnocky et al. 2018 ; Jost et al. 2018). Si la théorie de l'« *immunohandicap* » est souvent citée et prise pour acquise (e.g. Puts & Aung, 2019), la seule étude ayant été réalisée à ce sujet est celle de Nowak, Pawłowski, Borkowska, Augustyniak & Drulis-Kawa (2018), dont les résultats n'ont pas permis de vérifier cette hypothèse chez les hommes malgré une méthodologie rigoureuse visant à établir et à quantifier les liens entretenus entre taux d'androgènes, système immunitaire et divers traits dont le visage, le corps et la voix. De même, Skrinda et al. (2014) n'ont pas trouvé de lien entre le F0 et la capacité immunitaire. La seule étude ayant trouvé un résultat est celle d'Arnocky et al. (2018), qui a rapporté une corrélation négative entre le F0 et les résonances (mesurée via le Pf) avec l'immunoglobuline A (une molécule liée à l'immunité des muqueuses), seules corrélations significative sur une soixantaine n'ayant pas été corrigée par l'âge et la multiplicité des tests. Notons également, en marge de l'hypothèse de l'« *immunohandicap* » chez les hommes, que les mêmes voix graves jugées attractives par les

femmes sont également perçues comme appartenant à des individus socialement et physiquement plus dominants lorsque jugées par d'autres hommes (Hodges-Simeon et al., 2010 ; Puts et al., 2006/2007). La testostérone n'a pas seulement un effet immunosuppresseur, puisqu'elle varie aussi positivement avec la compétitivité, possiblement l'agression, et la recherche et le maintien du statut social (Eisenegger et al., 2011). Mais, étant donné les résultats mitigés entre paramètres vocaux et testostérone, de plus amples études sont nécessaires pour élucider ces aspects-là.

Chez les femelles, des compromis sont opérés entre l'allocation des ressources à la reproduction actuelle et la conservation des ressources pour la reproduction future (Stearns, 1992 ; Thornhill & Gangestad, 2008). Les œstrogènes jouent un rôle essentiel dans un certain nombre de fonctions de reproduction des femmes, telles que la fertilité et la grossesse (Ellison, 1988) et semblent également impliquées dans la production de signaux sexuels (Jasińska, Ziomkiewicz, Ellison, Lipson, & Thune, 2004 ; Law Smith, Perrett, Jones, Cornwell, Moore, Feinberg, Boothroyd, Durrani, Stirrat, Whiten, Pitman & Hillier, 2006). Par conséquent, l'effet de l'œstrogène sur la fonction immunitaire pourrait être lié à des compromis impliquant ces différentes fonctions. Cependant, la direction de l'effet des œstrogènes sur la fonction immunitaire n'est pas claire, et les études ont rapporté des résultats contradictoires (voir Foo et al., 2017 pour une méta-analyse et une revue exhaustive sur le sujet). Malgré l'absence de consensus, nous pouvons avancer l'hypothèse que si la voix des femmes reflète effectivement les taux d'œstrogènes et donc signale honnêtement la fertilité, cela expliquerait les préférences des hommes pour des voix féminines relativement aiguës. Néanmoins, si Vukovic, Feinberg, DeBruine, Smith & Jones (2010) rapportent une simple corrélation positive entre la hauteur vocale et un index de santé, rappelons que les deux seules études ayant tenté de quantifier le lien entre paramètres vocaux et œstrogènes ont observé une absence de corrélation (Jost et al., 2018 ; Puts et al., 2013). De plus amples études sont donc nécessaires pour éclaircir ces aspects.

Alternativement à l'hypothèse de l'« *immunohandicap* », il a aussi été suggéré que la voix constituerait un signal « *honnête* » de la taille corporelle et de la force physique. Ainsi, si la voix reflète la taille corporelle et, étant donné leur relation, possiblement la force physique d'un individu (Sell et al., 2010), le caractère « *honnête* » du signal fonctionne ici comme un indice contraint par d'autres aspects du phénotype (Fitch & Giedd, 1999 ; Fitch, 1997). Dans le contexte de la voix qui nous préoccupe ici, les qualités vocales d'un individu sont contraintes par les dimensions physio-anatomiques de l'appareil vocal, elles-mêmes imposées par des relations allométriques. Ceci justifierait que les individus mâles exhibant des voix de basses fréquences auraient été sélectionnés chez les hommes pour signaler la taille corporelle (et donc la force physique) afin d'éviter des situations de compétitions qui pourraient être potentiellement coûteuses (Ohala, 1996). De plus, la taille constitue un critère important pour le choix d'un partenaire (Courtiol, Raymond, Godelle, & Ferdy, 2010). Toutefois, si cette hypothèse est relativement bien validée pour d'autres mammifères (Bowling et al., 2017 ; Charlton & Reby, 2016), nous avons montré d'une part que la qualité vocale est un prédicteur limité de la taille corporelle (Pisanski et al., 2014) et d'autre part, que les résultats liés à la force physique sont divergents (e.g. Han et al., 2018 vs. Hodges-Simeon et al., 2015). Enfin, il a également été suggéré que la voix serait un indice de la qualité du développement foetal mais les résultats liant qualité vocale, le ratio 2D:4D et l'asymétrie fluctuante s'avèrent également contrastés (e.g. pour le 2D:4D voir Evans et al., 2008 et Hughes et al., 2008 vs. Levrero et al., 2018 ; pour l'asymétrie fluctuante voir Hill et al., 2017 vs. Hughes et al., 2008).

V.3. La voix : un biais sensoriel ?

Une hypothèse alternative peut permettre d'expliquer l'évolution du dimorphisme vocal : la préférence pour les voix masculines pourrait simplement refléter l'exploitation sensorielle d'un ancien biais évolutif concernant la réaction aux objets émettant des vibrations de basses fréquences (Armstrong, Lee, & Feinberg, 2019 ; Feinberg, Jones, & Armstrong, 2018

; Feinberg, Jones, & Armstrong, 2019). Dans notre contexte, cette hypothèse suggère que les mâles possédant des voix relativement graves, élicitant des taux élevés de stimulation dans les systèmes sensoriels des femelles, seront préférés par celles-ci et auraient *in fine* plus de succès (Ryan, 1990). Au fil du temps, la sélection augmente la fréquence et la taille de ce trait via la préférence des femelles. Surtout, les préférences des femelles n'ont pas besoin d'être adaptatives en elles-mêmes mais peuvent être des sous-produits de réponses neuronales ayant évolué pour d'autres fonctions. Cette hypothèse a été récemment proposée et développée par Feinberg et al. (2018).

Les auteurs suggèrent qu'une propriété physique générale de notre environnement est la tendance des objets larges à produire des sons de basses fréquences. De manière intéressante, les auteurs argumentent que les hommes et les femmes associent des bruits mécaniques de basses vs. hautes fréquences à des objets respectivement grands/larges vs. petits/fins, phénomène supposément transposables aux voix humaines. Ce phénomène semble d'ailleurs largement répandu dans le règne animal (Morton, 1977) et, de manière intéressante, il a été montré que même des adultes congénitalement aveugles effectuent ces associations (Pisanski, Oleszkiewicz, & Sorokowska, 2016). De plus, les caractéristiques acoustiques d'une voix expliquent relativement peu de variation dans la taille corporelle (Pisanski et al., 2014) et les individus réalisent des estimations relativement incorrectes de celle-ci (Rendall, Vokey, & Nemeth, 2007). Surtout, l'association générale entre sons de basses et/ou hautes fréquences avec des objets larges et/ou fins pourrait servir d'intermédiaire à la perception de dominance sociale et physique, étant donné la relation existant entre la taille corporelle et la dominance (Buunk, Park, Zurriaga, Klavina, & Massar, 2008 ; Watkins, Fraccaro, Smith, Vukovic, Feinberg, DeBruine, & Jones, 2010).

Dans ce contexte, cette association cognitive pourrait engendrer une sélection pour des voix relativement graves chez les hommes via la préférence des femmes si, toutes choses étant

égales par ailleurs, ces hommes possédant ce type de voix exploitent un biais sensoriel présent chez les femmes à être attiré par des hommes « *résonant* » comme plus corpulent. Même s'il reste à mettre directement en évidence que les femmes préfèrent de telles voix chez les hommes (i.e. « *corpulent* » = « *attractif* »), nous avons montré qu'elles considèrent les voix relativement graves comme plus attractives (Bruckert et al., 2006 ; Collins, 2000 ; Feinberg et al., 2005 ; Hodges-Simeon et al., 2010 ; Hughes et al., 2010 ; Jones et al., 2010 ; Pisanski & Rendall, 2011 ; Vukovic et al., 2008 ; Xu et al., 2013). De même, cette association pourrait aussi évoluer par une pression de sélection intrasexuelle si, toutes choses étant égales par ailleurs, les hommes possédant de telles voix ont plus de succès dans divers contextes de compétition en exploitant ce biais chez les compétiteurs pour les faire perceptuellement apparaître comme plus corpulents et donc dominants.

A l'heure actuelle, une seule étude corrélationnelle a abordé cette hypothèse afin de comprendre l'absence de relation biunivoque entre hauteur vocale et taille corporelle. Les résultats de Armstrong et al. (2019) suggèrent ainsi que l'existence du biais sensoriel pourrait expliquer ce décalage entre perception et réalité.

V.4. La voix : un signal « *redondant* » ?

Selon l'hypothèse des « *signaux redondants* » (i.e. « *backup signals* » en anglais) les traits d'un individu, tels que le visage, le corps et/ou la voix, fournissent des informations plus ou moins similaires et chevauchantes sur sa qualité phénotypique (Johnstone, 1997). Suivant ce modèle, les individus sont attentifs à plusieurs traits car, lorsque combinés, ils fournissent une meilleure estimation de la qualité d'un individu que n'importe quel trait pris isolément. L'existence des signaux redondants permettrait donc de réduire la probabilité de faire des évaluations inexactes. Ceci nécessite donc que ces signaux doivent être des indicateurs fiables de la qualité générale d'un individu, et doivent être plus susceptibles d'informer celle-ci sur le long-terme plutôt que d'être influencés par des états contextes-dépendants. Le visage et la voix

sont considérés comme de bons candidats pour valider cette hypothèse. En effet, il a été montré que l'attractivité faciale et vocale sont positivement corrélées dans les deux sexes (Abend et al., 2015 ; Feinberg et al., 2005 ; Valentova, Varella, Havlíček, & Kleisner, 2017 ; Wheatley, Apicella, Burriss, Cárdenas, Bailey, Welling, & Puts, 2014). Dans notre contexte, la voix aurait évolué afin de signaler les mêmes informations sur la qualité phénotypique d'un individu.

A l'inverse, certaines études suggèrent que les informations recueillies à partir de plusieurs modalités peuvent avoir des effets indépendants et additifs sur les jugements d'attractivité, de telle sorte que les voix, les visages, les corps et les odeurs corporelles peuvent fournir des informations partiellement redondantes – voire non redondantes – sur la qualité d'un partenaire (Groyecka, Pisanski, Sorokowska, Havlíček, Karwowski, Puts, Roberts, & Sorokowski, 2017). Toutefois, pour valider l'une ou l'autre de ces hypothèses, il est d'abord nécessaire de comprendre en quoi et dans quelle mesure la qualité vocale représente un indice fiable de la qualité d'un partenaire, ce qui, comme nous l'avons montré, ne fait pas encore consensus.

Enfin, il est à noter que l'attraction pour un trait particulier a tendance à augmenter la perception d'autres traits liés à l'attractivité (e.g. l'attractivité du visage augmente l'attractivité vocale, et inversement, Hughes & Miller, 2016 ; Zuckerman & Driver, 1988). De même, les individus ont naturellement tendance à attribuer aux personnes qu'ils jugent attractives d'autres qualités positives, comme l'intelligence ou la confiance en soi (Brand, Bonatsos, D'Orazio, & DeShong, 2012 ; Kanazawa, 2004).

V.5. Conclusion

Peu d'hypothèses évolutives ont été formulées afin d'expliquer l'évolution du dimorphisme vocal. Si la théorie des signaux « *honnêtes* » est prédominante, il est à noter que les recherches conduites sur ce sujet ont surtout tenté d'éprouver son bien-fondé en multipliant les corrélations entre qualité vocale et tout autre type de trait. L'hypothèse du biais sensoriel

qui a été récemment proposée reste encore à explorer dans la mesure où elle ne fournit que des explications partielles quant à l'émergence du dimorphisme vocal. En l'absence de causes ultimes clairement définies, nous nous autorisons donc à formuler ici une nouvelle hypothèse.

Typiquement, nous suggérons que la voix masculine (notamment la hauteur et les résonances) pourrait signaler le statut social. L'acquisition d'un statut social plus élevé permet l'accès à plus de ressources et, *in fine*, à une meilleure reproduction (Cummins, 2015). De plus, le statut social est - pour une femme - le trait le plus important à évaluer lors du choix d'un partenaire, et ceci dans toutes les cultures humaines (Buss, 1989 ; Cummins, 2015 ; Sugiyama, 2015), dans la mesure où il est étroitement lié aux ressources mobilisables par l'homme pour un investissement parental de qualité. Le statut social est évidemment affiché par des éléments culturels (e.g. habits), mais sa détection pourrait en être facilitée par des indices vocaux, pour lesquelles les femmes devraient y être particulièrement attentives.

Ainsi, l'idée qui consiste à considérer que la voix a pu être sélectionnée comme un moyen efficace d'afficher son statut social mérite d'être explorée. Quelques études suggèrent que la qualité vocale varie avec le statut social chez les hommes, mais de manière intéressante, aussi chez les femmes. Par exemple, une étude expérimentale a montré que les individus des deux sexes modifient la hauteur de leur voix (F0), leur intonation (F0-SD) et leur intensité vocale après manipulation de leur statut social (Ko, Sadler, & Galinsky, 2015). Il a aussi été montré aux cours d'expériences où les sujets ont pour tâche de coopérer, que les individus émergeant en haut vs. ceux émergeant en bas de la hiérarchie exhibent des qualités vocales distinctes avant, pendant et à la fin de l'expérience (Cheng, Tracy, Ho, & Henrich, 2016). De manière générale, il semblerait également que les individus s'ajustent vocalement (i.e. notion de convergence) selon le statut social de leur interlocuteur (Gregory & Webster, 1996 ; Leongómez, Mileva, Little, & Roberts, 2017). Bien qu'il y ait relativement peu d'études qui se soient intéressées aux avantages sociaux directs procurés par la voix, il a aussi été montré que

les gens préfèrent choisir comme leader un individu attestant une voix plus masculine (Anderson & Klobstad, 2012 ; Klobstad, Anderson, & Peters, 2012 ; Klobstad, Anderson, & Nowicki, 2015). Afficher de telles caractéristiques vocales semblent même influencer sur les comportements de vote (Tigue, Borak, O'Connor, Schandl, & Feinberg, 2012) et, de façon surprenante, il est possible de prédire les résultats des élections présidentielles en se basant uniquement sur la voix des candidats (Banai, Banai, & Bovan, 2017). Il a aussi été montré que la hauteur vocale est liée à des postes plus élevés dans les entreprises : les PDG avec des voix plus graves gèrent de plus grandes entreprises, gagnent plus d'argent et occupent leurs fonctions plus longtemps (Mayew, Parsons, & Venkatachalam, 2013). Enfin, chez les Hadzas (une population de chasseurs-cueilleurs de Tanzanie), il semblerait également que les hommes ayant une haute réputation en tant que chasseurs cueilleurs présentent des voix plus graves que ceux ayant une faible réputation (Smith, Olkov, Puts, & Apicella, 2017).

Il apparaît ainsi particulièrement intéressant de s'intéresser au lien de causalité : le changement de qualité vocale serait-il dû à une meilleure qualité de vie (alimentation, diminution du stress, etc.) ? Ou existe-t-il une base environnementale et génétique influençant le développement de cette qualité et qui, par le biais de l'hypothèse sensorielle, favoriserait ce type d'individu ? Enfin, les androgènes constitueraient-ils le facteur médiateur de cette relation ? De cette dernière question, rappelons qu'un statut social élevé modifie la physiologie et le comportement des individus concernés (Cummins, 2015), notamment via le cortisol et/ou la testostérone, hormones qui participent à la motivation à acquérir et à maintenir le statut social (Eisenegger et al., 2011). De manière intéressante, Puts et al. (2016) ont montré que l'interaction entre la testostérone et le cortisol prédit la hauteur vocale, de telle manière que la testostérone est négativement corrélée à celle-ci seulement chez les hommes présentant des taux de cortisol relativement peu élevés. En ce sens, le désaccord dans la littérature concernant les liens entre paramètres vocaux et la testostérone pourrait être confondu par le statut social. Enfin,

et de manière intéressante, la dominance physique et sociale, liées à une plus grande taille corporelle et force physique, devraient faciliter l'accès à un plus haut statut. Toutefois, nous suggérons que ces propriétés ne sont plus les seuls déterminants dans l'acquisition et le maintien du statut social dans le contexte des sociétés occidentales, puisque depuis la révolution industrielle, l'accès à la nourriture et les soins de haute qualité ne sont plus réservés aux individus dominants physiquement. L'accès à un plus haut statut doit ainsi reposer sur d'autres éléments, et la voix pourrait en constituer un indice. Dans ce contexte, l'étude entre paramètres vocaux et statut social nous permet d'envisager un champ de recherche intéressant pour l'avenir.

Article 3

The influence of sexual orientation and testosterone levels on speech acoustic features

Alexandre Suire, Arnaud Tognetti, Valérie Durand, Michel Raymond &

Melissa Barkat-Defradas

Article en revue dans *Archives of Sexual Behavior*

Résumé

Bien que cette étude ne semble pas directement s'inscrire dans le chapitre jusque-là décrit, les résultats qui en découlent nous permettent indirectement de proposer de nouvelles idées concernant le lien entre paramètres acoustiques et signal « *honnête* ». En ce sens, nous avons choisi de la présenter ici.

Les différences entre les hommes homosexuels et hétérosexuels ont été étudiées sur une diversité de traits sociobiologiques. Concernant les caractéristiques vocales, l'hypothèse de féminisation suppose que la qualité vocale des hommes homosexuels tend vers celle des femmes. Par exemple, les stéréotypes sociaux supposent que ces derniers exhiberaient des voix relativement plus aiguës que les hommes hétérosexuels. Toutefois, les études ayant tenté d'établir des différences acoustiques ont mené à des résultats relativement contradictoires. Enfin, aucune étude n'a encore été menée pour déterminer si les niveaux de testostérone pourraient jouer un rôle médiateur dans l'association entre l'orientation sexuelle et la qualité vocale.

Dans une étude exploratoire, nous avons mesuré la testostérone circulante (dite « *libre* ») et 4 paramètres acoustiques (i.e. le F0, le F0-SD, le jitter et le HNR) lors d'une récitation d'une histoire chez 49 hommes hétérosexuels, 58 hommes homosexuels et 54 femmes hétérosexuelles. En plus de répliquer les différences sexuellement dimorphiques pour les traits analysés, nos résultats ont révélé que les hommes homosexuels exhibent une intonation plus marquée (haut F0-SD) et présentent moins de souffle dans leur voix (haut HNR) que les hommes hétérosexuels, suggérant ainsi une légère féminisation de la qualité vocale. Concernant la testostérone circulante, nos résultats n'ont révélé aucune différence entre les hommes hétérosexuels et homosexuels et n'a prédit aucune variation pour tous les paramètres acoustiques. Enfin, le statut socio-économique (mesuré via les revenus mensuels et le niveau

d'éducation) n'a prédit aucune variation pour les paramètres acoustiques, même si cette absence d'effet résulte probablement d'une très faible variation du statut social des sujets considérés.

Premièrement, une des possibles interprétation pour expliquer ces différences acoustiques est qu'elles seraient socialement apprises pour signaler une appartenance à un groupe social. Par exemple, une intonation relativement marquée pourrait constituer un signal d'identité intra-groupe permettant d'identifier l'orientation sexuelle d'un homme. En effet, l'hypothèse du « *vocal gaydar* » a montré que l'orientation sexuelle était détectable sur la seule base de la voix, et ceci à la fois pour des auditeurs hommes et femmes, hétérosexuels et non-hétérosexuels. Si l'hypothèse de facteurs biologiques n'est évidemment pas à écarter (ni même une interaction entre facteurs biologiques et sociaux), l'idée d'un signal acoustique socialement « *appris* » suggère de nouvelles recherches prometteuses, au-delà même de l'orientation sexuelle. Deuxièmement, indépendamment de l'orientation sexuelle, aucun des paramètres acoustiques n'étaient liés au taux de testostérone circulant (hormis les différences hommes vs. femmes). Ces résultats suggèrent que la voix ne constitue pas un signal « *honnête* » de ce dernier, invalidant ainsi l'hypothèse de l'« *immunohandicap* ».

The influence of sexual orientation and testosterone levels on speech acoustic features

Authors: Alexandre Suire^{1*}, Arnaud Tognetti^{2*}, Valérie Durand¹, Michel Raymond¹ &

Melissa Barkat-Defradas¹

Corresponding author: Arnaud Tognetti

E-mail: arnaud.tognetti@ki.se

¹ Institut des Sciences de l'Évolution de Montpellier, Université de Montpellier, Centre National de la Recherche Scientifique, Institut pour la Recherche et le Développement, École Pratique des Hautes Études – Place Eugène Bataillon 34095 Montpellier, France.

² Division of Psychology, Department of Clinical Neuroscience, Karolinska Institutet, SE-171

77 Stockholm - Sweden

*Co-first authorship

Abstract

Potential differences between homosexual and heterosexual men have been studied on a diverse set of social and biological traits. Regarding acoustic features of speech, researchers have hypothesized a feminization of such characteristics in homosexual men, but previous investigations have so far produced mixed results. Moreover, most studies have been conducted with English speaking populations, which calls for further cross-linguistic examinations. Lastly, no studies investigated so far the potential role of testosterone in the association between sexual orientation and speech acoustic features. To fill these gaps, we explored potential differences in acoustic features of speech between homosexual and heterosexual native French men and investigated whether the former showed a trend towards feminization by comparing theirs to that of heterosexual native French women. Lastly, we examined whether testosterone levels mediated the association between speech acoustic features and sexual orientation. We studied four sexually dimorphic acoustic features relevant for the qualification of feminine vs. masculine voices: the fundamental frequency, its modulation, and two understudied acoustic features of speech, the Harmonics-to-Noise ratio (a proxy of vocal breathiness) and the jitter (a proxy of vocal roughness). Results showed that homosexual men displayed significantly higher pitch-modulations patterns and less breathy voices compared to heterosexual men, with values shifted towards those of heterosexual women. Lastly, testosterone levels did not influence any of the investigated acoustic features. Combined with the literature conducted in other languages, our findings bring new support for the feminization hypothesis and suggest that the feminization of some acoustic features could be shared across-languages.

Keywords: speech; voice; acoustics; homosexuality; testosterone levels; gender atypicality hypothesis.

Introduction

The gender atypicality hypothesis suggests that gender atypical traits in homosexuals could be used as cues to indicate sexual orientation. Differences between heterosexual and homosexual individuals have thus been studied on a diverse set of traits such as face (e.g., Freeman, Johnson, Ambady, & Rule, 2010; Lyons, Lynch, Brewer, & Bruno, 2014; Rieger, Linsenmeier, Gygax, Garcia, & Bailey, 2010), olfaction (e.g., Sergeant, Dickins, Davies, & Griffiths, 2007), behavior (e.g., Ambady, Hallahan, & Conner, 1999; Rieger, Linsenmeier, Gygax, & Bailey, 2008; Valentova, Rieger, Havlicek, Linsenmeier, & Bailey, 2011), cognition (e.g., Neave, Menaged, & Weightman, 1999; Xu, Norton, & Rahman, 2017) and voice (e.g., Gaudio, 1994; Munson, McDonald, DeBoe, & White, 2006; Pierrehumbert, Bent, Munson, Bradlow, & Bailey, 2004; Rendall, Vasey, & McKenzie, 2008). In addition to the fact that homosexuals exhibit traits that differ from those of heterosexuals, it has been shown that some of them, such as specific neural processes (LeVay, 1991; Savic, Berglund, & Lindstrom, 2005) or specific childhood behaviors (Alanko et al., 2010; Bailey & Zucker, 1995) displayed values shifted towards those of the opposite sex, i.e., a feminization in homosexual men and a masculinization in homosexual women (Pierrehumbert et al., 2004). Moreover, studies have shown that both men and women are able to accurately assess sexual orientation from both sexes from various features such as the face or body movements (Ambady et al., 1999; Rieger et al., 2010; Valentova et al., 2011). These findings emphasize the idea that specific phenotypic traits may be influenced by sexual orientation and may be used as cues to detect or advertise it.

Another important trait that seems to be influenced by sexual orientation and used as a cue to assess is speech (for a detailed review, see: Munson & Babel, 2007). For example, popular stereotypes regarding the speech of homosexual men generally attribute speech patterns characteristic of the opposite-sex, i.e., a broadly feminized speech, such as a higher fundamental frequency (i.e., F0, the acoustic correlate of voice pitch) and a greater variation in the intonation

(i.e., F0-SD, the local variations of F0 throughout speech, henceforth, referred to as vocal modulation) (Cartei & Reby, 2012; Munson & Babel, 2007). Although there is no clear evidence that the mean fundamental frequency differ between homosexual and heterosexual men (Gaudio, 1994; Lerman & Damsté, 1969; Munson, McDonald, et al., 2006; Rendall et al., 2008; Rogers, Jacobs, & Smyth, 2001; Smyth, Jacobs, & Rogers, 2003; but see: Baeck, Corthals, & Borsel, 2011), results towards differences in pitch-modulation patterns are more controversial: some studies have found that homosexual men displayed greater variations in intonation, with values shifted towards those of women (Baeck et al., 2011; Gaudio, 1994) while others did not find any difference (Levon, 2006; Rogers et al., 2001). Spectral measures of fricatives also seemed to be influenced by sexual orientation (Munson, Jefferson, & McDonald, 2006). For instance, homosexual men produce higher peak frequency and longer duration values for /s/ (Linville, 1998) and these speech characteristics are associated with “gayer-sounding” voices by listeners (Mack & Munson, 2012). Lastly, homosexual men seem to produce a more expanded vowel space than heterosexual men for some specific vowels (Rendall et al., 2008), hyper articulation being commonly found in female speech (Pierrehumbert et al., 2004).

Aside these acoustic speech features, other characteristics could vary with sexual orientation, such as vocal breathiness and roughness that are respectively captured by the Harmonics-to-Noise ratio and the jitter. Indeed, both components are sexually dimorphic as women exhibit significantly higher values of HNR (i.e., lower “breathy” voices) and lower values of jitter (i.e., lower “rougher” voices) than men (Graddol, 1989; Van Borsel, Janssens, & De Bodt, 2009). Although vocal breathiness has been suggested to be an important component of femininity for female voices (Van Borsel et al., 2009), significant relationships in vocal attractiveness for both sexes have been reported (e.g., Xu, Lee, Wu, Liu, & Birkholz, 2013), while vocal roughness has been found to be positively associated to male vocal

attractiveness (Hughes, Dispenza, & Gallup, 2004). Such results suggest that vocal breathiness and roughness may play a role in the qualification of masculine vs. feminine sounding voices, thus questioning homosexuals' vocal breathiness and roughness within this continuum. In line with the speech feminization hypothesis, homosexual men could indeed potentially exhibit higher values of HNR and lower values of jitter, but so far, no studies have tackled this issue.

Researches have tried to assess if the feminized traits in homosexual men can be attributable to proximate mechanisms such as the differences in sexual hormone levels. Testosterone, a male sexual hormone, has thus been intensively studied as it was found to be associated, for instance, with facial (e.g., Penton-Voak & Chen, 2004; Pound, Penton-Voak, & Surridge, 2009; Roney, Hanson, Durante, & Maestripieri, 2006) and behavioral masculinity (e.g., Apicella et al., 2008; Archer, 2006; Booth, Shelley, Mazur, Tharp, & Kittok, 1989). Concerning acoustic characteristics, several studies have found a negative relationship between fundamental frequency and testosterone levels in men (Dabbs & Mallinger, 1999; Evans, Neave, Wakelin, & Hamilton, 2008; Hodges-Simeon, Gurven, & Gaulin, 2015; Puts, Apicella, & Cardenas, 2012). Although little is known about their physiological mechanisms, both the HNR and jitter have also been suggested to be sensitive to hormonal influx as they both relate to the oscillations of the vocal folds, which possess receptors to circulating androgens (Pisanski et al., 2016). Although evidence of a difference in testosterone levels between homosexual and heterosexual men are inconsistent (Brodie, Gartrell, Doering, & Rhue, 1974; Jaffee, McCormack, & Vaitukaitis, 1980; Kolodny, Masters, Hendryx, & Toro, 1971; Neave et al., 1999; Pillard, Rose, & Sherwood, 1974), testosterone may still mediate the relationship between sexual orientation and the aforementioned vocal speech features, which has received little attention so far.

Finally, most of the studies that investigated the link between sexual orientation and speech characteristics have been conducted with native English speakers (e.g.: Gaudio, 1994;

Linville, 1998; Pierrehumbert et al., 2004; Rendall et al., 2008; although see Baeck et al., 2011; Valentova & Havlíček, 2013 for examples with Dutch and Czech men). This calls for further cross-linguistic examinations as numerous studies have unveiled important differences in vocal quality (i.e., the set of acoustic characteristics linked to a particular voice) across languages (e.g., Andreeva et al., 2014; Keating & Kuo, 2012; Traunmüller & Eriksson, 1995; Zimmerer, Jügler, Andreeva, Möbius, & Trouvain, 2014). Consequently, communities of homosexual men could potentially differ in their specific vocal speech features across different languages.

In this context, the goal of the present study is to provide further details on the potential differences between homosexual and heterosexual men's speech in an underrepresented population in the literature (i.e., French men). We investigated the effect of sexual orientation on four sexually dimorphic acoustic parameters (F0, F0-SD, jitter and HNR) and examined whether homosexual men's vocal characteristics showed a feminization by comparing theirs with that of heterosexual women. Lastly, we examined the potential role of testosterone in the association between speech acoustic features and sexual orientation.

Methods

Participants

The French National Commission of Informatics and Liberties approved all protocols used in this study (CNIL number 1261003). Participants were recruited by means of flyers handed out as well as advertisements posted on public and private locations in the city of XXX, France.

In order to recruit as much as possible homosexual males, we contacted the local LGBTQ community to help advertise the study as well as directly advertising it in known local gay bars. All participants gave a written consent prior to the study and were given a financial compensation for their participation. In total, 150 women and 181 men participated in the study. All participants self-reported their sexual orientation and filled out a questionnaire assessing their nationality, age, relationship status (single vs. in relationship), socio-economic

status (i.e., level of education and monthly income) as well as country of birth of their parents and grandparents.

Speech Samples and Acoustic Analysis Procedure

Recordings

Recordings took place in a quiet room in our laboratory at the University of Montpellier. All recordings took place between 2:00 pm and 5:00 pm. Each participant heard the French version of the story “The North Wind and the Sun” from the International Phonetic Association and were asked to tell the story back to the research assistant. The rationale for using semi-spontaneous speech is that it is more ecologically valid than sustained vowels or read speech while controlling for semantic content, as the latter produce very different acoustic speech characteristics that do not represent how an individual vocally behave in social interactions (Laan, 1997; Suire, Raymond, & Barkat-Defradas, 2018). Speech samples were recorded using a linear PCM recorder (DR-O7 MKII, Tascam©) with a sampling rate of 22 kHz, 16-bit, mono, then saved as .wav files. To control for intensity, participants were asked to speak within a constant distance of 15 cm from the recorder.

Speech analyses

Because origin (e.g., Ordin & Mennen, 2017; Zimmerer et al., 2014) and language (e.g., Andreeva et al., 2014; Keating & Kuo, 2012; Traunmüller & Eriksson, 1995) influence speech characteristics and vocal quality parameters, we only analyzed participants who were native French speakers with European ascendants. We also only focused on participants who declared themselves as homosexual and heterosexual (we excluded those who declared to be bisexual). The final sample size resulted in 49 heterosexual men (mean age \pm SD = 26.12 \pm 5.37 years), 58 homosexual men (26.38 \pm 5.06 years), and 54 heterosexual women (24.85 \pm 4.34 years).

In total, we analyzed these 161 speech samples with the Praat© software (Paul Boersma and David Weenink, Phonetic Sciences, University of Amsterdam, www.praat.org). Pitch

floors were set to 75 Hz with a ceiling of 300 Hz for both heterosexual and homosexual men and 85-400 Hz for heterosexual women. All other settings were kept as default. For each participant, we extracted four acoustic parameters: mean fundamental frequency (F0, in Hz), its variations (F0-SD, in Hz), the jitter (%) and the Harmonics-to-Noise Ratio (HNR, in dB).

The mean fundamental frequency is the perceptual correlate of the vocal pitch while its variations are the perceptual correlate of micro-intonation patterns. HNR is the perceptual correlate of vocal breathiness, which corresponds to a ratio between periodic components (i.e., the harmonics) and a non-periodic component (i.e., noise) comprising a segment of voiced speech (Teixeira, Oliveira, & Lopes, 2013). More specifically, this ratio reflects the efficiency of speech production. The greater the flow of air expelled from the lungs into energy of vibration of the vocal cords, the greater the HNR, which is perceptually associated with a more sonorant and harmonic voice. On the contrary, a lower HNR is generally associated with a perceptually asthenic, dysphonic and breathier voice. Vocal roughness can be captured by the jitter, a measure of the F0 disturbance, which is defined as the parameter capturing the frequency variation from cycle to cycle in the sound wave (Rabinov, Kreiman, Gerratt, & Bielamowicz, 1995). More specifically, the jitter measures the control of the vocal folds during successive periods of oscillations. The higher the jitter, the “rougher” sounds the voice.

Saliva Collection and Testosterone Assays

Testosterone levels (henceforth T-levels) were measured in saliva samples (pg/ml). This non-invasive technique has been previously validated and yields T-levels that are highly correlated with serum levels (Ellison, 1988). At the beginning of the experiment, one labeled tube and straw (Salicaps kits, IBL-Hamburg) was given to each participant to collect saliva. Participants were asked not to eat, drink (except plain water), smoke, chew gum, or brush their teeth for one hour before each session so as to avoid saliva contamination. Samples were kept cold during the duration of the experiment then stored at -80°C before being analyzed by

Luminescence ImmunoAssay (LIA) technique, using LIA Testosterone kits (IBL, Hamburg). The assay of each sample was replicated twice and only measures for which inter-assay CV was lower than 10% were used.

Statistical Analysis

In order to examine the potential influence of T-levels and sexual orientation on men's speech, we performed four linear models, one for each acoustic parameter studied (i.e., F0, F0-SD, jitter, HNR). Each acoustic parameter was used as a response variable. To investigate the effects of sexual orientation and test the hypothesis of feminization on these vocal features, we used an explanatory variable called "SexOr" that considers both sex and sexual orientation with three modalities: heterosexual men, homosexual men and heterosexual women. T-level was added as another explanatory variable. To investigate whether T-level mediates the association between acoustic features and sexual orientation, we also added the interaction between T-level and "SexOr". When the interaction was not statistically significant, we removed it from the linear models. Age, monthly income, level of education and relationship status were added as confounding variables. All continuous variables (T-level, age, income and education) were standardized. Then, to assess if homosexual men displayed vocal features with values shifted towards those of heterosexual women, post-hoc analyses (Tukey HSD tests) were performed to compare which category (i.e., heterosexual men, homosexual men and heterosexual women) differ from one another. Thresholds of significance were corrected for the number of models and post-hoc comparisons using the Bonferroni method.

In order to assess the overall difference on speech acoustic features between heterosexual and homosexual men and to examine whether homosexual men's vocal features are shifted towards those of women, we conducted a linear discriminant analysis (LDA). LDA attempts to model whether a set of variables (here F0, F0-SD, Jitter, HNR) is effective in predicting category membership (here heterosexual men, homosexual men and heterosexual

women). Then, individuals' coordinates were computed from the two linear discriminant functions. Those coordinates were used to produce a continuous axis of vocal femininity and masculinity to determine where homosexual men were positioned within this axis. The coordinates of the three groups were then entered in linear models followed by post-hoc comparisons (Tukey HSD tests) to assess the overall difference in acoustic speech features.

All statistical analyses were performed under the R software (version 3.1.2).

Results

Descriptive statistics of all acoustic parameters and T-levels are given in Table 1 (all tables and figures are reported at the end of the article after the references).

The interactions between T-level and "SexOr" did not have a significant effect on mean F0 ($F(2, 151) = 2.06, p = .13$), F0-SD ($F(2, 151) = 0.07, p = .93$), HNR ($F(2, 151) = 0.09, p = .90$) and jitter ($F(2, 151) = 0.23, p = .79$). These interactions were thus subsequently removed from the linear models. "SexOr" showed a significant effect on mean F0 ($F(2, 153) = 194.44, p < .001$, Table 2), F0-SD ($F(2, 153) = 97.11, p < .001$, Table 3), jitter ($F(2, 153) = 6.97, p < .01$, Table 4) and HNR ($F(2, 153) = 55.78, p < .001$, Table 5). Post-hoc comparisons showed that all acoustic characteristics of heterosexual women were significantly different from both heterosexual and homosexual men (all $p < .01$). In addition, homosexual men displayed a significantly higher F0-SD and HNR than heterosexual men (respectively $t(158) = -4.46, p < .001$; $t(158) = 2.91, p < .05$), with values shifted towards those of heterosexual women (see Table 1). Mean F0 and jitter did not differ between homosexual and heterosexual men (respectively $t(158) = 0.98, p = .58$; $t(158) = 2.10, p = .09$). Age had a significant positive effect on F0-SD ($F(1, 153) = 7.75, p = .01$, Table 3). Other control variables had no influence on any of the acoustic features under study ($p > .05$, Table 2 to 5).

The LDA separated the three groups using two discriminant functions, each respectively explaining 97.72% and 2.28% of the between group variance. Coordinates were then computed

from these functions. Since the second function could not accurately discriminate between the two sexes using their coordinates ($F(1, 159) = 0.97, p = .32$), the overall acoustic difference between the three groups was assessed using the coordinates of the first function. Post-hoc comparisons revealed significant differences between heterosexual and homosexual men (mean difference = 0.74, $t = 3.86, p < .01$). Homosexual men showed a total of 11.20% differences in overall speech acoustic features compared to heterosexual men, slightly but significantly shifting towards those of heterosexual women (see Figure 1).

Discussion

This descriptive study offers an interesting take on the interaction between sexual orientation and acoustic features of speech in a French speaker sample. First, our analysis of different acoustic features revealed well-known patterns of sexual dimorphism in human voices (i.e., F0, F0-SD, jitter and HNR). Secondly, our findings showed that French homosexual men significantly displayed a more modulated and less breathy voice than French heterosexual men, thus supporting and extending previous studies conducted mostly with English speakers. Eventually, our results showed that French homosexual men attested a slight but significant vocal feminization (up to 11.20%), which support the feminization hypothesis (at least for specific vocal features). Lastly, testosterone levels did not mediate the association between vocal patterns and sexual orientation.

Consistent with previous findings in English speaking populations, no differences were observed in mean F0 between French speaking heterosexual and homosexual men (Gaudio, 1994; Lerman & Damsté, 1969; Munson, McDonald, et al., 2006; Rendall et al., 2008; Rogers et al., 2001; Smyth et al., 2003). Moreover, results showed a difference between homosexual and heterosexual men in intonation, the former displaying higher pitch variations than the latter. The relationship between pitch variations and sexual orientation was previously found in one Dutch (Baeck et al., 2011) and one American-English population (Gaudio, 1994), suggesting

that feminized pitch variations might be characteristic of male homosexual speech across languages (but see: Levon, 2006). In our study, the average difference in pitch variations reached to ~4.16 Hz, which is largely above the just noticeable difference for pitch (Pisanski & Rendall, 2011). Hence, our findings suggest that pitch variations could be one of the acoustic correlates of sexual orientation that is used by listeners when they correctly assessed sexual orientation through speech only (Gaudio, 1994; Linville, 1998; Smyth et al., 2003; Valentova & Havlíček, 2013). Further investigations are nevertheless needed to confirm if such a difference in pitch variations between homosexual and heterosexual men is enough to be used as a cue for assessing sexual orientation.

To our knowledge, this is the first study to report an association between men's vocal breathiness and sexual orientation. Interestingly, vocal breathiness has been suggested to be an important component of vocal femininity in female voices (Van Borsel et al., 2009) and significant relationships to vocal attractiveness have been reported in both sexes (Xu et al., 2013). Although the difference in vocal breathiness between homosexual and heterosexual men is rather low (mean average difference reached ~0.80 dB), further research should test whether it is perceptible by listeners to assess male sexual orientation and whether homosexual men's voices, which are richer in harmonics compared to those of heterosexuals, is perceived as more attractive among homosexual men.

In our study, T-levels did not influence any of the acoustic parameters investigated. The methods to measure T-level and the sample size used in this study are similar to those used in previous studies finding a significant negative link between T-levels and F0 (e.g., Dabbs & Mallinger, 1999; Evans et al., 2008). However, testosterone is a multiple-effect hormone under the influence of numerous biological and environmental factors and pathways. As such, it is generally difficult to correlate T-levels to other biological or behavioral traits especially with a unique measurement as realized here. Nevertheless, our results might suggest that other

underlying processes, different than basal T-level, are involved in vocal differences between homosexual and heterosexual men.

Although our study does not aim to provide an explanation for why vocal differences were found between homosexual and heterosexual men, several biological and social mechanisms can be invoked. For instance, exposure to prenatal testosterone has been suggested to be responsible for the differences between homosexual and heterosexual men on a large range of characteristics such as physiological and behavioral traits including speech characteristics (Balthazart, 2017; Ehrhardt & Meyer-Bahlburg, 1981). Several studies have thus tested whether the 2D:4D ratio (relative length of the 2nd and 4th digit), a proxy of testosterone prenatal exposure, differ between homosexual and heterosexual men (Balthazart, 2017; Ehrhardt & Meyer-Bahlburg, 1981). However, there is currently no consensus regarding whether the 2D:4D ratio differs between heterosexual and homosexual men as studies have yielded mixed results (e.g., Grimbos, Dawood, Burriss, Zucker, & Puts, 2010; Rahman & Wilson, 2003; Robinson, 2000; Williams et al., 2000). Regarding the social mechanisms, a social imitation of women's speech peculiarities by homosexual men could also explain the differences observed between homosexual and heterosexual men's speech characteristics (at least for F0-SD and HNR). The use of more feminine acoustic characteristics by homosexual men could reflect a selective adoption model of opposite-sex speech patterns or a selective use of acoustic features for signaling in-group identity (Pierrehumbert et al., 2004), an ability called "gaydar" (i.e., the detection of homosexuality based on a set of specific cues among which, speech). However, it has also been argued that such differences could have also arisen incidentally (e.g., due to variation of exposure to prenatal testosterone) rather than via explicit or implicit efforts to imitate opposite-sex speakers or other homosexuals (Rendall et al., 2008). Further work should investigate the relative importance of those different mechanisms for a deeper understanding of the feminization of homosexual men's speech.

To conclude, although our study did not aim to test specific hypotheses against a formal theoretical framework to understand the differences between homosexual and heterosexual men's speech, it provides some new descriptive findings. By examining for the first time native French speakers and some understudied acoustic features (i.e. namely, jitter and HNR), our results indicated that some vocal traits differed between heterosexual and homosexual men (i.e., variations of pitch and vocal breathiness) with values shifted towards heterosexual women's vocal characteristics. Combined with the literature conducted in other languages, our findings bring new support for the feminization hypothesis (at least for some acoustic features) and suggest that the feminization of some acoustic features could be shared across-languages. Further studies are needed to test whether intonation and vocal breathiness are perceptually salient to distinguish homosexual and heterosexual men, and whether overall differences are due to biological and/or sociolinguistic reasons.

Compliance with Ethical Standards

The French National Commission of Informatics and Liberties approved all protocols used in this study (CNIL number 1261003).

Conflicts of Interest

The authors have no conflicts of interest to declare.

References

- Alanko, K., Santtila, P., Harlaar, N., Witting, K., Varjonen, M., Jern, P., ... Sandnabba, N. K. (2010). Common Genetic Effects of Gender Atypical Behavior in Childhood and Sexual Orientation in Adulthood: A Study of Finnish Twins. *Archives of Sexual Behavior*, 39(1), 81–92. <https://doi.org/10.1007/s10508-008-9457-3>
- Ambady, N., Hallahan, M., & Conner, B. (1999). Accuracy of judgments of sexual orientation from thin slices of behavior. *Journal of Personality and Social Psychology*, 77(3), 538.
- Andreeva, B., Demenko, G., Wolska, M., Möbius, B., Zimmerer, F., Jügler, J., ... Trouvain, J.

- (2014). Comparison of Pitch Range and Pitch Variation in Slavic and Germanic Languages. In N. Campbell, D. Gibbon, & D. Hirst (Eds.), *Proceedings to the 7th Speech Prosody Conference*. (pp. 776–780). Dublin, Ireland: International Speech Communication Association.
- Apicella, C. L., Dreber, A., Campbell, B., Gray, P. B., Hoffman, M., & Little, A. C. (2008). Testosterone and financial risk preferences. *Evolution and Human Behavior*, *29*, 384–390.
- Archer, J. (2006). Testosterone and human aggression: an evaluation of the challenge hypothesis. *Neuroscience and Biobehavioral Reviews*, *30*, 319–345.
- Baeck, H., Corthals, P., & Borsel, J. V. (2011). Pitch Characteristics of Homosexual Males. *Journal of Voice*, *25*(5), e211–e214. <https://doi.org/10.1016/j.jvoice.2010.10.019>
- Bailey, J. M., & Zucker, K. J. (1995). Childhood sex-typed behavior and sexual orientation: A conceptual analysis and quantitative review. *Developmental Psychology*, *31*(1), 43.
- Balthazart, J. (2017). Genetic and Prenatal Components of Homosexuality. In T. K. Shackelford & V. A. Weekes-Shackelford (Eds.), *Encyclopedia of Evolutionary Psychological Science* (pp. 1–4). https://doi.org/10.1007/978-3-319-16999-6_3370-1
- Booth, A., Shelley, G., Mazur, A., Tharp, G., & Kittok, R. (1989). Testosterone, and Winning and Losing in Human Competition. *Hormones and Behavior*, *23*(4), 556–571.
- Brodie, H. K. H., Gartrell, N., Doering, C., & Rhue, T. (1974). Plasma testosterone levels in heterosexual and homosexual men. *American Journal of Psychiatry*, *131*(1), 82–82.
- Cartei, V., & Reby, D. (2012). Acting Gay: Male Actors Shift the Frequency Components of Their Voices Towards Female Values When Playing Homosexual Characters. *Journal of Nonverbal Behavior*, *36*(1), 79–93. <https://doi.org/10.1007/s10919-011-0123-4>
- Dabbs, J. M., & Mallinger, A. (1999). High testosterone levels predict low voice pitch among men. *Personality and Individual Differences*, *27*, 801–804.
- Ehrhardt, A. A., & Meyer-Bahlburg, H. F. L. (1981). Effects of Prenatal Sex Hormones on Gender-Related Behavior. *Science*, *211*(4488), 1312–1318.

- Ellison, P. T. (1988). Human Salivary Steroids - Methodological Considerations and Applications in Physical-Anthropology. *Yearbook of Physical Anthropology*, 31, 115–142.
- Evans, S., Neave, N., Wakelin, D., & Hamilton, C. (2008). The relationship between testosterone and vocal frequencies in human males. *Physiology & Behavior*, 93(4–5), 783–788. <https://doi.org/10.1016/j.physbeh.2007.11.033>
- Freeman, J. B., Johnson, K. L., Ambady, N., & Rule, N. O. (2010). Sexual Orientation Perception Involves Gendered Facial Cues. *Personality and Social Psychology Bulletin*, 36(10), 1318–1331. <https://doi.org/10.1177/0146167210378755>
- Gaudio, R. P. (1994). Sounding Gay: Pitch Properties in the Speech of Gay and Straight Men. *American Speech*, 69(1), 30. <https://doi.org/10.2307/455948>
- Graddol, D. (1989). *Gender voices*. Blackwell Publishing.
- Grimbos, T., Dawood, K., Burriss, R. P., Zucker, K. J., & Puts, D. A. (2010). Sexual orientation and the second to fourth finger length ratio: A meta-analysis in men and women. *Behavioral Neuroscience*, 124(2), 278–287. <https://doi.org/10.1037/a0018764>
- Hodges-Simeon, C. R., Gurven, M., & Gaulin, S. J. C. (2015). The low male voice is a costly signal of phenotypic quality among Bolivian adolescents. *Evolution and Human Behavior*, 36(4), 294–302. <https://doi.org/10.1016/j.evolhumbehav.2015.01.002>
- Hughes, S. M., Dispenza, F., & Gallup, G. G. (2004). Ratings of voice attractiveness predict sexual behavior and body configuration. *Evolution and Human Behavior*, 25(5), 295–304.
- Jaffee, W. L., McCormack, W. M., & Vaitukaitis, J. L. (1980). Plasma hormones and the sexual preferences of men. *Psychoneuroendocrinology*, 5(1), 33–38. [https://doi.org/10.1016/0306-4530\(80\)90006-2](https://doi.org/10.1016/0306-4530(80)90006-2)
- Keating, P., & Kuo, G. (2012). Comparison of speaking fundamental frequency in English and Mandarin. *The Journal of the Acoustical Society of America*, 132(2), 1050–1060.
- Kolodny, R. C., Masters, W. H., Hendryx, J., & Toro, G. (1971). Plasma testosterone and semen

- analysis in male homosexuals. *New England Journal of Medicine*, 285(21), 1170–1174.
- Laan, G. P. (1997). The contribution of intonation, segmental durations, and spectral features to the perception of a spontaneous and a read speaking style. *Speech Communication*, 22(1), 43–65.
- Lerman, J. W., & Damsté, P. H. (1969). Voice pitch of homosexuals. *Folia Phoniatica et Logopaedica*, 21(5), 340–346.
- LeVay, S. (1991). A Difference in Hypothalamic Structure between Heterosexual and Homosexual Men. *Science*, 253(5023), 1034–1037.
- Levon, E. (2006). Hearing “gay”: Prosody, interpretation, and the affective judgments of men’s speech. *American Speech*, 81(1), 56–78.
- Linville, S. E. (1998). Acoustic correlates of perceived versus actual sexual orientation in men’s speech. *Folia Phoniatica et Logopaedica*, 50(1), 35–48.
- Lyons, M., Lynch, A., Brewer, G., & Bruno, D. (2014). Detection of Sexual Orientation (“Gaydar”) by Homosexual and Heterosexual Women. *Archives of Sexual Behavior*, 43(2), 345–352. <https://doi.org/10.1007/s10508-013-0144-7>
- Mack, S., & Munson, B. (2012). The influence of /s/ quality on ratings of men’s sexual orientation: Explicit and implicit measures of the ‘gay lisp’ stereotype. *Journal of Phonetics*, 40(1), 198–212. <https://doi.org/10.1016/j.wocn.2011.10.002>
- Munson, B., & Babel, M. (2007). Loose Lips and Silver Tongues, or, Projecting Sexual Orientation Through Speech. *Language and Linguistics Compass*, 1(5), 416–449. <https://doi.org/10.1111/j.1749-818X.2007.00028.x>
- Munson, B., Jefferson, S. V., & McDonald, E. C. (2006). The influence of perceived sexual orientation on fricative identification. *The Journal of the Acoustical Society of America*, 119(4), 2427–2437.
- Munson, B., McDonald, E. C., DeBoe, N. L., & White, A. R. (2006). The acoustic and

perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics*, 34(2), 202–240. <https://doi.org/10.1016/j.wocn.2005.05.003>

Neave, N., Menaged, M., & Weightman, D. R. (1999). Sex Differences in Cognition: The Role of Testosterone and Sexual Orientation. *Brain and Cognition*, 41(3), 245–262. <https://doi.org/10.1006/brcg.1999.1125>

Ordin, M., & Mennen, I. (2017). Cross-linguistic differences in bilinguals' fundamental frequency ranges. *Journal of Speech, Language, and Hearing Research*, 60(6), 1493–1506.

Penton-Voak, I. S., & Chen, J. Y. (2004). High salivary testosterone is linked to masculine male facial appearance in humans. *Evolution and Human Behavior*, 25(4), 229–241. <https://doi.org/10.1016/j.evolhumbehav.2004.04.003>

Pierrehumbert, J. B., Bent, T., Munson, B., Bradlow, A. R., & Bailey, J. M. (2004). The influence of sexual orientation on vowel production (L). *The Journal of the Acoustical Society of America*, 116(4), 1905–1908. <https://doi.org/10.1121/1.1788729>

Pillard, R. C., Rose, R. M., & Sherwood, M. (1974). Plasma testosterone levels in homosexual men. *Archives of Sexual Behavior*, 3(5), 453–458. <https://doi.org/10.1007/BF01541165>

Pisanski, K., Jones, B. C., Fink, B., O'Connor, J. J., DeBruine, L. M., Röder, S., & Feinberg, D. R. (2016). Voice parameters predict sex-specific body morphology in men and women. *Animal Behaviour*, 112, 13–22.

Pisanski, K., & Rendall, D. (2011). The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. *The Journal of the Acoustical Society of America*, 129(4), 2201–2212.

Pound, N., Penton-Voak, I. S., & SurrIDGE, A. K. (2009). Testosterone responses to competition in men are related to facial masculinity. *Proceedings of the Royal Society B: Biological Sciences*, 276(1654), 153–159. <https://doi.org/10.1098/rspb.2008.0990>

Puts, D. A., Apicella, C. L., & Cardenas, R. A. (2012). Masculine voices signal men's threat

- potential in forager and industrial societies. *Proceedings of the Royal Society B: Biological Sciences*, 279(1728), 601–609. <https://doi.org/10.1098/rspb.2011.0829>
- Rabinov, C. R., Kreiman, J., Gerratt, B. R., & Bielamowicz, S. (1995). Comparing reliability of perceptual ratings of roughness and acoustic measures of jitter. *Journal of Speech, Language, and Hearing Research*, 38(1), 26–32.
- Rahman, Q., & Wilson, G. D. (2003). Sexual orientation and the 2nd to 4th finger length ratio: evidence for organising effects of sex hormones or developmental instability? *Psychoneuroendocrinology*, 28(3), 288–303. [https://doi.org/10.1016/S0306-4530\(02\)00022-7](https://doi.org/10.1016/S0306-4530(02)00022-7)
- Rendall, D., Vasey, P. L., & McKenzie, J. (2008). The Queen’s English: An Alternative, Biosocial Hypothesis for the Distinctive Features of “Gay Speech.” *Archives of Sexual Behavior*, 37(1), 188–204. <https://doi.org/10.1007/s10508-007-9269-x>
- Rieger, G., Linsenmeier, J. A. W., Gygax, L., & Bailey, J. M. (2008). Sexual orientation and childhood gender nonconformity: Evidence from home videos. *Developmental Psychology*, 44(1), 46–58. <https://doi.org/10.1037/0012-1649.44.1.46>
- Rieger, G., Linsenmeier, J. A. W., Gygax, L., Garcia, S., & Bailey, J. M. (2010). Dissecting “Gaydar”: Accuracy and the Role of Masculinity–Femininity. *Archives of Sexual Behavior*, 39(1), 124–140. <https://doi.org/10.1007/s10508-008-9405-2>
- Robinson, S. (2000). The ratio of 2nd to 4th digit length and male homosexuality. *Evolution and Human Behavior*, 21(5), 333–345. [https://doi.org/10.1016/S1090-5138\(00\)00052-0](https://doi.org/10.1016/S1090-5138(00)00052-0)
- Rogers, H., Jacobs, G., & Smyth, R. (2001). Searching for phonetic correlates of gay- and straight-sounding voices. *Toronto Working Papers in Linguistics*, 18, 46–64.
- Roney, J. R., Hanson, K. N., Durante, K. M., & Maestripieri, D. (2006). Reading men’s faces: women’s mate attractiveness judgments track men’s testosterone and interest in infants. *Proceedings of the Royal Society B: Biological Sciences*, 273(1598), 2169–2175. <https://doi.org/10.1098/rspb.2006.3569>

- Savic, I., Berglund, H., & Lindstrom, P. (2005). Brain response to putative pheromones in homosexual men. *Proceedings of the National Academy of Sciences*, *102*(20), 7356–7361. <https://doi.org/10.1073/pnas.0407998102>
- Sergeant, M. J. T., Dickins, T. E., Davies, M. N. O., & Griffiths, M. D. (2007). Women's Hedonic Ratings of Body Odor of Heterosexual and Homosexual Men. *Archives of Sexual Behavior*, *36*(3), 395–401. <https://doi.org/10.1007/s10508-006-9126-3>
- Smyth, R., Jacobs, G., & Rogers, H. (2003). Male voices and perceived sexual orientation: An experimental and theoretical approach. *Language in Society*, *32*(3), 329–350. <https://doi.org/10.1017/S0047404503323024>
- Suire, A., Raymond, M., & Barkat-Defradas, M. (2018). Human vocal behavior within competitive and courtship contexts and its relation to mating success. *Evolution and Human Behavior*, *39*(6), 684–691. <https://doi.org/10.1016/j.evolhumbehav.2018.07.001>
- Teixeira, J. P., Oliveira, C., & Lopes, C. (2013). Vocal acoustic analysis—jitter, shimmer and hnr parameters. *Procedia Technology*, *9*, 1112–1122.
- Traunmüller, H., & Eriksson, A. (1995). The frequency range of the voice fundamental in the speech of male and female adults. *Unpublished Manuscript*.
- Valentova, J. V., & Havlíček, J. (2013). Perceived Sexual Orientation Based on Vocal and Facial Stimuli Is Linked to Self-Rated Sexual Orientation in Czech Men. *PLoS ONE*, *8*(12), e82417. <https://doi.org/10.1371/journal.pone.0082417>
- Valentova, J. V., Rieger, G., Havlicek, J., Linsenmeier, J. A. W., & Bailey, J. M. (2011). Judgments of Sexual Orientation and Masculinity–Femininity Based on Thin Slices of Behavior: A Cross-Cultural Comparison. *Archives of Sexual Behavior*, *40*(6), 1145–1152. <https://doi.org/10.1007/s10508-011-9818-1>
- Van Borsel, J., Janssens, J., & De Bodt, M. (2009). Breathiness as a Feminine Voice Characteristic: A Perceptual Approach. *Journal of Voice*, *23*(3), 291–294.

<https://doi.org/10.1016/j.jvoice.2007.08.002>

Williams, T. J., Pepitone, M. E., Christensen, S. E., Cooke, B. M., Huberman, A. D., Breedlove, N. J., ... Breedlove, S. M. (2000). Finger-length ratios and sexual orientation. *Nature*, *404*(6777), 455–456. <https://doi.org/10.1038/35006555>

Xu, Y., Lee, A., Wu, W.-L., Liu, X., & Birkholz, P. (2013). Human Vocal Attractiveness as Signaled by Body Size Projection. *PLoS ONE*, *8*(4), e62397. <https://doi.org/10.1371/journal.pone.0062397>

Xu, Y., Norton, S., & Rahman, Q. (2017). Sexual orientation and neurocognitive ability: A meta-analysis in men and women. *Neuroscience & Biobehavioral Reviews*, *83*, 691–696. <https://doi.org/10.1016/j.neubiorev.2017.06.014>

Zimmerer, F., Jügler, J., Andreeva, B., Möbius, B., & Trouvain, J. (2014). Too cautious to vary more? A comparison of pitch variation in native and non-native productions of French and German speakers. In N. Campbell, D. Gibbon, & D. Hirst (Eds.), *Proceedings to the 7th Speech Prosody Conference*. (pp. 1037–1041). Dublin, Ireland: International Speech Communication Association.

Table 1: Descriptive statistics of mean F0, F0-SD, jitter, HNR, speaking time and T-levels for heterosexual men and women and homosexual men. Values reported are mean \pm standard-deviation.

	Heterosexual men (n = 49)	Homosexual men (n = 58)	Heterosexual women (n = 54)
<i>Mean F0 (Hz)</i>	120.13 \pm 19.68	116.52 \pm 13.91	205.89 \pm 17.58
<i>F0-SD (Hz)</i>	14.06 \pm 4.39	18.22 \pm 3.88	32.35 \pm 6.60
<i>Jitter (%)</i>	2.84 \pm 1.63	2.43 \pm 0.57	1.81 \pm 0.22
<i>HNR (dB)</i>	10.07 \pm 1.41	10.86 \pm 1.31	13.68 \pm 1.12
<i>Speaking time (s)</i>	69.24 \pm 30.26	63.49 \pm 33.41	67.58 \pm 1.12
<i>T-levels (pg/ml)</i>	137.71 \pm 61.32	136.65 \pm 53.50	33.08 \pm 21.24

Table 2: Linear model examining the influence of sexual orientation and sex on mean F0¹.

<i>(R² = 85.2%)</i>	β	SE	<i>F</i>	<i>p</i>
Intercept	117.41	2.94		
SexOr			194.44	<.001
Heterosexual men / <u>Homosexual men</u>	3.38	3.44		
Heterosexual women / <u>Homosexual men</u>	85.28	4.57		
Age	-0.24	0.29	0.66	.42
Testosterone	-0.03	0.03	1.46	.22
Monthly income	0.18	0.66	0.07	.78
Education	-0.01	0.43	0.001	.97
Relationship status			0.11	.73
Yes/ <u>No</u>	0.95	2.80		

¹For each variable, the estimate (β), standard error of the mean (SE), the *F* and the *p* values associated from the Fisher test of the comparison between the full model and the model without the factor are given. For the categorical variables “SexOr” and “Relationship status”, the estimates are given for one category compared to the reference category (underlined terms). *R*² is the variance explained by the model. Sample size: *N*_{Heterosexual men} = 49; *N*_{Homosexual men} = 58, *N*_{women} = 54.

Table 3: Linear model examining the influence of sexual orientation and sex on F0-SD₁.

<i>R</i> ² = 71.5%	β	SE	<i>F</i>	<i>p</i>
Intercept	17.67	0.84		
SexOr			97.11	<.001
Heterosexual men / <u>Homosexual men</u>	-4.45	0.99		
Heterosexual women / <u>Homosexual men</u>	13.88	1.31		
Age	0.24	0.08	7.75	<.01
Testosterone	<-0.01	0.01	0.18	.66
Monthly income	0.17	0.19	0.83	.36
Education	0.05	0.12	0.17	.68
Relationship status			2.48	.11
Yes/ <u>No</u>	1.27	0.80		

Table 4: Linear model examining the influence of sexual orientation and sex on jitter₁.

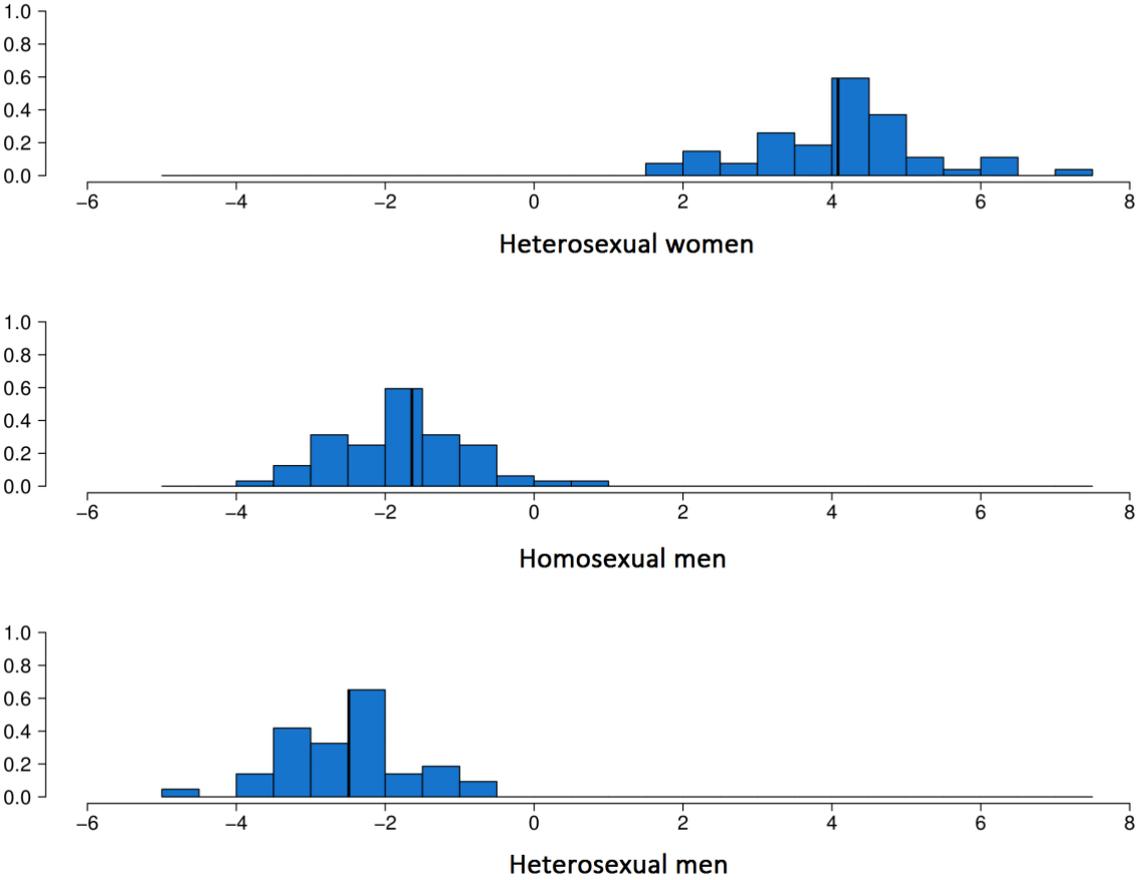
$R^2 = 14.8\%$	β	SE	F	p
Intercept	2.52	0.16		
SexOr			6.97	< .01
Heterosexual men / <u>Homosexual men</u>	0.41	0.19		
Heterosexual women / <u>Homosexual men</u>	-0.53	0.26		
Age	0.02	0.01	1.40	.23
Testosterone	<0.001	0.001	0.22	.63
Monthly income	-0.03	0.04	0.90	.34
Education	0.03	0.02	1.61	.20
Relationship status			1.66	.19
Yes/ <u>No</u>	-0.20	0.16		

Table 5: Linear model examining the influence of sexual orientation and sex on HNR₁.

$R^2 = 58.7\%$	β	SE	F	p
Intercept	10.80	0.22		
SexOr			55.78	<.001
Heterosexual men / <u>Homosexual men</u>	-0.74	0.25		
Heterosexual women / <u>Homosexual men</u>	2.84	0.34		
Age	-0.04	0.02	3.32	.07
Testosterone	<-0.001	<0.01	0.02	.88
Monthly income	0.02	0.05	0.16	.68
Education	-0.04	0.03	1.52	.21
Relationship status			0.11	.73
Yes/ <u>No</u>	0.07	0.21		

Caption

Figure 1: Distributions' histograms of heterosexual women and homosexual and heterosexual men computed from the coordinates of the first linear discriminant function. The Y axis represents the frequency and the X axis the coordinates. Vertical solid lines represent the mean of each group ($N_{\text{Heterosexual men}} = 49$; $N_{\text{Homosexual men}} = 58$, $N_{\text{women}} = 54$).



VI. Succès copulatoire et reproducteur

Après avoir mis en évidence les préférences pour certains aspects de la qualité vocale dans les deux sexes et présentés leurs différentes interprétations évolutives, nous discutons dans ce chapitre les avantages directs (i.e. en termes de valeur sélective) que procurent les qualités vocales jugées comme attractives. En effet, nous avons montré d'une part que les femmes préfèrent les hommes aux hauteurs vocales graves (bas F0) avec une intonation relativement marquée (haut F0-SD), et d'autre part que les hommes préfèrent les femmes aux hauteurs vocales graves (bas F0) avec une intonation relativement marquée (haut F0-SD), aux timbres aiguës (haut Df) et présentant plus de souffle et de raucité vocales. De plus, indépendamment du sexe, nous avons aussi suggéré que les individus s'exprimant à des débits relativement élevés devraient être perçus comme plus attractifs. Dans ce contexte, nous pouvons penser que les individus exhibant ces caractéristiques vocales jouissent d'un plus grand accès aux partenaires sexuels (i.e. « *succès copulatoire* » ou « *mating success* » en anglais), ce qui peut se traduire par un nombre plus élevé d'enfants (i.e. « *succès reproducteur* » ou « *reproductive success* »). Indépendamment du succès copulatoire, les individus exhibant ces qualités vocales permettraient aussi aux autres individus aussi de mieux choisir un partenaire de qualité, ce qui peut augmenter la qualité phénotypique des enfants indépendamment de leur nombre. De manière générale, un petit nombre d'études a ainsi tenté d'établir des liens directs entre qualité vocale et succès copulatoire/reproducteur dans deux environnements, c'est-à-dire auprès de populations industrialisées et non-industrialisées.

VI.1. Populations non-industrialisées

La première étude ayant réalisé une quantification du lien entre la hauteur vocale et le succès reproducteur est celle de Apicella, Feinberg & Marlowe (2007). Chez les Hadza, une population de chasseurs-cueilleurs de Tanzanie, il a été montré que le F0 prédit de manière négative le nombre d'enfants, en contrôlant pour l'âge des sujets ; ces deux facteurs expliquant

à eux seuls près de 50% de la variance du succès reproducteur. À l'inverse, aucune relation n'a été trouvée chez les sujets de sexe féminin. Toutefois, ces résultats sont à prendre avec précaution. En effet, il a récemment été montré que lorsque l'on contrôle pour la réputation en tant que chasseur – trait particulièrement important dans le contexte des populations non-industrialisées car directement lié au nombre de femmes et d'enfants pour un homme – le F0 ne permet plus de prédire efficacement le succès reproducteur des individus (Smith, Olkhov, Puts, & Apicella, 2017). Chez les Himba, une population semi-nomade pastoraliste vivant au nord de la Namibie, il a été observé que la hauteur vocale des femmes permet de prédire le succès reproducteur : les femmes ayant des voix relativement hautes (haut F0, supposément liée à une plus grande fertilité) ont plus d'enfants et de petits-enfants. Toutefois, la relation entre hauteur vocale et succès reproducteur n'a pas été observée chez les hommes appartenant à la même ethnie (Atkinson, Pipitone, Sorokowska, Sorokowski, Mberira, Bartels, Gallup, & O'Rourke, 2012). Enfin, chez les Tsimané, une population de glaneurs-horticulteurs vivant en Bolivie, il a été montré que la hauteur vocale des voix masculines permet de prédire le nombre d'enfants et le succès reproducteur des femmes (Rosenfield, Sorokowska, Sorokowski, & Puts, 2019). Enfin, les auteurs de cette étude ont montré que les hommes ayant des voix relativement basses semblent avoir un succès reproducteur plus élevé s'ils effectuent des mariages avec des femmes ayant des grossesses rapprochées.

VI.2. Populations industrialisées

Dans les populations industrialisées, les études ont tenté d'établir des relations entre la qualité vocale et le succès copulatoire plutôt que reproducteur. Ces travaux ont pour la plupart été conduits auprès de populations d'étudiants et en contexte anglophone.

Dans ces travaux, le succès copulatoire est généralement mesuré en utilisant l'Inventaire de l'Orientation Sociosexuelle, conçu pour mesurer les différences interindividuelles en termes de relations sexuelles occasionnelles et sans engagements (Penke & Asendorpf, 2008). Afin de

mesurer le succès copulatoire, la question d'intérêt est celle du nombre de partenaires sexuels au cours des douze derniers mois. Cette méthodologie a été utilisée par Puts et al. (2005/2006) pour mesurer le succès copulatoire des hommes via un protocole pensé comme plus écologique que les études menées en laboratoire. Les auteurs ont en effet acquis leurs données par le biais d'un (faux) speed-dating. Dans la première étude, un participant masculin observe la vidéo d'une femme pour laquelle on lui indique qu'il est en compétition avec d'autres hommes pour obtenir un rendez-vous. Il doit ensuite produire une réponse à l'attention de la femme, en expliquant en quoi elle aurait tout intérêt à le choisir lui plutôt que son rival. Les auteurs trouvent une relation négative – quoique modeste (-0.17) – entre la hauteur vocale et le succès copulatoire (i.e. nombre de partenaire déclaré au cours des douze derniers mois). Dans la seconde étude, les auteurs tentent de voir si la qualité vocale en situation de compétition permet aussi de prédire le succès copulatoire des sujets. Les participants masculins à cette étude ont produit un enregistrement vocal en situation de compétition après avoir écouté le stimuli produit par un pseudo-rival. Les auteurs ont trouvé une corrélation négative (bien que non significative) entre le F0 et le succès copulatoire.

Plus tard, Hodges-Simeon et al. (2011) se sont attachés à répliquer le même protocole en tentant cette fois-ci de comprendre si le succès copulatoire est plutôt prédit par des discours produits en situation de compétition ou en situation de séduction. L'idée est de produire deux modèles pour prédire le succès copulatoire, soit par la qualité vocale produite au cours de l'enregistrement de séduction, soit par l'enregistrement obtenu à l'issue de la situation de compétition, et de comparer les deux approches. Les résultats montrent que les variations de l'intonation (F0-SD) observées dans les stimuli produits en situation de compétition prédisent le succès copulatoire des locuteurs : les hommes présentant des voix plutôt monotones (bas F0-SD) déclarent plus de partenaires sexuels que ceux dont les voix sont plus mélodieuses. En revanche, la hauteur vocale et la distribution formantique (Df) mesurée dans les échantillons

vocaux extraits des situations de compétition n'expliquent pas la variance du succès copulatoire, tout comme la qualité vocale produite en situation de séduction. Les auteurs ont ainsi conclu que la sélection intrasexuelle est plus forte que la sélection intersexuelle chez les hommes, et que cette première est encore effective dans les populations industrialisées.

Pour finir, soulignons l'étude de Hughes, Dispenza & Gallup (2004). Les auteurs ont enregistré des participants hommes et femmes produisant des stimuli vocaux issus d'une récitation de nombres et qu'ils ont ensuite fait évalués en termes d'attractivité par des juges des deux sexes. Les auteurs ont montré que l'attractivité vocale est négativement corrélée à l'âge déclaré du premier rapport sexuel mais positivement corrélé avec le nombre de partenaire sexuels, le nombre de copulations extra-conjugales et le nombre de copulations avec des partenaires engagés dans une autre relation. En outre, lorsque l'attractivité vocale est évaluée par les membres du même sexe, elle n'est positivement corrélée qu'au nombre de partenaires sexuels et au nombre de copulations extra-conjugales chez les hommes ; tandis que chez les femmes, l'attractivité vocale s'avère être positivement corrélée au nombre de partenaires sexuels et au nombre de rapports sexuels avec des individus en couple.

Indiquons toutefois que les études fondées sur les déclarations des sujets peuvent être biaisées car celles-ci peuvent être sujettes à l'oubli voire au mensonge, en particulier dans des sociétés où les stéréotypes et les attentes sociales relatives au couple sont très fortes. Par ailleurs, les individus – notamment les hommes – ont tendance à rapporter des chiffres ne reflétant pas la réalité. Il est en effet plutôt bien établi que les hommes ont tendances à surévaluer le nombre de partenaire sexuels, tandis que les femmes le sous-évaluent (pour une référence récente sur ces aspects, consulter Mitchell, Mercer, Prah, Clifton, Tanton, Wellings, & Copas, 2019). Dans les faits, le nombre de partenaire moyen par individu devrait être égal entre les deux sexes.

VI.3. Conclusion

En dressant un rapide portrait des études en lien entre la qualité vocale et le succès copulatoire/reproducteur, il apparaît d'une part que les résultats ne sont pas similaires pour les deux sexes au sein et entre les populations industrialisées et non-industrialisées. Aussi, nous remarquons une fois de plus qu'une attention particulière a été attribuée au critère de la hauteur vocale et du timbre, alors même que d'autres aspects de la qualité vocale pourraient être liés au succès copulatoire et reproducteur des sujets, mis en évidence par l'étude des préférences vocales.

Une critique supplémentaire peut être émise au sujet des études menés dans des populations industrialisés, ce qui nous permet de formuler une nouvelle perspective de recherche (Hodges-Simeon et al., 2011 ; Puts et al., 2005/2006). En effet, seuls deux stimuli ont été enregistrés auprès des participants, une femme « *cible* » et un homme « *compétiteur* », ces deux personnages étant joués par des étudiants. Or, les préférences interindividuelles dans le choix d'un partenaire peuvent être très variables. Logiquement, au sein de chaque sexe, les individus ne sont pas tous attirés pour les mêmes raisons par un partenaire, ce qui peut grandement affecter la qualité vocale et la manière de se comporter (Hughes et al., 2014 ; Pisanski et al., 2018). Il apparaît donc clairement que les conditions expérimentales mises en place en laboratoire sont beaucoup trop éloignées des situations écologiques où s'effectue la recherche de partenaires sexuels. Le caractère peu naturel des interactions créées à des fins scientifiques affecte très probablement la qualité vocale. Néanmoins, les résultats observés au niveau des différences interindividuelles démontrent l'existence d'un phénomène de « *modulation vocale* », défini comme la capacité d'un individu à modifier sa voix en fonction de son interlocuteur et plus généralement selon la situation sociale dans lequel il se trouve (e.g. compétition, séduction). Il apparaît donc comme particulièrement stimulant de comprendre dans quels contextes et de quelles manières les individus modulent leur voix. Certains auteurs

ont suggéré que cette capacité puisse résulter de l'effet de la sélection sexuelle (Pisanski, Cartei, McGettigan, Raine, & Reby, 2016), et c'est cet aspect que nous développerons dans le chapitre suivant.

Article 4

**Human vocal behavior within competitive and courtship contexts and
its relation to mating success**

Alexandre Suire, Michel Raymond & Melissa Barkat-Defradas

Article publié dans *Evolution and Human Behavior*

Résumé

Hommes et femmes semblent exhiber des préférences particulières pour certaines qualités vocales, même si un consensus général n'existe pas. Nos deux études sur les préférences ont montré d'une part que les femmes préféraient des voix masculines exhibant de bas F0 et de haut F0-SD et, d'autre part, que les hommes préféraient des voix féminines exhibant des débits plus rapides, de bas F0 et de haut F0-SD, jitter et HNR. Les individus exhibant de telles qualités vocales devraient ainsi voir augmenter leur probabilité d'attirer plus de partenaires sexuels, toutes choses étant égales par ailleurs. Enfin, s'il existe peu d'études sur le succès copulatoire et la voix (notamment pour les femmes), il est aussi intéressant de voir si ce lien est prédit par la compétition intra- ou intersexuelle, du moins dans nos sociétés modernes.

Afin d'explorer ces hypothèses, nous avons recruté 58 hommes et 68 femmes et les avons fait participer à un jeu de speed-dating « *simulé* », dont le protocole est identique à celui de Hodges-Simeon et al. 2011 et Puts et al. 2005/2006. La tâche consistait à gagner un rendez-vous galant avec une personne du sexe opposé, tout en étant en compétition avec une autre personne du même sexe. Il a été demandé aux participants de séduire une personne du sexe opposé pour un rendez-vous galant après avoir visionné une vidéo où un acteur et/ou une actrice se présentait (i.e. enregistrement de « *séduction* »). Il a aussi été demandé aux participants d'expliquer en quoi ils seraient plus à même de gagner ce rendez-vous qu'un autre compétiteur après l'écoute d'un enregistrement audio compétitif de l'acteur/actrice (i.e. enregistrement de « *compétition* »). Les participants ont ensuite complété un questionnaire où ils devaient indiquer en particulier leur âge, s'ils étaient engagés dans une relation, et le nombre de partenaires sexuels durant les 12 derniers mois. Les enregistrements des participants hommes et femmes lors des conditions de « *compétition* » et de « *séduction* » ont ensuite été analysés pour les composantes acoustiques suivantes : le F0, le F0-SD, le jitter, le HNR, le débit (nombre

de syllabes par secondes en incluant les pauses), le taux d'articulation (nombre de syllabes par secondes en excluant les pauses), la durée et l'intensité.

Nos résultats ont montré que la qualité vocale en compétition ne prédit pas le succès copulatoire dans les deux sexes, au contraire de celle produite en séduction¹. Les hommes exhibant un taux d'articulation plus élevés et des voix plus intenses ont déclaré plus de partenaires sexuels, au contraire de ceux avec une intonation élevée (haut F0-SD), dont l'effet de leur souffle et de leur raucité (bas HNR et haut jitter) était d'autant plus négatif sur leur succès copulatoire. Concernant les femmes, celles exhibant moins de souffles (haut HNR) et parlant moins ont déclaré moins de partenaires sexuels. Ces résultats suggèrent que la qualité vocale est surtout fonctionnelle en situation intersexuelle, possiblement due au fait que les sociétés industrialisées ne reposent plus sur une confrontation intrasexuelle directe pour accéder à un partenaire. Toutefois, ces résultats ne soutiennent pas ceux présentés dans les études sur les préférences. De plus, de nombreux aspects du protocole sont critiquables, notamment la production vocale en situation de laboratoire et le déclaratif qui n'est pas nécessairement véridique.

¹Notons que deux participants hommes ont été exclus de l'analyse due à une déclaration significativement élevée de partenaires sexuels par rapport à la moyenne, et n'ont malheureusement pas été reportés dans l'article. Toutefois, l'ajout ou l'exclusion de ces deux individus n'a pas changé qualitativement les résultats.



Contents lists available at ScienceDirect

Evolution and Human Behavior

journal homepage: www.elsevier.com/locate/ens

Human vocal behavior within competitive and courtship contexts and its relation to mating success

Alexandre Suire*, Michel Raymond, Melissa Barkat-Defradas

Institut des Sciences de l'Evolution de Montpellier, University of Montpellier, Centre National de la Recherche Scientifique, Institut pour la Recherche et le Développement, Ecole Pratique des Hautes Etudes – Place Eugène Bataillon, 34095 Montpellier, France

ARTICLE INFO

Keywords:

Sexual selection
Human voice
Vocal behavior
Mating success
Courtship
Competition

ABSTRACT

Beyond the linguistic content of their speech, speakers of both sexes convey diverse biological and psychosocial information through their voices, which are important when assessing potential mates and competitors. However, studies investigating the relationships between mating success and acoustic inter-individual differences are scarce. In this study, we investigated such relationships in both sexes in courtship and competitive interactions—as they correspond to the two different types of sexual selection—using an experimental design based on a simulated dating game. We assessed which type of sexual selection best predicted mating success, here defined as the self-reported number of sexual partners within the past year. Our results show that only acoustic inter-individual differences in the courtship context for both men and women predicted their mating success. Men displaying faster articulation rate and louder voices reported significantly more sexual partners; in contrast, men displaying higher intonation reported a greater negative effect of roughness and breathiness on their mating success. Women who displayed relatively less breathy voices and shorter speech duration reported significantly fewer sexual partners. These novel findings are discussed in light of the mate choice context and the relative contribution of both types of sexual selection shaping acoustic features of speech.

1. Introduction

Beyond the linguistic content it conveys, voice is one of the fundamental aspects of human communication, as it enables the expression of a wide range of emotional and affective states.

A large body of work has shown that voices convey an array of different biological and social information such as sex (Puts, Gaulin, & Verdolini, 2006), age (Linville & Fisher, 1985; Ptacek & Sander, 1966), sexual orientation (Lyons, Lynch, Brewer, & Bruno, 2014; Munson, McDonald, DeBoe, & White, 2006), hormone levels (Dabbs Jr & Mallinger, 1999), physical strength (Sell et al., 2010), body configuration (Hughes, Dispenza, & Gallup, 2004) and social status (Cheng, Tracy, Ho, & Henrich, 2016). Such information is crucial to assess potential mates and competitors, as it reflects indexical cues of attractiveness, dominance, masculinity and femininity or impressions of size-related features (Puts, Jones, & DeBruine, 2012). For instance, men exhibiting a relatively lower voice pitch are perceived as more attractive by members of the opposite sex and more dominant by same-sex individuals (Collins, 2000; Puts et al., 2006; Puts, Hodges, Cárdenas, & Gaulin, 2007). Some studies have found that women with a relatively

higher voice pitch are perceived as more attractive by men while being judged as more promiscuous by other women (Collins & Missing, 2003; Puts, Barndt, Welling, Dawood, & Burriss, 2011), although other evidence suggests the opposite relationship (Hughes, Farley, & Rhodes, 2010; Hughes, Mogilski, & Harrison, 2014; Tuomi & Fisher, 1979).

In addition to attractiveness and dominance, several other auditory impressions can be conveyed through voice. For instance, one study from O'Connor, Re, and Feinberg (2011) showed that the perception of infidelity risk increases with more feminine voices in women (i.e., relatively higher voice pitch) and more masculine voices in men (i.e., relatively lower voice pitch). Voice can also play upon perceptions of cooperativeness, where a higher voice pitch is associated with increased perception of cooperation (Knowles & Little, 2016) and trustworthiness, and individuals with lower voice pitch are preferentially selected when judges are asked to pick a leader (Tigue, Borak, O'Connor, Schandl, & Feinberg, 2012). However, other evidence has shown the opposite relationship (Montano, Tigue, Isenstein, Barclay, & Feinberg, 2017; Oleszkiewicz, Pisanski, Lachowicz-Tabaczek, & Sorokowska, 2017). Concerning impressions of the size-related feature, listeners regularly associate deeper voice pitch to larger and taller individuals, and

* Corresponding author.

E-mail addresses: alexandre.suire@umontpellier.fr (A. Suire), michel.raymond@umontpellier.fr (M. Raymond), melissa.barkat-defradas@umontpellier.fr (M. Barkat-Defradas).

<https://doi.org/10.1016/j.evolhumbehav.2018.07.001>

Received 31 January 2018; Received in revised form 27 June 2018; Accepted 9 July 2018

1090-5138/© 2018 Elsevier Inc. All rights reserved.

conversely, higher voice pitch to thinner and smaller individuals (Pisanski & Rendall, 2011; Rendall, Vokey, & Nemeth, 2007); despite this, vocal features explain little variation in body size (Pisanski et al., 2014). Although voice pitch and its resonant frequencies signal little information on the latter, it has been suggested that they could rather be a reliable signal of hormonal quality; e.g., lower voice pitch in men correlated to circulating testosterone levels (Evans, Neave, Wakelin, & Hamilton, 2008) or a signal of fertility in women (Feinberg et al., 2005; Pisanski, Bhardwaj, & Reby, 2018).

Morton (1977) first suggested that in many birds and mammal species, a common structural convergence of acoustic features exists where low-frequency sounds are used in “hostile” and “harsh” contexts; conversely, higher-frequency sounds are used when “frightened” or “approaching in a friendly manner”. Later, Ohala (1983) built upon this idea by suggesting that cross-language patterns in the use of vocal height in natural languages and vocalizations in other species serve purposes of threatening (or not) conspecifics. Listeners associate higher acoustic frequencies with the impression of a “small vocalizer” perceived as subordinate, submissive, non-threatening and lower acoustic frequencies with a “large vocalizer” perceived as dominant, aggressive and threatening. This theoretical framework is supported from recent comparative approaches studying dimorphism in the vocalizations of mammals, which is of great importance in the context of conspecific encounters within intra-sexual competition (Bowling et al., 2017; Charlton & Reby, 2016; Puts et al., 2016). Indeed, many males of mammal species use their vocalizations to assess and repel competitors, with deeper frequencies associated with higher mating and reproductive success. For humans, it is commonly accepted that vocal sexual dimorphism, which is due to anatomical and physiological differences between men and women (Fitch & Giedd, 1999), has also been shaped by sexual selection and is argued to be, at least in men, mostly shaped by intra-sexual competition (Hill et al., 2013; Kordsmeyer, Hunt, Puts, Ostner, & Penke, 2018; Puts, 2010).

Among the multiple acoustic components of human speech, fundamental frequency is the most studied and has been linked to the aforementioned indexical cues (i.e., F0, rate of vocal fold vibration corresponding to the acoustic correlate of voice pitch) and the formants (i.e., the acoustic resonances of the vocal tract); these cues are perceived as the most salient features of the human voice. However, other vocal traits can also convey information on indexical cues of speakers. For instance, variations of F0 (i.e., its standard-deviation, hereafter F0-SD, the acoustic correlate of intonation) appears to be a dimorphic acoustic characteristic, where men's speech is more monotonous in comparison with women's, who tend to vary their intonation more. Therefore, a more dynamic speech is more likely to be perceived as a reliable cue of femininity in women, while rather monotonous intonation can be linked to self-confident, relaxed and more masculine men (Hodges-Simeon, Gaulin, & Puts, 2011; Leongómez et al., 2014). Another vocal feature of interest is breathiness, which correlates with a lower harmonics-to-noise ratio (hereafter HNR) and quantifies the relative amount of additive noise in the voice signal (De Krom, 1995; Hillenbrand & Houde, 1996). Additive noise arises from turbulent air-flow generated at the glottis during phonation, and it is believed to result from an inadequate closure of the vocal folds. HNR reflects breathy voice quality, which is considered to be a signal of femininity (De Krom, 1995; Van Borsel, Janssens, & De Bodt, 2009), and has recently been found to be correlated to perceived attractiveness (Šebesta et al., 2017). In addition, vocal roughness (or hoarseness), which is related to irregular patterns of vocal folds vibration (Coleman & Wendahl, 1967), can be captured by the jitter, a measure of the cycle-to-cycle variation of the pitch period (with higher jitter values being associated with rougher voices; see Jones, Trabold, Plante, Cheetham, & Earis, 2001). Though small irregularities in the acoustic wave are considered normal variation associated with physiologic body function and voice production, it has also been shown that roughness may increase attractiveness, as higher jitter values result in a voice quality

known as vocal fry and is associated with impressions of desirability, more authoritative, educated, urban-oriented and upwardly mobile women (Greer & Winters, 2015; Oliveira, Davidson, Holczer, Kaplan, & Paretzky, 2016; Yuasa, 2010). Lastly, some other prosodic features of continuous speech, such as the speaking or the articulation rate (i.e., number of syllables produced per seconds, including or excluding pauses, respectively) may also play a role within courtship and competitive interactions; these have been previously shown to play upon perception of the speaker's competence, benevolence, trustworthiness, persuasiveness and social attractiveness (Brown, Strong, & Rencher, 1973; Miller, Maruyama, Beaber, & Valone, 1976; Smith, Brown, Rencher, & Strong, 1975; Street Jr & Brady, 1982).

Studies investigating the relationships between acoustic inter-individual differences and mating/reproductive success are scarce, and the results vary from one study to another. Apicella, Feinberg, and Marlowe (2007) showed that hunter-gatherers' voice pitch could reliably predict their reproductive success where a relatively low F0 is correlated to higher fitness, but it has been recently reported that this relationship does not hold when controlling for reputation (Smith, Olkhov, Puts, & Apicella, 2017). Through a simulated dating game, two studies from Puts (2005) and Puts et al. (2006) showed that lower F0 negatively correlated to men's mating success, although the latter study found that it was not significant. Using a similar approach, Hodges-Simeon et al. (2011) found that men who spoke in a more monotonic manner (i.e., lower F0-SD) when speaking to competitors declared more sexual partners over the past year. Moreover, Hughes et al. (2004) reported that female and male vocal attractiveness (when rated by members of the opposite sex) could predict their mating success, their declared number of extra-pair copulations and their age at first sexual intercourse. Lastly, Atkinson et al. (2012) found that F0 significantly predicted several measures of reproductive success in a group of Namibian females; higher voice pitch was associated with overall higher fitness.

However, methodologies varied concerning speech samples used in previous studies; some studies used the recordings of spoken vowels and read speech without any contextual background (Apicella et al., 2007; Atkinson et al., 2012; Hughes et al., 2004; Smith et al., 2017), which may not properly reflect how an individual vocally behaves in real ecological and social interactions (Hodges-Simeon et al., 2011; Puts, 2005; Puts et al., 2006; Puts et al., 2007). Using the former approach seems problematic, as it has been regularly shown that studies conducted on read/reciting vs. spontaneous speech produce quite different results (Blaauw, 1992; Daly & Zue, 1992; Howell & Kadi-Hanifi, 1991), and therefore, a focus on the latter approach is needed when one is interested in voice and its association to mating and reproductive success. In addition, particular attention has been given to F0 and its resonance frequencies, but other, understudied acoustic parameters are perceptible and are known to potentially affect the listeners' perceptions, such as the aforementioned breathiness, hoarseness and speech tempo. In addition, most of the work devoted to the study of voice and sexual selection has been conducted with native English speakers (in other cases in hunter-gatherer societies), but some evidence suggest that cultural variations may affect acoustic and prosodic features of speech and that preferences for a certain range of vocal features may vary socially, culturally and ecologically (Everett, Blasi, & Roberts, 2015; Šebesta et al., 2017; Van Bezooijen, 1995). Finally, it remains unclear which particular type of sexual selection has shaped acoustic features of speech; to our knowledge, only one study formally tested this hypothesis for acoustic features (Hodges-Simeon et al., 2011). The latter stated that intra-sexual competition was the main force driving the evolution of acoustic features. However, mate choice may also have led to the selection of different vocal qualities, i.e., a set of acoustic features that defines a particular individual's tone of voice (Laver, 1980) in men and women.

The purpose of this study is thus to determine the acoustic and prosodic features of speech that best predict men's and women's mating success, here defined as the self-reported number of sexual partners

within the past year, by studying spontaneous speech in both competitive and seductive contexts while taking a closer look at a set of understudied vocal parameters. Finally, we wish to assess which context best predicts the relationship between mating success and acoustic features.

2. Methods

2.1. Participants

In all, 68 female (mean age = 22.9; standard deviation = 1) and 56 male (mean age = 23; standard deviation = 3.36) participants were recruited by social networks and advertising in the university campus and other public places in Montpellier, France. All participants were self-declared heterosexual and native French speakers. For each participant, the general purpose of the study was explained ('a study on attractive vocal behavior'), and a written consent was requested for a statistical use of data (private information and recordings). All participants received financial compensation for their participation. The French National Commission on Informatics and Liberty approved protocols for this study (CNIL number 2-17029).

2.2. Stimuli

E-mails were sent to professional comedians who graduated from the National School of Dramatic Art of Montpellier (France). From those who responded positively, we selected one male and one female actor, respectively aged 28 and 25 years old. Both were French native speakers with European ascendants and were chosen because they should be able to act naturally to depict an attractive and dominant person, as they are professional comedians. They were first video recorded presenting themselves in front view with a blank background and facing the camera, then audio recorded speaking in a competitive manner stating why they should be better at winning the dating game (see below). The video and audio recordings were scripted to depict a friendly and a competitive picture of the person, respectively. The female and male video recordings lasted 84 s and 115 s, respectively, and the audio recordings lasted 12.5 s and 14 s, respectively (see Supplementary Materials for additional information).

2.3. Procedure and measures

Upon arrival at the laboratory, participants were seated in a quiet, anechoic, soundproof room equipped with a Sennheiser™ BF 515 microphone connected to a PC located in another room. All recordings were encoded using the Adobe© Audition CS6 at a sampling rate of 44 kHz – 32 bit – mono then saved as .wav files.

All participants played a simulated dating game using the same protocol as in (Hodges-Simeon et al., 2011; Puts, 2005; Puts et al., 2006). The game consisted of winning a date with a person of the opposite sex while being in competition with another same-sex person. Participants were first recorded producing a given sentence in their natural voice by repeating it after the examiner. Then, participants were either asked to seduce the potential date after seeing the video (courtship recording) or to explain why they were better mates after hearing the competitor (competitive recording). The rationale for using a video and an audio recording for courtship and competitive context, respectively, is that participants should be more motivated and involved when seeing the lunch date and less intimidated if they had seen the competitor, which in both cases could affect their vocal behavior during the dating game segment. The order of recordings was randomized between participants, and the same actors' recordings were used in all trials. To control for intensity, participants were asked to speak at a constant distance of 15 cm from the microphone.

After the dating game segment, participants completed a questionnaire assessing personal information: date and place of birth,

parents' and grandparents' origins, sexual orientation, relationship status (categorical variable stating whether the participants were in a relationship or not at the moment of the study) and the number of sexual partners they had over the past year (i.e., mating success). First, the number of past-year sexual partners was chosen because it represents an interval over which participants' recollections were expected to be accurate and the measured voice characteristics were likely to be stable (Hodges-Simeon et al., 2011). Second, it is, among other measures, a valid way of measuring mating success in humans, as it has been used in similar previous studies and was shown to be accurate (Faurie, Pontier, & Raymond, 2004; Hodges-Simeon et al., 2011; Puts, 2005; Puts et al., 2006; Puts et al., 2007). Moreover, human mating success should be an important component of expected fitness in past environments, as it represents their potential fertility (Perusse, 1993). In both sexes, potential fertility varies considerably with age, so reducing the time interval should give a more accurate proxy of their potential current fitness at the moment of the study.

2.4. Acoustic analyses

The courtship and competitive recordings were analyzed using Praat© voice analysis software (version 6.0.31, Boersma & Weenink, 2018). For each recording, six acoustic and two prosodic parameters were extracted: mean F0 (Hz), F0-SD (i.e., variation of voice pitch, Hz), local jitter (i.e., the average absolute difference between consecutive periods, divided by the average period, calculated in percentage), harmonics-to-noise ratio (dB), intensity (dB), duration (i.e., duration of the recording, in seconds), speaking rate (defined as the number of syllables produced per seconds, including pauses) and articulation rate (defined as the number of syllables produced per seconds, excluding pauses). Pitch was extracted using the autocorrelation method with pitch settings set to 85 to 400 Hz for women and 75 to 300 Hz for men. Although Praat's recommendation is a ceiling of 600 Hz for women, reducing it allows a more thorough extraction of vocal parameters and has been common in previous studies (e.g., Babel & McGuire, 2015; Escudero, Boersma, Rauber, & Bion, 2009; Lortie, Rivard, Thibeault, & Tremblay, 2017). All other settings were kept as default.

2.5. Data analysis

Men and women for each context were analyzed separately. Mating success was considered a dependent variable in a linear regression with the vocal variables as the explanatory variables. As the measure of mating success consists of a number of discrete events occurring in a fixed interval of time, a generalized linear regression was used with a quasi-Poisson error structure to increase the robustness of possible findings since slight over- and under-dispersion were present (i.e., scale factor slightly higher or lower than 1). Each acoustic (i.e., mean F0, F0-SD, jitter, HNR and intensity) and prosodic parameter (i.e., speaking and articulation rate and speech duration) was added as an explanatory variable. Additionally, four interactions were added between F0 and F0-SD with both jitter and HNR, as some evidence suggest that vocal breathiness and roughness could affect pitch and intonation (e.g., Oriikoff & Baken, 1990; Šebesta et al., 2017; Xu, Lee, Wu, Liu, & Birkholz, 2013). Finally, age and relationship status (categorical variable) were added as control variables, as they may have confounding effects on the response variable. All continuous variables were standardized. The significance of each term was assessed from the comparison of the model excluding the term with the model including all the other variables. Pseudo R^2 were computed for each model and adjusted for the number of parameters and observations. All statistical analyses were performed using R software (version 3.4.0).

3. Results

Descriptive statistics of all variables used in the models are given in

Table 1

Descriptive statistics for A) men and B) women of all acoustic and prosodic parameters for the courtship and competitive recordings as well as for age, number of past-year sexual partners and individuals involved in a relationship. Values given correspond to mean ± standard-deviations.

A.		
Men (n = 56)	Courtship recording	Competitive recording
Mean F0 (Hz)	109.76 ± 10.65	111.05 ± 12.25
F0-SD (Hz)	13.14 ± 3.58	12.90 ± 4.10
Jitter (%)	2.58 ± 0.60	2.68 ± 0.64
HNR (dB)	12.34 ± 1.75	11.67 ± 1.83
Speaking rate (syl/s with pauses)	3.65 ± 0.61	3.69 ± 0.87
Articulation rate (syl/s without pauses)	5.33 ± 0.56	5.67 ± 0.80
Intensity (dB)	62.25 ± 3.50	63.11 ± 3.50
Speech duration (s)	127.59 ± 77.13	68.70 ± 75.97
Age (years)	23 ± 3.36	
	(range = 18–33)	
Number of past-year sexual partners	2.03 ± 1.78	
In a relationship	Yes: 11	No: 45

B.		
Women (N = 68)	Courtship recording	Competitive recording
Mean F0 (Hz)	195.70 ± 17.38	196.93 ± 18.00
F0-SD (Hz)	30.18 ± 7.29	29.23 ± 8.46
Jitter (%)	1.97 ± 0.71	1.94 ± 0.45
HNR (dB)	15.43 ± 2.14	15.07 ± 2.33
Speaking rate (syl/s with pauses)	3.45 ± 0.61	3.56 ± 0.73
Articulation rate (syl/s without pauses)	4.79 ± 0.53	4.96 ± 0.74
Intensity (dB)	62.02 ± 4.61	62.36 ± 4.93
Speech duration (s)	129.27 ± 62.84	69.11 ± 43.25
Age (years)	22.9 ± 1	
	(range = 19–36)	
Number of past-year sexual partners	1.70 ± 1.58	
In a relationship	Yes: 25	No: 43

Table 1. A Mann-Whitney-Wilcoxon test was first conducted to ensure that no discrepancy was observed between men and women in the self-reported past-year mating success. As we expected, the test revealed no significant difference ($W = 1722, p = 0.34$).

3.1. Men

The results of the generalized linear models for men are reported in **Table 2A**. In the courtship context, articulation rate and intensity positively predicted mating success (respectively $\chi^2 = 5.19, df = 1, p < 0.05$; respectively $\chi^2 = 6.76, df = 1, p < 0.05$); i.e., men displaying relatively faster speech tempo and louder speech reported significantly more past-year sexual partners. In addition, the interactions F0-SD with jitter and F0-SD with HNR both negatively predicted their mating success (respectively $\chi^2 = 4.96, df = 1, p < 0.05$; $\chi^2 = 4.55, df = 1, p < 0.05$). A negative value for the effect of the interactions term implies that for individuals displaying higher F0-SD, the greater the negative effect of jitter and HNR is on mating success. The model yielded a R^2 of 0.14. None of the other explanatory and control variables were significant ($p > 0.05$). Moreover, no variables explained the number of self-reported sexual partners in the competitive context ($p > 0.05$).

3.2. Women

The results of the generalized linear models for women are reported

Table 2

Results of the generalized linear models predicting self-reported past-year mating success in A) men and B) women in the courtship and competitive contexts. For each factor, the estimate, standard error of the mean (SE), the χ^2 and the p values associated from the likelihood ratio test of the comparison between the full model and the model without the factor are given. For the categorical variable “relationship status”, the estimates are given for one category compared to the reference category (underlined term; “yes” refer to individuals in a current relationship at the moment of the study). R^2 is the variance explained by the model. Significance code: “****” $p < 0.001$; “***” $p < 0.01$; “**” $p < 0.05$.

A.				
Predicting past-year mating success (courtship recording)				
$R^2 = 0.14$	Estimate	SE	χ^2	p -Value
Intercept	0.45	0.18	/	/
Mean F0	-0.41	0.18	3.02	0.08
F0-SD	0.20	0.19	1.04	0.30
Jitter	-0.44	0.22	2.75	0.09
HNR	-0.01	0.18	0.02	0.87
Speaking rate	-0.18	0.14	1.66	0.19
Articulation rate	0.34	0.15	5.19	0.02*
Intensity	0.37	0.14	6.76	0.009**
Speech duration	-0.26	0.17	2.46	0.11
Mean F0: jitter	0.04	0.19	0.04	0.82
Mean F0: HNR	0.23	0.13	3.26	0.07
F0-SD: jitter	-0.62	0.29	4.96	0.02*
F0-SD: HNR	-0.42	0.20	4.55	0.03*
Age	-0.008	0.04	0.03	0.84
Relationship status	0.043	0.31	0.02	0.89
Yes/No				

Predicting past-year mating success (competitive recording)				
$R^2 = -0.12$	Estimate	SE	χ^2	p -Value
Intercept	0.64	0.18	/	/
Mean F0	-0.18	0.20	0.28	0.59
F0-SD	0.12	0.22	0.29	0.58
Jitter	-0.14	0.20	0.08	0.76
HNR	0.07	0.19	0.11	0.73
Speaking rate	-0.12	0.21	0.34	0.55
Articulation rate	0.10	0.18	0.37	0.54
Intensity	0.07	0.15	0.21	0.64
Speech duration	-0.38	0.25	2.99	0.08
Mean F0: jitter	-0.17	0.28	0.39	0.53
Mean F0: HNR	0.09	0.21	0.18	0.66
F0-SD: jitter	0.03	0.35	0.008	0.92
F0-SD: HNR	-0.04	0.22	0.03	0.85
Age	0.01	0.04	0.08	0.77
Relationship status	-0.33	0.40	0.72	0.39
Yes/No				

B.				
Predicting past-year mating success (courtship recording)				
$R^2 = 0.26$	Estimate	SE	χ^2	p -Value
Intercept	0.53	0.15	/	/
Mean F0	0.16	0.15	0.37	0.54
F0-SD	-0.13	0.13	0.82	0.36
Jitter	0.05	0.33	1.14	0.28
HNR	-0.43	0.12	14.68	0.0001****
Speaking rate	-0.02	0.13	0.03	0.86
Articulation rate	0.02	0.14	0.02	0.87
Speech duration	-0.31	0.11	7.45	0.006**
Intensity	-0.06	0.11	0.35	0.55
Mean F0: jitter	0.26	0.29	0.81	0.36
Mean F0: HNR	-0.01	0.16	0.009	0.92
F0-SD: jitter	-0.22	0.23	0.94	0.33
F0-SD: HNR	-0.14	0.09	2.42	0.12
Age	0.12	0.09	1.66	0.19

(continued on next page)

Table 2 (continued)

B.				
Predicting past-year mating success (courtship recording)				
R ² = 0.26	Estimate	SE	χ^2	p-Value
Relationship status Yes/No	−0.18	0.23	0.68	0.40
Predicting past-year mating success (competitive recording)				
R ² = 0.06	Estimate	SE	χ^2	p-Value
Intercept	0.50	0.15	/	/
Mean F0	−0.02	0.13	0.05	0.82
F0-SD	−0.12	0.13	0.76	0.38
Jitter	−0.013	0.18	0.08	0.77
HNR	−0.31	0.18	2.89	0.09
Speaking rate	−0.14	0.13	1.04	0.31
Articulation rate	−0.02	0.14	0.01	0.89
Intensity	0.10	0.13	0.61	0.43
Speech duration	−0.01	0.13	0.01	0.90
Mean F0: jitter	0.07	0.17	0.19	0.65
Mean F0: HNR	−0.007	0.15	0.002	0.95
F0-SD: jitter	−0.24	0.18	1.86	0.17
F0-SD: HNR	−0.04	0.22	2.86	0.09
Age	0.22	0.10	4.46	0.03*
Relationship status Yes/No	−0.02	0.26	0.008	0.92

Significant *p* values are in bold.

in Table 2B. In the courtship context, HNR and speech duration negatively predicted mating success (respectively $\chi^2 = 14.68$, $df = 1$, $p < 0.05$; $\chi^2 = 7.45$, $df = 1$, $p < 0.05$). Women who displayed less breathy voices (i.e., higher HNR values) and spoke less during the game segment reported significantly fewer past-year sexual partners. None of the other explanatory and control variables were significant ($p > 0.05$). In the courtship context, only age was significantly correlated to mating success ($\chi^2 = 4.46$, $df = 1$, $p < 0.05$).

4. Discussion

The present findings show that only acoustic inter-individual differences within the courtship context are correlated to self-reported past-year mating success in both men and women.

Men displaying faster articulation rate reported significantly more sexual partners over the past-year. Articulation rate determines the pace at which speech segments are actually produced. It has been previously reported that increased speech tempo also increases the perception of competence and social attractiveness (Smith et al., 1975; Street Jr & Brady, 1982; Street Jr, Brady, & Putman, 1983). Conversely, slower speech tempo is being perceived as less truthful, less fluent, less persuasive and more passive (Apple, Streeter, & Krauss, 1979; Smith & Shaffer, 1995); moreover, it has been associated with anxiety, sadness and depression (Siegman & Boyle, 1993; Smith & Shaffer, 1995). Others have also reported that it diminishes the perception of benevolence (Brown et al., 1973; Ray, 1986). It has been systematically reported that individuals in poor health conditions suffering from physiological and/or psychological disorders have slower speaking rates as it affects cognitive capacities (e.g., Caligiuri, 1989; France, Shiavi, Silverman, Silverman, & Wilkes, 2000; Turner, Tjaden, & Weismer, 1995). Moreover, faster speech tempo may constitute an honest cue of mate quality, since for these individuals, it requires less cognitive effort and time to produce speech. As earlier research indicates, liars pause frequently, speak slowly, and make many mistakes when they are giving cognitively complex speeches (Knapp, Hall, & Horgan, 2013). This is corroborated by the fact that deception is often inferred from these paralinguistic complexity cues because they imply that the speaker is

concocting a lie (Vrij, 2000). Faster speech tempo may also generally indicate a more forceful and, possibly, a more frequent and skillful use of the vocal organs (Hillman, Holmberg, Perkell, Walsh, & Vaughan, 1990; Jiang & Titze, 1994).

Furthermore, men who displayed louder voices reported significantly more sexual partners over the past year. Zuckerman and Miyake (1993) found that perceived louder voices were positively associated with attractiveness. Scherer (1978) showed that male speakers were rated as emotionally stable and extroverted by their peers when they spoke with a louder and possibly more nasal voice; it appears that louder speech (but more intermediate in loudness) using a greater range of loudness variation increases credibility. Indeed, in a later study, Ketrow (1990) found that moderately louder voices lead to greater perception of credibility and social attractiveness, which in turn should play upon persuasiveness and compliance of the target (although see Burgoon, Birk, & Pfau, 1990). This corroborates results found by Scherer, London, and Wolf (1973) where confidence was expressed by increased loudness of voice, rapid rate of speech, and infrequent, short pauses. Moreover, another study showed that in ratings of short samples of music and speech, loud excerpts were judged as more pleasant, energetic, and tense than soft excerpts (Ilie & Thompson, 2006). We suggest that relatively louder voices should be correlated to displays of power, authoritarian and higher social statuses in men, as deep loud voices are expected to be associated with larger and taller individuals, thereby increasing attention and acting as a possible marker of self-confidence or dominance. Nonetheless, louder speech (and thus loudness) is also associated with emotional states such as anger and anxiety (Laukka et al., 2008), which emphasizes its role in conveying information about the current emotional and affective state of the speaker.

As suggested and discussed in Hodges-Simeon et al. (2011) for intonation, men displaying faster articulation rate and relatively louder voice in this study had higher previous success with women; these men would presumably feel more self-assured when placed in a mating or competitive context. This self-confidence may have relaxed male subjects, stabilizing their tempo and vocal intensity across speech. If so, this suggests that when men feel confident, they speak with faster and louder speech. Both these prosodic and acoustic components may be one modality through which confidence is communicated to others, since both are commonly used to express emotion and affective states.

The results showed that in men with greater intonation (i.e., F0-SD), the more negative the effect of higher roughness and lower breathiness on their mating success. On one hand, higher jitter values can be detrimental in the context of courtship, leading to a perception of creaky voices, which is a marker of pathological voices (Wolk, Abdelli-Beruh, & Slavin, 2012). Although not common in men, it has been previously reported that in women, this leads to the perception of hesitancy (Yuasa, 2010). Coupled with a greater intonation, this may be even more disadvantageous; it has been previously reported that more monotonous speeches in men have been found to be positively correlated to mating success (Hodges-Simeon et al., 2011). As previously stated, intonation is one modality through which one can express self-confidence, dominance and experience. On the other hand, although breathiness has been recently correlated to perceived attractiveness (Šebesta et al., 2017; Xu et al., 2013), its interaction with intonation can also lead to unattractive voice as breathiness diminishes in higher intonation. In any case, perceptual studies are needed to better understand such interactions when men are trying to attract potential mates.

The results show that women displaying less breathy voices (i.e., higher HNR values) reported significantly fewer past-year sexual partners. Two studies indicated that vocal breathiness was positively related to perceived vocal attractiveness (Fraccaro et al., 2013; Xu et al., 2013) and seemed to be restricted to female voices only (Babel, McGuire, & King, 2014). Moreover, it has been argued that breathiness is a female attribute (Henton & Bladon, 1985; Van Borsel et al., 2009) because it “softens” some other aspects of speech such as F0 and

formants, although in this study, it did not interact with voice pitch or intonation. Breathiness has also been associated with youth; Ferrand (2002) and Gorham-Rowan and Laures-Gore (2006) showed that young women are perceived as significantly more breathy than elderly women. Thus, HNR could reflect a signal of fertility, since it is correlated to youth in women (Hawkes, O'Connell, Jones, Alvarez, & Charnov, 1998).

Several authors have recently argued that intra-sexual competition has mainly driven the evolution of several morphological traits in men, including voice pitch and its resonant frequencies (Hill et al., 2013; Kordsmeyer et al., 2018; Puts, 2010). However, the present findings suggest that mate choice may be at least as important as same-sex competition in shaping speech acoustic features in both men and women. In general, the different types of sexual selection may have shaped acoustic and prosodic speech features differently, leading to different vocal qualities that can be important in one context and not in another (i.e., courtship and same-sex competition). Moreover, in the human species, mate choice should still be an important type of selection, as the sex that invests more in reproduction (i.e., female) has the right of scrutiny when choosing a mate. The fact that we did not replicate the findings of Hodges-Simeon et al. (2011), where intra-sexual competition is more important in shaping men's acoustic features, may be due to a difference in sample size; our study had half the number of participants as their study, which potentially results in less power to detect the observed effect. Furthermore, such differences can be attributed to the fact that variation in courting mates and repelling competitors can be language-dependent. Nonetheless, further studies are needed to better understand such aspects.

The lack of findings for women within the competitive context may lie in the fact that this type of sexual selection was not strong enough to shape their vocal behavior. Indeed, past research has shown that competition among women relies very little on physical combat or aggression; women are assumed to be more prone to use indirect aggression (i.e., attempting to cause harm while simultaneously trying to make it appear as though there was no harmful intention). Such attempts may include social manipulation and shenanigans, for instance by spreading false information about one's reputation or interfering with friendships and group inclusion of competitors (see Fisher, 2015 for an extensive review). Therefore, it seems rather logical that this kind of competition does not lead to larger, taller and stronger statures in women, and thus, women do not need to convey impressions of dominance or largeness through their vocal features against competitors.

A potential limitation of the present study is that we relied on self-reported mating success, which is subject to imperfect memory or intentional distortion. Another measure would have been to ask the subjects about their lifetime numbers of sexual partners; however, this measure is more prone to memory error than recollections from the past year. In theory, we should expect men and women to have the same mean number of sexual partners over a lifetime. However, it has been regularly shown that lifetime self-reported number of sexual partners leads to considerable discrepancies between men and women (Brewer et al., 2000; Brown & Sinclair, 1999; Morris, 1993; Wiederman, 1997), where men tend to overestimate and women to underestimate such numbers. Thus, such biases can be partly accounted for by reducing the interval over which participants recollect their number of sexual partners, producing an equal mean number of sexual partners. In this study, no discrepancy was observed between our sampled men and women. Another potential limitation is that the interest of each participant for the potential lunch date was not evaluated. Although mates' preferences can vary from one individual to another, the use of only two actors allows for control of potential changes in acoustic parameters due to interest varying if several stimuli were used.

In summary, this study provides some original findings on vocal behavior within different contexts related to sexual selection. It is, to the best of our knowledge, the first study to report evidence that men's articulation rates and vocal intensity can reliably predict their mating

success when courting, while women may use vocal breathiness to potentially signal attractiveness. Further studies are needed to understand the underlying proximate mechanisms linking articulation rate, intensity and breathiness to phenotypic quality.

Acknowledgements

The authors are grateful to Patrick Khaldy for his valuable help in the collection of acoustic data. We would also like to thank the professional actors Blanche Adilon and Charly Breton for playing the role of the potential dates and competitors in our experimental design. This is a contribution of the Institut des Sciences de l'Evolution de Montpellier (UMR CNRS 5554).

Declaration of interest

The authors have no competing interests to declare.

Funding

This study received financial support from the project HumanWay (ANR-12-BSV7-0008).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.evolhumbehav.2018.07.001>.

References

- Apicella, C. L., Feinberg, D. R., & Marlowe, F. W. (2007). Voice pitch predicts reproductive success in male hunter-gatherers. *Biology Letters*, 3(6), 682–684.
- Apple, W., Streeter, L. A., & Krauss, R. M. (1979). Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology*, 37(5), 715.
- Atkinson, J., Pipitone, R. N., Sorokowska, A., Sorokowski, P., Mberira, M., Bartels, A., & Gallup, G. G., Jr. (2012). Voice and handgrip strength predict reproductive success in a group of indigenous African females. *PLoS One*, 7(8), e41811.
- Babel, M., & McGuire, G. (2015). Perceptual fluency and judgments of vocal aesthetics and stereotypicality. *Cognitive Science*, 39(4), 766–787.
- Babel, M., McGuire, G., & King, J. (2014). Towards a more nuanced view of vocal attractiveness. *PLoS One*, 9(2), e88616.
- Blaauw, E. (1992). Phonetic differences between read and spontaneous speech. *Second international conference on spoken language processing*.
- Boersma, P., & Weenink, D. (2018). Praat: Doing phonetics by computer [computer program] (Version 6.0.39, retrieved from) <http://www.praat.org/>.
- Bowling, D. L., Garcia, M., Dunn, J. C., Ruprecht, R., Stewart, A., Frommolt, K. H., & Fitch, W. T. (2017). Body size and vocalization in primates and carnivores. *Scientific Reports*, 7, 41070.
- Brewer, D. D., Potterat, J. J., Garrett, S. B., Muth, S. Q., Roberts, J. M., Kasprzyk, D., ... Darrow, W. W. (2000). Prostitution and the sex discrepancy in reported number of sexual partners. *Proceedings of the National Academy of Sciences*, 97(22), 12385–12388.
- Brown, N. R., & Sinclair, R. C. (1999). Estimating number of lifetime sexual partners: Men and women do it differently. *Journal of Sex Research*, 36(3), 292–297.
- Brown, B. L., Strong, W. J., & Rencher, A. C. (1973). Perceptions of personality from speech: Effects of manipulations of acoustical parameters. *The Journal of the Acoustical Society of America*, 54(1), 29–35.
- Burgoon, J. K., Birk, T., & Pfau, M. (1990). Nonverbal behaviors, persuasion, and credibility. *Human Communication Research*, 17(1), 140–169.
- Caligiuri, M. P. (1989). The influence of speaking rate on articulatory hypokinesia in Parkinsonian dysarthria. *Brain and Language*, 36(3), 493–502.
- Charlton, B. D., & Reby, D. (2016). The evolution of acoustic size exaggeration in terrestrial mammals. *Nature Communications*, 7, 12739.
- Cheng, J. T., Tracy, J. L., Ho, S., & Henrich, J. (2016). Listen, follow me: Dynamic vocal signals of dominance predict emergent social rank in humans. *Journal of Experimental Psychology: General*, 145(5), 536.
- Coleman, R. F., & Wendahl, R. W. (1967). Vocal roughness and stimulus duration. *Communication Monographs*, 34(1), 85–92.
- Collins, S. A. (2000). Men's voices and women's choices. *Animal Behaviour*, 60(6), 773–780.
- Collins, S. A., & Missing, C. (2003). Vocal and visual attractiveness are related in women. *Animal Behaviour*, 65(5), 997–1004.
- Dabbs, J. M., Jr., & Mallinger, A. (1999). High testosterone levels predict low voice pitch among men. *Personality and Individual Differences*, 27(4), 801–804.
- Daly, N. A., & Zue, V. (1992). Statistical and linguistic analyses of F0 in read and spontaneous speech. *Second international conference on spoken language processing*.

- De Krom, G. (1995). Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments. *Journal of Speech, Language, and Hearing Research*, 38(4), 794–811.
- Escudero, P., Boersma, P., Rauber, A. S., & Bion, R. A. (2009). A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *The Journal of the Acoustical Society of America*, 126(3), 1379–1393.
- Evans, S., Neave, N., Wakelin, D., & Hamilton, C. (2008). The relationship between testosterone and vocal frequencies in human males. *Physiology & Behavior*, 93(4–5), 783–788.
- Everett, C., Blasi, D. E., & Roberts, S. G. (2015). Climate, vocal folds, and tonal languages: Connecting the physiological and geographic dots. *Proceedings of the National Academy of Sciences*, 112(5), 1322–1327.
- Faurie, C., Pontier, D., & Raymond, M. (2004). Student athletes claim to have more sexual partners than other students. *Evolution and Human Behavior*, 25(1), 1–8.
- Feinberg, D. R., Jones, B. C., DeBruine, L. M., Moore, F. R., Smith, M. J. L., Cornwell, R. E., ... Perrett, D. I. (2005). The voice and face of woman: One ornament that signals quality? *Evolution and Human Behavior*, 26(5), 398–408.
- Ferrand, C. T. (2002). Harmonics-to-noise ratio: An index of vocal aging. *Journal of Voice*, 16(4), 480–487.
- Fisher, M. L. (2015). Women's competition for mates: Experimental findings leading to ethological studies. *Human Ethology Bulletin*, 30(1), 53–70.
- Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *The Journal of the Acoustical Society of America*, 106(3), 1511–1522.
- Fraccaro, P. J., O'Connor, J. J., Re, D. E., Jones, B. C., DeBruine, L. M., & Feinberg, D. R. (2013). Faking it: Deliberately altered voice pitch and vocal attractiveness. *Animal Behaviour*, 85(1), 127–136.
- France, D. J., Shiyavi, R. G., Silverman, S., Silverman, M., & Wilkes, M. (2000). Acoustical properties of speech as indicators of depression and suicidal risk. *IEEE Transactions on Biomedical Engineering*, 47(7), 829–837.
- Gorham-Rowan, M. M., & Laures-Gore, J. (2006). Acoustic-perceptual correlates of voice quality in elderly men and women. *Journal of Communication Disorders*, 39(3), 171–184.
- Greer, S. D., & Winters, S. J. (2015). The perception of coolness: Differences in evaluating voice quality in male and female speakers. *Proceedings of the 18th ICPhS*.
- Hawkes, K., O'Connell, J. F., Jones, N. B., Alvarez, H., & Charnov, E. L. (1998). Grandmothering, menopause, and the evolution of human life histories. *Proceedings of the National Academy of Sciences*, 95(3), 1336–1339.
- Henton, C. G., & Bladon, R. A. (1985). Breathiness in normal female speech: Inefficiency versus desirability. *Language & Communication*, 5(3), 221–227.
- Hill, A. K., Hunt, J., Welling, L. L., Cárdenas, R. A., Rotella, M. A., Wheatley, J. R., ... Puts, D. A. (2013). Quantifying the strength and form of sexual selection on men's traits. *Evolution and Human Behavior*, 34(5), 334–341.
- Hillenbrand, J., & Houde, R. A. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech, Language, and Hearing Research*, 39(2), 311–321.
- Hillman, R. E., Holmberg, E. B., Perkell, J. S., Walsh, M., & Vaughan, C. (1990). Phonatory function associated with hyperfunctionally related vocal fold lesions. *Journal of Voice*, 4(1), 52–63.
- Hodges-Simeon, C. R., Gaulin, S. J., & Puts, D. A. (2011). Voice correlates of mating success in men: Examining "contests" versus "mate choice" modes of sexual selection. *Archives of Sexual Behavior*, 40(3), 551–557.
- Howell, P., & Kadi-Hanifi, K. (1991). Comparison of prosodic properties between read and spontaneous speech material. *Speech Communication*, 10(2), 163–169.
- Hughes, S. M., Dispenza, F., & Gallup, G. G. (2004). Ratings of voice attractiveness predict sexual behavior and body configuration. *Evolution and Human Behavior*, 25(5), 295–304.
- Hughes, S. M., Farley, S. D., & Rhodes, B. C. (2010). Vocal and physiological changes in response to the physical attractiveness of conversational partners. *Journal of Nonverbal Behavior*, 34(3), 155–167.
- Hughes, S. M., Mogilski, J. K., & Harrison, M. A. (2014). The perception and parameters of intentional voice manipulation. *Journal of Nonverbal Behavior*, 38(1), 107–127.
- Ilie, G., & Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception: An Interdisciplinary Journal*, 23(4), 319–330.
- Jiang, J. J., & Titze, I. R. (1994). Measurement of vocal fold intraglottal pressure and impact stress. *Journal of Voice*, 8(2), 132–144.
- Jones, T. M., Trabold, M., Plante, F., Cheetham, B. M. G., & Earis, J. E. (2001). Objective assessment of hoarseness by measuring jitter. *Clinical Otolaryngology and Allied Sciences*, 26(1), 29–32.
- Ketrow, S. M. (1990). Attributes of a telemarketer's voice and persuasiveness. A review and synthesis of the literature. *Journal of Interactive Marketing*, 4(3), 7–21.
- Knapp, M. L., Hall, J. A., & Horgan, T. G. (2013). *Nonverbal communication in human interaction*. Cengage Learning.
- Knowles, K. K., & Little, A. C. (2016). Vocal fundamental and formant frequencies affect perceptions of speaker cooperativeness. *The Quarterly Journal of Experimental Psychology*, 69(9), 1657–1675.
- Kordsmeyer, T. L., Hunt, J., Puts, D. A., Ostner, J., & Penke, L. (2018). The relative importance of intra- and intersexual selection on human male sexually dimorphic traits. *Evolution and Human Behavior*, 39(4), 424–436.
- Laukka, P., Linnman, C., Åhs, F., Pissioti, A., Frans, Ö., Faria, V., ... Furmark, T. (2008). In a nervous voice: Acoustic analysis and perception of anxiety in social phobics' speech. *Journal of Nonverbal Behavior*, 32(4), 195.
- Laver, J. (1980). *The phonetic description of voice quality*. 31, Cambridge Studies in Linguistics London1–186.
- Leongómez, J. D., Binter, J., Kubicová, L., Stolařová, P., Klapilová, K., Havlíček, J., & Roberts, S. C. (2014). Vocal modulation during courtship increases proceptivity even in naive listeners. *Evolution and Human Behavior*, 35(6), 489–496.
- Linville, S. E., & Fisher, H. B. (1985). Acoustic characteristics of perceived versus actual vocal age in controlled phonation by adult females. *The Journal of the Acoustical Society of America*, 78(1), 40–48.
- Lortie, C. L., Rivard, J., Thibeault, M., & Tremblay, P. (2017). The moderating effect of frequent singing on voice aging. *Journal of Voice*, 31(1), 112–e1.
- Lyons, M., Lynch, A., Brewer, G., & Bruno, D. (2014). Detection of sexual orientation ("gaydar") by homosexual and heterosexual women. *Archives of Sexual Behavior*, 43(2), 345–352.
- Miller, N., Maruyama, G., Beaver, R. J., & Valone, K. (1976). Speed of speech and persuasion. *Journal of Personality and Social Psychology*, 34(4), 615.
- Montano, K. J., Tigue, C. C., Isenstein, S. G., Barclay, P., & Feinberg, D. R. (2017). Men's voice pitch influences women's trusting behavior. *Evolution and Human Behavior*, 38(3), 293–297.
- Morris, M. (1993). Telling tails explain the discrepancy in sexual partner reports. *Nature*, 365(6445), 437.
- Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist*, 111(981), 855–869.
- Munson, B., McDonald, E. C., Deboe, N. L., & White, A. R. (2006). The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics*, 34(2), 202–240.
- O'Connor, J. J., Re, D. E., & Feinberg, D. R. (2011). Voice pitch influences perceptions of sexual infidelity. *Evolutionary Psychology*, 9(1), 64–78.
- Ohala, J. J. (1983). Cross-language use of pitch: An ethological view. *Phonetica*, 40(1), 1–18.
- Oleszkiewicz, A., Pisanski, K., Lachowicz-Tabaczek, K., & Sorokowska, A. (2017). Voice-based assessments of trustworthiness, competence, and warmth in blind and sighted adults. *Psychonomic Bulletin & Review*, 24(3), 856–862.
- Oliveira, G., Davidson, A., Holczer, R., Kaplan, S., & Paretzky, A. (2016). A comparison of the use of glottal fry in the spontaneous speech of young and middle-aged American women. *Journal of Voice*, 30(6), 684–687.
- Orlikoff, R. F., & Baken, R. J. (1990). Consideration of the relationship between the fundamental frequency of phonation and vocal jitter. *Folia Phoniatrica et Logopaedica*, 42(1), 31–40.
- Perusse, D. (1993). Cultural and reproductive success in industrial societies: Testing the relationship at the proximate and ultimate levels. *Behavioral and Brain Sciences*, 16(2), 267–283.
- Pisanski, K., Bhardwaj, K., & Reby, D. (2018). Women's voice pitch lowers after pregnancy. *Evolution and Human Behavior*, 39(4), 457–463.
- Pisanski, K., Fraccaro, P. J., Tigue, C. C., O'Connor, J. J., Röder, S., Andrews, P. W., ... Feinberg, D. R. (2014). Vocal indicators of body size in men and women: A meta-analysis. *Animal Behaviour*, 95, 89–99.
- Pisanski, K., & Rendall, D. (2011). The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. *The Journal of the Acoustical Society of America*, 129(4), 2201–2212.
- Ptacek, P. H., & Sander, E. K. (1966). Age recognition from voice. *Journal of Speech, Language, and Hearing Research*, 9(2), 273–277.
- Puts, D. A. (2005). Mating context and menstrual phase affect women's preferences for male voice pitch. *Evolution and Human Behavior*, 26(5), 388–397.
- Puts, D. A. (2010). Beauty and the beast: Mechanisms of sexual selection in humans. *Evolution and Human Behavior*, 31(3), 157–175.
- Puts, D. A., Barnrd, J. L., Welling, L. L., Dawood, K., & Burriss, R. P. (2011). Intrasexual competition among women: Vocal femininity affects perceptions of attractiveness and flirtatiousness. *Personality and Individual Differences*, 50(1), 111–115.
- Puts, D. A., Gaulin, S. J., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior*, 27(4), 283–296.
- Puts, D. A., Hill, A. K., Bailey, D. H., Walker, R. S., Rendall, D., Wheatley, J. R., ... Ramos-Fernandez, G. (2016). Sexual selection on male vocal fundamental frequency in humans and other anthropoids. *Proceedings of the Royal Society of London B: Biological Sciences*, 283(1829), 2015–2830.
- Puts, D. A., Hodges, C. R., Cárdenas, R. A., & Gaulin, S. J. (2007). Men's voices as dominance signals: Vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior*, 28(5), 340–344.
- Puts, D. A., Jones, B. C., & DeBruine, L. M. (2012). Sexual selection on human faces and voices. *Journal of Sex Research*, 49(2–3), 227–243.
- Ray, G. B. (1986). Vocally cued personality prototypes: An implicit personality theory approach. *Communication Monographs*, 53(3), 266–276.
- Rendall, D., Vokey, J. R., & Nemeth, C. (2007). Lifting the curtain on the Wizard of Oz: Biased voice-based impressions of speaker size. *Journal of Experimental Psychology: Human Perception and Performance*, 33(5), 1208.
- Scherer, K. R. (1978). Personality inference from voice quality: The loud voice of extroversion. *European Journal of Social Psychology*, 8(4), 467–487.
- Scherer, K. R., London, H., & Wolf, J. J. (1973). The voice of confidence: Paralinguistic cues and acoustic evaluation. *Journal of Research in Personality*, 7(1), 31–44.
- Šebesta, P., Kleisner, K., Tureček, P., Kočnar, T., Akoko, R. M., Třebický, V., & Havlíček, J. (2017). Voices of Africa: Acoustic predictors of human male vocal attractiveness. *Animal Behaviour*, 127, 205–211.
- Sell, A., Bryant, G. A., Cosmides, L., Tooby, J., Szyner, D., Von Rueden, C., ... Gurven, M. (2010). Adaptations in humans for assessing physical strength from the voice. *Proceedings of the Royal Society of London B: Biological Sciences*, 277, 3509–3518.
- Siegmán, A. W., & Boyle, S. (1993). Voices of fear and anxiety and sadness and depression: The effects of speech rate and loudness on fear and anxiety and sadness and depression. *Journal of Abnormal Psychology*, 102(3), 430.
- Smith, B. L., Brown, B. L., Rencher, A. C., & Strong, J. (1975). Effects of speech rate on personality perception. *Language and Speech*, 18(2), 145–152.

- Smith, K. M., Olkhov, Y. M., Puts, D. A., & Apicella, C. L. (2017). Hadza men with lower voice pitch have a better hunting reputation. *Evolutionary Psychology*, 15(4), 147470491774046.
- Smith, S. M., & Shaffer, D. R. (1995). Speed of speech and persuasion: Evidence for multiple effects. *Personality and Social Psychology Bulletin*, 21(10), 1051–1060.
- Street, R. L., Jr., & Brady, R. M. (1982). Speech rate acceptance ranges as a function of evaluative domain, listener speech rate, and communication context. *Communication Monographs*, 49(4), 290–308.
- Street, R. L., Jr., Brady, R. M., & Putman, W. B. (1983). The influence of speech rate stereotypes and rate similarity on listeners' evaluations of speakers. *Journal of Language and Social Psychology*, 2(1), 37–56.
- Tigue, C. C., Borak, D. J., O'Connor, J. J., Schandl, C., & Feinberg, D. R. (2012). Voice pitch influences voting behavior. *Evolution and Human Behavior*, 33(3), 210–216.
- Tuomi, S. K., & Fisher, J. E. (1979). Characteristics of simulated sexy voice. *Folia Phoniatrica et Logopaedica*, 31(4), 242–249.
- Turner, G. S., Tjaden, K., & Weismer, G. (1995). The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research*, 38(5), 1001–1013.
- Van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*, 38(3), 253–265.
- Van Borsel, J., Janssens, J., & De Bodt, M. (2009). Breathiness as a feminine voice characteristic: A perceptual approach. *Journal of Voice*, 23(3), 291–294.
- Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and implications for professional practice*. Wiley.
- Wiederman, M. W. (1997). The truth must be in here somewhere: Examining the gender discrepancy in self-reported lifetime number of sex partners. *Journal of Sex Research*, 34(4), 375–386.
- Wolk, L., Abdelli-Beruh, N. B., & Slavin, D. (2012). Habitual use of vocal fry in young adult female speakers. *Journal of Voice*, 26(3), e111–e116.
- Xu, Y., Lee, A., Wu, W. L., Liu, X., & Birkholz, P. (2013). Human vocal attractiveness as signaled by body size projection. *PLoS One*, 8(4), e62397.
- Yuasa, I. P. (2010). Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech*, 85(3), 315–337.
- Zuckerman, M., & Miyake, K. (1993). The attractive voice: What makes it so? *Journal of Nonverbal Behavior*, 17(2), 119–135.

VII. La modulation vocale

Jusque-ici, les études qui se sont intéressées à la question de l'attractivité vocale se sont surtout attachées à la décrire à l'aide d'indices statiques presque exclusivement étudiés hors contexte (i.e. parole de laboratoire). Pourtant, si la qualité vocale a pu être façonnée par la sélection sexuelle, il se peut que la capacité à manipuler sa voix puisse avoir été sélectionnée pour exploiter la tendance naturelle des auditeurs à associer certains stéréotypes vocaux à certains traits typiquement attribués à l'un ou l'autre des deux sexes, tels que la féminité, la masculinité, la dominance, l'agressivité ou encore la soumission. Dans ce contexte, il a été suggéré que la « *modulation vocale* » – définie comme la manipulation contrôlée, orientée et volontaire de toute propriété vocale incluant mais non limitée à sa hauteur et ses résonances – permettrait aux individus de tirer avantage du dimorphisme vocal (Pisanski, Cartei, McGettigan, Raine, & Reby, 2016). Cette capacité interviendrait tout particulièrement dans une variété de contextes tels que celui de la séduction, les entretiens d'embauches, les débats, le chant, la comédie, le cinéma, et de façon plus ou moins subtile selon la familiarité ou le lien social entretenu avec l'interlocuteur.

Bien que les notions de « *volontaire* » et « *orientée* » soient discutables selon ce qu'elles recouvrent, le concept de modulation vocale a été observée chez de nombreuses espèces (Charlton & Reby, 2016 ; Morton, 1977). Plus précisément, les mâles exagèrent perceptivement leur taille corporelle via un contrôle de l'oscillation des plis vocaux et une modification des cavités de résonance situées au niveau supra-laryngé, ce qui augmenterait la perception de dominance et de menace. Ainsi, nous pouvons nous attendre à ce que les hommes aggravent volontairement leurs voix lorsqu'ils souhaitent paraître plus masculins, dominants ou menaçants, et inversement que les femmes s'expriment dans des registres vocaux plus hauts pour signaler la féminité, la petitesse et la soumission, comme cela est prédit par la théorie du « *code-fréquence* » (Ohala, 1984, ou plus récemment le « *gender code* », Cartei, Cowles, &

Reby, 2012). Cette capacité à modifier son propre signal vocal en fonction du contexte et de l'interlocuteur serait également apprise très tôt au cours de l'enfance (autour des 6 à 9 ans), suggérant que les enfants pré-pubères intègrent relativement tôt les caractéristiques vocales propres à leur sexe (Cartei, Cowles, Banerjee, & Reby, 2014).

Dans ce chapitre, nous nous attacherons à présenter les études mettant en évidence l'habileté à moduler sa voix en contexte de séduction vs. de compétition afin de vérifier s'il existe une tendance générale sur les comportements vocaux des hommes et des femmes en fonction du contexte. Dans un second temps, nous discuterons des avantages et des coûts associés à la modulation vocale, aspect qui n'a encore jamais été abordé dans la littérature. Enfin, nous discuterons de la pertinence du contrôle vocal pour l'émergence du langage articulé.

VII.1. État de l'art

Si la modulation vocale apparaît évidente dans des situations tel que le chant (e.g. Joliveau, Smith, & Wolfe, 2004), les jeux d'acteurs (e.g. Cartei et al., 2012 ; Eriksson & Wretling, 1997) et le langage adressé à l'enfant (Grieser & Kulh, 1988 ; Kuhl et al., 1997), les études qui s'y sont intéressées d'un point de vue évolutif l'ont surtout décrite dans les contextes de séduction et de compétition.

i. Dominance vocale

Les études s'accordent plutôt bien sur le fait que les individus diminuent et/ou augmentent la fréquence de leur hauteur vocale (F0) et leur timbre (formants) pour signaler respectivement des projections de dominance et masculinité ou de soumission et féminité.

Par exemple, dans les études fondées sur des speed-datings simulés évoquées précédemment, les hommes qui se perçoivent comme physiquement plus dominants par rapport à un compétiteur ont tendance à aggraver leur voix, tandis que les hommes se percevant comme plutôt soumis s'expriment dans des registres vocaux plus élevés, c'est-à-dire plus féminins (Puts et al., 2006). Les hommes possédant des traits relativement masculins, tels que, par

exemple, une haute stature et de hauts taux d'androgènes, semblent également articuler les voyelles moins clairement que les hommes perçus comme plus féminins (Kempe et al., 2013). De même, des juges ayant pour tâche d'évaluer en termes de dominance vs. soumission la voix de sujets à qui l'on a préalablement demandé de manipuler leurs voix pour sonner plus masculins ou plus féminins, considèrent effectivement les voix graves comme plus dominantes par rapport à leur version féminisée (haut F0) et/ou neutre (Fraccaro, O'Connor, Re, Jones, DeBruine, Feinberg, 2013). En reconstituant expérimentalement des situations d'embauches, Leongómez, Mileva, Little & Roberts (2017) ont montré que les participants hommes et femmes se percevant comme dominants ont tendance – après avoir été stimulés visuellement par des profils d'hommes exhibant un haut statut social – à abaisser leur hauteur vocale (F0 bas) et à s'exprimer avec des intonations plus monotones (faible F0-SD). La tendance inverse a été observée pour les participants se percevant comme peu dominants. Notons toutefois qu'une étude a rapporté des résultats contraires pour les deux sexes : les hommes comme les femmes ont – contre toutes attentes – élevés leur hauteur vocale lorsqu'il leur a été demandé de paraître dominant et ont été effectivement perçus comme tel par des auditeurs (Hughes et al., 2014).

La seule étude ayant été menée en dehors des conditions de laboratoire est celle de Sorokowski, Puts, Johnson, Żółkiewicz, Oleszkiewicz, Sorokowska, Kowal, Borkowska & Pisanski (2019), dont le protocole original mérite d'être souligné. Les auteurs ont fait rencontrer deux assistants chercheurs à 27 hommes et 24 femmes âgés de 30 à 55 ans, tous chercheurs et/ou membres de l'Université de Varsovie. Lors de cette rencontre, deux assistants de recherche se sont présentés comme des employés d'un service de programme radio et ont engagé deux types de conversations avec leurs interlocuteurs, une consistant à demander où étaient les services d'administration de la faculté (i.e. condition « *contrôle* »), et l'autre où ils ont demandé à leurs interlocuteurs de donner leur avis sur un programme radio du campus en

répondant à la question suivante « *Comment devenir un chercheur, et est-ce intéressant de l'être ?* ». Les hommes et les femmes interrogées ont aggravé de façon significative leurs voix dans la deuxième condition par rapport à la première, avec une diminution plus marquée chez les femmes que chez les hommes pour le F0 (33 Hz vs. 14 Hz de différence entre les deux conditions). Des expériences de playbacks montrèrent par la suite que des juges étrangers (i.e. ne parlant pas le polonais) ont perçu les stimuli vocaux issus de la deuxième condition comme appartenant à des individus plus compétents et plus autoritaires que ceux extraits de la première condition.

Enfin, concernant la projection de la taille corporelle, il a été montré dans une étude interculturelle que des hommes et des femmes Canadiens, Cubains et Polonais aggravent leurs hauteurs vocales (bas F0) et usent également d'un timbre relativement bas (bas Df) pour projeter une impression de grande taille corporelle ; à l'inverse, ils augmentent leurs hauteurs vocales et optent pour un timbre plus élevé pour signaler une petite taille corporelle (Pisanski, Mora, Pisanski, Reby, Sorokowski, Frackowiak, & Feinberg, 2016). Cette étude suggère ainsi que l'utilisation de la hauteur vocale et des résonances est employée de la même manière pour projeter la taille corporelle, quel que soit la culture/langue considérée, tel que prédite par la théorie « *code-fréquence* » (Ohala, 1984).

ii. Attractivité vocale

Concernant les modifications vocales liées à la perception de l'attractivité, les résultats sont bien moins consensuels.

Dans une des premières études mené à ce sujet, Fraccaro, Jones, Vukovic, Smith, Watkins, Feinberg, Little & DeBruine (2011) ont demandé à un groupe de femmes de laisser un message téléphonique (i.e. texte lu) à un homme présenté à elles à partir de deux photographies manipulée expérimentalement, l'une avec un visage masculinisé et l'autre avec un visage féminisé. Les auteurs ont montré que les femmes ont laissé des messages avec des

voix plus aiguës (F0 élevé) face à la version masculinisée de leur pseudo-interlocuteur, et que la différence de F0 entre les deux messages laissés aux deux types de stimuli visuels était positivement corrélée avec leur préférence pour les visages masculins.

Si cette dernière étude conforte les prédictions attendues concernant la voix des femmes (i.e. voix aiguës pour augmenter la perception de féminité, Pisanski et al., 2018), deux études ont rapporté des résultats contraires. Dans une étude semi-écologique, Hughes et al. (2010) ont utilisé un protocole original pour mesurer la modulation vocale en contexte de séduction. Chaque sujet avait pour tâche d'appeler trois personnes du sexe opposé et de leur poser plusieurs questions concernant le domaine de la psychologie. Avant d'entrer en contact téléphonique avec leurs interlocuteurs, les participants pouvaient voir le prétendu visage de la personne avec lesquelles ils allaient entrer en interaction verbale. Les trois visages ont précédemment été évalués en termes d'attractivité, et pour chaque sexe ont été sélectionné un visage jugé en amont comme attractif, un visage jugé comme non-attractif et un visage jugé comme moyennement attractif. En réalité, les personnes à appeler étaient fictives, et les participants tombaient systématiquement sur un répondeur sur lequel il leur était demandé de laisser un message qui leur avait été dicté par les investigateurs. Les messages laissés par les participants à l'étude aux personnes prétendument attractives, hommes comme femmes, abaissèrent leurs voix de façon significative comparativement aux messages laissés aux deux autres types de personnes (i.e. moyennement ou pas attractifs). Un autre groupe de juges ont également jugés la qualité des messages laissés aux personnes physiquement attractives comme plus attractifs par rapport aux autres messages vocaux. Dans une autre étude également menée par Hughes et al. (2014), il a été demandé à des sujets de compter de 1 à 4, puis de répéter l'opération en manipulant cette fois-ci leurs voix afin de paraître, entre autres, attractif. Les chercheurs ont ensuite demandé à des juges d'évaluer si l'impression recherchée était plus effective qu'en situation neutre (voix non modifiée pour répondre aux consignes des

expérimentateurs). Les résultats montrèrent que dans la condition d'attractivité, les deux sexes ont abaissé leur voix afin de la rendre plus grave. Notons qu'ils ont aussi diminué leur débit et joué sur le caractère rauque de leur voix. Ici, seules les voix féminines ont été considérées comme plus attractives en situation de manipulation qu'en situation normale (i.e. neutre). Dans ce contexte, ces deux études remettent en question l'idée générale qu'une plus haute hauteur vocale chez les femmes est considérée comme plus attractif (Pisanski et al., 2018). Une des explications donnée par Hughes et al. (2010/2014) est qu'un changement de paradigme a actuellement lieu, dans lequel la société actuelle permet aux femmes d'accéder à de plus hautes études et à des sphères professionnelles jusque-là réservé aux hommes. Dans ce contexte, la relative indépendance, assurance et confiance en soi pourrait ainsi être exprimé au niveau vocal via une modulation dans les graves de la hauteur vocale et du timbre.

Ces derniers résultats sont confortés par ceux obtenus dans la seule étude écologique menée sur la modulation vocale de l'attractivité. Dans cette étude, Pisanski et al. (2018) ont enregistré des sujets hommes et femmes ayant participé à un speed-dating en condition réelle. Sur un échantillon de 15 hommes et 15 femmes appariés en âge, chaque membre d'un sexe a rencontré tous les autres membres du sexe opposé et a engagé une conversation sur un temps limité avant de noter à la fin s'il/elle souhaitait revoir cette personne. Les résultats montrèrent que les hommes ont abaissé la valeur du $F0_{min}$ (i.e. plus petite valeur du $F0$ observée dans la chaîne parlée) lorsqu'ils discutaient avec des femmes hautement désirées par l'ensemble des participants hommes, abaissant également leur voix lorsqu'ils se sont trouvés face aux femmes qu'ils ont préféré et encore plus lorsqu'il s'agissait d'une attirance partagée (i.e. « *match* »). De manière intéressante, les femmes ont systématiquement aggravé leur voix lorsqu'elles communiquaient avec les hommes qu'elles préféraient et/ou ceux qui étaient hautement désirés par les autres femmes. Toutefois, notons l'absence de condition contrôle dans cette expérience,

ainsi que l'absence de contrôle de tous les autres paramètres pouvant fortement influencer l'attractivité d'un partenaire (i.e. visage, corps, statut social, etc.).

Concernant les variations de l'intonation, le peu de résultats rapportés ne fait également pas consensus. Dans une étude avec des sujets anglais et tchèques des deux sexes (Leongómez et al., 2014), les sujets ont été stimulés à l'aide d'une série de vidéos dans laquelle un homme et une femme se présentaient mutuellement, ceux-ci ayant été préalablement évalués en termes d'attractivité par un autre groupe de juges indépendants. Après le visionnement de toutes les vidéos, chaque participant devait laisser un message à la personne du sexe opposé expliquant pourquoi ils/elles aimeraient obtenir un rendez-vous, et un autre message à la personne du même sexe en expliquant pourquoi ils devraient être choisis plutôt qu'eux. Les auteurs ont rapporté que les hommes parlaient de façon plus expressive (F0-SD élevé) lorsqu'ils interagissent avec une personne attractive, tandis que les femmes ne l'ont augmenté qu'après avoir été exposées à des images de compétitrices attractives. Par la suite, des juges indépendants ont également évalué les voix les plus expressives comme plus attractives et ceci en contrôlant pour la langue parlée. A l'inverse, Hodges-Simeon et al. (2010) ont rapporté des résultats opposés : les hommes testés ont eu tendance à s'exprimer avec des voix plus monotones (F0-SD bas), caractéristiques des voix masculines, lorsqu'ils se décrivent à une potentielle partenaire par rapport à une condition neutre.

Enfin, notons l'étude de Fraccaro et al. (2013) qui questionne même l'efficacité perceptive de la modulation vocale en situation d'attraction et de séduction. Les auteurs ont demandé à des sujets masculins et féminins de produire des voyelles dans trois conditions articulatoires différentes : (i) en voix normale, (ii) en voix féminisée et (iii) en voix masculinisée. Des juges ont ensuite évalué les voix pour leur attractivité. Les résultats ont montré que les femmes préférèrent la voix normale des hommes par rapport aux versions féminisées des voix masculines, mais aucune différence de préférence entre la voix normale et

celle masculinisée n'a été observée. Les participants hommes ont montré une préférence pour la voix normale des femmes par rapport à leur version masculinisée, mais aucune différence de préférence entre la voix normale et la voix féminisée n'a été observée.

V.II. Avantages et coûts de la modulation vocale

Toute modification acoustique, volontaire ou non, doit être suffisamment perceptible par les auditeurs pour qu'elle forme ou renforce une impression spécifique selon l'intention du locuteur (cf. notion de « *différence juste notable* », cf section II.1.viii). En ce sens, l'avantage de la capacité à moduler sa voix devrait *in fine* correspondre à une plus forte probabilité d'être choisi en tant que partenaire ou d'avoir accès à une ressource particulière dans un contexte de compétition. En effet, même si certains résultats semblent contradictoires, nous avons indiqué précédemment que la modulation vocale, qu'elle soit élicitée en laboratoire ou produite en condition naturelle, modifie la perception de certains traits dont l'attractivité et/ou la dominance (Fraccaro et al., 2011, Hughes et al., 2010/2014, Sorokowski et al., 2019). Il est aussi important de noter que la modulation suppose que chaque individu exhibe une qualité vocale modale ou neutre qui serait la voix des contextes sociaux avec peu ou pas d'enjeux.

Néanmoins, il doit exister un coût à la modulation vocale sans lequel les individus se maintiendraient dans un état de modulation permanente. Par exemple, si aucun coût n'avait lieu, tous les hommes abaisseraient constamment leur F0 étant donné les bénéfices associés aux voix relativement graves (e.g. Klofstad et al., 2012 ; Mayew et al., 2013 ; Puts et al., 2005). Si les avantages perceptifs sont relativement bien compris, celui du coût de production est en revanche peu – voire pas – discuté dans la littérature (Pisanski et al., 2016). Les coûts peuvent être de différentes natures et se traduire en termes physiologiques et énergétiques. Une première approche pour le comprendre est d'établir la durée pendant laquelle un individu est capable de se maintenir dans un état de modulation et jusque quand un tel comportement vocal reste efficace au plan perceptif. Il est ainsi primordial d'évaluer la variation intra-individuelle,

d'établir avec précision la plage acoustique dans laquelle les individus peuvent moduler et surtout si les mêmes efforts sont nécessaires pour atteindre les extrémités de cette plage par rapport à des variations mineures autour de la moyenne habituelle. En conséquence, il est tout aussi important de savoir si la modulation a un effet linéaire au niveau perceptif : la diminution du F0, par tranche de « *différence juste notable* », augmente-t-elle linéairement avec la perception d'attractivité ? Notons également que les coûts liés à la modulation peuvent simplement se traduire par des facteurs limitant les plages de la modulation. En effet, puisque le conduit vocal des mammifères est constitué des cavités de résonance pharyngale, buccale et nasale, lesquelles sont fortement contraintes par les os du crâne, et que la taille du crâne est elle-même étroitement liée à la taille générale du corps (Fitch & Giedd, 1999 ; Fitch, 1997), la longueur du conduit est contrainte par des relations allométriques qui *in fine*, vont limiter l'espace acoustique sur lequel repose la modulation. Quant aux possibilités oscillatoires des plis vocaux (qu'il faut faire vibrer lentement pour abaisser sa voix et/ou rapidement pour l'élever), la fréquence de vibration dépend directement de la longueur et de l'épaisseur des plis vocaux, eux-mêmes possiblement influencés par les nombreux facteurs que nous avons précédemment évoqués. Dans ce contexte, nous comprenons pourquoi la question du coût associé à la modulation vocale est un problème complexe et multiparamétriques encore peu exploré. Mais plus généralement, nous pouvons avancer l'idée qu'un individu ne peut moduler sa voix au-delà de certaines limites imposées par son phénotype. La variation vocale observée au niveau intra-individuel peut ainsi constituer un signal « *honnête* » en soi dans la mesure où la plage acoustique des variations possibles pour un locuteur donné est un indice de sa qualité phénotypique.

Notons également deux dernières remarques concernant la modulation vocale. Premièrement, la question de la variation interindividuelle pour les préférences pourrait constituer un facteur important dans l'évolution de la modulation. S'il existe des avantages et

des coûts associés à celle-ci, il semble évident que cette habileté phonétique aurait évolué afin d'optimiser les efforts alloués aux interactions sociales, par exemple, tournées vers des comportements de séduction (i.e. concentrer ses efforts vocaux envers les individus attractifs). Deuxièmement, si la modulation vocale permet effectivement d'éliciter ou de manipuler une impression particulière chez un auditeur, les indices visuels fournissent déjà des informations saillantes relatifs à la qualité d'un partenaire (i.e. visage, corps), selon la multiplicité des signaux permettant de mieux estimer la qualité phénotypique (Johnston, 1997).

VII.3. Modulation vocale et émergence du langage

La modulation vocale, telle que définie dans notre travail, semble revêtir une importance particulière pour projeter chez autrui des impressions perceptives. Mais cette habileté vocale est également importante pour la production de la parole, puisque les mouvements des articulateurs sont aussi nécessaires à la production des différents sons qui font le langage articulé propre à l'espèce humaine. Nous nous posons donc la question de savoir si la modulation vocale a émergé pour servir initialement l'une ou l'autre de ces fonctions, ou si cette capacité a évolué conjointement pour servir ces deux objectifs.

Pour tenter de répondre à cette question, il nous semble particulièrement intéressant d'adopter une approche comparative. Depuis les années 2000, un nombre croissant d'études a mis en évidence que de nombreuses espèces de mammifères non parlantes sont capables de moduler leurs vocalisations – tant la hauteur que les résonances vocales – lors des interactions sociales via des adaptations comportementales et/ou anatomiques (Charlton & Reby, 2016). Le cas le plus emblématique est celui du cerf rouge, où les mâles en rut abaissent leurs larynx afin d'augmenter la longueur de leurs conduits vocaux, ceci dans le but d'aggraver leur « voix » et donner l'impression d'exagérer leur taille corporelle (Fitch & Reby, 2001). Nous pouvons citer d'autres espèces parmi lesquelles : les mâles koalas qui sont capables de modifier la taille de leurs « *soufflets* » (i.e. caisses de résonances) et donc des fréquences sonores qu'ils produisent

en fonction de la taille corporelle des compétiteurs (Charlton, Frey, McKinnon, Fritsch, Fitch, & Reby, 2013), les daims européens (Pitcher, Briefer, & McElligott, 2015) ou encore les pandas (Charlton, Zhihe, & Snyder, 2010).

Une récente étude concernant un large panel d'espèces de mammifères a montré que la sélection sexuelle ne favorise pas systématiquement l'utilisation du F0 (i.e. modification de la hauteur vocale) comme unique indice acoustique pertinent de l'exagération de la taille corporelle. Cette étude montre ainsi qu'au-delà de la hauteur vocale c'est surtout la manipulation des formants des vocalisations qui serait utilisé pour signifier l'impression d'une taille supérieure à celle réellement constatée (Charlton & Reby, 2016). Il a été suggéré que cette exagération de la taille corporelle via la manipulation des fréquences formantique serait initialement apparue pour servir des fonctions paralinguistiques. Elle aurait par la suite servi de base à des vocalisations plus volontaires et intentionnelles (Pisanski et al., 2016). Afin de réfléchir à la question de l'émergence du langage, il est ainsi avant tout nécessaire de montrer que la modulation d'une vocalisation particulière résulte bien d'un contrôle cognitif volontaire et motivé, et non d'un comportement réflexe provoqué par l'affect et/ou les émotions (« *environmental trigger* »).

Toutefois, le contrôle des vocalisations chez des espèces phylogénétiquement proches tels que les primates non-humains ne semblent pas chose aisée (Owren, Amoss, & Rendall, 2011). En effet, deux systèmes neuronaux sont impliqués dans la production et la modulation des vocalisations (« *dual neural pathways* »). Un premier fait intervenir une circuiterie neuronale liée à l'affect et aux émotions (i.e. système limbique). Cette structure, très similaire entre les humains et les autres primates, serait impliquée dans la production des vocalisations non-linguistiques chez les petits d'hommes (pleurs, rires et autres vagissements réflexes) et aux vocalisations spécifiques, propres à chaque espèce de primates non-humains (par exemple les « *hoos* » des chimpanzés). Ces vocalisations incontrôlées qui permettent l'expression des

émotions ne nécessitent pas d'apprentissage ni de stimulations extérieures pour leur production. A l'inverse, la production de vocalisations contrôlées et motivées au niveau cognitif (i.e. signaux intentionnels) repose sur l'intervention d'un second système cortical. L'apprentissage et la stimulation extérieure sont des éléments cruciaux de l'acquisition de ce système de communication où il ne s'agit plus d'exprimer uniquement des états émotionnels. Bien que les deux systèmes puissent entrer en interaction, les humains sont également capables de produire des vocalisations en absence de contexte et/ou de stimuli particuliers (nous parlons de communication non-référentielle dans la mesure où celle-ci peut avoir lieu en l'absence d'éléments extérieurs déclencheurs). La production de nouvelles vocalisations chez certains primates comme les orangs-outans et les chimpanzés captifs suggère qu'ils en sont capables mais dans une moindre mesure (Hopkins, Taglialatela, & Leavens, 2007 ; Lameira, Hardus, Kowalsky, de Vries, Spruijt, Sterck, Shumaker, & Wich, 2013 ; Lameira, Hardus, Mielke, Wich, & Shumaker, 2016). Ainsi, la capacité rudimentaire à produire des sons sans intention communicative – c'est-à-dire en l'absence d'une information à partager avec le reste du groupe – a pu être présente chez les premiers hominidés, elle a très certainement servi de substrat aux formes plus intentionnelles de la communication verbale dans l'espèce humaine.

V.II.4. Conclusion

Dans l'espèce humaine, un large champ de recherche reste ouvert pour explorer les diverses facettes de la modulation vocale. D'une part, des études s'avèrent nécessaires pour caractériser la modulation vocale dans des situations écologiquement valides, de même que les facteurs liés à ses coûts doivent être mis en évidence. D'autre part, de plus amples études sont nécessaires à propos de la variation intra-individuelle des différents types de vocalisations chez les primates non-humains afin de vérifier l'existence d'une communication intentionnelle. Plus généralement, il conviendrait également de réfléchir à la question des facteurs de

déclenchement de la modulation vocale afin de voir si celle-ci correspond à une réponse à un stimuli extérieur ou si, au contraire, elle procède d'un contrôle volontaire.

Étude exploratoire 1¹

Human vocal modulation within mate choice: a preliminary study in a speed-dating context.

Alexandre Suire, Valérie Durand, Michel Raymond & Melissa Barkat-Defradas

¹ Les deux études exploratoires de la thèse sont présentées en anglais avec une synthèse et une courte description du Matériels et Méthodes et des Résultats. Ce sont des études pour lesquelles il reste une part substantielle de travail à accomplir mais que nous avons tout de même choisi de présenter.

Résumé

La modulation vocale humaine reste un champ de recherche encore largement inexploré. De plus, les quelques études existantes ont été conduites en conditions de laboratoire, ce qui questionne leur validité écologique. Dans ce contexte, nous avons tenté de mesurer la modulation vocale dans des conditions réelles de speed-dating, tout en contrôlant pour l'aspect physique des locuteurs. De plus, nous avons comparé diverses mesures de la hauteur vocale (i.e. F0) en situation de séduction vs. une condition contrôle.

Si la modulation des paramètres du F0 semble bien se produire étant donné les différences observées entre les conditions de séduction et de contrôle, certaines ne s'effectuent pas nécessairement dans une direction facilement interprétable. Par exemple, les femmes ont augmenté leurs F0 et leurs F0-SD lors des conversations de speed-dating, ce qui concorde relativement bien avec l'idée d'un comportement vocal visant à signaler la féminité. Mais de manière surprenante, les hommes ont aussi augmenté leurs F0 lors des conversations de speed-dating par rapport à la condition contrôle (par ailleurs sans modification de leurs F0-SD). De plus, les femmes n'ont pas modifié leurs F0 entre les hommes qu'elles ont préféré vs. ceux moins préférés, tandis que les hommes ont diminué leurs F0 lorsqu'ils conversaient avec les femmes préférées. Hormis ces résultats, en contrôlant pour les paramètres physiques des locuteurs, les hommes ayant exhibé un F0 relativement bas ont été préférentiellement choisis par les femmes et ont obtenu des scores d'attractivité plus élevés. De manière plus intéressante, les hommes ont moins préféré les femmes exhibant un F0 relativement élevée, et pour lesquelles les scores d'attractivité attribuées étaient plus faibles.

Ce dernier résultat conforte relativement ceux présentés sur les préférences vocales des hommes français, pour lesquels les préférences peuvent être culturellement distinctes. Toutefois, le fait que les femmes aient augmenté leurs F0 alors qu'il est peu préféré par les hommes semble paradoxal, mais cela pourrait s'expliquer par le choix de la condition contrôle

dont la validité écologique peut être remise en question. Enfin, notons que cette étude reste exploratoire étant donné les tailles d'échantillons relativement faibles. Néanmoins, ces quelques résultats préliminaires encouragent la poursuite de ce type d'étude.

**Human vocal modulation within mate choice: a preliminary study in a speed-dating
context.**

Authors: Alexandre Suire¹, Valérie Durand¹, Michel Raymond¹, Melissa Barkat-

Defradas¹

¹ Institut des Sciences de l'Évolution de Montpellier, Université de Montpellier, Centre
National de la Recherche Scientifique, Institut pour la Recherche et le Développement, École
Pratique des Hautes Études – Place Eugène Bataillon 34095 Montpellier, France.

Summary

Previous studies have shown that humans are able to volitionally modify the acoustic properties of their voices in various social situations. More specifically, a particular attention has been given to the modulation of voice pitch (F0) and formant frequencies in sexually competitive contexts. However, most studies have been conducted within laboratory settings, which question the relative validity of potential findings in real ecological and social settings.

The only study that tested vocal modulation in a socially valid courtship context is from Pisanski, Oleszkiewicz, Plachetka, Gmiterek, & Reby (2018), in which the authors recorded conversations from real speed-dating sessions. However, this study suffers from two relatively problematic shortcomings: the lack of a control recording and a measure of physical aspects to control for choosiness and perceived attractiveness. The present study aims at filling these two gaps and offers further insight into the vocal modulation of another real speed-dating setting.

We found that the modulation of F0 occurred between courtship and control recordings, whereby most male and female individuals raised their F0. However, differences in patterns of modulation for the other parameters were found between males and females. For instance, F0-SD did not significantly change in males between the two conditions, whereas females raised theirs. Surprisingly, females did not modify their F0 properties as a function of preferred vs. non-preferred males. However, males did lower their F0 when speaking to females they preferred. Furthermore, males who exhibited a relatively lower F0 during courtship interactions seem to be preferred by females and were given higher attractiveness scores. But most importantly, our results suggest the same patterns for females. This corroborates some of the previous findings of Pisanski et al. (2018). This finding goes against the general view that higher F0s in females are attractive in females. A possible explanation for this observation is that, more generally, lower F0s express confidence and self-assurance, which may reinforce or

project impressions of independency, to which men may be more attracted to in modern and industrialized society.

Methods

Recruitment procedure

The study was advertised as both a real speed-dating event and a scientific experiment on romantic relationships (designed for heterosexual individuals), by means of flyers and posters handed out at the University of Montpellier and the University of Paul Valéry (Montpellier, France). All interested individuals had to register through a website designed for the specific purpose of this study. They had to fill a questionnaire regarding age, sex, various socio-economic aspects, answer the Sociosexual Orientation Inventory (SOI-R, Penke & Asendorpf, 2008), and indicate on which day they would be available to participate (socio-economic and SOI-R variables were not used in the analyses). Then, two of the experimenters formed out groups balanced in age and sex, with in mind to have a maximum of 7 male and 7 female participants, while assuring that the former was always at least slightly more aged than the latter. After this, emails were sent to the chosen participants to confirm their presence for the chosen date. Another email was sent the day before the session as a reminder.

Due a surprisingly high dropout rate, we ultimately conducted 7 sessions of speed-dating with a maximum of 8 participants and a minimum of 2 participants. Across all sessions, 23 men (Age: mean \pm SD = 24.31 \pm 4.29) and 14 women (Age: mean \pm SD = 24.08 years old \pm 4.24) participated in this study.

Experimental procedure

Upon arrival at the laboratory, male and female participants were seated in separate rooms. After obtaining a written consent from every participant, an instructor explained in details how the speed-dating will take place. The speed-dating was segmented into rounds, in which every participant did not necessarily converse with someone else due in imbalanced groups in sex.

Thus, each participant was given two sheets, one explaining where they needed to go after each round, and one where they could rate their conversational partners. During each round, participants could have thus been doing one of these four different things: converse with a person of the opposite sex in one of the two rooms dedicated to the speed-dating, getting his/her body measured by an experimenter in another room, converse with one of the two experimenters in another quiet room (control recordings), or wait in the waiting room.

Every participant from one sex conversed with every participant of the other sex. Participants were told not to start speaking until some music indicated so, and the conversation ended with that same music. Conversations were timed between 5 and 6 minutes. On each speed-dating table, two microphones (CMC 6+MK 41, Schoeps©), one in front of each participant, recorded the whole conversation and were encoded using ProTools© at a sampling rate of 44 kHz – 32 bit – mono then saved as .wav files. After each conversation, participants were asked to write down whether they would like to see that person again by indicating either ‘yes’ or ‘no’ (i.e., preferences or choosiness score) and to rate their attractiveness on a scale of 1 to 7 (i.e., 1 being ‘very not attractive’ – 7 being ‘very attractive’, which are later referred to as attractiveness scores).

Conversations with one of the experimenters lasted the same amount of time and consisted on a presentation of the participant (i.e., ‘Could you please introduce yourself and talk a little about your life’). Males were recorded with a linear PCM recorder (H1n, Zoom©), females were recorded with a linear PCM recorder (DR-O7 MKII, Tascam©) with both a sampling rate of 44 kHz, 32-bit, mono, then saved as .wav files.

After the speed-dating segment, individuals who reported a mutual preference (i.e. ‘match’) were put in contact via email.

Body measurement and facial photographs

Several physical measures were taken from every participant: height (cm), weight (kg), the circumference (cm) of the waist, hip, chest, shoulder and neck, and the strength of right and left arms using a dynamometer. Left or right handedness was also recorded, as well as the length of the second and four digits. Facial photographs were also taken. Facial images, handedness and strength measures were not used in the analyses. Ultimately, we computed the body mass index (BMI), the waist-to-hip ratio (WHR), the chest-to-hip ratio (CHR) and the waist-to-shoulder ratio (WSR), as these measurements are known to influence attractiveness assessments.

Voice analysis

All speed-dating conversations and control recordings were analyzed using the Praat© voice analysis software (version 6.0.31, Boersma and Weenink, 2018). The mean fundamental frequency (F0), its variation (F0-SD), the minimum (F0min) and maximum (F0max) were measured using the autocorrelation method with a pitch floor of 75 Hz and a ceiling of 300 Hz for male recordings and 100-500 Hz for female recordings (Praat's recommendation), with other settings kept as default.

Statistical analyses

Firstly, we analyzed if F0 parameters differed between the speed-dating conversations and the control recordings using t-tests (two tailed). We also analyzed if male or females changed their F0 parameters between preferred and non-preferred partners (i.e., F0 parameters in 'yes' vs. 'no' conversational partners). To do so, we used a generalized linear mixed model (GLMM) with preferences as the response variable and the F0 parameters as the explanatory variables.

Secondly, we analyzed how the different physical and vocal parameters of the participants predicted their choosiness and attractiveness scores given by their conversational partners. In order to analyze the probability of being chosen by the conversational partner, a

GLMM with the response variable being preferences (i.e., ‘yes’ vs. ‘no’) was constructed. In order to analyze perceived attractiveness, a linear mixed model (LMM) was used with the response variable being the attractiveness scores. In both models, the explanatory variables under study were: F0, F0-SD, F0min, F0max, BMI, WHR, CHR, WSR and age of the participant. All these variables were standardized before being entered in the models. Both conversational partners were entered as random factors. Males and females were analyzed separately.

All analyses were performed using the R software. Statistical results were considered significant at the 0.1 level.

Results

In total, we recorded 51 speed-dating conversations. Out of these interactions, men reported ‘yes’ on 44 occasions and ‘no’ on 7 occasions, while females reported ‘yes’ on 9 occasions vs. ‘no’ in 42 occasions. Mean attractiveness scores given to males was 3.65 ± 1.42 ; mean attractiveness scores given to females was 4.67 ± 1.27 .

Vocal modulation in courtship and control recordings

Males significantly spoke with a higher F0 in speed-dating conversations ($t = 3.783, p < .001$) and also exhibited a higher F0min and F0 max (respectively $t = 2.313, p = .024$; $t = 6.887, p < .001$). However, their F0-SD did not change between contexts ($t = 1.097, p = .27$). Women did not change their F0 and F0max between contexts (respectively $t = 1.222, p = .224$; $t = -0.113, p = .910$). However, females spoke with a significantly higher F0-SD and a lower F0min in speed-dating conversations (respectively $t = 2.949, p < .001$; $t = -7.812, p < .001$).

More specifically, males spoke with a lower F0 to females they were attracted to ($\chi_1^2 = 3.859, p = .049$). They also raised their F0max ($\chi_1^2 = 5.048, p = .024$). Females did not modulate their F0 parameters as a function of their preferences (all $p > .1$).

Correlations between F0 parameters in courtship conditions and body measurements

Table 1 reports all the correlations between F0 parameters and physical measurements. In both male and female participants, F0 was highly and positively correlated to F0-SD, F0min and F0max. F0max was also positively correlated to F0-SD and F0min in males, while it was only correlated to F0min in females. In males, F0-SD was positively associated to the waist-to-shoulder ratio, and F0min and F0max were positively associated to the chest-to-hip ratio. In females, only the F0 and F0min were positively associated to the chest-to-hip ratio.

Table 1. A) Males F0 and physical parameters; B) Females F0 and physical parameters. Reported correlations were computed from the Spearman method. Significant correlations are in bold (0.05 level). BMI = Body Mass Index; WHR = Waist-to-Hip Ratio; CHR = Chest-to-Hip Ratio; WSR = Waist-to-Shoulder Ratio.

A.

	<i>F0</i>	<i>F0-SD</i>	<i>F0min</i>	<i>F0max</i>	<i>BMI</i>	<i>WHR</i>	<i>CHR</i>	<i>WSR</i>	<i>Age</i>
<i>F0</i>	/								
<i>F0-SD</i>	0.30	/							
<i>F0min</i>	0.33	-0.15	/						
<i>F0max</i>	0.69	0.48	0.46	/					
<i>BMI</i>	-0.04	0.01	0.06	-0.01	/				
<i>WHR</i>	-0.23	-0.02	-0.02	-0.01	0.58	/			
<i>CHR</i>	0.04	0.01	0.30	0.25	-0.29	-0.06	/		
<i>WSR</i>	0.02	0.30	-0.05	0.16	0.63	0.73	-0.39	/	
<i>Age</i>	-0.12	0.04	-0.14	-0.21	0.37	0.40	-0.36	0.41	/

B.

	<i>F0</i>	<i>F0-SD</i>	<i>F0min</i>	<i>F0max</i>	<i>BMI</i>	<i>WHR</i>	<i>CHR</i>	<i>WSR</i>	<i>Age</i>
<i>F0</i>	/								
<i>F0-SD</i>	0.44	/							
<i>F0min</i>	0.71	-0.04	/						
<i>F0max</i>	0.40	0.57	-0.02	/					
<i>BMI</i>	-0.05	-0.10	0.13	0.01	/				
<i>WHR</i>	-0.15	-0.26	0.00	-0.02	0.71	/			
<i>CHR</i>	0.28	0.24	0.32	0.06	0.13	0.04	/		
<i>WSR</i>	-0.12	-0.12	-0.08	0.01	0.70	0.79	-0.08	/	
<i>Age</i>	-0.01	0.01	-0.01	0.18	0.24	0.06	0.18	0.20	/

Female choosiness and perceived attractiveness of males

Results of male choosiness and perceived attractiveness of females are respectively given in Table 2 and 3. During speed-dating conversations, females who exhibited a relatively higher *F0* had less chance to be chosen ($\chi_1^2 = 3.478, p = .062$). Females who exhibited a higher *F0max* had higher chance of being chosen ($\chi_1^2 = 5.045, p = .024$). Females who spoke longer also had less chance of being chosen ($\chi_1^2 = 3.467, p = .062$). Lastly, females with relatively higher waist-to-hip ratios had less chance to be chosen ($\chi_1^2 = 3.527, p = .060$).

The attractiveness scores given by males for females followed the same tendency. Male gave lower attractiveness scores to females exhibiting relatively higher *F0s* ($\chi_1^2 = 4.898, p = .060$), and those with higher waist-to-hip ratios ($\chi_1^2 = 3.960, p = .046$). However, male gave higher attractiveness scores to females who exhibited a higher *F0max* ($\chi_1^2 = 4.120, p = .042$).

Table 2. GLMM predicting the probability of choosing a female¹.

¹ For each variable, the χ_2 and the p values associated from the likelihood-ratio chi-square test of the comparison between the full model and the model without the variables are given (ANOVA type II). For the choosiness models, the reference value is 'yes'. The degrees of freedom is 1 for every test. P values are considered significant at the 0.1 threshold (in bold). BMI = Body

	Estimate	SE	χ^2	<i>p</i> value
Intercept	5.257	2.053	/	/
F0	-3.689	1.978	3.478	.062
F0-SD	0.061	0.680	0.008	.928
F0min	2.488	1.919	1.680	.194
F0max	5.788	2.576	5.045	.024
Duration	-2.482	1.333	3.467	.062
BMI	0.124	1.346	0.008	.926
WHR	-2.276	1.211	3.527	.060
CHR	0.092	0.931	0.009	.921
WSR	-0.672	1.922	0.122	.726
Age	0.681	0.810	0.706	.400

Table 3. LMM predicting male's attractiveness scores given to females.

	Estimate	SE	χ^2	<i>p</i> value
Intercept	-0.004	0.142	/	/
F0	-0.726	0.328	4.898	.026
F0-SD	0.164	0.186	0.772	.379
F0min	0.265	0.189	1.960	.161
F0max	0.734	0.361	4.120	.042
Duration	-0.145	0.149	0.940	.332
BMI	0.053	0.257	0.042	.835
WHR	-0.372	0.187	3.960	.046
CHR	-0.005	0.160	0.001	.973
WSR	-0.157	0.345	0.208	.647
Age	-0.028	0.164	0.030	.862

Female choosiness and perceived attractiveness of males

Results of female choosiness and perceived attractiveness of males are respectively given in Table 4 and 5. Although both the choosiness and the attractiveness scores models' output reveal the same tendency, the only two significant result was that males who exhibited a relatively higher F0 had less chance to be chosen by their female conversational partners ($\chi^2_1 = 2.804$, $p = .093$), and older males had higher attractiveness scores ($\chi^2_1 = 2.865$, $p = .090$).

Table 3. GLMM predicting the probability of choosing a male.

Mass Index; WHR = Waist-to-Hip Ratio; CHR = Chest-to-Hip Ratio; WSR = Waist-to-Shoulder Ratio. **All subsequent tables follow the same presentation.**

	Estimate	SE	χ^2	<i>p</i> value
Intercept	2.781	0.910	/	/
F0	-1.570	0.937	2.804	.093
F0-SD	0.363	1.173	0.096	.756
F0min	0.120	0.783	0.023	.877
F0max	-1.096	0.969	1.279	.258
Duration	0.148	0.558	0.070	.789
BMI	0.448	0.787	0.324	.568
WHR	0.648	0.958	0.457	.498
CHR	0.406	0.624	0.423	.515
WSR	0.262	0.881	0.088	.765
Age	-0.675	0.570	1.400	.236

Table 4. LMM predicting male's attractiveness scores given to males.

	Estimate	SE	χ^2	<i>p</i> value
Intercept	0.005	0.226	/	/
F0	-0.319	0.220	2.096	.147
F0-SD	-0.122	0.222	0.305	.580
F0min	0.183	0.156	1.376	.240
F0max	0.164	0.191	0.742	.389
Duration	-0.118	0.159	0.550	.458
BMI	-0.133	0.338	0.155	.693
WHR	-0.359	0.392	0.840	.359
CHR	0.022	0.261	0.007	.930
WSR	0.274	0.401	0.466	.494
Age	0.290	0.171	2.865	.090

References

- Penke, L., & Asendorpf, J. B. (2008). Beyond global sociosexual orientations: A more differentiated look at sociosexuality and its effects on courtship and romantic relationships. *Journal of Personality and Social Psychology*, 95(5), 1113–1135. <https://doi.org/10.1037/0022-3514.95.5.1113>
- Pisanski, K., Oleszkiewicz, A., Plachetka, J., Gmiterek, M., & Reby, D. (2018). Voice pitch modulation in human mate choice. *Proceedings of the Royal Society B: Biological Sciences*, 285(1893), 20181634. <https://doi.org/10.1098/rspb.2018.1634>

Étude exploratoire 2

Modulation of the acoustic features of ‘wahoo’ calls in wild chacma

baboons (*Papio ursinus*)

Alexandre Suire, Margaux Lefèvre & Alice Baniel

Résumé

A notre connaissance, peu d'études se sont intéressées à la modulation d'une vocalisation particulière chez les primates. Comme nous l'avions précédemment indiqué, certaines vocalisations sociales pourraient être sujettes à des ajustements selon le contexte dans lesquelles elles sont produites, en fonction du ou des conspécifiques présents.

Les vocalisations de dominance des mâles babouins chacmas (*Papio ursinus*) constituent une excellente opportunité pour explorer cette idée. Ces babouins vivent dans des groupes multi-mâles où l'établissement du rang de dominance pour monopoliser les opportunités d'accouplements fait fréquemment intervenir des démonstrations de dominance vocales, appelés « *wahoos* ». Ces derniers peuvent être produits dans différents contextes, lors d'agressions inter- ou intra-groupe et peuvent impliquer simultanément plusieurs mâles. Des études précédentes ont suggéré que ces vocalisations fonctionnent en tant que signal de capacité compétitive et de rang social : les mâles de hauts rangs produisent des vocalisations avec des F0 plus élevés et des syllabes « *hoo* » plus longues. Néanmoins, si leur fonction semble relativement bien comprise, aucune étude à notre connaissance ne s'est intéressée à leur modulation selon la situation et l'auditoire dans lesquelles ils sont produits.

Dans ce contexte, nous avons étudié plus de 800 de ces vocalisations produites par 19 mâles appartenant à deux troupes différentes. Nous nous sommes surtout intéressés à trois composantes de la syllable « *wa* » et « *hoo* » (durée, F0 et intensité), et à leur modulation possible en situation d'agressions inter- vs. intra-groupe et en fonction du ou des mâles impliqués dans celles-ci. Premièrement, nos résultats montrent que le rang de dominance est effectivement prédit par le F0 et la longueur de la syllabe « *hoo* ». De manière intéressante, nos résultats suggèrent que seules les longueurs des deux syllabes varient entre les situations inter- et intra-groupes, et que le F0 et la longueur du « *hoo* » semblent bien se moduler en fonction du ou des mâles présents.

Si ces quelques résultats suggèrent que la modulation vocale est effective chez les babouins chacmas, cette étude est seulement exploratoire. Il reste notamment une large partie des formants à analyser et d'autres données telles que l'âge, la taille corporelle et l'asymétrie fluctuante, pouvant influencer cette modulation, n'ont pas encore été exploités.

Modulation of the acoustic features of ‘wahoo’ calls in wild chacma baboons (*Papio ursinus*)

Authors: Alexandre Suire¹, Margaux Lefèvre¹ & Alice Baniel

¹ Institut des Sciences de l'Évolution de Montpellier, Université de Montpellier, Centre National de la Recherche Scientifique, Institut pour la Recherche et le Développement, École Pratique des Hautes Études – Place Eugène Bataillon 34095 Montpellier, France.

² Stony Brook University, New York

Summary

Chacma baboons (*Papio ursinus*) live in multi-male multi-female groups, where males compete aggressively with each other to establish their dominance rank and monopolize mating opportunities (Bulger, 1993). In this species, males frequently engage in conspicuous vocal dominance displays, involving two-syllable ‘wahoo’ calls (Fischer, Kitchen, Seyfarth, & Cheney, 2004; Kitchen, Seyfarth, Fischer, & Cheney, 2003a). These ‘wahoo’ contests are triggered during both inter-group encounters and intra-group aggressive interactions, and can be performed alone or involve multiple males (Kitchen, Cheney, & Seyfarth, 2005a; Kitchen et al., 2003a; Kitchen, Seyfarth, Fischer, et al., 2003b). ‘Wahoo’ bouts involving several males can last up to one hour and frequently involve running, chasing and physical fight among males, or between males and females (Fischer et al., 2004; Kitchen et al., 2003b). Such contests typically aim to defend resources of high fitness value, such as establishment or maintenance of dominance rank, monopolization of estrous females or protection of infants vulnerable to infanticide.

Previous research has shown that ‘wahoo’ calls function as honest indicators of males competitive abilities and stamina: high-ranking males participate more often and call at higher rates than low-ranking males during contests (Kitchen et al., 2003b) and acoustic features of the calls correlate with dominance rank (Fischer, Hammerschmidt, Cheney, & Seyfarth, 2002; Fischer et al., 2004; Kitchen, Cheney, Engh, Fischer, Moscovice, & Seyfarth, 2013). Nonetheless, while the function of wahoo vocalizations in mediating male-male competition is well understood (Kitchen, Cheney, & Seyfarth, 2005b; Dawn M. Kitchen et al., 2005a; Kitchen et al., 2003a), the way in which males adjust their vocalizations according to the social context (inter- vs. intra-group aggressions) and to the social rank of their rivals remains poorly understood.

In this exploratory study, we aim to offer some insights regarding this matter. Preliminary results suggest that the rank of males is correlated with several acoustical components of their ‘wahoo’ vocalizations. Namely, high-ranking males produced longer ‘hoo’ syllables with higher F0s, which corroborates previous findings (Fischer et al., 2004). Most importantly, results show that males seem to modulate differently the length of both ‘wa’ and ‘hoo’ syllables during inter- and intra-group contests. Males also seem to modulate some acoustic components according to the identity of the other males calling in the same bout, relative to their own rank.

Such findings are interesting in regards to recent studies who have highlighted vocal modulation in other primates (Gustison & Bergman, 2017; Lameira, Hardus, Mielke, Wich, & Shumaker, 2016). Those results also bear relative importance regarding the emergence of human speech since vocal modulation has been suggested to be a pre-requisite to the evolution of articulated language (Pisanski, Cartei, McGettigan, Raine, & Reby, 2016).

Methods

Study site and data collection

The studied chacma baboon population consists of two habituated troops (called J and L) living at Tsaobis Nature Park in Central Namibia (22°22’S 15°44’E). Tsaobis is a semi-arid area characterised by mountains and rocky foothills, and scattered woodland patches surrounding the ephemeral Swakop River. In 2013 and 2014, observers followed both troops daily on foot from dawn to dusk to collect *ad libitum* and behavioural data. All animals were individually identifiable and fully habituated to human observers. Wahoo vocalizations of resident adult males were recorded opportunistically during inter- or intragroup contests using a Sennheiser MKH 416T directional microphone and a Marantz CP230 tape recorder, at a sampling rate of 44.1 kHz – 16 bit – mono then saved as .wav files. In total, 19 males were recorded across different wahoo bouts, producing 834 ‘wahoos’ (720 intra-group and 114 inter-group ‘wahoos’)

that were further analyzed acoustically (number of observations per male: range: 2-155, mean \pm s.d.: 49 ± 42.6). The recorder indicated for each wahoo the identity of the calling male, the habitat (hills versus woodland), the social context (intra versus inter-group aggression), the other males involved in the wahoo bout and the distance of recording (only for a subset of calls, $N = 542$).

Outside of wahoo contests, observers conducted routinely 1h-focal animal sampling (Altmann, 1974) on all adult males to record occurrences and directions of approach-retreat interactions (supplants, displacements) and aggressive behaviours (threats, chases, attacks) between the focal male and any other male. A total of 538 focal observations of 24 males were collected (number of observations per male: range: 10-37, mean \pm s.d.: 22.4 ± 8.6). Agonistic interactions between males we also collected *ad libitum* throughout the day. Using these data, male dominance ranks were established with an Elo-rating procedure (Albers and de Vries 2001) implemented in the R package EloRating (Neumann et al. 2011). We obtained a score for each individual male for each day of observation. To obtain comparable ranks across the two troops (that differ in the number of males), we derived a daily standardized rank by scaling the Elo-rating score of each individual proportionally between 0 (corresponding to the lowest-ranking individual) and 1 (corresponding to the highest-ranking individual).

Acoustic analyses

The ‘wa’ and ‘hoo’ syllables were analyzed separately as they both display very different acoustic structure (Boë et al., 2017; Fischer et al., 2004). We used the Praat© voice analysis software to extract acoustic parameters (version 6.0.31, Boersma and Weenink, 2018). Based on previous studies and from J.L. Böe (personal communication), we retrieved via the autocorrelation method the fundamental frequency (F_0 , Hz), its standard deviation (F_0 -SD), the minimum (F_{0min}) and maximum (F_{0max}) using a floor of 200 Hz and 500 Hz for the ‘wa’ syllable and a floor of 50 Hz and a ceiling of 120 Hz for the ‘hoo’ syllable, with other settings

kept as default. The voicing threshold was sometimes adjusted between 0.1 and 0.2 when the ‘wahoo’ was not considered as voiced by the software. Intensity (dB) and duration (s) of each syllables were also retrieved.

Statistical analyses

Preliminary analyses reveal strong correlations between the different F0 parameters (see Table 1) and multicollinearity checks revealed strong variation inflation factors within models, so we only analyzed the duration, the F0 and the intensity of both the ‘wa’ and ‘hoo’ syllables. The distance of recording did not affect the chosen acoustic parameters (Table 1) and was subsequently ignored in the following models to increase the sample size.

First, we analyzed if the duration and acoustic properties of the ‘wahoos’ could predict male dominance ranks. To this end, we used a linear mixed model with male rank as the response variable and the acoustic components as the explanatory variables. Secondly, we analyzed if the acoustic structure of ‘wahoos’ generally differed between inter- and intra-group contests (hereafter referred to as social context). We used a generalized linear mixed model with social context as the response variable. Acoustic components of both the ‘wa’ and ‘hoo’ syllables were entered as the explanatory variables. Lastly, we analyzed if any acoustic components of the ‘wahoos’ was modulated in the presence of particular males in the context of intra-group aggressions, relative to the rank of the male vocalizing. To this end, we build several linear mixed models with the acoustic components of ‘wahoos’ as the response variables. Male rank and the identity of other males involved in the contest were entered as the explanatory variables. Both variables were also fitted in interaction to assess whether the modulation of any acoustic component was mediated by the rank of male and the specific rival males involved in the bout.

Identity of the male vocalizing was added as a random factor in all models and all quantitative variables were standardized. All analyses were performed using the R software. Statistical results were considered significant at the .05 level.

Results

Correlations among the acoustic components and male dominance ranks are given in Table 1. Among the variables of interests, rank was highly and negatively correlated to the duration and the F0, F0-SD, F0min and F0 max of the ‘wa’ syllable (all $p < .001$), whereas it was highly and positively correlated to the duration and the F0, F0-SD, F0max and intensity of the ‘hoo’ syllable (all $p < .001$).

Table 1. Spearman correlations of the acoustic components between and within the ‘wa’ and ‘hoo’ syllables and with male dominance rank, the successive order of ‘wahoos’ in a sequence (N) and the distance of recordings (D). Significant values are in bold: ‘***’ $p < .001$; ‘**’ $p < .01$; ‘*’ $p < .05$

		‘wa’						‘hoo’						Rank	N	D
		Duration	F0	F0-SD	F0min	F0max	Intensity	Duration	F0	F0-SD	F0min	F0max	Intensity			
‘w a’	Duration	/														
	F0	0.02	/													
	F0-SD	0.12***	0.44***	/												
	F0min	-0.13***	0.49***	-0.19***	/											
	F0max	0.12***	0.74***	0.80***	0.17***	/										
	Intensity	0.08*	0.16***	0.11***	0.04	0.10**	/									
‘h o o’	Duration	-0.01	0.19***	0.07*	0.07*	0.16***	-0.14***	/								
	F0	-0.12***	-0.12***	-0.13***	-0.05	-0.15***	-0.05	-0.11***	/							
	F0-SD	-0.06	0.01	0.00	-0.07*	0.00	0.02	0.14***	0.29***	/						
	F0min	-0.08*	-0.10**	-0.10**	0.02	-0.11***	-0.03	-0.22***	0.69***	-0.28***	/					
	F0max	-0.11**	-0.05	-0.06	-0.05	-0.07*	-0.05	0.04	0.83***	0.63***	0.35***	/				
	Intensity	-0.24***	-0.18***	-0.14***	-0.07*	-0.16***	0.06	0.16***	0.15***	0.20***	0.04	0.19***	/			
Rank	-0.19***	-0.18***	-0.24***	-0.01	-0.22***	-0.17***	0.22***	0.14***	0.05	0.06	0.12***	0.30***	/			
N	0.15***	-0.07*	-0.04	0.00	-0.05	-0.20***	-0.02	0.01	-0.09**	0.02	-0.03	-0.28	0.15***	/		
D	0.03	-0.05	-0.09*	0.02	-0.08	0.06	-0.05	0.05	-0.09	0.08	0.01	-0.01	-0.05	0.12**	/	

Predictions of male hierarchy by the acoustic features of their ‘wahoos’ revealed that high-ranking males produced longer ‘hoo’ syllables with higher F0s (respectively $\chi_1^2 = 14.3$, $p < .001$; $\chi_1^2 = 4.6$, $p = .031$ see Table 2). The ‘wa’ however did not differ according to male dominance rank.

Table 2. Male dominance rank predicted by the acoustic features of their ‘wahoos’¹.

	Estimate	SE	χ^2	<i>p value</i>
Intercept	0.609	0.056	/	/
‘wa’				
<i>Duration</i>	0.012	0.009	1.938	.163
<i>F0</i>	0.003	0.009	0.150	.698
<i>Intensity</i>	-0.011	0.008	2.122	.145
‘hoo’				
<i>Duration</i>	0.030	0.007	14.348	< .001
<i>F0</i>	0.016	0.007	4.613	.031
<i>Intensity</i>	0.007	0.009	0.746	.387

Modulation of acoustic features as a function of social context (see Table 3) revealed that, independently of their rank, males produced shorter ‘wa’ but longer ‘hoo’ syllables in inter-group contests, compared to intra-group aggression (respectively $\chi_1^2 = 4.4$, $p = .035$; $\chi_1^2 = 11.6$, $p < .001$), with louder ‘hoo’ syllables ($\chi_1^2 = 4.4$, $p = .035$).

¹ For each variable, the χ^2 and the p values associated from the likelihood-ratio chi-square test of the comparison between the full model and the model without the variables are given (ANOVA type II). The degrees of freedom are 1 for every tests. All subsequent tables follow the same presentation.

Table 3. Variation in the acoustic features of ‘wahoos’ as a function of social context (intra-group ‘wahoos’ is the reference factor).

	Estimate	SE	χ^2	<i>p value</i>
Intercept	-3.709	0.982	/	/
‘wa’				
<i>Duration</i>	-0.328	0.156	4.430	.035
<i>F0</i>	0.202	0.162	1.549	.213
<i>Intensity</i>	-0.232	0.126	3.400	.065
‘hoo’				
<i>Duration</i>	0.468	0.137	11.575	< .001
<i>F0</i>	0.241	0.128	3.509	.061
<i>Intensity</i>	0.408	0.176	5.382	.020

Lastly, F0 and intensity of the ‘wa’ syllable were not modulated as a function of male rank and the presence of other males (all $p > .05$). However, the duration of the ‘wa’ syllable was significantly influenced by the interaction between male rank and the presence of particular males ($\chi^2_{15} = 62.7, p < .001$). The interaction between the rank of the male vocalizing and the presence of particular males significantly influenced the duration of the ‘hoo’ syllable ($\chi^2_{14} = 26.2, p < .001$). Independently of male rank, F0 was modulated depending on the presence of specific males ($\chi^2_{33} = 50.4, p = .027$). Intensity of both the ‘wa’ and ‘hoo’ syllables was not modulated according to male rank and/or the presence of other males (all $p > .05$).

References

- Altmann, J. (1974). Observational Study of Behavior: Sampling Methods. *Behaviour*, 49(3/4), 227–267.
- Boë, L.-J., Berthommier, F., Legou, T., Captier, G., Kemp, C., Sawallis, T. R., ... Fagot, J. (2017). Evidence of a Vocalic Proto-System in the Baboon (*Papio papio*) Suggests Pre-

Hominin Speech Precursors. *PLOS ONE*, 12(1), e0169321.
<https://doi.org/10.1371/journal.pone.0169321>

Bulger, J. B. (1993). Dominance Rank and Access to Estrous Females in Male Savanna Baboons. *Behaviour*, 127(1/2), 67–103.

Fischer, J., Hammerschmidt, K., Cheney, D. L., & Seyfarth, R. M. (2002). Acoustic features of male baboon loud calls: influences of context, age, and individuality. *The Journal of the Acoustical Society of America*, 111(3), 1465. <https://doi.org/10.1121/1.1433807>

Fischer, J., Kitchen, D. M., Seyfarth, R. M., & Cheney, D. L. (2004). Baboon loud calls advertise male quality: acoustic features and their relation to rank, age, and exhaustion. *Behavioral Ecology and Sociobiology*, 56(2), 140–148. <https://doi.org/10.1007/s00265-003-0739-4>

Gustison, M. L., & Bergman, T. J. (2017). Divergent acoustic properties of gelada and baboon vocalizations and their implications for the evolution of human speech. *Journal of Language Evolution*, 2(1), 20–36. <https://doi.org/10.1093/jole/lzx015>

Kitchen, D. M., Cheney, D. L., Engh, A. L., Fischer, J., Moscovice, L. R., & Seyfarth, R. M. (2013). Male baboon responses to experimental manipulations of loud “wahoo calls”: testing an honest signal of fighting ability. *Behavioral Ecology and Sociobiology*, 67(11), 1825–1835. <https://doi.org/10.1007/s00265-013-1592-8>

Kitchen, Dawn M., Cheney, D. L., & Seyfarth, R. M. (2005a). Contextual factors mediating contests between male chacma baboons in Botswana: effects of food, friends and females. *International Journal of Primatology*, 26(1), 105–125.

Kitchen, Dawn M., Cheney, D. L., & Seyfarth, R. M. (2005b). Male chacma baboons (*Papio*

hamadryas ursinus) discriminate loud call contests between rivals of different relative ranks. *Animal Cognition*, 8, 1–6. <https://doi.org/10.1007/s10071-004-0222-2>

Kitchen, Dawn M, Seyfarth, R. M., Cheney, D. L., & Seyfarth, R. M. (2003). Factors mediating inter-group encounters in savannah baboons (*Papio cynocephalus ursinus*). *Behaviour*, 141(2), 197–218. <https://doi.org/10.1163/156853904322890816>

Kitchen, Dawn M, Seyfarth, R. M., Fischer, J., & Cheney, D. L. (2003). Loud calls as indicators of dominance in male baboons (*Papio cynocephalus ursinus*). *Behavioral Ecology and Sociobiology*, 53, 374–384.

Lameira, A. R., Hardus, M. E., Mielke, A., Wich, S. A., & Shumaker, R. W. (2016). Vocal fold control beyond the species-specific repertoire in an orang-utan. *Scientific Reports*, 6(1). <https://doi.org/10.1038/srep30315>

Pisanski, K., Cartei, V., McGettigan, C., Raine, J., & Reby, D. (2016). Voice Modulation: A Window into the Origins of Human Vocal Control? *Trends in Cognitive Sciences*, 20(4), 304–318. <https://doi.org/10.1016/j.tics.2016.01.002>

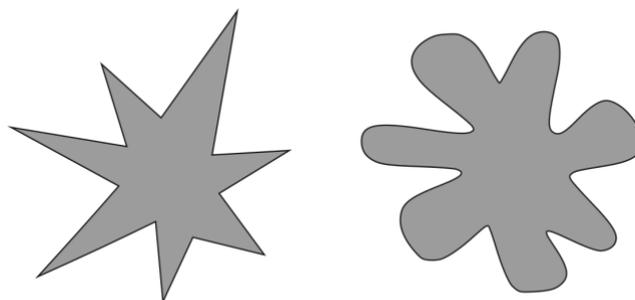
VIII. Symbolisme phonétique et sélection sexuelle

Dans la dernière partie de ce travail, nous nous sommes intéressés à des unités supérieures du langage – typiquement les sons (ou phonèmes) constituant les mots – en faisant intervenir le concept de symbolisme phonétique et de la théorie du « *code-fréquence* » dans le cadre de la sélection sexuelle.

VIII.1. Principe du symbolisme phonétique

Contrairement au principe d'arbitraire du signe linguistique selon lequel il n'existe aucun rapport naturel entre le signifié (i.e. la représentation mentale de l'objet, ou le concept) et le signifiant (i.e. l'image acoustique des mots), la notion de symbolisme phonétique suggère qu'il existe une relation naturelle et motivée entre la forme sonore des signes linguistiques et leurs sens (Svantesson, 2017). Ce principe peut être illustré à travers l'expérience classique de Köhler (1929). Dans cette étude, l'auteur a montré que les sujets avaient tendance à associer les séquences sonores « *takete* » et « *baluba* » (devenu « *maluma* » dans une étude postérieure en 1947) respectivement à une forme angulaire et circulaire (Figure 11).

Figure 11. La forme étoilée (gauche) auquel le non-mot « *takete* » est préférentiellement associé, et la forme arrondie (droite) auquel le non-mot « *baluba* » est préférentiellement associé.



Bien plus tard, Ramachandran & Hubbard (2001) ont répliqué cette expérience avec la paire de non-mots « *kiki* » et « *bouba* » et ont montré que plus de 90% des sujets anglophones américains et tamouls (locuteurs du Tamil, une langue dravidienne du sud de l'Inde) associaient la forme arrondie à la séquence « *bouba* » et la forme étoilée à « *kiki* ». De manière intéressante,

les travaux de Maurer, Pathman & Mondloch (2006) ont également révélé que les enfants sont déjà capables d'effectuer ce type d'associations dès 2.5 ans. Ces deux études suggèrent ainsi que ces associations ne sont pas dépendantes de la langue (ou socialement apprises), et qu'elles seraient plutôt universelles au niveau transculturelle.

Cette dernière décennie a vu un nombre croissant d'études mettre en évidence que ce principe d'association est effectif dans de très nombreuses langues naturelles, même lorsqu'elles sont phylogénétiquement très éloignées (Dautriche, Mahowald, Gibson, & Piantadosi, 2017 ; Levickij, 2013). Par exemple, dans une étude exploratoire sur plus de 4000 langues actuelles appartenant à plus de 350 familles linguistiques différentes, une importante similarité dans les associations sons-sens a été observée (Blasi, Wichmann, Hammarström, Stadler, & Christiansen, 2016). Par exemple, les mots renvoyant au concept de « *petit* » contiennent fréquemment la voyelle haute /i/ (e.g. « *petit* » en français, « *sghiiir* » en arabe), le mot correspondant au « *nez* » contient lui-même souvent la consonne nasale /n/ (e.g. « *niif* » en arabe, « *nose* » en anglais), de même que « *langue* » (ici l'organe) contient souvent la consonne latérale alvéolaire /l/ (e.g. « *lsen* » en arabe).

Plusieurs mécanismes peuvent expliquer ces associations sons-sens (pour une revue exhaustive à ce sujet, consulter Sidhu & Pexman, 2018). Afin d'en illustrer quelques-uns, reprenons l'exemple de « *bouba* » et « *kiki* ». Un premier mécanisme se réfère à la position des articulateurs recrutés lors de la production de ces sons, en plus particulier celle des lèvres. En effet, l'association de « *bouba* » avec la forme arrondie viendrait du fait que la bouche prend une forme effectivement plus arrondie lors de la production de la consonne bilabiale /b/ et de la voyelle arrondie /u/, tandis que l'association de « *kiki* » à la forme étoilée viendrait du fait que la bouche est étirée pour la réalisation de la voyelle d'avant non arrondie /i/. Un deuxième mécanisme consisterait en une association entre la forme des lettres et les formes visuelles (Cuskley, Simner, & Kirby, 2017). En d'autres termes, la forme des graphèmes qui

interviennent pour l'écriture graphique des mots « *bouba* » et « *kiki* » pourrait influencer la probabilité de les associer à différentes formes, en tous cas chez des sujets lettrés et connaissant l'alphabet latin. Ainsi, la forme arrondie serait associée à la séquence *bouba* grâce aux lettres « *b* » et « *o* », dont les formes sont relativement rondes, tandis que la forme pointue serait associée à « *kiki* » du fait de la forme verticale des lettres « *i* » et « *k* ». Un dernier mécanisme possible pour expliquer ces relations repose sur les propriétés articulatoires et acoustique des voyelles et des consonnes composant les mots en question, tel que prédit par la théorie « *code-fréquence* » : les sons /u/ et /i/ sont respectivement produits dans de basses et hautes fréquences, renvoyant ainsi à des représentations mentales de formes respectivement arrondie/large/grande et petite/fine/pointue (Ohala, 1984).

Bien que ces mécanismes ne soient pas mutuellement exclusifs, nous nous concentrerons ici sur celui décrit par la théorie du « *code-fréquence* », laquelle peut, à la lumière de la sélection sexuelle, offrir un cadre original pour mieux comprendre l'absence de relation arbitraire entre les formes sonores des mots et les concepts auxquels ils renvoient. Dans ce dernier chapitre, nous tenterons donc de montrer que ces deux paradigmes peuvent être appliquées de façon originale à un corpus onomastique¹ composé des prénom. Nous verrons que leur fonction sociale relativement importante permet de tester l'hypothèse d'une traduction phonétique du dimorphisme sexuel au niveau de leur composition sonore. Pour terminer, nous verrons dans quelle mesure le symbolisme phonétique permet également de traduire les caractéristiques d'un corpus onomastique de créatures fictives issues d'un jeu vidéo populaire.

VIII.2. La théorie « *code-fréquence* »

Dans le cadre du symbolisme phonétique, les propriétés articulatoires et acoustiques des voyelles et des consonnes conduisent à des associations sons-sens non arbitraires. La théorie du « *code-fréquence* » telle que formulée par Ohala (1984) repose ainsi sur l'idée que les mots qui

¹ L'onomastique correspond à l'étude des noms propres, dont leur étymologie, leur formation et leur usage à travers le temps.

renvoient au concept ou à l'image mentale de « *petitesse* » présentent une incidence élevée de voyelles et de consonnes caractérisées par des fréquences acoustiques élevées. A l'inverse, les mots désignant la « *grandeur* » et/ou la « *largeur* » contiennent plus souvent des voyelles et des consonnes caractérisées par des fréquences acoustiques basses. Quelques-uns de ces exemples incluent l'anglais (« *teeny* », « *wee* », « *itsy bitsy* » contre « *large* », « *humongous* »), le français (« *petit* », « *fin* », « *mince* » contre « *grand* », « *large* », « *gros* »), le grec (« *mikros* » contre « *makros* ») ou encore le japonais (« *tchiisai* » contre « *ooki* »). En outre, Ohala (1984) suggère que les fondements du symbolisme phonétique proviennent des associations vocalisations-perceptions observées dans de nombreuses espèces (y compris l'humain) dans lesquelles l'agresseur produit des vocalisations de basses fréquences pour apparaître plus large, plus grand et menaçant, à l'inverse du subordonné qui produit des vocalisations de hautes fréquences afin de paraître plus petit et soumis. Ohala (1984) suggère ainsi que les mots possédant des sons de hautes et basses fréquences pourraient être utilisés pour projeter ces mêmes impressions.

Au regard des caractéristiques articulatoires et acoustiques des sons de la parole humaine (cf. section II.1), nous pouvons ainsi nous attendre à ce que les mots contenant respectivement des voyelles et des consonnes produites dans les basses fréquences, seront perceptuellement associés à la largeur, la masculinité et à la dominance, et à l'inverse des voyelles et des consonnes produites dans les hautes fréquences seront perceptuellement associés à la petitesse, la féminité et la soumission.

VIII.3. Dimorphisme sexuel et caractéristiques phonétiques des prénoms

Les prénoms pourraient constituer de bons candidats pour étudier le symbolisme phonétique dans le cadre de la théorie du « *code-fréquence* » et de la sélection sexuelle. Selon les hypothèses que nous avons formulées, nous suggérons que les prénoms masculins devraient exhiber une incidence plus élevée de phonèmes produits dans les basses fréquences, du fait des impressions auditives de masculinité et de dominance associées aux voix graves,

caractéristiques attractives bien établies chez les hommes (Bruckert et al., 2006 ; Collins, 2000 ; Feinberg et al., 2005/2008 ; Hodges-Simeon et al., 2010 ; Hughes et al., 2010 ; Jones et al., 2010 ; Pisanski & Rendall, 2011 ; Puts et al., 2006/2007 ; Vukovic et al., 2008 ; Xu et al., 2013). A l'inverse, les prénoms féminins devraient exhiber une incidence plus élevée de phonèmes produits dans les hautes fréquences, via les liens observés entre voix aiguës, attractivité et féminité chez les femmes (du moins telle que suggéré par une majorité des études, Borkowska & Pawlowski, 2011 ; Collins & Missing, 2003 ; Feinberg, 2008 ; Jones et al., 2010 ; O'Connor et al., 2013 ; Puts et al., 2013 ; Re et al., 2012).

Toutefois, pourquoi devrions-nous nous attendre à un dimorphisme phonétique marqué entre les prénoms masculins et féminins ? Bien que les parents ne recherchent pas volontairement un prénom sonnante comme grand, petit ou dominant pour leur progéniture, ils pourraient toutefois manifester une préférence inconsciente pour un prénom plutôt masculin ou féminin en fonction du sexe de leur enfant. Nous suggérons que ce comportement peut s'expliquer par deux phénomènes. Premièrement, la dénomination des objets de l'environnement est un outil important de catégorisation dans l'espèce humaine (Rosch, 1999). Dans le contexte où le sexe d'un individu est l'une des caractéristiques les plus saillantes immédiatement traitée lors des interactions sociales, son marquage via une utilisation différentielle du matériel phonétique pour les prénoms des hommes et des femmes peut présenter des avantages en termes d'efficacité cognitive. En effet, si l'existence d'un dimorphisme phonétique au niveau des prénoms attribués aux petites filles vs. petits garçons est vérifiée, les individus peuvent rapidement (et en l'absence de différences morphologiques évidentes) déduire les propriétés biologiques et sociales associés à la catégorie « *sexe* ». Ce savoir implicite permettrait en retour aux individus d'adapter leurs attentes et leurs comportements par rapport aux propriétés de l'individu (et d'autant plus lorsque la personne n'a pas encore été rencontrée). Deuxièmement, les prénoms masculins et féminins revêtent une

importance cruciale dans le renforcement de l'identité sexuelle d'un individu et plus globalement de son rôle attendu au sein d'une société (Pilcher, 2016).

En outre, les prénoms ont une influence relativement importante sur de nombreux aspects de la vie sociale des individus. Par exemple, il a été rapporté que les prénoms pouvaient impacter la perception de divers aspects physiques dont le visage et le corps (Erwin, 1993 ; Hassebrauck, 1988 ; Hensley & Spencer, 1985 ; Klenovsak, Hartung, Santiago, & Zaefferer, 2018 ; Perfors, 2004 ; Zwebner, Sellier, Rosenfeld, Goldenberg, & Mayo, 2017), l'attribution de certains traits de personnalité comme l'intelligence, la compétence et l'extraversion (Leirer, Hamilton, & Carpenter, 1982 ; Mehrabian, 2001 ; Young, Kennedy, Newhouse, Browne, & Thiessen, 1993), modifier les attitudes et les comportements des conspécifiques (Figlio, 2007 ; Pelham, Mirenberg, & Jones, 2002), influencer la désirabilité sociale (Busse & Seraydarian, 1978; Gebauer, Leary, & Neberich, 2012) et le succès dans différents contextes sociaux (Cotton, O'Neill, & Griffin, 2008 ; D. Figlio, 2005 ; Harari & McDavid, 1973). Plus précisément, plusieurs études ont montré que le simple fait d'avoir seulement accès à l'information d'un prénom masculin ou féminin influence la probabilité d'être recruté – avec un avantage significatif pour les individus porteurs d'un prénom masculin (Kasof, 1993 ; Moss-Racusin, Dovidio, Brescoll, Graham, & Handelsman, 2012 ; Steinpreis, Anders, & Ritzke, 1999). De même, Bruning, Polinko, Zerbst, & Buckingham, (2000) ont rapporté que des juges estiment que la réussite professionnelle est plus probable chez les hommes dont les prénoms sonnent comme plus masculins et chez les femmes dont les prénoms présentent des sonorités plus féminines.

Quelques études suggèrent également que la composition phonétique d'un prénom peut influencer la perception physique que l'on se fait d'une personne. Par exemple, deux études ont examiné comment certains phonèmes spécifiques aux prénoms allemands et anglais peuvent influencer l'attractivité faciale. Dans le cas des prénoms allemands, les visages masculins sont

perçus comme plus attractifs lorsqu'ils sont associés à des prénoms possédant des voyelles d'arrière (i.e. voyelles graves) (Klenovsak et al., 2018). À l'inverse, les prénoms féminins composés de voyelles d'avant (i.e. voyelles aiguës) augmentent l'attractivité perçue d'un visage féminin. Des résultats contraires ont été trouvés pour les prénoms anglais (Perfors, 2004), mais montre toutefois que la composition phonétique de ceux-ci ont une influence sur l'attractivité d'un visage. Enfin, il a été montré que certains phonèmes peuvent aussi influencer la perception de la forme du corps. Par exemple, en considérant les consonnes /b/, /l/, /m/, /n/ et les voyelles /u/, /o/ et /ɒ/ (e.g. « *Molly* ») comme des voyelles « rondes » (analogie liée à leur propriété articulatoire impliquant les lèvres), et les consonnes /k/, /p/ et /t/ et les voyelles /i/, /e/, /ɛ/ et /ʌ/ (e.g. « *Luana* ») comme « pointues », Sidhu & Pexman (2015) ont montré que des sujets canadiens sont plus susceptibles d'associer une silhouette ronde avec un prénom contenant des sons arrondis (i.e. labialisés ; e.g. « *Molly* », « *Bob* »), et à l'inverse une silhouette plutôt fine et longue avec des prénoms composés de sons « pointus » (e.g. « *Kate* », « *Kirk* »). Les mêmes résultats ont été observés avec des locuteurs français (Sidhu, Pexman, & Saint-Aubin, 2016). Enfin, en utilisant la même nomenclature de phonèmes, Barton & Halberstadt (2018) ont montré que des sujets américains considèrent les prénoms contenant des voyelles « rondes » et « pointues » comme convenant respectivement mieux à des visages ronds et anguleux. De manière plus intéressante encore, les auteurs ont également montré que les candidats aux élections sénatoriales américaines gagnaient 10% de vote en plus sur la seule base d'une congruence prénom-visage très forte vs. très faible.

Dans ce contexte, quelques études préliminaires ont mis en évidence un biais sexuel phonétique dans la composition des prénoms anglophones, qui dépend en partie de la syllabe accentuée (Cutler, McQueen, & Robinson, 1990 ; Pitcher, Mesoudi, & McElligott, 2013 ; Slater & Feinman, 1985 ; Slepian & Galinsky, 2016). En effet, l'accentuation permet de mettre en relief une des syllabes (i.e. la syllabe tonique) d'un mot laquelle est dès lors perceptivement

proéminente par rapport aux autres (celles-ci sont appelées syllabes atones). En anglais, la syllabe tonique tombe généralement sur la première syllabe des noms dissyllabiques, ou sur la deuxième syllabe des noms trissyllabiques ; en revanche, dans le français, c'est toujours la dernière syllabe des mots qui est accentuée. Dans ce contexte, Cutler et al. (1990) ont ainsi montré que les prénoms féminins anglophones sont plus susceptibles de contenir la voyelle haute /i/ (i.e. voyelle aiguë) dans la première syllabe des prénoms dissyllabiques (e.g., « *Michelle* », « *Tina* ») et/ou sur la deuxième syllabe (e.g. « *Christine* », « *Elizabeth* », « *Patricia* »). Cette étude a également révélé que les voyelles de basses fréquences (voyelles d'arrière comme /u/ et/ou voyelles basses comme /o/) sont beaucoup moins fréquentes dans les prénoms féminins. Ces résultats sont soutenus par ceux de Pitcher et al. (2013) qui ont procédé à une analyse d'un corpus de prénoms britanniques, australiens et américains les plus populaires entre 2001 et 2010. Les auteurs ont montré que les voyelles antérieures telles que /i/ ou /e/ (i.e. voyelles aiguës) sont principalement attestées dans les prénoms féminins (e.g. « *Emily* ») et les voyelles d'arrière telles que /u/ ou /o/ (i.e. voyelles graves) dans les prénoms masculins (e.g. « *Thomas* »). Enfin, Slepian & Galinsky (2016) ont établi que les prénoms masculins anglophones et indiens ont plus souvent tendance à exhiber en position initiale un son voisé (i.e. son plus grave qu'un son non-voisé) comparativement aux prénoms féminins. Ils ont également montré que la présence d'un phonème voisé en position initiale du prénom augmente significativement la perception de masculinité, et ceci indépendamment du genre communément associé à celui-ci.

Enfin, notons que l'évolution culturelle joue un rôle évidemment important dans l'augmentation et la diminution des fréquences d'utilisation des différents prénoms masculins et féminins (Bentley, Hahn, & Shennan, 2004 ; Berger, Bradlow, Braunstein, & Zhang, 2012). Bien que les contraintes culturelles, la tradition familiale, la mode et l'euphonie jouent un rôle primordial dans le choix des prénoms que les parents attribuent à leur enfant, ces facteurs ne

semblent déterminer que le choix à l'intérieur d'un ensemble essentiellement marqué pour chaque sexe. A ce propos il convient de remarquer qu'il n'existe pas – à notre connaissance – de société faisant un usage exclusif de prénoms non marqués sexuellement (i.e. prénoms épicènes ou unisexes).

VIII.4. Vers un autre corpus onomastique : le cas des Pokémon

Si le symbolisme phonétique se retrouve dans certains corpus onomastiques tels que les prénoms (Pitcher et al., 2013), le noms latins des espèces (e.g. le /i/, le /t/ et le /k/ sont préférentiellement associés aux noms d'écureuils dans les langues amérindiennes, Berlin, 2006) ou encore les surnoms des serveuses japonaises (Shinohara & Kawahara, 2013)¹, un autre corpus intéressant est celui des noms attribués aux créatures fictives provenant du jeu vidéo Pokémon, créé en 1996 par la compagnie Nintendo. L'objectif principal du jeu est de capturer toutes sortes de créatures fictives portant toutes des noms uniques et des caractéristiques spécifiques en termes de puissance (ou force), de poids et de taille. Certaines de ces créatures ont également la capacité d'évoluer en de nouvelles versions d'elles-mêmes, ce qui accroît leurs caractéristiques et conduit à l'émergence d'une nouvelle dénomination (e.g. Stade 1 : « *Fantominus* » > Stade 2 : « *Spectrum* » > Stade 3 : « *Ectoplasma* »).

Extrêmement populaire au début des années 2000, le jeu a donné lieu à de nombreuses traductions dont le français, l'anglais, l'allemand ou encore le mandarin, ce qui offre une excellente opportunité d'étudier le symbolisme phonétique dans une perspective interculturelle. En effet, si les facteurs non linguistiques sont maintenus constants (i.e. les propriétés physiques de force, de taille et de poids), les facteurs linguistiques varient de manière significative d'une langue à l'autre (i.e. noms uniques pour chaque langue, Shih, Ackerman, Hermalin, Inkelas, & Kavitskaya, 2018). Bien qu'il existe une petite proportion de créatures iconiques dont le nom

¹ Les serveuses dites « *tsun* », caractérisées comme plutôt « *inaccessibles* » et « *crues* » ont des surnoms présentant une plus grande incidence d'obstruantes (e.g. /s/, /t/, /k/). Les serveuses dites « *honwaka*, *moe* », caractérisées comme plutôt « *accessibles* » et « *douces* » ont des noms présentant une plus grande incidence de sonantes (e.g. /m/, /n/).

ne change pas (notamment « *Pikachu* »), chaque créature possède un nom différent selon la langue/culture considérée. Par exemple, la créature représentant un fantôme se nomme « *Fantominus* » en français, « *Gastly* » en anglais, « *Nebulak* » en allemand et « *Gōsu* » en japonais. De plus, la traduction des noms ne semble pas résulter directement d'une simple transcription littérale du japonais. En effet, leur construction repose surtout sur le champ lexical général auquel la créature peut être associée et/ou certaines expressions figées de la langue considérée afin de refléter au mieux la culture sous-jacente à la langue en question (pour une discussion détaillée des liens langues/cultures et de leurs conséquences en traduction voir Brougère, 2004). Par exemple, « *Tartard* » en français est une créature aquatique dont le nom provient de la fusion du mot « *têtard* » et du mot « *tarte* » (au sens de « *claque* »). En anglais, cette même créature se nomme « *Poliwrath* », mot valise construit à partir des termes « *polliwog* » (têtard) et « *wrath* » (colère) ; en allemand il s'appelle « *quappo* », construit à partir « *kaulquappe* » (têtard) et de « *protzen* » (frimer) ; enfin, en japonais, il se nomme « *Nyorobon* », qui est une référence aux monstres géants de la fiction japonaise. Un autre exemple est celui de « *Canarticho* », un oiseau dont le nom français est construit à partir de « *canard* » et de « *artichaut* », « *Farfetch'd* » en anglais (une expression signifiant « *tiré par les cheveux* »), « *Porenta* » en allemand, construit à partir de « *porree* » (poireau) et de « *ente* » (canard), et « *Kamonegi* » en japonais, construit à partir de « *kamo* » (canard) et « *negi* » (oignon vert ou ciboule). Dans ce contexte, si le phonétisme symbolique est effectif au niveau des noms afin de renvoyer à des créatures plus ou moins fortes, grandes et larges, certains motifs sonores (e.g. présence de sons de basses fréquences pour référer à des créatures particulièrement fortes) pourraient être communs aux différentes langues. L'intérêt de ce corpus réside aussi dans l'existence d'une longue liste d'éléments lexicaux (plus de 800 noms propres par langue).

Plus récemment, deux études ont montré que le symbolisme phonétique était effectif dans les noms des Pokémon en japonais et en anglais (Kawahara, Noto, & Kumagai, 2018 ;

Shih et al. , 2018). Dans l'étude japonaise, les auteurs ont constaté que le nombre de mores (i.e. unités rythmiques comparable aux syllabes mais dont la nature repose sur le poids) et les consonnes voisées (e.g. /b/, /d/, /g/, /z/) sont positivement corrélées à la puissance, la taille, le poids et le niveau d'évolution du Pokémon. De même, les créatures dont les noms présentent des voyelles fermées (e.g. /i/, /u/) ont tendance à être plus petites et plus légères. Dans l'étude conduite en langue anglaise, les auteurs ont constaté que le nombre de phonèmes et les voyelles d'arrière et/ou ouvertes (e.g. /a/, /ɔ/) sont positivement corrélé au poids. La taille du Pokémon étant quant à elle est positivement corrélée avec la présence de consonnes alvéolaires (e.g. /s/) et négativement associée à la présence de voyelles fermées (e.g. /i/, /u/). En outre, les caractéristiques de puissance semblent positivement corrélées à la longueur du nom et à la présence de plusieurs consonnes alvéolaires. Enfin, les Pokémon les plus évolués ont également tendance à avoir des noms plus longs en termes de nombre de phonèmes.

L'ensemble des résultats présentés met ainsi en évidence des similarités mais également des différences pertinentes dès lors que l'on cherche à mieux comprendre le principe du symbolisme phonétique dans une perspective translinguistique et transculturelle. Par exemple, en japonais, les consonnes voisées et les voyelles fermées semblent véhiculer les mêmes impressions physiques que les voyelles postérieures et ouvertes de l'anglais. De même, la longueur des noms semble en être un bon prédicteur de la force des Pokémon et ce, indépendamment de la langue. Mais les résultats des deux études montrent également un nombre important de différences, suggérant que d'autres contraintes, liées ou non aux systèmes phonologiques des langues étudiées, entrent également en jeu dans les processus de créations des noms.

VIII.5. Conclusion

L'étude du symbolisme phonétique des prénoms et de ses possibles conséquences sur la perception de divers aspects physiques et sociaux des individus qui les « *arborent* » apparaît

comme un champ de recherche très stimulant (Sidhu & Pexman, 2019). Plusieurs lignes de recherches originales pourraient être développées sur ce sujet précis pour notamment mettre en évidence l'existence universelle d'un dimorphisme sexuel dans la composition sonore des prénoms. D'autres recherches visant à mesurer l'effet et le poids de cette information phonétique sur la perception de l'attractivité, et plus globalement sur la fitness d'un individu, pourraient être également envisagée. Enfin, le corpus onomastique de l'univers Pokémon constitue une excellente opportunité d'étudier le symbolisme phonétique dans une perspective interculturelle.

Article 5

Sex-biased sound symbolism in French first names

Alexandre Suire, Alba Bossoms Mesa, Michel Raymond & Melissa

Barkat-Defradas

Article publié dans *Evolutionary Human Sciences*

Résumé

Plusieurs études ont mis en évidence que le symbolisme phonétique était effectif dans les prénoms anglophones. De plus, la mise en lumière de ces motifs semble être compatible avec la théorie du « *code-fréquence* », dans le sens où les prénoms masculins et féminins exhibent respectivement des phonèmes de basses et hautes fréquences pour projeter perceptivement des impressions de grande taille corporelle/masculinité et de petite taille corporelle/féminité. Toutefois, des études supplémentaires dans d'autres langues/cultures sont nécessaires pour mettre en évidence que les prénoms suivent cette même tendance.

Dans ce contexte, nous avons étudié le symbolisme phonétique dans les prénoms français. Nous nous attendions à ce que les prénoms masculins exhibent une incidence plus élevée de phonèmes de basses fréquences (i.e. voyelles d'arrière et consonnes voisées), tandis que les prénoms féminins exhibent une incidence plus élevée de phonèmes de hautes fréquences (i.e. voyelles d'avant et consonnes non voisées), une telle différence devant se situer dans la syllabe perceptuellement proéminente, la dernière en français.

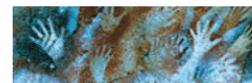
A cette fin, nous avons étudié les 100 prénoms féminins et masculins les plus fréquemment attribués pour chaque décennie entre 1900 et 2010. Après la transcription phonétique, nous avons exclu les prénoms épïcènes (donné au deux sexes), ceux composés (construit à partir de deux prénoms féminins/masculins ou de l'un avec l'autre), les monosyllabiques (afin de comparer la première et la dernière syllabe), et conservé une seule version phonétique d'un prénom (homophones non-homographes)¹.

Toutes périodes confondues, nos résultats ont montré que sur la dernière syllabe les prénoms masculins sont plus susceptibles de contenir des voyelles d'arrière telles que /o/ (e.g. « *Arnaud* ») et des voyelles nasales tel que /ã/ (e.g. « *Roland* »), tandis que les prénoms

¹ Les prénoms épïcènes, composés, homophones non-homographes et monosyllabiques représentent en cumulé environ ~21% de l'ensemble des prénoms étudiés.

féminins sont plus susceptibles d'exhiber des voyelles d'avant telles que /i/ (e.g. « *Marie* », « *Cécile* ») ou /ɛ/ (e.g. « *Claire* », « *Odette* »). De manière inattendue, les prénoms masculins ont exhibé une incidence plus élevée de constrictives non voisées dans leur dernière syllabe, telle que /s/ (e.g. « *Francis* ») et des occlusives voisées dans la première syllabe telle que /b/ (e.g. « *Bernard* »). Nos résultats semblent plutôt bien compatibles avec la théorie du « *code-fréquence* ».

Enfin, une analyse temporelle de la variation de ces différents motifs a révélé que les prénoms féminins avaient tendance à se « *masculiniser* » phonétiquement à partir des années 1960 (i.e. augmentation des prénoms avec /ɔ̃/ comme « *Marion* », ou le /u/ de « *Lou* »). Dans ce même contexte, un cas peut être plus spécifique concerne la proportion des prénoms féminins se terminant par /a/, voyelle dont la fréquence fondamentale intrinsèque et la dispersion formantique sont relativement bas. De manière intéressante, ce phénomène concorde temporellement avec la « *masculinisation* » de plusieurs autres traits, notamment les préférences des hommes pour le ratio taille-hanche, et la perception individuelle de masculinité/féminité et d'assurance chez les femmes.



RESEARCH ARTICLE

Sex-biased sound symbolism in French first names

Alexandre Suire*[†], Alba Bossoms Mesa[†], Michel Raymond and Melissa Barkat-Defradas

ISEM, University of Montpellier, CNRS, EPHE, IRD, Montpellier, France

*Corresponding author. E-mail: alexandre.suire@umontpellier.fr

Summary

Low- and high-frequency vowels in the stressed syllable of French first names may respectively project impressions of largeness/masculinity and smallness/femininity.

Abstract

Given that first names can have a lifelong impact on the bearer, parents should choose a name based on the impressions they want their offspring to evoke in other people. This name-to-mental-image association can be mediated through sound symbolism: a natural link between the sounds and meaning of a word. From an evolutionary perspective, parents should pick names which sounds convey traits advantageous in human sexual selection: largeness and masculinity for males through lower-frequency sounds as opposed to smallness and femininity for females through higher-frequency sounds. Using a database of French first names from 1900 to 2009, we observed a sex-biased sound symbolism pattern in the last syllable, which is the perceptually prominent one in French. Male names were more likely to include lower-frequency vowels (e.g. /o/ , /ã/) and female names higher-frequency vowels (e.g. /i/, /e/). Unexpected patterns in consonants were observed in masculine names with higher-frequency sounds (e.g. /s/, /ʃ/) in the last syllable and lower-frequency sounds (e.g. /b/, /g/) in the first syllable. However, little variance was explained and the modest size effect suggests that cultural traits influence these sex differences. Lastly, exploratory analyses revealed a phonetic masculinization in women's first names that has increased since the 1960s.

Keywords: Sound symbolism; first names; femininity; masculinity; voice

Introduction

Arbitrariness, the notion that the sound and the meaning of a word are independent, has long been considered one of the most widely shared principles in linguistics. However, a growing body of evidence challenges this view, stating that there is a natural link between the sound units of a word – known as phonemes – and the mental image they evoke (see Svantesson 2017 for an overview). This principle, referred to as sound symbolism, is well illustrated by the ‘kiki-bouba’ and ‘maluma-takete’ experiments, in which participants are asked to associate such non-words to two figures of different shapes: results show above-chance matchings of ‘bouba’ and ‘maluma’ with a round silhouette, and ‘kiki’ and ‘takete’ with a sharp one (Ramachandran and Hubbard 2001; Werner 1957; Köhler 1947). Although it is uncertain to generalize the ‘kiki-bouba’ effect across cultures (see Bremner et al. 2013 and Cuskley et al. 2017), other similar sound-meaning mappings have been recorded in thousands of the world's languages, suggesting an underlying universal cognitive association mechanism (Blasi et al. 2016). Sexual selection for body size offers one possible explanation for why sound symbolism might be so ubiquitously distributed.

[†]A.S. and A.B.M. are both first authors.

© The Author(s) 2019. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

The first clue was provided by the ‘motivation-structural rules’ theory (Morton 1977), after observing that many animals modulate their vocalizations during competitive encounters: they use low-pitched vocalizations when their intention is to be threatening and dominant, and high-pitched vocalizations if they wish to appear conciliatory or submissive. The hypothesized reason is that the frequency of vocalizations reflects a projection of the individual’s body size, a key determinant in the outcome of physical contests but also courtship interactions (Bradbury and Vehrencamp 2011). This notion was then extended to humans in the ‘frequency-code’ theory (Ohala 1984), which provides a plausible explanation for the observed vocal dimorphism in human voices. Before puberty, boys and girls exhibit similar vocal frequencies, until males experience a significant enlargement of their larynx and vocal folds under the influence of androgens, which lowers their vocal pitch and resonant frequencies to the point that they practically do not overlap with those of adult females (Titze 1989). Such findings hint towards the action of sexual selection and can be interpreted as a result of different selective pressures acting on each sex (Puts 2010). In males, lower-frequency voices could have been favoured within intra-sexual contests because they are perceptually associated with largeness (Pisanski *et al.* 2014; Xu *et al.* 2013; Pisanski and Rendall 2011; Rendall *et al.* 2007; van Dommelen and Moxness 1995), more masculine and more socially and physically dominant men (Hodges-Simeon *et al.* 2014; Puts *et al.* 2006, 2007; Xu *et al.* 2013; although see Armstrong *et al.* 2019 for why voice pitch may not be an honest signal of dominance). In contrast, higher frequencies in female voices could have been selected in mate-choice dynamics as such frequencies were shown to be associated with perceived smallness, femininity and more attractive women (Xu *et al.* 2013; Fraccaro *et al.* 2011; Puts *et al.* 2011; Jones *et al.* 2010; Feinberg *et al.* 2008; Collins and Missing 2003).

Although naming practices are assumed to be highly driven by sociocultural factors, few studies have underpinned the ultimate causes that have driven most male and female names to not overlap phonetically (Pitcher *et al.* 2013). As observed for other dimorphic traits in humans such as the body size and stature (Geary 1998; Puts 2010), one can reasonably assume that these two different sexual selective pressures on human voices could have driven the attested sexual phonetic dimorphism. Preliminary evidence has shown that across languages as diverse as English, Japanese, Chinese, Korean and several Native American and Australian languages, high- and low-frequency vowels are respectively associated with perceived smallness and largeness (Haynie *et al.* 2014; Shinohara and Kawahara 2010; Ultan 1978; Newman 1933; Sapir 1929), as well as perceived femininity and masculinity (Wu *et al.* 2013; Klink 2000). Thus, indexical cues that are known to be relevant to human mating (e.g. body size, masculinity and femininity) may be conveyed or projected in first names through sound symbolism, using an array of different phonemes that can differ in their intrinsic fundamental frequency (i.e. the perceptual correlate of pitch), formant frequencies (i.e. resonances of the vocal tract) and their dispersion (i.e. a proxy of the vocal tract length) (Knoeferle *et al.* 2017; Ohala 1994; Ultan 1978).

Although parents may not volitionally seek a large or small, dominant and attractive sounding name for their offspring, they might display an unconscious preference for either a more masculine or feminine name to suit their child’s sex. This behaviour can be explained by the fact that gendered naming is an important tool of categorization in humans. Indeed, sex is one of the most pervasive characteristics individuals first infer when interacting with others: distinguishing it by using different phonetic material for first names may find benefits in that it increases cognitive efficiency by allowing individuals to rapidly infer properties of sex category, even with little or no first-hand experience with that person. In turn, it enables individuals to tailor their expectations about the behaviours and capacities linked to the biological composition of that individual. Additionally, masculine and feminine names take on great importance in the reinforcement of an individual’s sexual identity and gender role (Pilcher 2016). Although first names are not inherited and no studies have yet tackled the issue of their influence on reproductive success, it has been reported that first names can impact their bearers on several aspects: their physical perception (Zwebner *et al.* 2017; Hartung 2018; Perfors 2004; Erwin 1993; Hassebrauck 1988; Hensley and Spencer 1985), inferences on personality (Mehrabian 2001; Mehrabian and Piercy 1993; Leirer *et al.* 1982), attitudes and behaviours (Figlio 2007; Pelham *et al.* 2002), social desirability (Gebauer *et al.* 2012; Busse and Seraydarian 1978) and

social outcomes (Cotton et al. 2008; Figlio 2005; Hodson and Olson 2005; Harari and McDavid 1973). Thus, it can be suggested that this cognitive bias could interfere during the naming process, since the phonetic peculiarities of forenames may underline and reinforce the perceptual associations of the biological and social characteristics linked to each sex through sound symbolism, which ultimately might be relatively important towards competitors and potential mates. Furthermore, to our knowledge, no societies (industrialized or not) currently use, or have been using, the same set of names for males and females. Lastly, it is worth noting that even though cultural evolution drives popularity and the emergence of novel names (e.g. Berger et al. 2012), it merely explains why individuals primarily perceive them as either male or female.

Sound symbolism has already been observed in the phonetic composition of English first names (Sidhu and Pexman 2015; Pitcher et al. 2013; Cassidy et al. 1999; Cutler et al. 1990). So far, only one study has formally tested these evolutionary hypotheses through the lens of sexual selection using a database of the thousand most popular English, American and Australian first names between 2001 and 2010 (Pitcher et al. 2013). In accordance with the evolutionary predictions, high-frequency vowels such as /i/ or /e/ were mostly attested in female names and low-frequency ones such as /u/ or /o/ in male names. Such differences were found on the first syllable, where stress is generally located and which is consequently perceptually prominent in English. However, the authors did not investigate consonant patterns nor take a look on the last syllable to ensure that no phonetic dimorphism was also present there.

The goal of the present study is to quantify the hypothesized phonetic dimorphism of male and female names, using a large sample size of popular first names in France that extends over the last century. In this context, this study extends the results that have already been observed in English first names. However, two major differences exist between French and English. First the lexical stress falls on the last syllable in French and most of the time on the first syllable in English. Secondly, all phonological units are not equally represented in French and English. For example, nasal vowels are attested in the former but absent in the latter. Moreover, analyses can be expanded by including consonants, for which patterns of sound symbolism have been previously reported (Nielsen and Rendall 2013; Maurer et al. 2006). Consequently, we expect to find sex-bias sound symbolic patterns in the phonemes of the stressed syllable in French names, namely back and nasal vowels and voiced consonants in male names, as they are produced at lower frequencies, as opposed to front vowels as well as voiceless consonants in female names, since their articulation produces noise in relatively higher frequencies (Knoeferle et al. 2017; Ohala 1994; Ultan 1978). Lastly, we will conduct exploratory analyses of the temporal variations of these sound symbolic patterns from 1900 to 2009 in order to examine whether they have remained constant or have evolved over time for each sex.

Material and methods

Data pre-treatment

Data was retrieved on September 2014 from the Institut National de la Statistique et des Études Économiques. We selected the most popular 100 female and 100 male names for each decade, ranging from 1900–1909 to 2000–2009. In order to control for population size, popularity was estimated by calculating the annual ranking position of each name and adding these up per decade. Although this approach excludes rare names, it captures naming practices properly for a given decade (Pitcher et al. 2013).

All retrieved names were subsequently transcribed independently by two native French-speaking phoneticians, following the International Phonetic Alphabet principles. When no agreement arose for certain transcriptions or when pronunciation was unknown, different web sources were used (e.g. <https://fr.wiktionary.org/wiki>). For each syllable of a name, we recorded the following articulatory features:

- The vowel place of articulation, which corresponds to the position of the tongue in the oral cavity during its articulation. As the tongue is closer to the lips, the sounds produced have an overall

higher frequency spectrum (i.e. front vowels such as /i/). Conversely, sounds that are produced with the tongue retracted at the back of the mouth (i.e. back vowels such as /u/) have an overall lower spectral distribution. Central vowels (i.e. /a/) correspond to a position where the tongue is placed in the middle of the mouth. Acoustically, vocalic frontness and/or backness correspond to the frequencies of the second formant (i.e. the spectral peaks of the sound spectrum). The vowel height, which corresponds to the degree of aperture of the mandible (i.e. the open/close dimension, corresponding acoustically to the first formant), was not retained here, as it would produce redundant information with vowel articulation (i.e. multicollinearity in the statistical analyses).

- The vowel's nasality, which is determined by the low position of the velum during articulation, leads the air to flow through the nose as well as the mouth. This extra resonance, which results from the intervention of the nasal cavity during phonation, lowers the frequency of the sound in comparison to its non-nasal counterpart. Note that only one type of vowel (oral or nasal) can be found in each syllable.
- The consonant's manner of articulation, which is determined by the way the airflow escapes from the vocal tract during articulation. Here, we focused on plosives, which are produced by a complete closing of the airflow that causes its blocking before the air is suddenly released. This type of sound produces a burst noise that is typical of consonantal stops. We also focused on fricatives, which are produced with a major constriction of the airflow, which acoustically causes a turbulent noise. Owing to their manner of articulation, plosives generally produce lower frequencies than fricatives.
- The consonant's voicing, which is determined by whether the vocal folds vibrate or not during articulation. This new source of laryngeal noise explains why voiced consonants are lower in frequencies than voiceless ones.

All phonemes coded with examples of first names are given in [Table 1](#).

Statistical analyses

Analysis on sound symbolism

The aim of this analysis is to study and quantify sex differences in first names' phonetic composition. According to our predictions, we expect to find in the stressed syllable of male names either back or nasal vowels and voiced consonants, as opposed to front, non-nasal vowels and unvoiced consonants in female names. In order to test these predictions, we aggregated all of the first names spanning over the century, giving only one list of first names (e.g. 'Marie' was found in several decades). Only one version of phonetically equivalent names in each sex was collated (e.g. 'Danielle' and 'Daniele', homophones non-homographs, i.e. names pronounced alike but not written alike). Compound names (e.g. 'Jean-Marie', 'Marie-Pierre') were discarded as they represent a particular set of names mostly composed of a masculine name joint to a feminine name. Monosyllabic names were also discarded from the analysis because it would preclude comparing the first and last syllable. This resulted in a sample size of 275 female and 197 male popular unique names distributed across the century. A generalized linear model was then used to investigate the existence of sex-biased sound symbolic patterns in French male vs. female names. Because the response variable 'sex' was binary, a binomial distribution with a logit link function was specified. The explanatory variables were the articulatory features aforementioned, each repeated for the first and the stressed last syllable:

- The vowel's place of articulation – a fixed factor with three modalities (i.e. front, central or back vowel).
- The vowel's nasality – a fixed factor with two modalities (i.e. nasal and non-nasal vowel).
- Counts of voiced and unvoiced consonants (plosives and fricatives) – covariates that were standardized.

Table 1. Examples of first names for each phoneme investigated (underlined)

Type of phoneme	Phonemes	Frequency domain	Name examples
Front vowels	/i/, /y/, /e/, /ø/, /ɛ/	High	Marie, <u>Luc</u> , Cécile, <u>Eugène</u> , Odette
Central vowels	/a/, /ə/	Central	<u>Jeanne</u> , <u>Denise</u>
Back vowels	/u/, /o/, /ɔ/	Low	<u>Lou</u> , <u>Renaud</u> , <u>Paul</u>
Nasal vowels	/ã/, /ẽ/, /õ/	Low	<u>Antoine</u> , <u>Sylvain</u> , <u>Raymond</u>
Voiced plosives	/b/, /d/, /g/	Low (voicing) Low (manner of articulation)	Nor <u>bert</u> , <u>Claude</u> , <u>Guy</u>
Voiced fricatives	/ʒ/, /v/, /ʁ/, /z/	Low (voicing) High (manner of articulation)	<u>Jean</u> , <u>Valérie</u> , <u>Suzanne</u> , <u>Claire</u>
Voiceless plosives	/p/, /t/, /k/	High (unvoiced) Low (manner of articulation)	<u>Pierre</u> , <u>Thibault</u> , <u>Nicolas</u>
Voiceless fricatives	/ʃ/, /f/, /s/	High (unvoiced) High (manner of articulation)	<u>Charlotte</u> , <u>Fabrice</u> , <u>Solange</u>

Finally, post-hoc comparisons (Tukey's range test) with a Bonferroni correction were performed for the vowel's place of articulation in order to assess comparisons between the sexes in each syllable. The general size effect was computed using Cohen's f^2 . A symbolic representation of the regression formula is given in the Supplementary Material (Figure S1).

Temporal analyses

We assessed if the potential significant sound symbolic patterns found in the previous analysis have evolved or remained constant over time between male and female French first names. Pseudo-replication was allowed but phonetically equivalent, compound and monosyllabic names were still excluded, as the aim was to study temporal variations in both the first and last syllable. This resulted in a sample size of 897 female and 790 male names distributed across all decades. To address the time series nature of the data, we first calculated all autocorrelations and partial correlations between each time lag in order to assess if the frequency of a given phonetic variable is dependent on its previous frequency. Vowel articulation was counted as the number of each type of vowel in each syllable and was centred around 0; with 0 corresponding to central vowels, 1 to front vowels and -1 to back vowels. For vowel nasality, it was counted as the proportion of each vowel type: if values are close to 0, first names contain overall fewer nasal vowels, and conversely, if values are close to 1, they contain more nasal vowels. For voiced and voiceless consonants, the mean number in each syllable was studied. Linear models were then used to describe all of the temporal trends. To study possible non-linear effects of time, we modelled a cubic and quadratic effect of decade. Sex was included as another explanatory variable and was put in interaction with time. Model comparisons using the Akaike Information Criterion were then used to assess the best describing model of the temporal variations. All statistical analyses were performed using the R software (version 3.4.4).

Results

Sex-biased sound symbolism

We found a sex-biased sound structure in first names as a function of the syllable under study (Table 2). In the last stressed syllable, significant clues of masculinity were given by the vowel place of articulation ($\chi^2 = 11.82$, $p < 0.01$), nasality ($\chi^2 = 65.41$, $p < 0.001$) and voiceless fricatives ($\chi^2 = 13.23$, $p < 0.001$). Namely, male names were significantly more prone to contain back vowels like /o/ or /ɔ/ (e.g. 'Enzo', 'Renaud'), instead of front or central ones such as /i/, /y/ or /a/ (respectively $t = 1.17$, $p < 0.01$; $t = 1.35$, $p < 0.01$; e.g. 'Jackie', 'Luc', 'Bernard'). Although back vowels can be

Table 2. Results of the generalized linear model. For each predictor, the estimate, standard error of the mean, the χ^2 , the degrees of freedom and the p -values associated from the likelihood ratio test of the comparison between the full model and the model without the predictor are given. For the categorical variables ‘Vowel place of articulation’ and ‘Nasality’, the estimates are given compared to the reference category (front and non-nasal vowels, respectively) for both syllables. Pseudo- R^2 is the variance explained by the model (adjusted by the number of predictors) and Cohen’s f^2 the overall size effect. Significant p -values are in bold

Pseudo- $R^2 = 0.14$ Cohen’s $f^2 = 0.17$ N total = 472 n female = 275 n male = 197					
	Estimate	Standard error	χ^2	d.f.	p
Intercept	-0.69	0.18			
<i>First syllable</i>					
Vowel place of articulation			0.27	2	0.87
Central vowel	-0.12	0.25			
Back vowel	-0.11	0.29			
Nasality			0.33	1	0.56
Nasal vowel	0.31	0.54			
Voiced plosives	0.38	0.11	12.59	1	<0.001
Voiced fricatives	0.16	0.10	2.33	1	0.12
Voiceless plosives	0.10	0.10	0.96	1	0.32
Voiceless fricatives	-0.09	0.11	0.74	1	0.38
<i>Last syllable</i>					
Vowel place of articulation			11.82	2	<0.01
Central vowel	-0.18	0.24			
Back vowel	1.17	0.38			
Nasality			65.41	1	<0.001
Nasal vowel	2.62	0.38			
Voiced plosives	0.14	0.10	1.83	1	0.17
Voiced fricatives	0.12	0.10	0.41	1	0.23
Voiceless plosives	0.04	0.10	0.12	1	0.72
Voiceless fricatives	0.39	0.10	13.23	1	<0.001

found in female names (e.g. ‘Simone’, ‘Laure’), front and central vowels are more common (e.g. ‘Emilie’, ‘Julie’, ‘Léa’, ‘Maria’) along with mid-front vowels such as / ϵ / (e.g. ‘Claire’, ‘Hélène’). Male names were also significantly more likely to contain nasal vowels such as / \tilde{a} / or / \tilde{o} / (e.g. ‘Roland’, ‘Raymond’; female counter-examples, ‘Fernande’, ‘Marion’) and voiceless fricatives such as / s / or / f / (e.g. ‘Fabrice’, ‘Michel’; female counter-examples, ‘Clemence’, ‘Blanche’). Probabilities of being a male name as a function of the type of vowel (oral and nasal) are given in [Figure 1](#).

Unexpectedly, in the first syllable, the probability of being a male name significantly increased as a function of the number of voiced plosives ($\chi^2_1 = 12.59$, $p < 0.001$) such as / b /, / d / or / g / (e.g. ‘Bernard’, ‘Dimitri’, ‘Gustave’; female counter-examples, ‘Brigitte’, ‘Deborah’, ‘Gwenaëlle’). Within the first syllable, vowel articulation and nasality did not differ between sexes, nor did the number of voiceless fricatives (all $p > 0.05$). Eventually, articulatory features explained 14% of the variation in sex differences and the Cohen’s f^2 (0.17) suggests a moderate size effect (Cohen 2013).

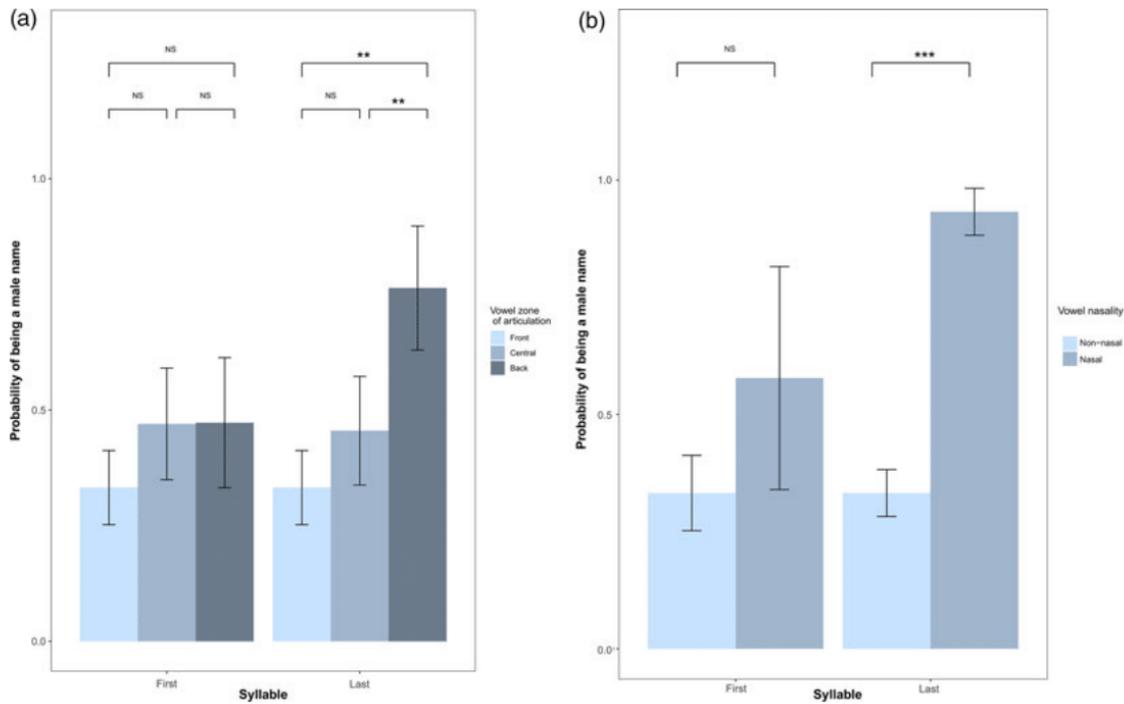


Figure 1. Estimates of the generalized linear model, log back-transformed to provide the probabilities of a name belonging to a male in function of the presence of a particular (a) oral vowel and (b) nasal vowel. Bars represent the mean probability associated with 95% confidence intervals. Significance code from the *post-hoc* comparisons: *** $p < 0.001$; ** $p < 0.01$; NS, non-significant.

Temporal analyses from 1900 to 2009

Trends investigated were the vowel's place of articulation, vowel's nasality, the number of voiced plosives and voiceless fricatives in both the first and last syllable. All trends are shown in Figure 2.

Analyses of the autocorrelations and partial correlations revealed that the frequency of each articulatory feature at a given timepoint is mostly independent of its previous frequency (most $p > 0.05$, all autocorrelations and partial correlations are given in the Supplementary Material, Table S1).

The proportion of oral vowels across time in the last syllable of both male and female names showed a cubic change ($F_{1,1686} = 14.01$, $p < 0.01$, Figure 2a) and the overall difference in proportion between the sexes was significant ($F_{1,1686} = 33.41$, $p < 0.001$). Interestingly, female names tended to be 'masculinized' (i.e. contained more central and back vowels, especially the former) over time starting from the 1960s with convergent values between male and female names towards 2009. In the first syllable, no overall difference in proportion was observed between the sexes ($F_{1,1686} = 1.62$, $p = 0.22$), but both followed a quadratic temporal change ($F_{1,1686} = 38.71$, $p < 0.001$, Figure 2b). In the last syllable, the difference in proportion of names with nasal vowels was different between male and female names ($F_{1,1686} = 117.25$, $p < 0.001$) and both remained more or less constant over time ($F_{1,1686} = 1.46$, $p = 0.24$, Figure 2c). In the first syllable, a slight difference in proportion was observed ($F_{1,1686} = 6.34$, $p < 0.05$), and both sexes followed a quadratic change over time ($F_{1,1686} = 51.59$, $p < 0.001$, Figure 2d).

In the last syllable, no sex difference and no temporal change in the mean number of voiced plosives were observed (respectively $F_{1,1686} = 1.11$, $p = 0.30$; $F_{1,1686} = 4.24$, $p = 0.054$, Figure 2e). In the first syllable, overall difference in voiced plosives between the sexes was significant ($F_{1,789} = 87.81$, $p < 0.001$), but no change was observed over time (Figure 2f), although the interaction between sex and a quadratic effect of time was significant ($F_{1,1686} = 8.48$, $p < 0.01$). Overall differences in the mean number of voiceless fricatives between the sexes was found in the last syllable ($F_{1,1686} = 60.09$, $p < 0.001$). In both sexes, the mean number of voiceless fricatives followed a cubic evolution through time ($F_{1,1686} = 12.46$, $p = 0.023$, Figure 2g), and an interaction between sex and time revealed to be

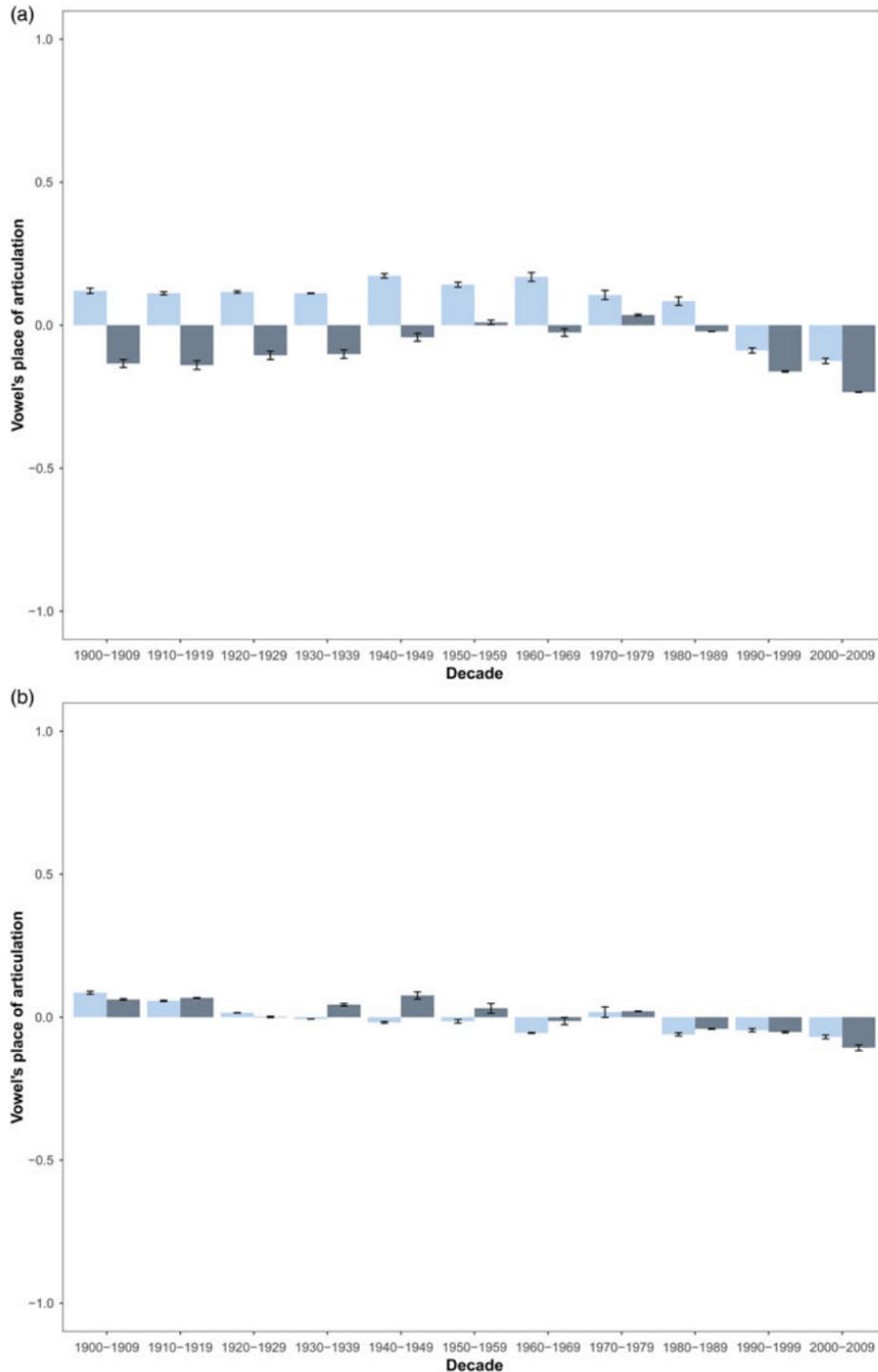


Figure 2. Barplots (mean \pm standard-error) of the temporal variations for each decade from 1900 to 2009 of each articulatory feature that revealed significant in the sound symbolic patterns analysis. Female first names are represented in light blue and male first names in dark blue. The vowel's place of articulation is represented in (a) last syllable and (b) first syllable. Vowel's nasality in the (c) last syllable and (d) first syllable. Mean number of voiced plosives are represented in the (e) last syllable and (f) first syllable. Lastly, mean number of voiceless fricatives are represented in the (g) last syllable and (h) first syllable. Vowel articulation accounts for the number of each type of vowel in each syllable and were centred around 0; with 0 more central vowels, 1 more front vowels and -1 more back vowels. For vowel nasality, it accounts for the number of each vowel type: if values are close to 0, first names contain fewer nasal vowels, and conversely, if values are close to 1, they contain more nasal vowels.

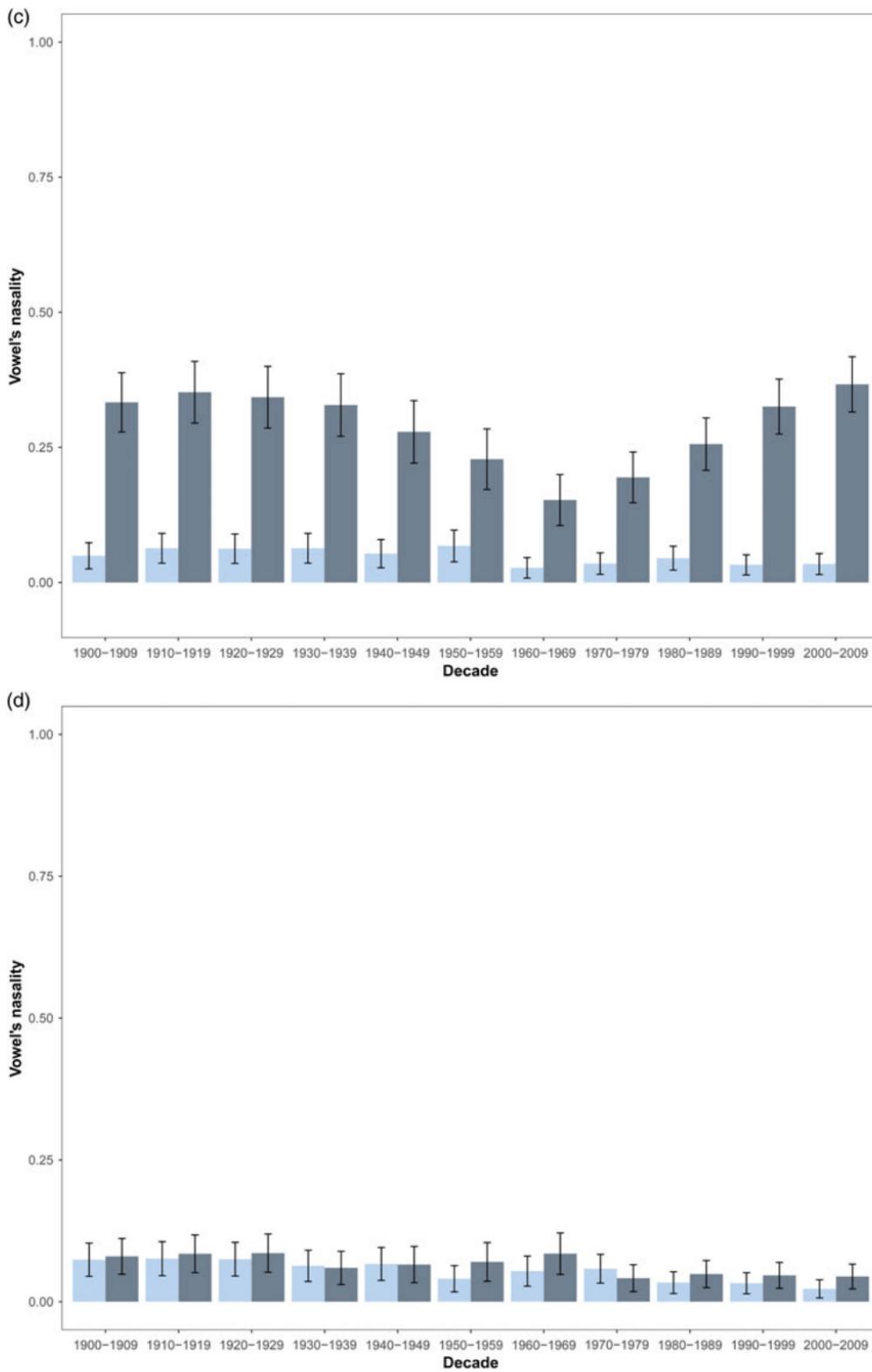


Figure 2. (continued)

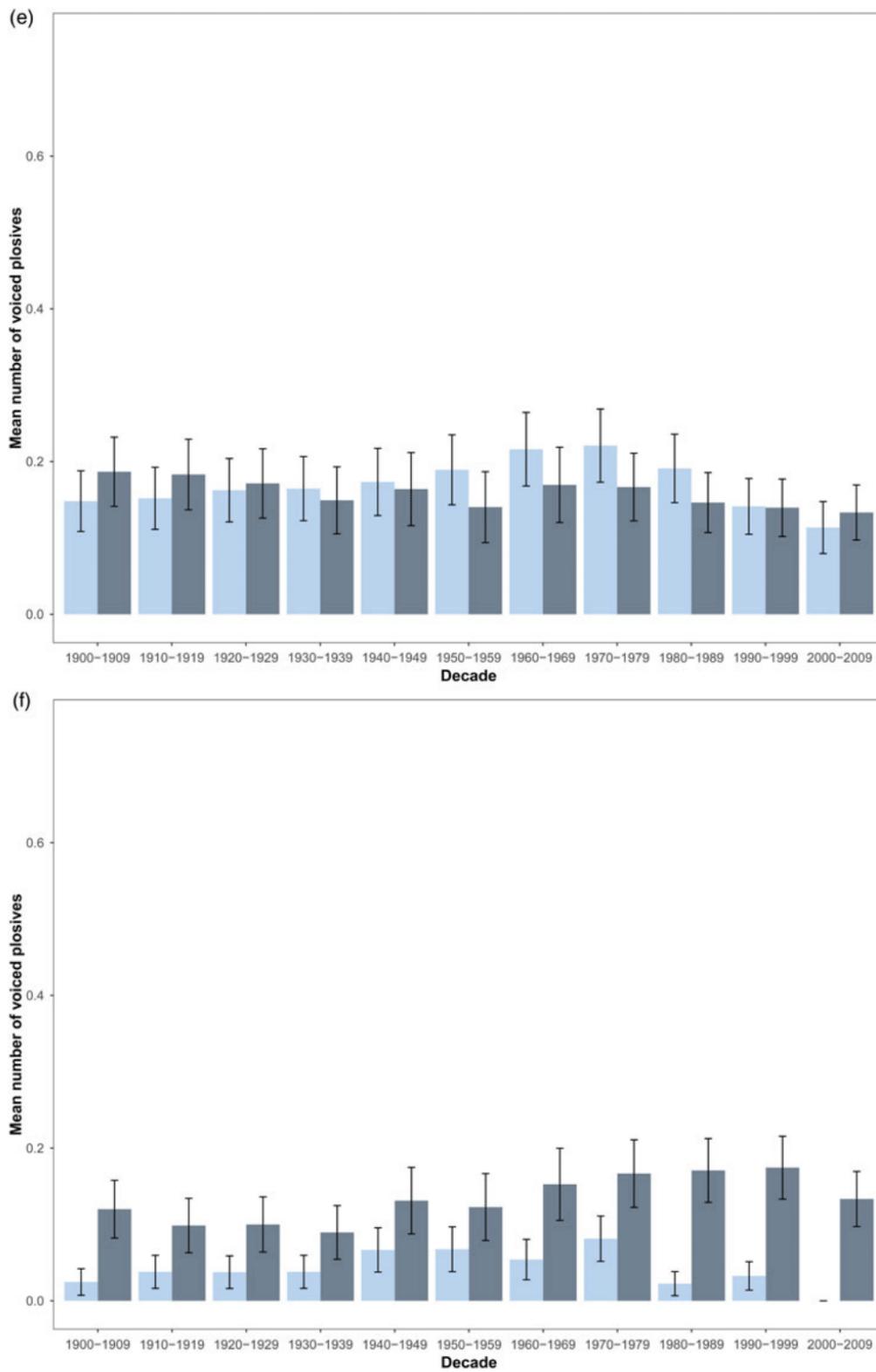


Figure 2. (continued)

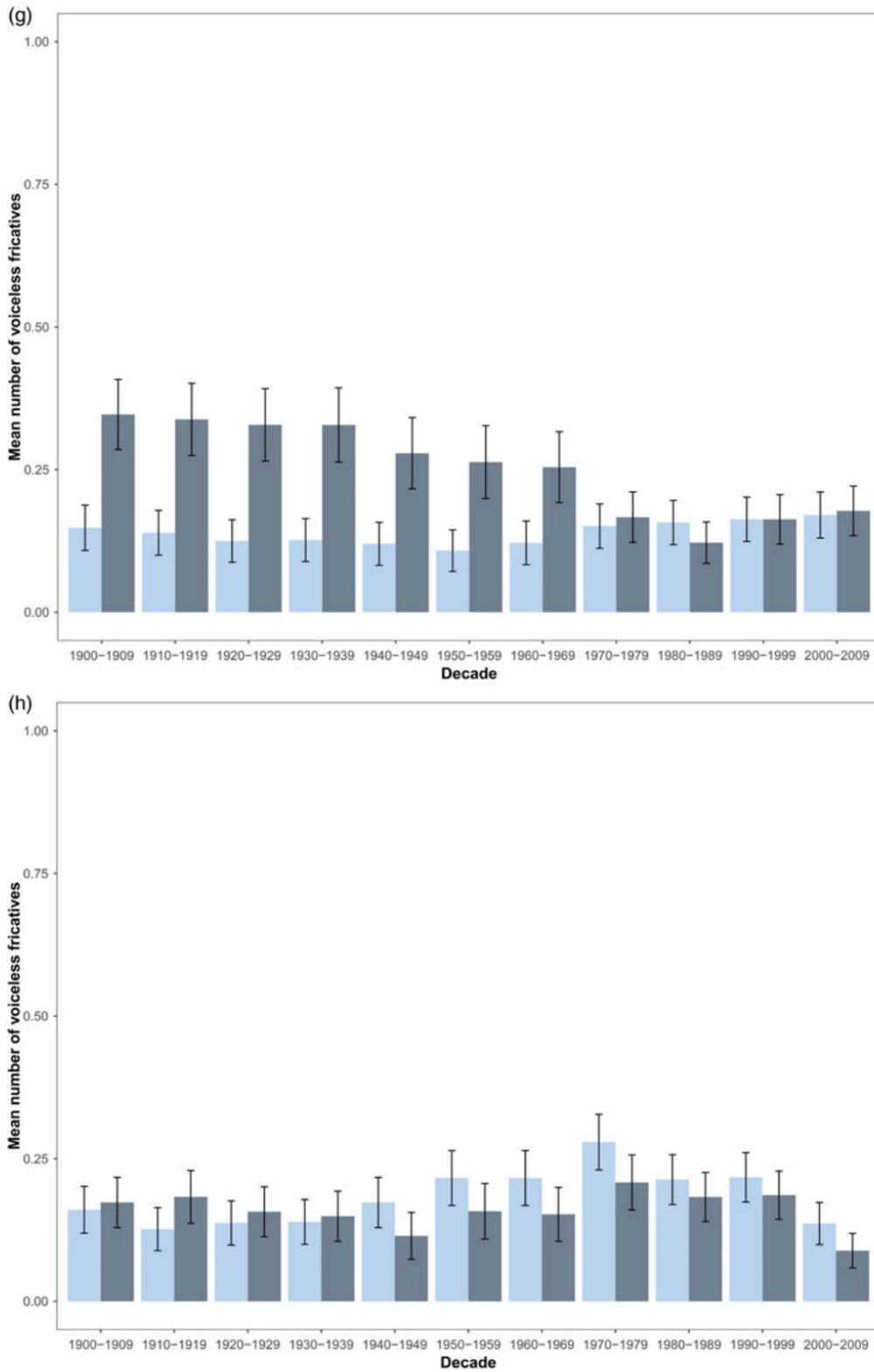


Figure 2. (continued)

significant ($F_{1,1686} = 30.66$, $p < 0.001$). Lastly, the mean number of voiceless fricatives in the first syllable for both sexes linearly varied over time ($F_{1,1686} = 31.50$, $p < 0.001$, Figure 2h) and an overall difference between the sexes was observed (respectively $F_{1,1686} = 103.32$, $p < 0.001$). The interaction between sex and time was also significant ($F_{1,1686} = 55.59$, $p < 0.001$).

Discussion

French first names exhibited sex differences in the distribution of vocalic sounds: low-frequency vowels (i.e. back and nasal) were more likely to be found in masculine names while higher frequency vowels (i.e. front and non-nasal) as well as central vowels (i.e. /a/) were more frequent in female names.

This sex-biased sound symbolism pattern was found in the last syllable, which is perceptually prominent in French, while in English, a similar sex-biased symbolism was reported for the first stressed syllable (Pitcher *et al.* 2013). However, regarding consonants, our results were more unexpected. Indeed, the mean number of voiceless fricatives (i.e. /f/, /s/ and /ʃ/; e.g. ‘Joseph’, ‘Alexis’, ‘Michel’) was higher in male than female names within the final stressed syllable (e.g. female names: ‘Delphine’, ‘Clarisse’). This is surprising according to the ‘frequency-code’ theory since their higher domain of frequency, relatively to voiced consonants, would rather be associated with indexical cues of smallness. The second unexpected finding was the presence of voiced plosives in the first syllable (i.e. /b/, /d/ and /g/; e.g. ‘Bernard’, ‘David’, ‘Gabriel’; e.g. female names, ‘Brigitte’, ‘Geraldine’), which is theoretically perceptually non-prominent in French. A possible explanation is that these consonantal patterns may perceptually compensate for each other, by which the presence in masculine names of voiceless fricatives in the last stressed syllable is perceptually counterbalanced by the presence of voiced consonants in the unstressed one. Otherwise, in a more general manner, vowels and consonants in the first and last syllable may be perceptually associated with different physical qualities. In this sense, while oral and nasal vowels could refer to body size, consonants might evoke other qualities such as shape or speed (Berlin 2006). For instance, it has been shown that people perceive a form as rounder if its signifier contains voiced consonants (such as /b/, /m/, /l/ or /n/) and as sharper if it contains voiceless stops (such as /k/, /p/, /t/) (Sidhu and Pexman 2015; Nielsen and Rendall 2013; Maurer *et al.* 2006). In the case of voiced plosives in the first syllable of male names, another explanation can be invoked as it is in accordance with results observed in American and Indian forenames (Slepian and Galinsky 2016). The authors showed a voiced gender naming effect, whereby the initial phonemes of masculine first names were voiced, as opposed to unvoiced in feminine names. They argued that voiced phonemes would sound ‘harder’ as a consequence of the vocal folds vibrating during pronunciation, whereas unvoiced phonemes will sound ‘softer’ to the ear as a consequence of unmodulated airflow, which in both cases would perceptually reinforce the stereotyped representations of males and females having respectively ‘tougher’ vs. ‘tender’ personalities and behaviours. Interestingly, the endorsement of these traditional gender stereotypes related to these ‘tougher/harder’ vs. ‘softer/tender’ dimensions moderated the influence of voiced and unvoiced phonemes on masculine vs. feminine judgments.

The name selected by parents for their offspring is, most of the time, linked to the assigned sex at birth, probably because such an information takes on great importance in both the perception of the bearer’s sex properties by conspecifics in the social environment, and in the bearer’s reinforcement of sexual identity and gender role (Pilcher 2016). In human societies, males and females have distinct roles and different reproductive strategies (Schmitt 2015). Owing to the associated sex-sound symbolism, giving a masculine or feminine name to conform to sex stereotypes could thus be seen as a form of parental investment with a lifelong-lasting effect. Although these effects have not been measured yet in reproductive value, it remains to be shown whether or not they influence fitness-related traits. However, the fact that most first names are sex-specific suggests that they are not fully socially neutral, and many studies have disclosed the influence of given names on some social trait, such as social desirability (Gebauer *et al.* 2012; Busse and Seraydarian 1978) and social outcomes (Cotton *et al.* 2008;

Figlio 2005; Hodson and Olson 2005; Harari and McDavid 1973). For instance, several studies have shown that having only the information of a masculine or feminine name already influences the bearer's job and career outcomes (Kasof 1993; Moss-Racusin et al. 2012; Steinpreis et al., 1999).

Yet while our results support the idea that humans possess a cognitive bias to assign different phonetic material to either sex, the relatively small amount of variance explained in sex differences ($\sim 14\%$) and the relatively modest size effect (Cohen's $f^2 = 0.17$) suggest that other factors than sexually sound symbolic patterns need to be considered when parents choose a particular name for their child. Evidence shows that the cultural environment is undeniably one of them (Acerbi and Bentley 2014; Barucca et al. 2015; Bentley et al. 2004; Berger et al. 2012; Xi et al. 2014). For instance, Bentley et al. (2004) have shown that name distributions and changes over time followed power laws, which were predicted by a simple mechanism of cultural drift and random copying between individuals, assuming that names are value-neutral in regards to fitness. Other models have been used to describe their distributions across time and space, the rate of innovation and their diversity, such as activation–inhibition processes (Zanette 2012), individual preferences and social influence (Xi et al. 2014), and spatial–temporal homogeneity (Bentley and Ormerod 2012). Most interestingly, Berger et al. (2012) have shown that names are more likely to be chosen when similar-sounding names in terms of phonetic similarity (i.e. sharing phonemes and their position within the name) have been popular the previous year, regardless of the names' gender. For instance, their model predicted that the popularity of the name 'Karen' depended on popular names that possessed the same first phoneme (i.e. /k/), such as 'Carl' (a male name) and 'Katie' (a female name). Predicted popularity was also correlated with other cultural items such as hurricanes' names (i.e. 'Katrina'), suggesting a strong effect of other cultural items on naming processes.

In this context, the temporal variations of the articulatory features suggest a strong effect of culture, given the somewhat stochastic variations of some phonetic variants, such as the frequency of occurrence of voiced plosives and voiceless fricatives. Nonetheless, we feel that particular attention should be given to the vowel's place of articulation. Its evolution in the stressed syllable of female first names suggests that high-frequency sounds were considered as most feminine in the 1960s, a period after which we notice an increase in phonetic masculinization that continues up to 2009. For instance, names with front vowels (e.g. 'Marie') in the early 1900s were more frequent than those with central and back vowels (e.g. 'Léa', 'Manon'), which increased in frequency in the 1960s up to the 2000s. Interestingly, an earlier study dealing with the evolution of feminization across the last century has shown that the 'ideal' waist-to-hip ratio (WHR), an important component of men's mate preferences, seemed to have followed the same trend in Western society. This 'ideal' WHR, as assessed through Playboy models and Miss pageants from 1920 to 2014, is most feminine in the 1960s (lower WHR values) then becomes less and less feminine until the 2010s (higher WHR values) (Bovet and Raymond 2015). Additionally, a meta-analysis on the self-perception of femininity and masculinity, as assessed through the Bem Sex-Role Inventory and the Personal Attributes Questionnaire, showed that American women perceived themselves as more masculine over time from the early 1970s to the mid 1990s (Twenge 1997), with additional findings demonstrating a decrease in endorsing feminine traits in women after the 2000s (Donnelly and Twenge 2017). Two other meta-analyses investigating women's own assertiveness from 1931 to 1993 showed that it decreased from 1946 to 1967, but increased from 1968 to 1993 (Twenge 2001). Such changes from the 1960s might be closely linked to historical political feminists' movements particularly active in this era during which awareness of inequalities in civil rights and social positions has been increasing. We hypothesize that one possible strategy to compensate for such inequalities is to masculinize some traits in women in order to compete against men for the same rights and privileges, at least in industrialized and traditionally male-dominated societies.

Conclusions

Overall, the present study offers some promising opportunities for follow-up studies that would lead to a better understanding of naming processes. An interesting avenue for further research would be to

model the relative importance of different selective pressures (sexual and cultural, or a joint effect) acting on the phonetic dimorphism, names' frequency and the emergence of novel names. Most importantly, to fully acknowledge the action of sexual selection on the phonetic dimorphism, a study on names and their relationship to reproductive value is required. One limitation is that not all names from each decade were analysed and a particular attention should be given to rare names in order to strengthen the present results. Moreover, particular attention should also be given to syllables between the first and last ones, as they can potentially play a particular role. Further inquiries in sound symbolic patterns in first names in dead and modern languages should be performed, so as to find some universal components in vowel quality to convey perceived masculinity and femininity.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/ehs.2019.7>

Author ORCIDs.  Alexandre Suire, 0000-0002-1852-3083

Acknowledgments. We would like to thank the Institut National de la Statistique et des Études Économiques for providing the data on French first names ('Fichier des prénoms' – Edition 2011 (electronic data), INSEE (producer) and Centre Maurice Halbwachs (distributor)). This is a contribution of the Institut des Sciences de l'Évolution 2019-101.

Author contributions. A.S. and A.B.M. wrote the paper. A.S. and M.B.D. transcribed the data. A.S., A.B.M and M.R. analysed the data. M.B.D. and M.R. supervised the study.

Financial Support. This study received a 'Equipe Recherche Junior' funding from the Centre Méditerranéen de l'Environnement et de la Biodiversité under the project name 'Symphosex'.

Conflict of Interest. None.

Data access. The data and the R code are available at <https://figshare.com/s/5618e367fae4c5774272>

References

- Acerbi A and Alexander Bentley R (2014) Biases in cultural transmission shape the turnover of popular traits. *Evolution and Human Behavior* 35(3), 228–236. <https://doi.org/10.1016/j.evolhumbehav.2014.02.003>
- Armstrong MM, Lee AJ and Feinberg DR (2019) A house of cards: bias in perception of body size mediates the relationship between voice pitch and perceptions of dominance. *Animal Behaviour* 147, 43–51. <https://doi.org/10.1016/j.anbehav.2018.11.005>
- Barucca P, Rocchi J, Marinari E, Parisi G and Ricci-Tersenghi F (2015) Cross-correlations of American baby names. *Proceedings of the National Academy of Sciences* 112(26), 7943–7947. <https://doi.org/10.1073/pnas.1507143112>
- Bentley RA and Ormerod P (2012) Accelerated innovation and increased spatial diversity of US popular culture. *Advances in Complex Systems* 15(01n02), 1150011. <https://doi.org/10.1142/S0219525911003232>
- Bentley RA, Hahn MW and Shennan SJ (2004) Random drift and culture change. *Proceedings of the Royal Society B: Biological Sciences* 271(1547), 1443–1450. <https://doi.org/10.1098/rspb.2004.2746>
- Berger J, Bradlow ET, Braunstein A and Zhang Y (2012) From Karen to Katie: using baby names to understand cultural evolution. *Psychological Science* 23(10), 1067–1073. <https://doi.org/10.1177/0956797612443371>
- Berlin B (2006) The first congress of ethnozoological nomenclature. *Journal of the Royal Anthropological Institute* 12(s1), S23–S44. <https://doi.org/10.1111/j.1467-9655.2006.00271.x>
- Blasi DE, Wichmann S, Hammarström H, Stadler PF and Christiansen MH (2016) Sound–meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences* 113(39), 10818–10823. <https://doi.org/10.1073/pnas.1605782113>
- Bovet J and Raymond M (2015) Preferred women's waist-to-hip ratio variation over the last 2,500 years. *PLoS One* 10(4), e0123284. <https://doi.org/10.1371/journal.pone.0123284>
- Bradbury JW and Vehrencamp SL (2011) Chapter 4: Conflict resolution and Chapter 12: Mate attraction and courtship. In *Principles of Animal Communication* (2nd edn) Sunderland: Sinauer Associates.
- Bremner AJ, Caparos S, Davidoff J, de Fockert J, Linnell KJ and Spence C (2013) 'Bouba' and 'kiki' in Namibia? A remote culture make similar shape–sound matches, but different shape–taste matches to Westerners. *Cognition* 126(2), 165–172. <https://doi.org/10.1016/j.cognition.2012.09.007>
- Busse TV and Seraydarian L (1978) The relationships between first name desirability and school readiness, IQ, and school achievement. *Psychology in the Schools* 15(2), 297–302. [https://doi.org/10.1002/1520-6807\(197804\)15:2<297::AID-PITS2310150229>3.0.CO;2-9](https://doi.org/10.1002/1520-6807(197804)15:2<297::AID-PITS2310150229>3.0.CO;2-9)
- Cassidy KW, Kelly MH and Sharoni LAJ (1999) Inferring gender from name phonology. *Journal of Experimental Psychology: General* 128(3), 362.

- Cohen J** (2013) *Statistical Power Analysis for the Behavioral Sciences*. Hoboken, NJ: Taylor and Francis.
- Collins SA and Missing C** (2003) Vocal and visual attractiveness are related in women. *Animal Behaviour* **65**(5), 997–1004. <https://doi.org/10.1006/anbe.2003.2123>
- Cotton JL, O'Neill BS and Griffin A** (2008) The 'name game': affective and hiring reactions to first names. *Journal of Managerial Psychology* **23**(1), 18–39. <https://doi.org/10.1108/02683940810849648>
- Cuskley C, Simner J and Kirby S** (2017) Phonological and orthographic influences in the bouba–kiki effect. *Psychological Research* **81**(1), 119–130. <https://doi.org/10.1007/s00426-015-0709-2>
- Cutler A, McQueen J and Robinson K** (1990) Elizabeth and John: sound patterns of men's and women's names. *Journal of Linguistics* **26**(02), 471. <https://doi.org/10.1017/S0022226700014754>
- Donnelly K and Twenge JM** (2017) Masculine and feminine traits on the Bem Sex–Role Inventory, 1993–2012: a cross-temporal meta-analysis. *Sex Roles* **76**(9–10), 556–565. <https://doi.org/10.1007/s11199-016-0625-y>
- Erwin PG** (1993) First names and perceptions of physical attractiveness. *The Journal of Psychology* **127**(6), 625–631. <https://doi.org/10.1080/00223980.1993.9914901>
- Feinberg DR, DeBruine LM, Jones BC and Perrett DI** (2008) The role of femininity and averageness of voice pitch in aesthetic judgments of women's voices. *Perception* **37**(4), 615–623. <https://doi.org/10.1068/p5514>
- Figlio D** (2005) *Names, Expectations and the Black–White Test Score Gap (no. w11195)*. Cambridge, MA: National Bureau of Economic Research, Available at: <https://www.nber.org/papers/w11195>.
- Figlio DN** (2007) Boys named Sue: Disruptive children and their peers. *Education Finance and Policy* **2**(4), 376–394.
- Fraccaro PJ, Jones BC, Vukovic J, Smith FG, Watkins CD, Feinberg DR, Little AC and DeBruine LM** (2011) Experimental evidence that women speak in a higher voice pitch to men they find attractive. *Journal of Evolutionary Psychology* **9**(1), 57–67. <https://doi.org/10.1556/JEP.9.2011.33.1>
- Geary DC** (1998) *Male, Female: The Evolution of Human Sex Differences*. American Psychological Association.
- Gebauer JE, Leary MR and Neberich W** (2012) Unfortunate first names: Effects of name-based relational devaluation and interpersonal neglect. *Social Psychological and Personality Science* **3**(5), 590–596. <https://doi.org/10.1177/1948550611431644>
- Harari H and McDavid JW** (1973) Name stereotypes and teachers' expectations. *Journal of Educational Psychology* **65**(2), 222–225. <https://doi.org/10.1037/h0034978>
- Hartung F** (2018) Are Tims hot and Toms not? <https://doi.org/10.31234/osf.io/9ctq5>
- Hassebrauck M** (1988) Beauty is more than 'name' deep: the effect of women's first names on ratings of physical attractiveness and personality attributes. *Journal of Applied Social Psychology* **18**(9), 721–726. <https://doi.org/10.1111/j.1559-1816.1988.tb02350.x>
- Haynie H, Bowern C and LaPalombara H** (2014) Sound symbolism in the languages of Australia. *PLoS One* **9**(4), e92852. <https://doi.org/10.1371/journal.pone.0092852>
- Hensley WE and Spencer BA** (1985) The effect of first names on perceptions of female attractiveness. *Sex Roles* **12**(7–8), 723–729. <https://doi.org/10.1007/BF00287866>
- Hodges-Simeon CR, Gurven M, Puts DA and Gaulin SJC** (2014) Vocal fundamental and formant frequencies are honest signals of threat potential in peripubertal males. *Behavioral Ecology* **25**(4), 984–988. <https://doi.org/10.1093/beheco/aru081>
- Hodson G and Olson JM** (2005) Testing the generality of the name letter effect: name initials and everyday attitudes. *Personality and Social Psychology Bulletin* **31**(8), 1099–1111. <https://doi.org/10.1177/0146167205274895>
- Jones BC, Feinberg DR, DeBruine LM, Little AC and Vukovic J** (2010) A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. *Animal Behaviour* **79**(1), 57–62. <https://doi.org/10.1016/j.anbehav.2009.10.003>
- Kasof J** (1993) Sex bias in the naming of stimulus persons. *Psychological Bulletin* **113**(1), 140–163. <https://doi.org/10.1037//0033-2909.113.1.140>
- Klink RR** (2000) Creating brand names with meaning: the use of sound symbolism. *Marketing Letters* **11**(1), 5–20.
- Knoeferle K, Li J, Maggioni E and Spence C** (2017) What drives sound symbolism? Different acoustic cues underlie sound-size and sound-shape mappings. *Scientific Reports* **7**(1) <https://doi.org/10.1038/s41598-017-05965-y>
- Köhler W** (1947) *Gestalt Psychology* (2nd edn) New York: Liveright.
- Leirer VO, Hamilton DL and Carpenter S** (1982) Common first names as cues for inferences about personality. *Personality and Social Psychology Bulletin* **8**(4), 712–718.
- Maurer D, Pathman T and Mondloch CJ** (2006) The shape of boubas: sound-shape correspondences in toddlers and adults. *Developmental Science* **9**(3), 316–322. <https://doi.org/10.1111/j.1467-7687.2006.00495.x>
- Mehrabian A** (2001) Characteristics attributed to individuals on the basis of their first names. *Genetic, Social, and General Psychology Monographs* **127**(1), 59.
- Mehrabian A and Piercy M** (1993) Affective and personality characteristics inferred from length of first names. *Personality and Social Psychology Bulletin* **19**(6), 755–758.
- Morton ES** (1977) On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist* **111**(981), 855–869.
- Moss-Racusin CA, Dovidio JF, Brescoll VL, Graham MJ and Handelsman J** (2012) Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences* **109**(41), 16474–16479. <https://doi.org/10.1073/pnas.1211286109>

- Newman SS (1933) Further experiments in phonetic symbolism. *The American Journal of Psychology* **45**(1), 53. <https://doi.org/10.2307/1414186>
- Nielsen AKS and Rendall D (2013) Parsing the role of consonants versus vowels in the classic Takete–Maluma phenomenon. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale* **67**(2), 153–163. <https://doi.org/10.1037/a0030553>
- Ohala JJ (1984) An ethological perspective on common cross-language utilization of F₀ of voice. *Phonetica* **41**(1), 1–16. <https://doi.org/10.1159/000261706>
- Ohala J (1994) The frequency codes underlies the sound symbolic use of voice pitch. In *Sound Symbolism* (pp. 325–347). Cambridge: Cambridge University Press.
- Pelham BW, Mirenberg MC and Jones JT (2002) Why Susie sells seashells by the seashore: implicit egotism and major life decisions. *Journal of Personality and Social Psychology* **82**(4), 469–487. <https://doi.org/10.1037//0022-3514.82.4.469>
- Perfors A (2004) What's in a Name? The effect of sound symbolism on perception of facial attractiveness. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. **26**.
- Pilcher J (2016) Names, bodies and identities. *Sociology* **50**(4), 764–779. <https://doi.org/10.1177/0038038515582157>
- Pisanski K and Rendall D (2011) The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. *The Journal of the Acoustical Society of America* **129**(4), 2201–2212. <https://doi.org/10.1121/1.3552866>
- Pisanski K, Fraccaro PJ, Tigue CC, O'Connor JJM and Feinberg DR (2014) Return to Oz: voice pitch facilitates assessments of men's body size. *Journal of Experimental Psychology: Human Perception and Performance* **40**(4), 1316–1331. <https://doi.org/10.1037/a0036956>
- Pitcher BJ, Mesoudi A and McElligott AG (2013) Sex-biased sound symbolism in English-language first names. *PLoS One* **8**(6), e64825. <https://doi.org/10.1371/journal.pone.0064825>
- Puts DA (2010) Beauty and the beast: mechanisms of sexual selection in humans. *Evolution and Human Behavior* **31**(3), 157–175. <https://doi.org/10.1016/j.evolhumbehav.2010.02.005>
- Puts DA, Gaulin SJC and Verdolini K (2006) Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior* **27**(4), 283–296. <https://doi.org/10.1016/j.evolhumbehav.2005.11.003>
- Puts DA, Hodges CR, Cárdenas RA and Gaulin SJC (2007) Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior* **28**(5), 340–344. <https://doi.org/10.1016/j.evolhumbehav.2007.05.002>
- Puts DA, Barndt JL, Welling LLM, Dawood K and Burriss RP (2011) Intrasexual competition among women: vocal femininity affects perceptions of attractiveness and flirtatiousness. *Personality and Individual Differences* **50**(1), 111–115. <https://doi.org/10.1016/j.paid.2010.09.011>
- Ramachandran VS and Hubbard EM (2001) Synaesthesia—A window into perception, thought and language. *Journal of consciousness studies* **8**(12), 3–34.
- Rendall D, Vokey JR and Nemeth C (2007) Lifting the curtain on the Wizard of Oz: biased voice-based impressions of speaker size. *Journal of Experimental Psychology: Human Perception and Performance* **33**(5), 1208–1219. <https://doi.org/10.1037/0096-1523.33.5.1208>
- Sapir E (1929) A study in phonetic symbolism. *Journal of Experimental Psychology* **12**(3), 225–239. <https://doi.org/10.1037/h0070931>
- Schmitt DP (2015) Fundamentals of human mating strategies. In *The Handbook of Evolutionary Psychology*. Hoboken, NJ, US: John Wiley & Sons, pp. 1–23.
- Shinohara K and Kawahara S (2010) A cross-linguistic study of sound symbolism: the images of size. *Annual Meeting of the Berkeley Linguistics Society* **36**(1), 396. <https://doi.org/10.3765/bls.v36i1.3926>
- Sidhu DM and Pexman PM (2015) What's in a name? Sound symbolism and gender in first names. *PLoS One* **10**(5), e0126809. <https://doi.org/10.1371/journal.pone.0126809>
- Slepian ML and Galinsky AD (2016) The voiced pronunciation of initial phonemes predicts the gender of names. *Journal of Personality and Social Psychology* **110**(4), 509–527. <https://doi.org/10.1037/pspa0000041>
- Steinpreis RE, Anders KA and Ritzke D (1999) The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: a national empirical study. *Sex Roles* **41**(7–8), 509–528.
- Svantesson J.-O (2017) Sound symbolism: the role of word sound in meaning: sound symbolism. *Wiley Interdisciplinary Reviews: Cognitive Science* **8**(5), e1441. <https://doi.org/10.1002/wcs.1441>
- Titze IR (1989) Physiologic and acoustic differences between male and female voices. *The Journal of the Acoustical Society of America* **85**(4), 1699–1707. <https://doi.org/10.1121/1.397959>
- Twenge JM (1997) Changes in masculine and feminine traits over time: a meta-analysis. *Sex Roles* **36**(5–6), 305–325. <https://doi.org/10.1007/BF02766650>
- Twenge JM (2001) Changes in women's assertiveness in response to status and roles: a cross-temporal meta-analysis, 1931–1993. *Journal of Personality and Social Psychology* **81**(1), 133–145. <https://doi.org/10.1037//0022-3514.81.1.133>
- Ullian R (1978) Size-sound symbolism. In *Universals of Human Language* (Vol. **2**, p. 525–568) Stanford, CA: Stanford University Press.

- van Dommelen WA and Moxness BH** (1995) Acoustic parameters in speaker height and weight identification: sex-specific behaviour. *Language and Speech* **38**(3), 267–287. <https://doi.org/10.1177/002383099503800304>
- Werner H** (1957) *Comparative Psychology of Mental Development* (revised edn). New York: International Universities Press.
- Wu L, Klink RR and Guo J** (2013) Creating gender brand personality with brand names: the effects of phonetic symbolism. *Journal of Marketing Theory and Practice* **21**(3), 319–330. <https://doi.org/10.2753/MTP1069-6679210306>
- Xi N, Zhang Z.-K, Zhang Y.-C, Ge Z, She L and Zhang K** (2014) Cultural evolution: the case of babies' first names. *Physica A: Statistical Mechanics and its Applications* **406**, 139–144. <https://doi.org/10.1016/j.physa.2014.03.042>
- Xu Y, Lee A, Wu W.-L, Liu X and Birkholz P** (2013) Human vocal attractiveness as signaled by body size projection. *PLoS One* **8**(4), e62397. <https://doi.org/10.1371/journal.pone.0062397>
- Zanette DH** (2012) Dynamics of fashion: the case of given names. *ArXiv Preprint ArXiv:1208.0576*, 7.
- Zwebner Y, Sellier A.-L, Rosenfeld N, Goldenberg J and Mayo R** (2017) We look like our names: the manifestation of name stereotypes in facial appearance. *Journal of Personality and Social Psychology* **112**(4), 527–554. <https://doi.org/10.1037/pspa0000076>

Article 6¹

**Perceptions of masculinity and femininity in first names can be
predicted by sound symbolic patterns**

Alexandre Suire, Michel Raymond & Melissa Barkat-Defradas

Article en préparation

¹ La présentation de cette étude suit le même format que les deux études exploratoires : une synthèse suivie du Matériel et Méthodes et des Résultats.

Résumé

Après avoir mis en évidence plusieurs motifs phonétiquement symboliques dans les prénoms Français, il reste à mettre en évidence que ceux-ci influencent de manière effective les perceptions de masculinité et de féminité. En d'autres termes, nous avons testé l'hypothèse que dans la syllabe accentuée, les voyelles d'arrière et nasales, les constrictives non voisées et dans la syllabe non accentuée les occlusives voisées, devraient augmenter la perception de masculinité des prénoms masculins. A l'inverse, dans les prénoms féminins, ceux contenant dans leur syllabe accentuée des voyelles d'avant ou centrales devraient être perçus comme plus féminins.

A cette fin, nous avons mis en place une étude en ligne par lequel nous avons recruté plus de 2000 participants ayant jugé la masculinité et la féminité de 363 prénoms (183 prénoms masculins et 180 prénoms féminins) via un paradigme de choix forcé. Globalement, nos résultats valident notre hypothèse, mais quelques résultats inattendus ont été trouvés. Notamment, les juges masculins ont considéré, comparativement aux juges féminins, les prénoms féminins contenant des voyelles d'arrière et nasales dans la dernière syllabe et des occlusives voisées dans la première syllabe, comme plus féminins. Ceci est d'autant plus intrigant que les prénoms féminins récemment populaires ont été considérés comme plus féminins, alors que la direction opposée a été trouvée pour les prénoms masculins. Ces résultats sont à considérer à la lumière des prénoms féminins présentant une incidence étonnamment plus élevée de la voyelle centrale /a/ (un phonème de fréquence relativement basse) dans leur dernière syllabe à partir des années 1960.

Notons que cette étude est encore exploratoire et que les modèles présentés peuvent être largement améliorés. Notamment, nous avons considéré les variables phonétiques comme numériques et non comme des facteurs, ce qui peut masquer certains des effets. De plus, quelques indices suggèrent que des interactions entre les composantes phonétiques au sein d'un

prénom peuvent influencer les différentes perceptions. Enfin, un certain nombre de données sur le statut socio-économique des juges ont été récupérées mais n'ont pas encore été analysées. Toutefois, ces quelques résultats préliminaires nous permettent d'envisager deux nouvelles études sur le symbolisme phonétique dans les prénoms : mesurer leur éventuelle influence sur les dimensions attractives et dominantes d'un visage et, sur le succès social comme la probabilité d'être embauché.

**Perceptions of masculinity and femininity in first names can be predicted by sound
symbolic patterns.**

Authors: Alexandre Suire¹, Michel Raymond¹, Melissa Barkat-Defradas¹

Corresponding author: Alexandre Suire

E-mail: alexandre.suire@umontpellier.fr

Fax: +33 4 67 14 36 22

Tel: +33 4 67 14 49 66

¹ Institut des Sciences de l'Évolution de Montpellier, Université de Montpellier, Centre
National de la Recherche Scientifique, Institut pour la Recherche et le Développement, École
Pratique des Hautes Études – Place Eugène Bataillon 34095 Montpellier, France.

Summary

Previous studies have unveiled sound symbolic patterns in first names (Cutler, McQueen, & Robinson, 1990; Pitcher, Mesoudi, & McElligott, 2013; Slepian & Galinsky, 2016). More recently, a study on French first names showed that males names were more likely to contain on their stressed syllable back and nasal vowels (e.g. /o/, /ɔ̃/), voiceless fricatives (e.g. /s/) and on their first syllable voiced plosives (e.g. /b/) compared to feminine names (Suire, Bossoms Mesa, Raymond, & Barkat-Defradas, 2019). It has been suggested that such symbolic patterns may have arose to respectively project impressions of largeness/masculinity and smallness/femininity in male and females first names. If so, individuals from both sexes should perceive these names as such. However, no study so far has shown whether such sound symbolic patterns actually reinforce perceptions of masculinity and femininity (except one study from Slepian & Galinsky (2016) who tested such associations with initial voiced phonemes) .

In this study, we show that coordinates produced from a linear discriminant analysis, used in order to assess scores of phonetic masculinity and femininity, overall predicted in the same direction perceived masculinity and femininity. In details, most of the phonetic variables that predicted sound symbolic patterns in first names influenced such perceptions, although a few non-predicted results were uncovered. Namely, masculine names with higher backness on their stressed syllable and voiced plosives on their first syllable were considered more masculine; but those with nasal vowels and those with voiceless fricatives on their stressed syllable were considered less masculine. No difference in masculinity judgments was found between male and female judges. For feminine names, those with nasal vowels on their stressed syllable and with voiced plosives on their first syllable were considered less feminine, and those with voiceless fricatives were considered more feminine. Interestingly, male judges considered phonetically more masculine female names as more feminine compared to female judges.

This latter finding is further intrigued by other results by the fact that relatively newer popular female names were considered as more feminine (while the opposite direction was found for male names). This is interesting as relatively newer feminine names exhibit a surprisingly higher incidence of the central vowel /a/ (a relatively low-frequency phoneme) in their last syllable starting from the 1960's (Suire et al., 2019).

Such findings may have implications on a broader scale for the social life of these names' bearer. Although further studies are also needed to assess perceptions of largeness/smallness to fully acknowledge the effect of sound symbolism, one particular study we are interested in is to test whether such names influence the attractiveness/dominance dimensions of facial appearance in individuals, and whether they influence probabilities of success in different social outcomes, for instance career and job opportunities.

Methods

Name stimuli

We retained 183 male and 180 female names, which represent the most popular first names given between 1900 and 2009 in France (Suire et al., 2019). Then, we used a linear discriminant analysis (LDA) to assess for each name a score of phonetic masculinity and femininity. The LDA modelled how phonetic variables can predict if a name is masculine or feminine. These variables were, each for the first and last syllable, the vowel's backness and nasality as well as voiced and voiceless obstruents. From the linear discriminant function, we computed the coordinates for each name, which we used as an axis of phonetic masculinity/femininity. Male names had higher and positive coordinates, conversely, female names had lower and negative coordinates.

Experimental procedure

The experiment was automated on a web interface program, which we advertised through social networks and mail lists. Before starting the experiment, each judge was asked to fill a

questionnaire regarding age, sex and a number of socio-economic variables (which are not reported here). Then, each judge was presented with 25 pairs of names. Each name was randomly drawn from the pool of name stimuli and randomly assigned a left or right position on the screen. Male names were compared against male names and female names against female names. For each pair, judges were either asked to choose the more masculine or feminine name by clicking on it (i.e., “Which name do you consider as more masculine?”, “Which name do you consider as more feminine?”). Masculinity or femininity questions remained the same throughout the whole test and did not depend on the judge’s sex. Judges could take as long as they wanted to pick a name, but could not go back to their previous choices. The first and last three pairs were the same so as to assess a judge’s reliability (e.g. random clicks).

In the end, we retained 2465 judges who had a virgin ip and who were at least consistent for 2 out of the 3 same choices. This resulted in a sample of 1842 unique female judges (mean age \pm SD = 32.70 \pm 11.55 years old; range = 18 – 78) and 623 unique male judges (mean age \pm SD = 35.12 \pm 12.97 years old; range = 18 – 76), for a total of 60 209 observations (30 616 observations for masculinity questions, 29 593 for femininity questions).

Statistical analyses

We conducted exploratory analyses of the data, for which masculinity and femininity preferences were analyzed separately. To this end, we used generalized linear mixed models (GLMM) with the response variable being if the judge chose or not the name presented to him/her on the left position. The GLMMs were fitted with a binomial error structure since the response variable consisted in a discrete probability distribution of the number of successes in a sequence of several independent trials.

First, two GLMMs were constructed to assess whether the scores of phonetic masculinity/femininity computed from the LDA function predicted perceived

masculinity/femininity. To this end, we computed the difference of scores between the two names.

Secondly, in order to explore in more detail which and how phonetic variables explained masculinity or femininity preferences, we constructed two additional GLMMs. In these models, LDA scores were replaced by the phonetic variables: the backness and nasality and the presence/absence of voiceless fricatives in the last syllable, and the presence/absence of voiced plosives in the first syllable. Vowel's backness was coded as 0 for front vowels, 1 for central vowels and 2 for back vowels. Vowel's nasality was coded as 0 for oral vowels and 1 for nasal vowels. Voiceless fricatives were coded as 0 if they were none in the syllable, or 1 or 2 (the last syllable can have 2 voiceless fricatives). Voiced plosives were coded as 0 if they were none, or 1. For easier understanding of the models' output, we transformed these variables as numerical and then computed their differences for each name.

In both the LDA scores and phonetic variables models, we also computed the difference in length name (i.e., syllables) and the difference in popularity between the two names. Popularity corresponds to a variable that considers when the name was the most popular within any decade between 1900 and 2009. For instance, names with the values 1 and 2 were respectively most popular between 1900-1909 and 1910-1919, whereas names with the values with 10 and 11 were respectively most popular in 1990-1999 and 2000-2009 (see Suire et al., 2019 for more details of this aspect). Judges' age and sex were put in interaction with the LDA scores and the phonetic variables to assess whether they influence masculinity and femininity preferences. Lastly, judges' identity and names were added as random factors on the intercept only.

The significance of each variable in all GLMMs was assessed from the comparison of the model excluding the variable with the model including all the other variables (i.e., likelihood-ratio chi-square tests, ANOVA type III).

Results

Perceptions of masculinity

Table 1 and Table 2 respectively reports the results with LDA scores and the phonetic variables. Overall, names whose LDA scores was higher were considered on average more masculine, i.e., phonetic masculinity predicted perceived masculinity ($\chi^2_1 = 121.853, p < 0.001$). Moreover, relatively older names were considered as more masculine compared to relatively new names ($\chi^2_1 = 438.271, p < 0.001$). Relatively older individuals judged older names as more masculine ($\chi^2_1 = 78.614, p < 0.001$), and no sex differences was found in the judgment of masculinity (both $p > .05$).

In details, results show that male names whose backness was higher in their stressed syllable (i.e., containing back vowels) and who had voiced plosives in their first syllable were considered more masculine (both $p < .001$). Names who contained nasal vowels and voiceless fricatives in their stressed syllable were considered less masculine (both $p < .001$). Older individuals considered longer and older names as more masculine (both $p < .01$); such individuals also preferred names with higher backness ($\chi^2_1 = 9.887, p < 0.01$). Lastly, male judges preferred longer names ($\chi^2_1 = 4.309, p < 0.05$).

Table 1. Results of the GLMM for the LDA scores¹.

	Estimate	SE	χ^2	<i>p value</i>
Intercept	0.038	0.0127	/	/
Difference in LDA scores	0.151	0.013	121.853	< .001
Difference in popularity	-0.289	0.013	438.271	< .001
Interactions with age				
<i>Difference in LDA scores</i>	0.003	0.012	0.064	.800
<i>Difference in popularity</i>	-0.111	0.012	78.614	< .001
Interactions with sex				
<i>Difference in LDA scores</i>	0.005	0.027	0.034	.852
<i>Difference in popularity</i>	-0.028	0.028	1.035	.308

¹For each variable, the χ^2 and the *p* values associated from the likelihood-ratio chi-square test of the comparison between the full model and the model without the variables are given (ANOVA type III). For the categorical variable 'sex', the estimates are given compared to the reference category (1 = Male judges). The degrees of freedom are 1 for every test. *P* values are considered significant at the 0.05 threshold (in bold). **All subsequent tables follow the same presentation.**

Table 2. Results of the GLMM with phonetic variables (numerical).

	Estimate	SE	χ^2	<i>p</i> value
Intercept	0.038	0.012	/	/
Difference of backness	0.127	0.013	84.792	< .001
Difference of nasality	-0.056	0.013	16.462	< .001
Difference of voiceless fricatives	-0.129	0.021	35.990	< .001
Difference of voiced plosives	0.083	0.013	36.251	< .001
Difference of syllable	-0.038	0.013	7.788	< .01
Difference of popularity	-0.313	0.013	502.361	< .001
Interactions with age				
<i>Difference of backness</i>	0.039	0.012	9.887	< .01
<i>Difference of nasality</i>	-0.002	0.012	0.030	.860
<i>Difference of voiceless fricatives</i>	0.028	0.019	2.081	.149
<i>Difference of voiced plosives</i>	0.004	0.013	0.116	.733
<i>Difference of syllable</i>	0.039	0.011	10.707	< .01
<i>Difference of popularity</i>	-0.112	0.012	80.186	< .001
Interactions with sex				
<i>Difference of backness</i>	0.008	0.027	0.102	.748
<i>Difference of nasality</i>	-0.030	0.027	1.165	.280
<i>Difference of voiceless fricatives</i>	0.028	0.019	0.000	.989
<i>Difference of voiced plosives</i>	-0.001	0.013	0.004	.948
<i>Difference of syllable</i>	0.039	0.011	4.309	< .05
<i>Difference of popularity</i>	-0.112	0.012	1.344	.246

Perceptions of femininity

Table 3 and Table 4 respectively reports the results with LDA scores and the phonetic variables. Overall, names whose LDA scores was lower were considered on average more feminine, i.e., phonetic femininity predicted perceived femininity ($\chi^2_1 = 81.218, p < 0.001$). Moreover, relatively newer names were considered as more feminine ($\chi^2_1 = 553.499, p < 0.001$). However, relatively older individuals judged older names as more feminine ($\chi^2_1 = 39.848, p < 0.001$), and male judges preferred relatively more phonetically masculine female names ($\chi^2_1 = 31.453, p < 0.001$).

In details, names with voiceless fricatives in their stressed syllable were considered as more feminine ($\chi^2_1 = 140.866, p < 0.001$). Names with nasal vowels in their stressed syllable and/or voiced plosives in their first syllable were considered less feminine (both $p < .05$).

Backness and name length did not predict perceived femininity (both $p > .05$) although older individuals considered names whose backness and nasality was higher as more feminine ($\chi^2_1 = 21.343, p < 0.001$); such individuals also preferred older names ($\chi^2_1 = 23.519, p < 0.001$). Lastly, male judges considered female names with nasal vowels in their stressed syllable and voiced plosives in their first syllable as more feminine (both $p < .01$).

Table 3. Results of GLMM for the LDA scores.

	Estimate	SE	χ^2	<i>p</i> value
Intercept	0.017	0.012	/	/
Difference in LDA scores	-0.124	0.013	81.218	< 0.001
Difference in popularity	0.335	0.014	553.499	< 0.001
Interactions with age				
<i>Difference in LDA scores</i>	0.004	0.012	0.136	.712
<i>Difference in popularity</i>	-0.080	0.012	39.848	< 0.001
Interactions with sex				
<i>Difference in LDA scores</i>	0.155	0.027	31.453	< 0.001
<i>Difference in popularity</i>	-0.040	0.028	2.125	.144

Table 4. Results of the GLMM with phonetic variables (numerical).

	Estimate	SE	χ^2	<i>p</i> value
Intercept	0.018	0.012	/	/
Difference of backness	0.003	0.015	0.056	.812
Difference of nasality	-0.142	0.015	86.102	< .001
Difference of voiceless fricatives	0.323	0.027	140.866	< .001
Difference of voiced plosives	-0.048	0.014	11.271	< .001
Difference of syllable	0.030	0.014	4.736	< .05
Difference of popularity	0.318	0.014	458.694	< .001
Interactions with age				
<i>Difference of backness</i>	-0.060	0.013	21.343	< .05
<i>Difference of nasality</i>	0.025	0.012	3.804	< .05
<i>Difference of voiceless fricatives</i>	-0.059	0.023	6.594	.149
<i>Difference of voiced plosives</i>	-0.003	0.012	0.098	.754
<i>Difference of syllable</i>	-0.008	0.012	0.426	.513
<i>Difference of popularity</i>	-0.064	0.013	23.519	< .001
Interactions with sex				
<i>Difference of backness</i>	0.054	0.031	3.134	.076
<i>Difference of nasality</i>	0.097	0.030	10.123	< .01
<i>Difference of voiceless fricatives</i>	-0.009	0.053	0.034	.852
<i>Difference of voiced plosives</i>	0.096	0.028	11.756	< .001
<i>Difference of syllable</i>	-0.032	0.028	1.314	.251
<i>Difference of popularity</i>	-0.037	0.029	1.633	.201

References

- Cutler, A., McQueen, J., & Robinson, K. (1990). Elizabeth and John: sound patterns of men's and women's names. *Journal of Linguistics*, 26(02), 471.
<https://doi.org/10.1017/S0022226700014754>
- Pitcher, B. J., Mesoudi, A., & McElligott, A. G. (2013). Sex-Biased Sound Symbolism in English-Language First Names. *PLoS ONE*, 8(6), e64825.
<https://doi.org/10.1371/journal.pone.0064825>
- Slepian, M. L., & Galinsky, A. D. (2016). The voiced pronunciation of initial phonemes predicts the gender of names. *Journal of Personality and Social Psychology*, 110(4), 509–527.
<https://doi.org/10.1037/pspa0000041>

Suire, A., Bossoms Mesa, A., Raymond, M., & Barkat-Defradas, M. (2019). Sex-biased sound symbolism in French first names. *Evolutionary Human Sciences*, 1. <https://doi.org/10.1017/ehs.2019.7>

Article 7

**Lost in translation: sound symbolic patterns in French Pokémon
names**

Alexandre Suire, Michel Raymond & Melissa Barkat-Defradas

Article en revue dans *Phonetica*

Résumé

Le corpus onomastique des Pokémon constitue une excellente opportunité pour étudier le symbolisme phonétique dans une perspective interculturelle, afin de mettre en évidence la manière dont le système phonologique et la prosodie spécifique à une langue contraignent et façonnent certains motifs phonétiquement symbolique. En effet, si les facteurs non linguistiques sont maintenus constants (i.e. les propriétés physiques de force, de taille et de poids), les facteurs linguistiques varient d'une langue à l'autre (i.e. noms uniques pour chaque langue). Bien qu'il existe une petite proportion de créatures iconiques dont le nom ne change pas (notamment « *Pikachu* »), chaque créature possède un nom différent selon la langue/culture considérée. De plus, ce corpus fournit une longue liste d'éléments lexicaux (plus de 800 noms propres uniques par langue).

Dans ce contexte, nous avons étudié les relations entre la qualité de la voyelle dans la syllabe accentuée (ouverture vs. fermeture, avant vs. arrière), le nombre moyen d'obstruantes voisées et non voisées, le nombre total de syllabes et de phonèmes avec les caractéristiques de puissance (force), la taille, le poids et l'évolution de chaque Pokémon. A cette fin, nous avons transcrit phonétiquement tous les noms des Pokémon actuellement disponibles (dernièrement consulté en Décembre 2017). Nos résultats ont montré que les voyelles centrales et ouvertes (/a/) de la syllabe accentuée sont associées à des Pokémon plus forts, plus grands et plus lourds que ceux contenant des voyelles d'avant, d'arrière et fermées (e.g. /i/, /u/, /o/). En outre, les premiers présentaient aussi une plus grande quantité de phonème indépendamment du niveau d'évolution. Les analyses avec et sans les noms non traduits du Japonais ne révèlent aucune différence significative.

Nos résultats sont partiellement compatibles avec la théorie du « *code-fréquence* », mis en évidence par les Pokémon possédant la voyelle /a/, dont la fréquence fondamentale intrinsèque et la dispersion formantique est parmi les plus basses comparativement à d'autres

voyelles. De manière intéressante, le français, l'anglais et le japonais présentent quelques similarités notoires, notamment la relation longueur des mots - plus fort/grand/large, et certaines tendances observées avec les voyelles, dont le /i/ et le /u/. Notons toutefois que la taille d'effet est extrêmement faible (variance expliquée avoisinant les 3%), suggérant qu'au moins en Français, les motifs phonétiquement symboliques sont peu généralisables à l'ensemble du corpus. D'autres études systématiques peuvent être envisagées dans des langues encore inexplorées.

Lost in translation: sound symbolic patterns in French Pokémon names

Authors: Alexandre Suire¹, Michel Raymond¹, Melissa Barkat-Defradas¹

Corresponding author: Alexandre Suire

E-mail: alexandre.suire@umontpellier.fr

Fax: +33 4 67 14 36 22

Tel: +33 4 67 14 49 66

¹ Institut des Sciences de l'Évolution de Montpellier, Université de Montpellier, Centre National de la Recherche Scientifique, Institut pour la Recherche et le Développement, École Pratique des Hautes Études – Place Eugène Bataillon 34095 Montpellier, France.

Abstract

Sound symbolism or iconicity, i.e., the idea that vocal sounds carry meaning in and of themselves, directly challenges the arbitrariness of the sign in human languages. One particular way to tackle this issue is to study sound symbolic patterns in a cross-linguistic perspective. The relatively large Pokémon onomastic corpus offers an opportunity to question the topic as extra-linguistic factors are cross-culturally held constant (i.e., physical properties such as height and weight) while linguistic factors vary (i.e., unique proper names from one language to another). In this context, two studies already showed that sound symbolic patterns were effective within Japanese and English Pokémon names, highlighting particular similar attributes but also differences rooted in language-specific structural constraints. Here, we contribute to this line of research by examining the issue in French. We investigated the relationships between the vowel quality in the stressed syllable, the mean number of voiced and voiceless consonants, the total number of syllables and phonemes (i.e. name length) with the Pokémons' overall power characteristics, height, weight and evolution levels. Our results show that central and open vowels in the stressed syllable were associated to stronger, taller and heavier Pokémons compared to those containing front, back and close vowels. In addition, the former had significantly more phonemic segment counts in their names. We discuss our results in light of those observed in Japanese and English Pokémon names as well as their implications for cross-linguistic sound symbolic patterns studies.

Keywords: sound symbolism; iconicity; onomastics; frequency-code; Pokémon; language.

Introduction

Although there is a substantial debate in the language sciences over the best way to characterize the unique features of the human's system of communication as a symbolic system of representation, there is a general consensus that the relationship between the sound (i.e., the signifier) and the meaning (i.e., the signifiant) of a word is arbitrary (de Saussure, 1916; Hockett and Hockett, 1960). However, several studies suggest that this rule is not systematically verified: sound symbolism, i.e., the idea that vocal sounds carry meaning in and of themselves, are more common and widespread in the languages of the world than previously thought (Blasi et al. 2016; Dautriche et al., 2017; Levickij, 2013).

The famous examples of sound sequences [kiki] and [buba] well illustrate this phenomenon: subjects agree to match angular and small shapes to the former sequence and curved and large shapes to the latter in a proportion of 9 to 1 (Kohler and Emery, 1947). This result has been replicated in various languages such as English, Spanish, Himba (i.e., a variety of the Herrero language native to Namibia, Botswana and Angola), Swahili and the Bantu dialect of Kitwonge (Bremner et al. 2013; Davis, 1961; Köhler, 1929; Maurer et al., 2006; Ramachandran and Hubbard, 2001), and even in 2.5-year-old children (Maurer et al., 2006). A common explanation is that, in many natural languages, sounds can be interpreted in terms of shape and size-related features based on their acoustic properties: the sounds' intrinsic fundamental frequency (i.e., the perceptual correlate of vocal pitch), the formant frequencies (i.e., the resonances of the vocal tract, which is the perceptual correlate of the vocal timbre) and the formants' dispersion (i.e., the relative distance between two consecutive formants) (Knoeferle et al., 2017; Ohala, 1994; Ultan, 1978).

Some of these associations can be understood through the lens of the "Frequency-code" hypothesis (Ohala, 1984). Within this theoretical framework, words denoting or connoting smallness tend to exhibit a disproportionate incidence of vowels and consonants characterized

by high acoustic frequencies. Conversely, words denoting or connoting largeness use segments with low acoustic frequency components. Additionally, Ohala (1984) also points to the biological bases of sound symbolism by analyzing close-contact competitive encounters between vocalizing animals (including humans), in which the aggressor uses low frequency staccato to appear as large as possible (e.g. dog's growl) while the opponent produces high frequency vocalizations to be perceived as smaller, non-threatening (e.g. dog's whine, yelp). The author draws parallel conclusions for speech sounds, where high-frequency sounds could be used to signal not only smallness, but also defense, politeness, and submission whereas low-frequency sounds would convey impressions of largeness, authority, aggression, and self-confidence.

Furthermore, higher frequency vowels, such as /i/ or /e/, which are typically associated with small size impressions, were also reported to be associated to non-size related features such as sharpness, rapid movement, and physically tall objects (Haynie et al., 2014; Newman, 1933; Sapir, 1929; Shinohara and Kawahara, 2010; Ultan, 1978). Likewise, lower frequency vowels, like /u/ or /a/, which symbolize larger size features, are also related to perceived physical properties such as softness, slow movement, and physically short objects. Regarding consonants, acoustic differences may also be perceptually associated to different shapes and figures. For instance, voiced bilabial consonants (i.e., /b/ and /m/) and other voiced consonants (e.g., /l/ and /n/) tend to be associated with rounded shapes, while voiceless stops (i.e., /k/, /p/ and /t/) tend to be associated with sharp forms (Maurer et al., 2006; Nielsen and Rendall, 2013). More generally, voiced obstruents (i.e., speech sounds that are produced by obstructing – either completely or partially – the airflow at the glottis) produce lower frequency sounds than their voiceless counterpart due to the vibration of the vocal cords, which adds a new source of laryngeal noise. Likewise, fricatives (e.g. /s/ or /ʃ/) exhibit higher frequencies than plain plosives (e.g. /p/, /b/ or /d/), while dental (e.g. /θ/, /ð/), alveolar (e.g. /s/, /n/), palatal (e.g. /c/,

/j/) and front velars (e.g. /k/, /x/) exhibit higher frequency than labials (e.g. /b/, /m/) and back velar consonants (e.g. /g/, /ŋ/).

Motivation for iconicity can also rest on physiological argument, i.e., articulatory feedback is associated with proprioceptive sensation. For instance, the size-symbolic dichotomy between /a/ and /i/ as representing respectively “*big*” vs. “*small*” impressions can be explained by the differences in the aperture of the mouth or the size of the air channel. Therefore, according to the “Frequency-code” hypothesis, the results universally observed, for instance, in [kiki] vs. [buba] experiments, can be interpreted by suggesting an acoustic and visual synesthesia between the sound units that are the front vowels /i/ vs. the back vowel /u/ (i.e., its diffuse and acute quality vs. compact and deep timbre, respectively) and the shape of the associated figure (i.e., vertical, sharpened, small and thin for /i/ vs. horizontal, rounded and wide for /u/, Jakobson et al., 1978). Lastly, it has been suggested that size-related sound symbolic patterns may be confounded by the shapes and forms of letters rather than their acoustic properties, at least among literate subjects (Cuskley et al., 2017).

One particular way to better understand the fascinating issue of sound and meaning relationships is to study phonetic symbolism in a cross-linguistic perspective. In this context, the Pokémon universe offers an interesting, but yet largely unexplored cross-linguistic onomastic corpus where non-linguistic factors are held constant (i.e., physical properties such as height or weight) while linguistic factors (i.e., unique names) vary from one culture to another. Moreover, it provides a large list of lexical items (more than 800 unique proper names per language) that largely exceeds the relatively short lists of words used in previous studies. To draw a few examples, Blasi et al. (2016) used between 28 and 40 lexical items drawn from the Swadesh list, while Johansson (2017) based his work on 28 opposing word-pairs and Wichmann et al. (2010) on 40 basic vocabulary items.

Recently, two studies showed that sound symbolic patterns existed in Japanese and English Pokémon names (Kawahara et al., 2018; Shih et al., 2018). Pokémon is a video role-playing game created in 1996 by Nintendo© where the main goal is to catch all sorts of eponym fictional creatures that all display specific dedicated names and characteristics in terms of power, weight and height. Some Pokémon also have the ability to evolve into completely new creatures, which increase their overall inherent characteristics. In the Japanese study, voiced obstruents, mora counts (rhythmic units that are subtly different from syllables), and the vowel quality of initial syllables were correlated to the Pokémon's power characteristics, height, weight, and evolution levels. For instance, the authors found that the number of moras and voiced obstruents (e.g., /b/, /d/, /g/, /z/) were positively correlated with all these aspects. Vowel quality also seemed to have an influence on size and weight: Pokémon characters with initial high vowels tended to be smaller and lighter. In the English study, the number of segments and the frequency of occurrence of open and back vowels (e.g. /æ/, /a/, /aʊ/, /aɪ/ and /ɔ/) were positively correlated to increased weight while height was positively correlated with the presence of more alveolar consonants (such as /s/) and low vowels, and negatively associated with the presence of high vowels (e.g. /i/, /ɪ/, /i/, /u/ and /ʊ/). In addition, power characteristics were also positively correlated with name length and the presence of several alveolar consonants. Lastly, more evolved Pokémon also tended to have more segments in their names. Those general findings are interesting in how some are both compatible with the “Frequency-code” hypothesis, but most importantly in how they highlight some interesting cross-linguistic differences in naming processes.

For instance, the translation of Pokémon names from Japanese to French seems not always literal. Most French Pokémon names, as well as English ones, are “portmanteau” words that are a lexical blend of two words. For instance, “*Feunard*” (/fø.naʁ/), which results from the blending of “*feu*” (i.e., “fire”) and “*renard*” (i.e., “fox”) refers to a fox that spits fire. In both

English and Japanese, this creature is respectively named “*Ninetales*” and “*Kyukon*” which both refer to a fox with nine tails from the Japanese mythology. Another example of a French translation is “*Rondoudou*” (/ʁɔ̃.du.du/) formed from “*rond*” (i.e., “round”) and “*doudou*” (i.e., a wubby). In English, this Pokémon is named “*Jigglypuff*”, which is also a portmanteau word formed out of the words “*jiggle*” and “*puff*”, and in Japanese “*Purin*” (i.e., a Japanese custard pudding). Most French Pokémon names seem rethought to fit the French lexicon rather than being literally transcribed from Japanese (with a few exceptions, such as “*Pikachu*”). Indeed, as Brougère (2004) notes: “Nintendo, aware of the importance of naming, translated the creatures’ names into terms that artfully reflect the language and culture of French children. This is an essential aspect of cultural and linguistic localization that is necessary if children are to relate intimately to a product [...]. The characters’ names in French, as in Japanese and English and other languages, convey a core characteristic of each culture”.

However, it is important to note that it is difficult to know how these naming patterns are processed, as very little is known on the way such a work is realized. From the very few available general public resources, the names of the first 251 French Pokémon were created by J. Bardakoff, a French translator hired by Nintendo©, who used different means of lexical creation such as play on words, slang and youth language, portmanteau words formed out of Japanese, English and French words and neologisms. These names were based on semantic similarities existing between the visual identity of the creature itself and the auditory impression conveyed by its name’s sound structure (for a detailed review on how the translator created the French Pokémon’s names, see <https://www.20minutes.fr/societe/1792619-20160226-carapuce-bulbizarre-rondoudou-nom-pokemon> and <https://www.liberation.fr/apps/2016/06/pokemon/>). Thus, although in this specific case it can be argued that naming processes may only be demonstrations of influences of sound-symbolism in business’ communication (Klink, 2000; Wu et al., 2013) or an artifact of a small number of programmers’ or marketers’ imagination, it

is still interesting to see whether French Pokémon names are subjected to specific sound symbolic patterns and, if so, how these differ from those observed in Japanese and English.

In this context, we hypothesize that in the absence of literal translation in Pokémon names, if sound symbolic patterns are observed, they will most likely differ from those found in English and Japanese in order to fit the French lexical constraints and culture. Moreover, since translated Pokémon names are available in many languages, including, for instance, English, French, German, Spanish and Italian, such data provides a relevant corpus for cross-linguistic studies on sound symbolism. We expect to find sound symbolic patterns in French Pokémon names, namely low-frequency vowels in the last stressed syllable (i.e., French words being stressed on the final syllable) and a larger number of voiced consonants and higher segment counts (as observed in Japanese and English) in stronger, taller and heavier Pokémon, as opposed to high-frequency vowels, voiceless consonants and less segment counts in weaker, smaller and lighter Pokémon.

Methods

1. Data treatment

In total, 807 Pokémon names were retrieved from <https://www.pokepedia.fr/> (December, 2017) and compiled into a single database, stating for each Pokémon its French name, evolution level, power characteristics (i.e., hit points, attack, defense, special attack, special defense, speed and the total, which is the sum of these characteristics), height (meters) and weight (kilograms) from the current available games. Following Kawahara et al. (2018) method, Pokémon with no possible evolution or at their first stage of evolution were coded as 0; if a Pokémon is an evolution from a previous one, we coded it as 1, if there is another possible evolution, we coded it as 2 (note that the game is constrained to a maximum of two levels of evolution). Pre-evolution stages were introduced later in the games (i.e., Pokémon that did not have a previous

evolution) and were coded as -1. Typically, Pokémons' overall characteristics increase after each evolution level.

Two native French speakers (one familiar with the game and another independent and non-familiar) transcribed phonetically each French Pokémon name using the International Phonetic Alphabet. For each name, the total number of syllables and phonemes, the total number of different consonants (i.e., fricatives and plosives) within the full name and the type of vowel (i.e., vowel quality) contained in the stressed syllable (i.e., the last one) were retrieved.

For the vowel quality, we defined vowels along their F1 and F2 formants' dimensions: respectively the degree of aperture of the mouth (i.e., vowel's height) and their place of articulation (i.e., vowel's frontness vs. backness). For vowel's height, they were classified into four groups: close vowels (/i/, /y/, /u/), close-mid vowels (/e/, /ø/, /o/), open-mid vowels (/ɛ/, /ɛ̃/, /œ/, /ɔ/, /ɔ̃/) and open vowels (/a/, /ɑ̃/). For vowel's backness, they were classified into three groups: front vowels (/i/, /y/, /e/, /ɛ/, /ɛ̃/, /œ/, /ø/), central vowels (/a/, /ə/) and back vowels (/ɔ/, /o/, /u/, /ɑ̃/, /ɔ̃/).

Typically, open vowels have higher F1 frequencies than close vowels, while front vowels have higher F2 frequencies than back vowels. However, as vowels are defined on both dimensions, the formants' dispersion can be chosen as a more precise criteria that determines their perceptual correlate to low- or high-frequency sounds, as previous studies have shown that relatively small formants' dispersion are perceptually associated to larger, more powerful and dominant individuals (Pisanski et al., 2014; Puts et al., 2006; Puts et al., 2007; Sell et al., 2010). In this sense, we predict that low-frequency vowels that have the smallest formants' dispersion (i.e., /u/, /o/, /ɔ/, /ɑ̃/, /ɔ̃/ and /a/) will be associated to stronger, taller and heavier Pokémons, as opposed to high-frequency vowels that have higher formants' dispersion (i.e., /i/, /y/, /e/, /ɛ/, /ɛ̃/, /œ/, /ø/) to weaker, smaller and lighter Pokémons.

For consonants, we distinguished between voiced plosives (i.e., /b/, /d/ and /g/), voiceless plosives (i.e., /p/, /t/ and /k/), voiced fricatives (i.e., /v/, /z/, and /ʒ/) and voiceless fricatives (i.e., /f/, /s/, and /ʃ/). Plosives are produced by a complete closing of the airflow that causes its blocking before the air is suddenly released, producing an impulsive noise (i.e., the burst); this sound is typical of consonantal plosives. Fricatives are produced with a major constriction of the airflow, which acoustically causes a turbulent noise. Since the higher the velocity of the airflow, the higher the noise in the frequency spectrum. Contrastively, plosives produce sounds in a bandwidth that are generally lower in frequencies than fricatives do (Stevens, 1998). The consonant's voicing (voiced vs. voiceless), which is determined by whether the vocal folds vibrate or not during articulation, adds a new source of laryngeal noise which explains why voiced consonants are lower in frequencies than voiceless ones. Grouping consonants this way allows a more fine-grained analysis than simply classifying them into voiced and voiceless obstruents categories as previously done (Kawahara et al., 2018; Shih et al., 2018). In this context, we predict that voiced consonants, either fricatives or plosives, will be associated to stronger, taller and heavier Pokémons, while voiceless consonants will be associated to weaker, smaller and lighter Pokémons.

2. Statistical analyses

First, a multivariate analysis of variance was conducted to assess the relationship between the evolution levels with the vowel quality of the last syllable (height and backness), the mean number of voiced and voiceless fricatives and plosives and the mean number of syllables and phonemes in Pokémon names. This analysis was conducted to assess if phonetic material, or the vowel quality, increased or changed after each evolution level to depict stronger, taller and heavier Pokémons. If the evolution level is significantly related to any of these variables, post-hoc comparisons were performed in order to assess which level differs from one another (Tukey's range test with a Bonferroni correction).

Second, multiple linear models were performed to assess if the sound structure of Pokémon names could be predicted by their total power characteristics, height and weight. These three variables were thus used as the response variables. Since height and weight showed right-skewed distributions, log transformations were performed in order to produce a normal distribution. As the number of syllables and the number of phonemes both measure segment counts, and as a preliminary investigation revealed a highly significant correlation ($r(805) = .63, p < .001$), we ran separate analyses with models only including syllables and models including phonemes, which also avoids problems of multicollinearity. Moreover, in order to avoid other problems of multicollinearity, as the vowel's backness and vowel's height strongly overlap in categories, we also ran separate models with the former and those with the latter.

If the vowel quality affected any of the response variables (power, height or weight), post-hoc comparisons were performed to assess which type of vowels (i.e., vowel's height: close, close-mid, open-mid and open; vowel's backness: front, central or back) differed from one another (Tukey's range test with a Bonferroni correction). Furthermore, in all models, the mean number of voiced and voiceless plosives and fricatives as well as the mean number of syllables and phonemes in Pokémon names were standardized and also used as predictors.

Contrary to correlation tests used in (Kawahara et al., 2018; Shih et al., 2018), linear models allow to control for the presence of each phonetic variable with one another in a name. The evolution levels were not added as other explanatory variables in these linear models as the aim is to study the percentage of variance only explained by possible sound symbolic patterns. For all models, the significance of each explanatory variable was assessed from the comparison of the model excluding the explanatory variable with the model including all the other variables (Type II analysis of variance).

All results were considered significant at the 0.05 level. All statistical analyses were performed using the R software (version 3.4.0).

Results

Descriptive statistics for all the phonetic variables are given in Table 1.

Evolution level <i>N</i> = 807	-1 <i>n</i> = 16	0 <i>n</i> = 424	1 <i>n</i> = 276	2 <i>n</i> = 91
Vowel's backness				
Front (<i>n</i> = 328)	8	172	117	31
Central (<i>n</i> = 164)	2	84	54	24
Back (<i>n</i> = 315)	6	168	105	36
Vowel's height				
Close (<i>n</i> = 236)	7	133	75	21
Close-mid (<i>n</i> = 163)	4	87	55	17
Open-mid (<i>n</i> = 211)	2	106	79	29
Open (<i>n</i> = 197)	3	98	67	24
Number of syllables	2.81 ± 0.65	2.91 ± 0.65	2.96 ± 0.66	3.07 ± 0.67
Number of phonemes	6.62 ± 1.5	6.96 ± 1.40	7.38 ± 1.33	7.79 ± 1.26
Number of voiced fricatives	0.68 ± 0.87	0.80 ± 0.76	0.91 ± 0.79	0.95 ± 0.78
Number of voiced plosives	0.75 ± 1	0.47 ± 0.66	0.55 ± 0.67	0.62 ± 0.71
Number of voiceless fricatives	0.18 ± 0.54	0.53 ± 0.62	0.61 ± 0.63	0.54 ± 0.60
Number of voiceless plosives	0.87 ± 0.80	1.02 ± 0.90	0.95 ± 0.92	1.01 ± 0.92

Table 1. For the categorical variables “vowel’s backness” and “vowel’s height”, the number of names containing a specific vowel in their last syllable (backness: front, central or back; height: close, close-mid, open-mid, open) are given for each evolution level and independently of evolution level (in parentheses). For the other variables, values show mean ± standard-deviation.

1. The influence of evolution levels on Pokémons’ sound structure

Overall, the Pokémons’ names sound structure significantly varied as a function of the evolution levels ($F(24, 2394) = 2.48, p < .001$). Upon closer inspection, the evolution levels significantly influenced the mean number of voiceless fricatives ($F(3, 803) = 2.81, p < .05$) and the mean number of phonemes ($F(3, 803) = 12.27, p < .001$). Post-hoc comparisons revealed a modest but significant increase in the number of voiceless fricatives between the “-1” and “1” Pokémon levels ($t(803) = 2.88, p < .05$); while the mean number of phonemes was greater for “2” level

Pokémons compared to the “-1” level ($t(803) = 2.80, p < .05$) and the “0” level Pokémons ($t(803) = 4.43, p < .001$). The “1” level Pokémons also had significantly more phonemes than the “0” level Pokémons ($t(803) = 3.01, p < .05$). All other levels did not show differences in the mean number of voiceless fricatives and phonemes (all $p > .05$) and all the other phonetic variables did not significantly vary or changed between each evolution level (all $p > 0.05$).

2. Relationships between Pokémons’ total characteristics, height and weight with the sound structure of their names

The results of the linear models are reported in Table 2 for the model with vowel’s backness and Table 3 for vowel’s height.

	β	SE	F	df	p
<i>N</i> = 807					
R ₂ (Total) = 2.41%					
R ₂ (Height) = 2.63%					
R ₂ (Weight) = 2.63%					
Intercept	414.03	6.18	/	/	/
	-0.10	0.04			
	<i>3.15</i>	<i>0.08</i>			
Vowel's backness (Back)			4.10	2	< .05
			7.82		< .001
			7.22		< .001
Central	29.05	10.66			
	0.21	0.07			
	<i>0.35</i>	<i>0.15</i>			
Front	2.46	8.67			
	-0.07	0.06			
	<i>-0.21</i>	<i>0.12</i>			
Number of voiced fricatives	14.33	5.04	8.06	1	< .01
	0.09	0.03	7.89		< .01
	<i>0.19</i>	<i>0.07</i>	7.22		< .01
Number of voiced plosives	0.99	6.21	0.02	1	.87
	0.03	0.04	0.57		.44
	<i>0.07</i>	<i>0.08</i>	<i>0.76</i>		.38
Number of voiceless fricatives	8.36	6.55	1.62	1	.20
	0.04	0.04	0.80		.37
	<i>0.02</i>	<i>0.09</i>	<i>0.07</i>		0.78
Number of voiceless plosives	-4.06	4.60	0.77	1	.37
	0.01	0.03	0.19		0.65
	<i>-0.01</i>	<i>0.06</i>	<i>0.03</i>		.85
Number of syllables	14.16	5.91	5.74	1	< .05
	0.07	0.04	3.07		.08
	<i>0.15</i>	<i>0.08</i>	<i>3.60</i>		.06

Table 2. Results of the linear models with vowel's backness. Bold values correspond to the analysis with the total power characteristics; normal values to height; italic values to weight. For each variable, the F statistic, the degrees of freedom and the p values associated from the Fisher test of the comparison between the full model and the model without the variable are given (Type II analysis of variance). For the categorical variable "vowel's backness", the estimates are given for one category compared to the reference category (in parentheses). The adjusted R² are given for each model.

	β	SE	F	df	p
<i>N</i> = 807 R ₂ (Total) = 3.04% R ₂ (Height) = 2.64% R ₂ (Weight) = 2.78%					
Intercept	407.50	7.15	/	/	/
	-0.21	0.05			
	2.88	0.09			
Vowel quality (Close)			3.87	3	< 0.01
			5.56		< .001
			5.62		< .001
Close-mid	2.26	11.32			
	0.09	0.07			
	0.19	0.15			
Open-mid	19.05	10.45			
	0.11	0.07			
	0.25	0.14			
Open	32.71	10.71			
	0.29	0.07			
	0.61	0.15			
Number of voiced fricatives	13.09	5.12	6.52	1	< .05
	0.09	0.03	7.29		< .01
	0.18	0.07	6.33		< .05
Number of voiced plosives	0.89	6.20	0.02	1	.88
	0.03	0.04	0.47		.49
	0.07	0.08	0.67		.41
Number of voiceless fricatives	7.97	6.54	1.48	1	0.22
	0.03	0.04	0.69		.40
	0.02	0.09	0.05		0.82
Number of voiceless plosives	-4.43	4.62	0.91	1	.33
	0.01	0.03	0.18		.66
	-0.01	0.06	0.04		.84
Number of syllables	15.42	5.96	6.68	1	< 0.01
	0.07	0.04	2.86		.09
	0.15	0.08	3.27		0.07

Table 3. Results of the linear models with vowel's height. Bold values correspond to the analysis with the total power characteristics; normal values to height; italic values to weight. For each variable, the F statistic, the degrees of freedom and the p values associated from the Fisher test of the comparison between the full model and the model without the variable are given (Type II analysis of variance). For the categorical variable "vowel's height", the estimates are given for one category compared to the reference category (in parentheses). The adjusted R² are given for each model.

2.1 Pokémons' total power characteristics

Vowel's backness had a significant influence on Pokémons' power characteristics ($F(2, 799) = 4.10, p < .05$). Post-hoc comparisons revealed that Pokémons with central vowels were significantly stronger than those with front vowels ($t(799) = 10.65, p < .05$) and those with back vowels ($t(799) = 10.65, p < .05$). In the same model, the mean number of voiced fricatives and syllables significantly and positively predicted the Pokémons' power characteristics (respectively $\beta = 14.33, t(799) = 2.83, p < .01$; $\beta = 14.33, t(799) = 2.39, p < .05$).

Similarly, vowel's height had a significant influence on Pokémons' power characteristics ($F(2, 799) = 3.87, p < .01$). Post-hoc comparisons revealed that Pokémons with open vowels in their last syllable were significantly stronger than those with close vowels ($t(799) = 3.05, p < .05$) and those with close-mid vowels ($t(799) = 2.59, p < .05$). In the same model, the mean number of voiced fricatives and syllables significantly and positively predicted the Pokémons' total characteristics (respectively $\beta = 13.09, t(799) = 2.55, p < .05$; $\beta = 15.42, t(799) = 2.58, p < .01$).

In both the vowel's backness and height model, replacing the mean number of syllables by the mean number of phonemes showed that the latter significantly and positively predicted the Pokémons' power characteristics (respectively for backness and height: $\beta = 8.65, t(799) = 2.64, p < .05$; $\beta = 8.21, t(799) = 2.48, p < .05$) but the mean number of voiced fricatives was not significantly related to the Pokémons' power characteristics anymore (all $p > .05$). Lastly, replacing the mean number of syllables by the mean number of phonemes did not qualitatively change the results for vowel's backness and height ($p > .05$); with one exception being that Pokémons with open vowels in their last syllable were still stronger than those with close-mid vowels, but the difference was not significant anymore ($t(799) = 2.44, p = .07$).

2.2 Pokémons' height

Vowel's backness had a significant influence on Pokémons' height ($F(2, 799) = 7.82, p < .001$). Post-hoc comparisons revealed that Pokémons with central vowels were significantly taller than those with front vowels ($t(799) = 3.92, p < .001$) and those with back vowels ($t(799) = 2.99, p < .01$). In the same model, the mean number of voiced fricatives significantly and positively predicted the Pokémons' height ($\beta = .10, t(799) = 2.80, p < .01$).

Vowel's height had a significant influence on Pokémons' height ($F(2, 799) = 5.56, p < .001$). Post-hoc comparisons revealed that Pokémons with open vowels in their last syllable were significantly taller than those with close vowels ($t(799) = 4.05, p < .001$). In the same model, the mean number of voiced fricatives significantly and positively predicted the Pokémons' height (respectively $\beta = .09, t(799) = 2.70, p < .01$).

In both the vowel's backness and height model, replacing the mean number of syllables by the mean number of phonemes showed that the latter significantly and positively predicted the Pokémons' height (respectively for backness and height: $\beta = 0.06, t(799) = 2.89, p < .01$; $\beta = .06, t(799) = 2.63, p < .01$) but the mean number of voiced fricatives was not significantly related to any of the Pokémons' height anymore (all $p > .05$). Lastly, replacing the mean number of syllables by the mean number of phonemes did not qualitatively change the results for vowel's backness and height (all $p < .05$).

2.3 Pokémons' weight

Vowel's backness had a significant influence on Pokémons' weight ($F(2, 799) = 7.31, p < .001$). Post-hoc comparisons revealed that Pokémons with central vowels were significantly heavier than those with front vowels ($t(799) = 3.81, p < .001$) and those with back vowels ($t(799) = 2.37, p < .01$). In the same model, the mean number of voiced fricatives significantly and positively predicted the Pokémons' weight ($\beta = .19, t(799) = 2.68, p < .01$).

Vowel's height had a significant influence on Pokémons' weight ($F(2, 799) = 5.62, p < .001$). Post-hoc comparisons revealed that Pokémons with open vowels in their last syllable were significantly taller than those with close vowels ($t(799) = 4.06, p < .001$). In the same model, the mean number of voiced fricatives significantly and positively predicted the Pokémons' height ($\beta = .18, t(799) = 2.7, p < .01$).

In both the vowel's backness and height model, replacing the mean number of syllables by the mean number of phonemes showed that the latter significantly and positively predicted the Pokémons' weight (respectively for backness and height: $\beta = 0.13, t(799) = 2.80, p < .01$; $\beta = .12, t(799) = 2.54, p < .05$) but the mean number of voiced fricatives was not significantly related to the Pokémons' weight anymore (all $p > .05$). Lastly, replacing the mean number of syllables by the mean number of phonemes did not qualitatively change the results for vowel's backness and height (all $p < .05$).

Lastly, variance in power characteristics, height and weight explained by the phonetic features for all models did not exceed the 3% threshold, suggesting that factors other than sound structure are at play to depict the Pokémons' physical characteristics.

Discussion

The present results highlight some specific sound symbolic patterns that matched some findings observed in Japanese and English, but also revealed some key differences.

First, the mean number of phonemes increased as a function of the evolution level, as observed in Japanese for mora (Kawahara et al., 2018) and the number of phonemic segments in English (Shih et al., 2018). Therefore, a higher number of segments in French Pokémons' names also iconically indicate a greater level of evolution. Additionally, independently of evolution levels, Pokémons whose names were longer were significantly stronger, taller and heavier, with the number of phonemes being a better predictor than the number of syllables. This "longer-is-stronger-taller-heavier" relationship between word length and physical

characteristics, similar to their Japanese and English counterparts, can be explained through two linguistic phenomena: emphatic lengthening and/or reduplication. An example of emphatic lengthening in French is the adjective “*long*” that is usually pronounced /lɔ̃/. It can be written as “*loooong*” and pronounced /lɔ̃:/ with a stylistic lengthening of the vowel, which will be perceived as longer and inevitably interpreted as such. On the other hand, reduplication is the phenomenon that involves the repetition of a word or one of its morphemes for either a specific grammatical purpose or for creating new words (e.g. “*kirakira*” in Japanese, meaning sparkling or shining). However, this process of reduplication is absent in French. To our knowledge, we did not observe any French Pokémon names that presented either of these two morphological characteristics. Thus, we suggest a possible third explanation: on average, we observed that such Pokémons tended to exhibit longer words blending from one evolution to another, which could possibly give an impression of displaying “more complex” creatures for players. For instance, one could give a longer synonym for the following evolution: “*Kadabra*” (/ka.da.bʁa/, 3 syllables and 7 phonemes) becomes “*Alakazam*” (/a.la.ka.zam/, 4 syllables and 8 phonemes), both being a synonym to the term used by magicians during their shows. Other cases can be a switch of words that involve greater emphasis on a particular aspect. For instance, “*Roucops*” (/ʁu.kops/, 2 syllables and 6 phonemes) is formed with the words “*rouer*” (to “*beat up*”) and “*coups*” which refers to “*blows*”. Its evolution is “*Roucarnage*” (/ʁu.kɑ̃.naʒ/, 3 syllables and 8 phonemes), with “*coups*” being replaced by “*carnage*” (i.e., “*slaughter*”).

Secondly, the evolution levels merely but significantly affected the mean number of fricatives. This finding suggests that an increased number of fricatives (e.g., /s/, /v/ or /f/) may indicate a more powerful Pokémon, although this may have very limited perceptual efficiency to the player, as the effect is relatively small. Additionally, voiced fricatives were positively associated to simultaneously more powerful, taller and larger Pokémons. However, such relationships only held significant when controlling name length with the number of syllables

instead of phonemes. This raises the question of individuals perceiving Pokémon with these phonetic characteristics as stronger, taller and heavier (perception of syllables vs. perception of sounds). Interestingly, Shih et al. (2018) suggested that fricatives (either voiced or voiceless) may be associated to specific shapes and forms in Pokémon names, rather than power and size-related features. For instance, empirical evidence from other natural languages show that the /z/ and /v/ sounds are associated to broader, pointed, efficient and slower physical qualities (McCormick and Kim, 2015; Shinohara et al., 2016). An interesting avenue for further research is to test these kinds of associations with the specific morphological forms that Pokémon display.

Lastly, the vowel quality (either height or frontness/backness) did not change as a function of evolution levels, which highlights the fact that French translational process from one level to another primarily rests on new portmanteau words. For instance, the Pokémon “*Salamèche*” (/sa.la.mɛʃ/) is literally a small salamander with a wick (“*mèche*”) at the end of its tail. It evolves into “*Reptincel*” (/ʁɛp.tɛ̃.sɛl/) which is a portmanteau word composed of the French words reptile and spark (“*étincelle*”). It finally evolves into “*Dracaufeu*” (/dʁa.ko.fø/), with “*feu*” meaning fire, while “*draco*” refers to a dragon. Regardless of the evolution levels, Pokémon names containing central and open vowels, and most specifically the vowel /a/ in their stressed syllable, were stronger, taller and heavier when compared to those containing, front, back, mid-close and close vowels. Such findings are interesting regarding those unveiled in Japanese and English. Both studies reported association of low and high vowels to height and weight (although not significant in Japanese, Kawahara et al., 2018; Shih et al., 2018). Firstly, these results were observed on the first stressed syllable, while in French they are observable on the last stressed syllable, highlighting the fact that sound symbolic patterns are rooted in the prosodic peculiarities of the system under study. Secondly, results in Japanese and English, which are compatible with the “Frequency-code” hypothesis, also hold in the case of French:

the /a/ is a low-frequency vowel as it possesses one of the lowest formants' dispersion (i.e., the relative distance between F1 and F2) within the set of oral French vowels (i.e., ~608 Hz compared to ~2309 Hz for /i/ and ~488 Hz for /u/, (Georgeton et al., 2012). It is also one of the lowest-frequency vowels in terms of intrinsic fundamental frequency, compared to /i/ and /u/ (Whalen and Levitt, 1995). Interestingly, in humans, sounds with low pitch and formants' dispersion are also known to be perceptually associated to impressions of largeness and dominance (Borkowska and Pawlowski, 2011; Hodges-Simeon et al., 2010; Pisanski et al., 2014; Puts et al., 2006, 2007; Rendall et al., 2007).

Otherwise, another interesting explanation for such findings is that some vowels may project or convey specific impressions not related to physical characteristics, as stated above for voiced fricatives. For instance, two studies have investigated how specific vowels in English and German first names can influence facial attractiveness (Klenovsak et al., 2018; Perfors, 2004) Interestingly, a recent study showed an increase of French first names given to female babies since the 1960's that end with the vowel /a/ (e.g. "*Léa*", "*Alexandra*", "*Célia*", Suire et al., 2019), suggesting a higher attractive power (or some other traits related to success) for female individuals bearing such names. During translational processes, some phonetic variants may thus be considered more attractive or more suitable to depict the physical representation of a Pokémon perhaps in terms of colour or shape, rather than size or strength.

In this study, an important limitation is the general representativeness of the Pokémon datasets (whether in English, Japanese or in French) for the study of sound symbolic patterns in a cross-linguistic and cultural perspective. Indeed, they may have very limited relevance to broader linguistic tendencies. As stated earlier, these Pokémon names may well be a cultural artifact and may only reveal the cultural influences of the persons who created these fictional creatures. However, it is interesting to note that French translations were presented to the Japanese Nintendo developers' team, who held the final decision on which names were retained

for the French game version (see <https://www.20minutes.fr/societe/1792619-20160226-carapuce-bulbizarre-rondoudou-nom-pokemon>). Thus, one could hypothesize that the names retained for the French version could have sounded phonetically attractive to the Japanese developing team. However, as stated in the Introduction, the Nintendo's developing team held great importance of giving names that faithfully portrayed the French culture (Brougère, 2004).

In conclusion, the relevance of our findings is twofold: first in how general naming processes still highlight the underlying cognitive biases involved in sound symbolism when naming new objects, and secondly, to a lesser extent, as an interesting issue of translational processes of cultural artefacts between linguistic communities. Our study offers some account of the relationships between some phonetic variables and physical or abstract properties of fictional creatures. Although it may be limited to broader aspects of sound symbolism research, the Pokémon corpus offers an interesting database to investigate the question of sound symbolic patterns, and studies in other languages is a welcomed avenue for further research.

Data accessibility

Data can be accessible at [<https://figshare.com/s/1b4da4c1e54d9efd91eb>]

Funding statement

This study received no funding.

Competing interests

The authors have no competing interests.

Authors' contributions

AS and MBD did the phonetic transcriptions of Pokémon names. AS and MR carried the statistical analyses. AS drafted the manuscript. MBD and MR provided revisions on the manuscript.

Acknowledgments

The authors are thankful to S. Kawahara for comments on a previous draft of this manuscript.

References

- Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., & Christiansen, M. H. (2016). Sound–meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences*, *113*(39), 10818–10823. <https://doi.org/10.1073/pnas.1605782113>
- Borkowska, B., & Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Animal Behaviour*, *82*(1), 55–59. <https://doi.org/10.1016/j.anbehav.2011.03.024>
- Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., & Spence, C. (2013). “Bouba” and “Kiki” in Namibia? A remote culture make similar shape–sound matches, but different shape–taste matches to Westerners. *Cognition*, *126*(2), 165–172. <https://doi.org/10.1016/j.cognition.2012.09.007>
- Brougère, G. (2004). How much is a Pokémon worth? Pokémon in France. In *Pikachu’s global adventure: The rise and Fall of Pokémon* (Duke University Press, p. 193). Durham.
- Cuskley, C., Simner, J., & Kirby, S. (2017). Phonological and orthographic influences in the bouba–kiki effect. *Psychological Research*, *81*(1), 119–130. <https://doi.org/10.1007/s00426-015-0709-2>
- Dautriche, I., Mahowald, K., Gibson, E., & Piantadosi, S. T. (2017). Wordform Similarity Increases With Semantic Similarity: An Analysis of 100 Languages. *Cognitive Science*, *41*(8), 2149–2169. <https://doi.org/10.1111/cogs.12453>
- Davis, R. (1961). THE FITNESS OF NAMES TO DRAWINGS. A CROSS-CULTURAL STUDY IN TANGANYIKA. *British Journal of Psychology*, *52*(3), 259–268. <https://doi.org/10.1111/j.2044-8295.1961.tb00788.x>
- de Saussure, F. (1916). *Cours de linguistique générale*. Paris: Payot.

- Georgeton, L., Paillereau, N., Landron, S., Gao, J., & Kamiyama, T. (2012). Analyse formantique des voyelles orales du français en contexte isolé: à la recherche d'une référence pour les apprenants de FLE. *Conférence conjointe JEP-TALN-RECITAL 2012*, 8.
- Haynie, H., Bowern, C., & LaPalombara, H. (2014). Sound Symbolism in the Languages of Australia. *PLoS ONE*, 9(4), e92852. <https://doi.org/10.1371/journal.pone.0092852>
- Hockett, C. F., & Hockett, C. D. (1960). The Origin of Speech. *Scientific American*, 203(3), 88–97.
- Hodges-Simeon, C. R., Gaulin, S. J. C., & Puts, D. A. (2010). Different Vocal Parameters Predict Perceptions of Dominance and Attractiveness. *Human Nature*, 21(4), 406–427. <https://doi.org/10.1007/s12110-010-9101-5>
- Jakobson, R., Lévi-Strauss, C., & Mepham, J. (1978). *Six Lectures on Sound and Meaning*. Hassocks: Harvester Press.
- Johansson, N. (2017). Tracking Linguistic Primitives. *Dimensions of Iconicity*, 15(39), 91.
- Kawahara, S., Noto, A., & Kumagai, G. (2018). Sound Symbolic Patterns in Pokémon Names. *Phonetica*, 75(3), 219–244. <https://doi.org/10.1159/000484938>
- Klenovsak, D., Hartung, F., Santiago, L., & Zaefferer, D. (2018). Preprint Are Tims hot and Toms not? *Psyarxiv*. <https://doi.org/10.31219/osf.io/kj3b2>
- Klink, R. R. (2000). Creating Brand Names with Meaning: The Use of Sound Symbolism. *Marketing Letters*, 11(1), 5–20.
- Knoeferle, K., Li, J., Maggioni, E., & Spence, C. (2017). What drives sound symbolism? Different acoustic cues underlie sound-size and sound-shape mappings. *Scientific Reports*, 7(1). <https://doi.org/10.1038/s41598-017-05965-y>
- Köhler, W. (1929). *Gestalt Psychology (2nd. Ed.)*. New York: Liveright.

- Kohler, W., & Emery, D. A. (1947). Figural After-Effects in the Third Dimension of Visual Space. *The American Journal of Psychology*, 60(2), 159. <https://doi.org/10.2307/1417870>
- Levickij, V. V. (2013). Phonetic Symbolism in Natural Languages. *Glottology International Journal of Theoretical Linguistics*, 4(1), 20.
- Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: sound-shape correspondences in toddlers and adults. *Developmental Science*, 9(3), 316–322. <https://doi.org/10.1111/j.1467-7687.2006.00495.x>
- McCormick, K., & Kim, J. Y. (2015). Sound to Meaning Mappings in the Bouba-Kiki Effect. *CogSci*, 2015, 1565–1570.
- Newman, S. S. (1933). Further Experiments in Phonetic Symbolism. *The American Journal of Psychology*, 45(1), 53. <https://doi.org/10.2307/1414186>
- Nielsen, A. K. S., & Rendall, D. (2013). Parsing the role of consonants versus vowels in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 67(2), 153–163. <https://doi.org/10.1037/a0030553>
- Ohala, J. (1994). The frequency codes underlies the sound symbolic use of voice pitch. In *Sound Symbolism* (Cambridge University Press, pp. 325–347). Cambridge.
- Ohala, J. J. (1984). An Ethological Perspective on Common Cross-Language Utilization of F₀ of Voice. *Phonetica*, 41(1), 1–16. <https://doi.org/10.1159/000261706>
- Perfors, A. (2004). The effect of sound symbolism on perception of facial attractiveness. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 26(26), 2.
- Pisanski, K., Fraccaro, P. J., Tigue, C. C., O'Connor, J. J. M., & Feinberg, D. R. (2014). Return to Oz: Voice pitch facilitates assessments of men's body size. *Journal of Experimental*

- Psychology: Human Perception and Performance*, 40(4), 1316–1331.
<https://doi.org/10.1037/a0036956>
- Puts, D. A., Gaulin, S. J. C., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior*, 27(4), 283–296.
<https://doi.org/10.1016/j.evolhumbehav.2005.11.003>
- Puts, D. A., Hodges, C. R., Cárdenas, R. A., & Gaulin, S. J. C. (2007). Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior*, 28(5), 340–344.
<https://doi.org/10.1016/j.evolhumbehav.2007.05.002>
- Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia — A Window Into Perception, Thought and Language. *Journal of Consciousness Studies*, 8(12), 3–34.
- Rendall, D., Vokey, J. R., & Nemeth, C. (2007). Lifting the curtain on the Wizard of Oz: Biased voice-based impressions of speaker size. *Journal of Experimental Psychology: Human Perception and Performance*, 33(5), 1208–1219. <https://doi.org/10.1037/0096-1523.33.5.1208>
- Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, 12(3), 225–239. <https://doi.org/10.1037/h0070931>
- Sell, A., Bryant, G. A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., ... Gurven, M. (2010). Adaptations in humans for assessing physical strength from the voice. *Proceedings of the Royal Society B: Biological Sciences*, 277(1699), 3509–3518.
<https://doi.org/10.1098/rspb.2010.0769>
- Shih, S. S., Ackerman, J., Hermalin, N., Inkelas, S., & Kavitskaya, D. (2018). Pokémonikers: A study of sound symbolism and Pokémon names. *Proceedings of the Linguistic Society of America*, 3(1), 42. <https://doi.org/10.3765/plsa.v3i1.4335>

- Shinohara, K., & Kawahara, S. (2010). A Cross-linguistic Study of Sound Symbolism: The Images of Size. *Annual Meeting of the Berkeley Linguistics Society*, 36(1), 396. <https://doi.org/10.3765/bls.v36i1.3926>
- Shinohara, K., Yamauchi, N., Kawahara, S., & Tanaka, H. (2016). Takete and Maluma in Action: A Cross-Modal Relationship between Gestures and Sounds. *PLOS ONE*, 11(9), e0163525. <https://doi.org/10.1371/journal.pone.0163525>
- Stevens, K. N. (1998). *Acoustic phonetics*. Cambridge, Mass: MIT Press.
- Suire, A., Bossoms Mesa, A., Raymond, M., & Barkat-Defradas, M. (2019). Sex-biased sound symbolism in French first names. *Evolutionary Human Sciences*, accepted.
- Ulan, R. (1978). Size-sound symbolism. In *Universals of human language* (Vol. 2, pp. 525–568). Stanford University Press.
- Whalen, D. H., & Levitt, A. G. (1995). The universality of intrinsic Fo of vowels. *Journal of Phonetics*, 23, 18.
- Wichmann, S., Holman, E. W., & Brown, C. H. (2010). Sound Symbolism in Basic Vocabulary. *Entropy*, 12(4), 844–858. <https://doi.org/10.3390/e12040844>
- Wu, L., Klink, R. R., & Guo, J. (2013). Creating Gender Brand Personality with Brand Names: The Effects of Phonetic Symbolism. *Journal of Marketing Theory and Practice*, 21(3), 319–330. <https://doi.org/10.2753/MTP1069-6679210306>

IX. Conclusion générale

Au cours de cette thèse, nous avons vu que la sélection sexuelle constitue un cadre évolutif intéressant pour appréhender l'aspect fonctionnel de la voix humaine.

En premier lieu, elle permet de comprendre pourquoi les potentiels partenaires sexuels sont attentifs et attirés par certains indices vocaux plutôt que d'autres. Dans ce contexte, nous avons montré que les préférences vocales des locuteurs francophones sont différentes de celles observées dans les populations anglophones. Plusieurs de nos études suggèrent en effet que les locuteurs français sont attirés par des voix féminines plutôt graves. Ces résultats sont à considérer à la lumière de ceux observés dans les populations anglophones, où il a été plutôt bien montré que les hommes affichent des préférences pour des voix féminines relativement plus hautes. Ces résultats contradictoires soulèvent des questions intéressantes.

Premièrement, parce que les deux groupes de sujets étudiés vivent dans des environnements socio-culturels similaires. Dans les deux cas, il s'agit en effet de populations occidentales et industrialisées, plus récemment appelées « *WEIRD* » (i.e. « *Western, Educated, Industrialized, Rich and Democratic Societies* »). Si de telles différences sont observées au niveau des préférences, des analyses plus fines relatives à l'effet de l'environnement socio-culturel sur les représentations de féminité et de masculinité devraient être conduites afin de mieux comprendre l'origine des différences observées au niveau des préférences vocales de ces individus, comme précédemment suggéré par van Bezooijen (1995). Deuxièmement, parce que cela questionne l'idée selon laquelle la voix des femmes constitue un signal « *honnête* » de la fertilité (i.e. haute fertilité = voix plus aiguës).

En l'état actuel des choses, nous ne pouvons donc que spéculer sur les raisons pouvant expliquer les préférences des hommes francophones de notre échantillon pour les voix féminines plus graves. L'étude que nous avons conduite sur le dimorphisme de sonorité des prénoms féminins et masculins nous autorise à avancer une hypothèse.

Nous avons en effet montré que les prénoms féminins ont, depuis les années 1960, tendance à se masculiniser phonétiquement (i.e. présence de sons de basses fréquences). Et paradoxalement, ces prénoms – dont la popularité augmente d’années en années – sont perçus par les hommes comme étant particulièrement féminins. Ce phénomène de masculinisation de certaines caractéristiques féminines se retrouve dans l’étude de Bovet & Raymond (2015), laquelle a montré que le ratio taille-hanche (RTH) – une composante importante de la fertilité féminine – semble, au cours de la même période, avoir suivi la même tendance. En effet, si jusque dans les années 60 le corps idéal d’une femme correspondait à des valeurs de RTH très basses (i.e. valeurs de RTH plus basses = corps plus féminins), celui-ci augmente progressivement à partir de cette décennie et ce jusqu’aux années 2010 (valeurs de RTH plus élevées = corps plus masculins).

Une hypothèse intéressante pour expliquer ces changements est qu’ils sont liés aux modifications des conditions socio-économiques survenus au cours du siècle dernier, lesquelles ont peu à peu conduit à modifier les préférences des hommes pour des femmes arborant certaines caractéristiques masculines. En effet, lorsque les conditions de vie sont particulièrement stressantes et difficiles à gérer (e.g. concurrence pour des ressources rares), les hommes peuvent être amenés à préférer des femmes affichant des caractéristiques plus masculines, signifiant une meilleure capacité de celles-ci à faire face à des conditions difficiles. Par exemple, il a été montré à travers une analyse des changements séculaires entre 1960 et 2000 que les hommes préfèrent les silhouettes exhibant des RTH plus élevés pendant les périodes de difficultés socio-économiques (Pettijohn & Jungeberg, 2004). Ainsi, si dans des conditions peu stressantes les RTH relativement bas sont généralement préférés car ils indiquent de meilleurs états de santé et une plus grande fertilité, des RTH relativement élevés pourraient être préférés en conditions environnementales difficiles. Et, en effet, les femmes présentant des RTH élevés ont des profils hormonaux particuliers, caractérisés par des taux élevés

d'androgènes et de cortisol et des taux d'œstrogènes relativement bas, ce qui globalement augmente l'affirmation de soi, la masse musculaire et la force physique, traits utiles en cas de difficultés. De plus, il est intéressant de constater que ces femmes se déclarent également plus ingénieuses, plus puissantes, et comme étant plus orientées vers l'action et les carrières professionnelles (voir Cashdan, 2008 pour une discussion de ces différents aspects). De manière intéressante, elles déclarent également posséder un poste professionnel de statut supérieur. Dedy & Smith (2015) ont aussi souligné que lorsque les conditions de vie deviennent difficiles, les préférences des hommes se tournent vers des femmes de grande taille, ces dernières se décrivant elles-mêmes comme plus compétitives et moins désireuses d'enfants.

Par ailleurs, dans la première partie de notre travail nous avons également tenté de mesurer la capacité de modulation vocale en contexte de séduction. Notre expérience fait suite à celle de Pisanski et al. (2018), en montrant que les femmes ont tendance à parler avec une voix plus grave quand elles conversent avec des hommes qu'elles trouvent attractifs. Ce comportement vocal pourrait s'expliquer en partie par le fait que les voix relativement graves sont perçues comme plus dominantes et compétentes, pouvant par là-même projeter l'image d'une femme indépendante ayant confiance en elle (Anderson & Klofstad, 2012; Borkowska & Pawlowski, 2011; Klofstad et al., 2012).

Enfin, méta-analyses relatives à l'évolution de la perception qu'ont les femmes à propos d'elles-mêmes au cours du temps ont révélé d'importants changements de points de vue. Par exemple, une étude a montré que les femmes se sont progressivement perçues comme plus masculines entre les années 1970 et le milieu des années 1990 (Twenge, 1997). De même, l'affirmation de soi a augmenté de manière significative entre 1968 et 1993 (Twenge, 2001). Enfin, des résultats complémentaires ont montré que les femmes endossaient de moins en moins les stéréotypes sociaux typiquement attribués aux femmes après les années 2000 (Donnelly & Twenge, 2017).

En ce sens, la « masculinisation » générale des femmes et les changements dans les préférences des hommes occidentaux vers des profils féminins moins stéréotypés peuvent être interprétés comme une réponse adaptative aux fluctuations socio-économiques actuelles (ou peut-être à un changement culturel plus large, où les femmes dites « modernes » revêtent des traits jusque-là réservés aux hommes), qui se traduiraient par une préférence accrue pour des partenaires féminins plus compétentes dans l'acquisition des ressources.

En second lieu, la sélection sexuelle nous a aussi permis de considérer sous un angle évolutif les interprétations liées aux préférences. Dans ce contexte, le regard attentif que nous avons porté sur les raisons liées à celles-ci montrent qu'elles sont en réalité peu fondées, ou dans une moindre mesure, relativement peu consensuelles. D'une part, la voix n'est pas un signal de la taille corporelle et l'hypothèse que la voix des hommes refléterait leur niveau de testostérone circulant est à considérer avec précaution étant donné la relative fragilité des résultats rapportés. Notre propre étude menée sur les hommes hétérosexuels et homosexuels n'a, par exemple, trouvé aucun lien entre paramètres acoustiques et taux de testostérone. D'autre part, la théorie du signal « honnête » suppose que les voix relativement graves des hommes et/ou aiguës des femmes, perçues comme plus attractives par les membres du sexe opposé, devraient *in fine* augmenter les opportunités d'accouplement, étant donné les relations supposées entre caractéristiques vocales et qualité phénotypique. Pourtant, notre étude sur le lien entre qualité vocale et le nombre de partenaires sexuels a révélé que la hauteur de la voix ne pouvait directement le prédire. Dans ce contexte, nous avons proposé que la modulation vocale (i.e. ajustement volontaire de la qualité vocale selon les caractéristiques du ou des auditeurs) serait un meilleur candidat en tant qu'indice de la qualité phénotypique, étant donné les contraintes physio-anatomiques appliquées aux plis vocaux et au tractus vocal, ce qui *in fine*, contraint la plage des valeurs acoustiques atteignables par un individu. De manière

générale, la modulation vocale fonctionnerait plutôt comme un outil flexible contexte-dépendant.

Parallèlement, nous avons aussi suggéré que la voix pourrait constituer un indice du statut social. L'acquisition d'un plus haut statut entraîne généralement toute une palette de modifications physiologiques et comportementales plus ou moins permanentes (Cummins, 2015), ce qui pourrait influencer par la même les caractéristiques vocales. De manière intéressante, deux études suggèrent que la modulation vocale peut être directement liée à la recherche et à l'acquisition d'un plus haut statut social et ceci dans les deux sexes (Cheng et al., 2016; Ko et al., 2015). A ce sujet, deux questions intéressantes restent en suspens. Premièrement, il s'agit de savoir si les changements vocaux associés au statut social s'inscrivent de façon définitive ou s'ils ne sont pas temporaires et simplement induits par le changement de situation lui-même. Deuxièmement, si le statut social affecte la qualité vocale, il est nécessaire de comprendre le mécanisme qui unit ces deux composantes. Nous avançons l'idée que la testostérone pourrait jouer le rôle de médiateur dans ce processus, puisqu'elle semble être liée à la recherche ou au maintien du statut social (Eisenegger et al., 2011).

Dans notre travail, nous avons aussi montré que la hauteur vocale était directement liée au rang hiérarchique chez les babouins chacmas, où les mâles situés en haut de l'échelle de dominance produisent des vocalisations plus haut perchées que celles des mâles non dominants. Par ailleurs, notons que les rangs sont flexibles dans cette espèce. En effet, les individus peuvent garder leur position au sein de la hiérarchie entre plusieurs mois et quelques années, de même qu'après l'avoir perdu, ils peuvent le regagner (Fischer, Kitchen, Seyfarth, & Cheney, 2004). Lorsque les mâles descendent dans la hiérarchie, la durée de la syllabe « *hoo* » et la hauteur vocale semblent diminuer de façon significative, suggérant fortement que les paramètres acoustiques sont flexibles et peuvent signaler perceptuellement une progression – ou une chute – dans le rang. Il serait intéressant d'utiliser cette spécificité comportementale pour mieux

comprendre le rôle joué par la voix dans l'expression du statut social. Cette perspective de recherche future pourrait également être étendue à d'autres espèces de primates non-humains ainsi qu'à l'espèce humaine. De manière intéressante, notons enfin que cette espèce constitue une exception à la théorie « *code-fréquence* », et de plus amples investigations sont nécessaires pour le comprendre.

En dernier lieu, nous avons montré que le principe de « *code-fréquence* », inscrit au sein de la sélection sexuelle, offrent un cadre théorique original pour s'intéresser à des unités supérieures du langage – les phonèmes. Nous avons ainsi démontré l'existence d'un symbolisme phonétique dans la composition sonore des prénoms féminins et masculins faisant écho au dimorphisme sexuel physique et vocal des individus. Nous avons également montré que les perceptions de féminité et de masculinité (accrue et/ou diminuée) pouvaient être influencées par la composition sonore du prénom. Ceux-ci peuvent ainsi constituer des attributs pouvant avoir des répercussions sur de nombreuses dimensions sociales voire biologiques de ceux et/ou celles qui les arborent. Par exemple, il serait intéressant de regarder l'effet du timbre de la voyelle incluse dans la syllabe perceptivement proéminente sur la perception d'attractivité ou de dominance d'un visage. Nous pouvons aller plus loin en nous demandant si les prénoms pourraient affecter les opportunités d'accouplements et jouer un rôle dans la valeur reproductive des individus.

Bibliographie

- Abend, P., Pflüger, L. S., Koppensteiner, M., Coquerelle, M., & Grammer, K. (2015). The sound of female shape: A redundant signal of vocal and facial attractiveness. *Evolution and Human Behavior, 36*(3), 174-181.
- Abitbol, J., Abitbol, P., & Abitbol, B. (1999). Sex hormones and the female voice. *Journal of Voice, 13*(3), 424-446.
- Addington, D. W. (1968). The relationship of selected vocal characteristics to personality perception. *Speech Monographs, 35*(4), 492-503.
- Amir, O., & Biron-Shental, T. (2004). The impact of hormonal fluctuations on female vocal folds: *Current Opinion in Otolaryngology & Head and Neck Surgery, 12*(3), 180-184.
- Anderson, R. C., & Klofstad, C. A. (2012). Preference for Leaders with Masculine Voices Holds in the Case of Feminine Leadership Roles. *PLoS ONE, 7*(12), e51216.
- Andreeva, B., Demenko, G., Mobius, B., Zimmerer, F., Jugler, J., & Oleskowicz-Popiel, M. (2014). Differences of Pitch Profiles in Germanic and Slavic Languages. *Fifteenth Annual Conference of the International Speech Communication Association*.
- Apicella, C. L., Feinberg, D. R., & Marlowe, F. W. (2007). Voice pitch predicts reproductive success in male hunter-gatherers. *Biology Letters, 3*(6), 682-684.
- Apple, W., Streeter, L. A., & Krauss, R. M. (1979). Effects of Pitch and Speech Rate on Personal Attributions. *Journal of Personality and Social Psychology, 37*(5), 715.
- Araujo, A. B., Dixon, J. M., Suarez, E. A., Murad, M. H., Guey, L. T., & Wittert, G. A. (2011). Endogenous Testosterone and Mortality in Men: A Systematic Review and Meta-Analysis. *The Journal of Clinical Endocrinology & Metabolism, 96*(10), 3007-3019.
- Armstrong, M. M., Lee, A. J., & Feinberg, D. R. (2019). A house of cards: Bias in perception of body size mediates the relationship between voice pitch and perceptions of dominance. *Animal Behaviour, 147*, 43-51.
- Arnocky, S., Hodges-Simeon, C. R., Ouellette, D., & Albert, G. (2018). Do men with more masculine voices have better immunocompetence? *Evolution and Human Behavior*.
- Atkinson, J., Pipitone, R. N., Sorokowska, A., Sorokowski, P., Mberira, M., Bartels, A., & Gallup, G. G. (2012). Voice and Handgrip Strength Predict Reproductive Success in a Group of Indigenous African Females. *PLoS ONE, 7*(8), e41811.
- Auerbach, O., Hammond, E. C., & Garfinkel, L. (1970). Histologic changes in the larynx in relation to smoking habits. *Cancer, 25*(1), 92-104.
- Babel, M., McGuire, G., & King, J. (2014). Towards a More Nuanced View of Vocal Attractiveness. *PLoS ONE, 9*(2), e88616.
- Banai, I., Banai, B., & Bovan, K. (2017). Vocal characteristics of presidential candidates can predict the outcome of actual elections. *Evolution and Human Behavior, 38*(3), 309-314.
- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of personality and social psychology, 70*(3), 614.
- Barkat-Defradas, M., Busseuil, C., Chauvy, O., Hirsch, F., Fauth, C., Révis, J., & Bretèque, B. A. de la. (2012). Dimension esthétique des voix normales et dysphoniques : Approches perceptive et acoustique. *TIPA. Travaux interdisciplinaires sur la parole et le langage, (28)*.
- Bartholomew, G. A., & Collias, N. E. (1962). The role of vocalization in the social behaviour of the northern elephant seal. *Animal Behaviour, 10*(1-2), 7-14.
- Barton, D. N., & Halberstadt, J. (2018). A social Bouba/Kiki effect: A bias for people whose names match their faces. *Psychonomic Bulletin & Review, 25*(3), 1013-1020.
- Baudouin, J.-Y., & Tiberghien, G. (2004). Symmetry, averageness, and feature size in the facial attractiveness of women. *Acta Psychologica, 117*(3), 313-332.
- Bentley, R. A., Hahn, M. W., & Shennan, S. J. (2004). Random drift and culture change. *Proceedings of the Royal Society B: Biological Sciences, 271*(1547), 1443-1450.
- Berenbaum, S. A., Bryk, K. K., Nowak, N., Quigley, C. A., & Moffat, S. (2009). Fingers as a Marker of Prenatal Androgen Exposure. *Endocrinology, 150*(11), 5119-5124.
- Berger, J., Bradlow, E. T., Braunstein, A., & Zhang, Y. (2012). From Karen to Katie: Using Baby Names to Understand Cultural Evolution. *Psychological Science, 23*(10), 1067-1073.
- Berlin, B. (2006). The First Congress of Ethnozoological Nomenclature. *Journal of the Royal Anthropological Institute, 12*(s1), S23-S44.
- Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., & Christiansen, M. H. (2016). Sound-meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences, 113*(39), 10818-10823.
- Borkowska, B., & Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Animal Behaviour, 82*(1), 55-59.
- Bovet, J. (2014). *Les éléments de l'attractivité féminine et leurs variations* (Université Montpellier II - Sciences et Techniques du Languedoc). Consulté à l'adresse <https://tel.archives-ouvertes.fr/tel-01347735>
- Bovet, J., & Raymond, M. (2015). Preferred Women's Waist-to-Hip Ratio Variation over the Last 2,500 Years. *PLOS ONE, 10*(4), e0123284.

- Bovet, J. (2019). Evolutionary Theories and Men's Preferences for Women's Waist-to-Hip Ratio: Which Hypotheses Remain? A Systematic Review. *Frontiers in Psychology*, 10.
- Bowling, D. L., Garcia, M., Dunn, J. C., Ruprecht, R., Stewart, A., Frommolt, K.-H., & Fitch, W. T. (2017). Body size and vocalization in primates and carnivores. *Scientific Reports*, 7(1).
- Bradbury, J. W. (1977). Lek mating behavior in the hammer-headed bat. *Zeitschrift für Tierpsychologie*, 45(3), 225-255.
- Brand, R. J., Bonatsos, A., D'Orazio, R., & DeShong, H. (2012). What is beautiful is good, even online: Correlations between photo attractiveness and text attractiveness in men's online dating profiles. *Computers in Human Behavior*, 28(1), 166-170.
- Brougère, G. (2004). How much is a Pokémon worth? Pokémon in France. In *Pikachu's global adventure: The rise and Fall of Pokémon* (Duke University Press, p. 193). Durham.
- Brown, B. L., Strong, W. J., & Rencher, A. C. (1973). Perceptions of personality from speech: Effects of manipulations of acoustical parameters. *The Journal of the Acoustical Society of America*, 54(1), 29-35.
- Brown, W. S., Morris, R. J., & Michel, J. F. (1989). Vocal jitter in young adult and aged female voices. *Journal of Voice*, 3(2), 113-119.
- Bruckert, L., Lienard, J.-S., Lacroix, A., Kreutzer, M., & Leboucher, G. (2006). Women use voice parameters to assess men's characteristics. *Proceedings of the Royal Society B: Biological Sciences*, 273(1582), 83-89.
- Bruckert, L., Bestelmeyer, P., Latinus, M., Rouger, J., Charest, I., Rousselet, G. A., ... Belin, P. (2010). Vocal Attractiveness Increases by Averaging. *Current Biology*, 20(2), 116-120.
- Bruning, J. L., Polinko, N. K., Zerbst, J. I., & Buckingham, J. T. (2000). The Effect on Expected Job Success of the Connotative Meanings of Names and Nicknames. *The Journal of Social Psychology*, 140(2), 197-201.
- Bryant, G. A., & Haselton, M. G. (2009). Vocal cues of ovulation in human females. *Biology Letters*, 5(1), 12-15.
- Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences*, 12(1), 1-14.
- Busse, T. V., & Seraydarian, L. (1978). The relationships between first name desirability and school readiness, IQ, and school achievement. *Psychology in the Schools*, 15(2), 297-302.
- Buunk, A. P., Park, J. H., Zurriaga, R., Klavina, L., & Massar, K. (2008). Height predicts jealousy differently for men and women. *Evolution and Human Behavior*, 29(2), 133-139.
- Cartei, V., Cowles, H. W., & Reby, D. (2012). Spontaneous Voice Gender Imitation Abilities in Adult Speakers. *PLoS ONE*, 7(2), e31353.
- Cartei, V., Bond, R., & Reby, D. (2014). What makes a voice masculine: Physiological and acoustical correlates of women's ratings of men's vocal masculinity. *Hormones and Behavior*, 66(4), 569-576.
- Cartei, V., Cowles, W., Banerjee, R., & Reby, D. (2014). Control of voice gender in pre-pubertal children. *British Journal of Developmental Psychology*, 32(1), 100-106.
- Cashdan, E. (2008). Waist-to-Hip Ratio across Cultures: Trade-Offs between Androgen- and Estrogen-Dependent Traits. *Current Anthropology*, 49(6), 1099-1107.
- Cézilly, F., & Allainé, D. (2010). La sélection sexuelle. In *Biologie évolutive* (2e éd.). De Boeck.
- Chaiken, S. (1979). Communicator Physical Attractiveness and Persuasion. *Journal of Personality and Social Psychology*, 37(8), 1387.
- Charlton, B. D., Zhihe, Z., & Snyder, R. J. (2010). Giant pandas perceive and attend to formant frequency variation in male bleats. *Animal Behaviour*, 79(6), 1221-1227.
- Charlton, B. D., Frey, R., McKinnon, A. J., Fritsch, G., Fitch, W. T., & Reby, D. (2013). Koalas use a novel vocal organ to produce unusually low-pitched mating calls. *Current Biology*, 23(23), R1035-R1036.
- Charlton, B. D., & Reby, D. (2016). The evolution of acoustic size exaggeration in terrestrial mammals. *Nature Communications*, 7(1).
- Cheng, J. T., Tracy, J. L., Ho, S., & Henrich, J. (2016). Listen, follow me: Dynamic vocal signals of dominance predict emergent social rank in humans. *Journal of Experimental Psychology: General*, 145(5), 536-547.
- Collins, S.A. (2000). Men's voices and women's choices. *Animal Behaviour*, 60(6), 773-780.
- Collins, S. A., & Missing, C. (2003). Vocal and visual attractiveness are related in women. *Animal Behaviour*, 65(5), 997-1004.
- Cotton, J. L., O'Neill, B. S., & Griffin, A. (2008). The "name game": Affective and hiring reactions to first names. *Journal of Managerial Psychology*, 23(1), 18-39.
- Courtiol, A., Raymond, M., Godelle, B., & Ferdy, J.-B. (2010). Mate choice and human stature: homogamy as a unified framework for understanding mating preferences. *Evolution: International Journal of Organic Evolution*, 64(8), 2189-2203.
- Cummins, D. (2015). Dominance, Status, and Social Hierarchies. In D. M. Buss (Éd.), *The Handbook of Evolutionary Psychology* (p. 676-697). John Wiley & Sons, Ltd.
- Cuskley, C., Simner, J., & Kirby, S. (2017). Phonological and orthographic influences in the bouba-kiki effect. *Psychological Research*, 81(1), 119-130.
- Cutler, A., McQueen, J., & Robinson, K. (1990). Elizabeth and John: Sound patterns of men's and women's names. *Journal of Linguistics*, 26(02), 471.
- Dabbs, J. M. (1990). Salivary testosterone measurements: Reliability across hours, days, and weeks. *Physiology & Behavior*, 48(1), 83-86.
- Dabbs, J. M., & Mallinger, A. (1999). High testosterone levels predict low voice pitch among men. *Personality and Individual Differences*, 27(4), 801-804.
- Darwin, C. (1871). *The descent of man and selection in relation to sex*. London: John Murray.
- Darwin, C. (1890). *The expression of the emotions in man and animals*. London: John Murray.

- Dautriche, I., Mahowald, K., Gibson, E., & Piantadosi, S. T. (2017). Wordform Similarity Increases With Semantic Similarity: An Analysis of 100 Languages. *Cognitive Science*, 41(8), 2149-2169.
- de Mareüil, P. B., Rilliard, A., & Allauzen, A. (2012). A Diachronic Study of Initial Stress and other Prosodic Features in the French News Announcer Style: Corpus-based Measurements and Perceptual Experiments. *Language and Speech*, 55(2), 263-293.
- Saussure, F. (1916). *Cours de linguistique générale*. Paris: Payot.
- Deady, D. K., & Smith, M. J. L. (2015). Changing Male Preferences for Female Body Type in the US: An Adaptive Response to a Changing Socioeconomic Climate. *Journal of Behavioral and Brain Science*, 05(13), 570-577.
- DeBruine, L. M., Jones, B. C., Unger, L., Little, A. C., & Feinberg, D. R. (2007). Dissociating averageness and attractiveness: Attractive faces are not always average. *Journal of Experimental Psychology: Human Perception and Performance*, 33(6), 1420-1430.
- Decoster, W., & Debruyne, F. (1997). The ageing voice: Changes in fundamental frequency, waveform stability and spectrum. *Acta oto-rhino-laryngologica belgica*, 51(2), 105-112.
- Delvaux, V., & Metens, T. (2002). Propriétés acoustiques et articulatoires des voyelles nasales du français. *Actes des XXIVèmes Journées d'étude sur la parole*, 348-352.
- Di Cristo, A. (2000). Interpréter la prosodie. *Actes des XXIIIèmes Journées d'étude sur la Parole*, 13-29.
- Donnelly, K., & Twenge, J. M. (2017). Masculine and Feminine Traits on the Bem Sex-Role Inventory, 1993–2012: A Cross-Temporal Meta-Analysis. *Sex Roles*, 76(9-10), 556-565.
- Dunn, J. C., Halenar, L. B., Davies, T. G., Cristobal-Azkarate, J., Reby, D., Sykes, D., ... Knapp, L. A. (2015). Evolutionary Trade-Off between Vocal Tract and Testes Dimensions in Howler Monkeys. *Current Biology*, 25(21), 2839-2844.
- Eisenegger, C., Haushofer, J., & Fehr, E. (2011). The role of testosterone in social interaction. *Trends in Cognitive Sciences*, 15(6), 263-271.
- Ellison, P. T. (1988). Human salivary steroids: Methodological considerations and applications in physical anthropology. *American Journal of Physical Anthropology*, 31(S9), 115-142.
- Eriksson, A., & Wretling, P. (1997). How Flexible is the Human Voice? A case study of mimicry. *Fifth European Conference on Speech Communication and Technology*, 4.
- Erwin, P. G. (1993). First Names and Perceptions of Physical Attractiveness. *The Journal of Psychology*, 127(6), 625-631.
- Evans, S., Neave, N., Wakelin, D., & Hamilton, C. (2008). The relationship between testosterone and vocal frequencies in human males. *Physiology & Behavior*, 93(4-5), 783-788.
- Everett, C., Blasi, D. E., & Roberts, S. G. (2015). Climate, vocal folds, and tonal languages: Connecting the physiological and geographic dots. *Proceedings of the National Academy of Sciences*, 112(5), 1322-1327.
- Fant, G. (1960). *Acoustic Theory of Speech Production*. Berlin: De Gruyter.
- Feinberg, D.R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, 69(3), 561-568.
- Feinberg, D.R., Jones, B. C., DeBruine, L. M., Moore, F. R., Law Smith, M. J., Cornwell, R. E., ... Perrett, D. I. (2005). The voice and face of woman: One ornament that signals quality? *Evolution and Human Behavior*, 26(5), 398-408.
- Feinberg, D.R., Jones, B. C., Law Smith, M. J., Moore, F. R., DeBruine, L. M., Cornwell, R. E., ... Perrett, D. I. (2006). Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. *Hormones and Behavior*, 49(2), 215-222.
- Feinberg, D.R., DeBruine, L. M., Jones, B. C., & Little, A. C. (2008). Correlated preferences for men's facial and vocal masculinity. *Evolution and Human Behavior*, 29(4), 233-241.
- Feinberg, D.R., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2008). The Role of Femininity and Averageness of Voice Pitch in Aesthetic Judgments of Women's Voices. *Perception*, 37(4), 615-623.
- Feinberg, D.R., Jones, B. C., & Armstrong, M. M. (2018). Sensory Exploitation, Sexual Dimorphism, and Human Voice Pitch. *Trends in Ecology & Evolution*, 33(12), 901-903.
- Feinberg, D.R., Jones, B. C., & Armstrong, M. M. (2019). No Evidence That Men's Voice Pitch Signals Formidability. *Trends in Ecology & Evolution*, 34(3), 190-192.
- Figlio, D. (2005). Names, Expectations and the Black-White Test Score Gap. *National Bureau of Economic Research*, No w11195.
- Figlio, D. N. (2007). Boys named Sue: Disruptive children and their peers. *Education finance and policy*, 2(4), 376-394.
- Fischer, J., Kitchen, D. M., Seyfarth, R. M., & Cheney, D. L. (2004). Baboon loud calls advertise male quality: Acoustic features and their relation to rank, age, and exhaustion. *Behavioral Ecology and Sociobiology*, 56(2), 140-148.
- Fischer, J., Semple, S., Fickenscher, G., Jürgens, R., Kruse, E., Heistermann, M., & Amir, O. (2011). Do Women's Voices Provide Cues of the Likelihood of Ovulation? The Importance of Sampling Regime. *PLoS ONE*, 6(9), e24490.
- Fitch, W.T., & Hauser, M. D. (1995). Vocal production in nonhuman primates: Acoustics, physiology, and functional constraints on "honest" advertisement. *American Journal of Primatology*, 37(3), 191-219.
- Fitch, W.T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *The Journal of the Acoustical Society of America*, 102(2), 1213-1222.
- Fitch, W.T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *The Journal of the Acoustical Society of America*, 106(3), 1511-1522.
- Fitch, W.T., & Reby, D. (2001). The descended larynx is not uniquely human. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1477), 1669-1675.

- Fitch, W.T., & Hauser, M. D. (2003). Unpacking “Honesty”: Vertebrate Vocal Production and the Evolution of Acoustic Signals. In A. M. Simmons, R. R. Fay, & A. N. Popper (Éd.), *Acoustic Communication* (Vol. 16, p. 65-137).
- Folstad, I., & Karter, A. J. (1992). Parasites, bright males, and the immunocompetence handicap. *The American Naturalist*, *139*(3), 603-622.
- Foo, Y. Z., Nakagawa, S., Rhodes, G., & Simmons, L. W. (2017). The effects of sex hormones on immune function: A meta-analysis: Sex hormones and immune function. *Biological Reviews*, *92*(1), 551-571.
- Fouquet, M., Pisanski, K., Mathevon, N., & Reby, D. (2016). Seven and up: Individual differences in male voice fundamental frequency emerge before puberty and remain stable throughout adulthood. *Royal Society Open Science*, *3*(10), 160395.
- Fracarro, P. J., Jones, B. C., Vukovic, J., Smith, F. G., Watkins, C. D., Feinberg, D. R., ... DeBruine, L. M. (2011). Experimental evidence that women speak in a higher voice pitch to men they find attractive. *Journal of Evolutionary Psychology*, *9*(1), 57-67.
- Fracarro, P. J., O'Connor, J. J. M., Re, D. E., Jones, B. C., DeBruine, L. M., & Feinberg, D. R. (2013). Faking it: Deliberately altered voice pitch and vocal attractiveness. *Animal Behaviour*, *85*(1), 127-136.
- Frey, R., Volodin, I., Volodina, E., Soldatova, N. V., & Juldachev, E. T. (2011). Descended and mobile larynx, vocal tract elongation and rutting roars in male goitred gazelles (*Gazella subgutturosa* Güldenstaedt, 1780): Goitred gazelle vocal anatomy and rutting calls. *Journal of Anatomy*, *218*(5), 566-585.
- Garcia, M., Herbst, C. T., Bowling, D. L., Dunn, J. C., & Fitch, W. T. (2017). Acoustic allometry revisited: Morphological determinants of fundamental frequency in primate vocal production. *Scientific Reports*, *7*(1).
- Geary, D. C., Vigil, J., & Byrd-Craven, J. (2004). Evolution of human mate choice. *Journal of Sex Research*, *41*(1), 27-42.
- Gebauer, J. E., Leary, M. R., & Neberich, W. (2012). Unfortunate First Names: Effects of Name-Based Relational Devaluation and Interpersonal Neglect. *Social Psychological and Personality Science*, *3*(5), 590-596.
- Goy, H., Fernandes, D. N., Pichora-Fuller, M. K., & van Lieshout, P. (2013). Normative Voice Data for Younger and Older Adults. *Journal of Voice*, *27*(5), 545-555.
- Gregory, S. W., & Webster, S. (1996). A Nonverbal Signal in Voices of Interview Partners Effectively Predicts Communication Accommodation and Social Status Perceptions. *Journal of Personality and Social Psychology*, *70*(6), 10.
- Grieser, D. L., & Kulh, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental psychology*, *24*(1), 14.
- Groyecka, A., Pisanski, K., Sorokowska, A., Havlíček, J., Karwowski, M., Puts, D., ... Sorokowski, P. (2017). Attractiveness Is Multimodal: Beauty Is Also in the Nose and Ear of the Beholder. *Frontiers in Psychology*, *8*.
- Halberstadt, J., & Rhodes, G. (2000). The Attractiveness of Nonface Averages: Implications for an Evolutionary Explanation of the Attractiveness of Average Faces. *Psychological Science*, *11*(4), 285-289.
- Halliday, T. R. (1983). The study of mate choice. *Mate choice*, *1*, 462.
- Hampson, E., & Sankar, J. S. (2012). Re-examining the Manning hypothesis: Androgen receptor polymorphism and the 2D:4D digit ratio. *Evolution and Human Behavior*, *33*(5), 557-561.
- Han, C., Wang, H., Fasolt, V., Hahn, A., Holzleitner, I. J., Lao, J., ... Jones, B. (2018). No clear evidence for correlations between handgrip strength and sexually dimorphic acoustic properties of voices. *BioRxiv*.
- Hanley, T. D., Snidecor, J. C., & Ringel, R. L. (1966). Some acoustic differences among languages. *Phonetica*, *14*(2), 97-107.
- Harari, H., & McDavid, J. W. (1973). Name stereotypes and teachers' expectations. *Journal of Educational Psychology*, *65*(2), 222-225.
- Harries, M. L. L., Walker, J. M., Williams, D. M., Hawkins, S., & Hughes, I. A. (1997). Changes in the male voice at puberty. *Archives of Disease in Childhood*, *77*(5), 445-447.
- Hassebrauck, M. (1988). Beauty Is More than « Name » Deep: The Effect of Women's First Names on Ratings of Physical Attractiveness and Personality Attributes. *Journal of Applied Social Psychology*, *18*(9), 721-726.
- Heffernan, K. (2004). Evidence from HNR that /s/ is a social marker of gender. *Toronto Working Papers in Linguistics*, *14*.
- Hensley, W. E., & Spencer, B. A. (1985). The effect of first names on perceptions of female attractiveness. *Sex Roles*, *12*(7-8), 723-729.
- Hill, A. K., Hunt, J., Welling, L. L. M., Cárdenas, R. A., Rotella, M. A., Wheatley, J. R., ... Puts, D. A. (2013). Quantifying the strength and form of sexual selection on men's traits. *Evolution and Human Behavior*, *34*(5), 334-341.
- Hill, A. K., Cárdenas, R. A., Wheatley, J. R., Welling, L. L. M., Burriss, R. P., Claes, P., ... Puts, D. A. (2017). Are there vocal cues to human developmental stability? Relationships between facial fluctuating asymmetry and voice attractiveness. *Evolution and Human Behavior*, *38*(2), 249-258.
- Hillenbrand, J. (1987). A Methodological Study of Perturbation and Additive Noise in Synthetically Generated Voice Signals. *Journal of Speech, Language, and Hearing Research*, *30*(4), 448-461.
- Hillenbrand, J. (1988). Perception of aperiodicities in synthetically generated voices. *The Journal of the Acoustical Society of America*, *83*(6), 2361-2371.
- Hillenbrand, J., & Houde, R. A. (1996). Acoustic Correlates of Breathless Vocal Quality: Dysphonic Voices and Continuous Speech. *Journal of Speech, Language, and Hearing Research*, *39*(2), 311-321.
- Hodges-Simeon, C.R., Gaulin, S. J. C., & Puts, D. A. (2010). Different Vocal Parameters Predict Perceptions of Dominance and Attractiveness. *Human Nature*, *21*(4), 406-427.
- Hodges-Simeon, C.R., Gaulin, S. J. C., & Puts, D. A. (2011). Voice Correlates of Mating Success in Men: Examining

- “Contests” Versus “Mate Choice” Modes of Sexual Selection. *Archives of Sexual Behavior*, 40(3), 551-557.
- Hodges-Simeon, C. R., Gurven, M., Puts, D. A., & Gaulin, S. J. C. (2014). Vocal fundamental and formant frequencies are honest signals of threat potential in peripubertal males. *Behavioral Ecology*, 25(4), 984-988.
- Hodges-Simeon, C.R., Gurven, M., & Gaulin, S. J. C. (2015). The low male voice is a costly signal of phenotypic quality among Bolivian adolescents. *Evolution and Human Behavior*, 36(4), 294-302.
- Hollien, H., & Shipp, T. (1972). Speaking Fundamental Frequency and Chronologic Age in Males. *Journal of Speech and Hearing Research*, 15(1), 155-159.
- Holzleitner, I. J., Lee, A. J., Hahn, A., Kandrik, M., Bovet, J., Renoult, J. P., ... Jones, B. C. (2018). *Comparing theory-driven and data-driven attractiveness models using images of real women's faces*. Accessible à <https://doi.org/10.31234/osf.io/vhc5k>
- Hopkins, W. D., Tagliatalata, J. P., & Leavens, D. A. (2007). Chimpanzees differentially produce novel vocalizations to capture the attention of a human. *Animal Behaviour*, 73(2), 281-286.
- Hughes, S. M., Harrison, M. A., & Gallup, G. G. (2002). The sound of symmetry: Voice as a marker of developmental instability. *Evolution and Human Behavior*, 23(3), 173-180.
- Hughes, S. M., Dispenza, F., & Gallup, G. G. (2004). Ratings of voice attractiveness predict sexual behavior and body configuration. *Evolution and Human Behavior*, 25(5), 295-304.
- Hughes, S. M., Pastizzo, M. J., & Gallup, G. G. (2008). The Sound of Symmetry Revisited: Subjective and Objective Analyses of Voice. *Journal of Nonverbal Behavior*, 32(2), 93-108.
- Hughes, S. M., Farley, S. D., & Rhodes, B. C. (2010). Vocal and Physiological Changes in Response to the Physical Attractiveness of Conversational Partners. *Journal of Nonverbal Behavior*, 34(3), 155-167.
- Hughes, S. M., Mogilski, J. K., & Harrison, M. A. (2014). The Perception and Parameters of Intentional Voice Manipulation. *Journal of Nonverbal Behavior*, 38(1), 107-127.
- Hughes, S. M., & Miller, N. E. (2016). What sounds beautiful looks beautiful stereotype : The matching of attractiveness of voices and faces. *Journal of Social and Personal Relationships*, 33(7), 984-996.
- Jasińska, G., Ziolkiewicz, A., Ellison, P. T., Lipson, S. F., & Thune, I. (2004). Large breasts and narrow waists indicate high reproductive potential in women. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1545), 1213-1217.
- Jenkins, J. (1998). The voice of the castrato. *The Lancet*, 351(9119), 1877-1880.
- Johnstone, R. A. (1997). The evolution of animal signals. *Behavioural ecology: An evolutionary approach*, 155-178.
- Joliveau, E., Smith, J., & Wolfe, J. (2004). Tuning of vocal tract resonance by sopranos. *Nature*, 427(6970), 116-116.
- Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., & Vukovic, J. (2010). A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. *Animal Behaviour*, 79(1), 57-62.
- Jones, B. C., Hahn, A. C., Fisher, C. I., Wang, H., Kandrik, M., Han, C., ... DeBruine, L. M. (2018). No Compelling Evidence that Preferences for Facial Masculinity Track Changes in Women's Hormonal Status. *Psychological Science*, 29(6), 10.
- Jost, L., Fuchs, M., Loeffler, M., Thiery, J., Kratzsch, J., Berger, T., & Engel, C. (2018). Associations of Sex Hormones and Anthropometry with the Speaking Voice Profile in the Adult General Population. *Journal of Voice*, 32(3), 261-272.
- Jünger, J., Motta-Mena, N. V., Cardenas, R., Bailey, D., Rosenfield, K. A., Schild, C., ... Puts, D. A. (2018). Do women's preferences for masculine voices shift across the ovulatory cycle? *Hormones and Behavior*, 106, 122-134.
- Kadokia, S., Carlson, D., & Sataloff, R. T. (2013). The Effect of Hormones on the Voice. *Journal of Singing*, 69(5), 4.
- Kanazawa, S. (2004). Why beautiful people are more intelligent. *Intelligence*, 32(3), 227-243.
- Kasof, J. (1993). Sex bias in the naming of stimulus persons. *Psychological Bulletin*, 113(1), 140-163.
- Kawahara, S., Noto, A., & Kumagai, G. (2018). Sound Symbolic Patterns in Pokémon Names. *Phonetica*, 75(3), 219-244.
- Keating, P., & Kuo, G. (2012). Comparison of speaking fundamental frequency in English and Mandarin. *The Journal of the Acoustical Society of America*, 132(2), 1050-1060.
- Kempe, V., Puts, D. A., & Cárdenas, R. A. (2013). Masculine Men Articulate Less Clearly. *Human Nature*, 24(4), 461-475.
- Klemuk, S. A., Riede, T., Walsh, E. J., & Titze, I. R. (2011). Adapted to Roar : Functional Morphology of Tiger and Lion Vocal Folds. *PLoS ONE*, 6(11), e27029.
- Klenovsak, D., Hartung, F., Santiago, L., & Zaefferer, D. (2018). Preprint Are Tims hot and Toms not? *Psyarxiv*. Accessible à <https://doi.org/10.31219/osf.io/kj3b2>
- Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner : Voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B: Biological Sciences*, 279(1738), 2698-2704.
- Klofstad, C.A., Anderson, R. C., & Nowicki, S. (2015). Perceptions of Competence, Strength, and Age Influence Voters to Select Leaders with Lower-Pitched Voices. *PLOS ONE*, 10(8), e0133779.
- Ko, S. J., Sadler, M. S., & Galinsky, A. D. (2015). The Sound of Power: Conveying and Detecting Hierarchical Rank Through Voice. *Psychological Science*, 26(1), 3-14.
- Koda, H., Murai, T., Tuuga, A., Goossens, B., Nathan, S. K. S. S., Stark, D. J., ... Matsuda, I. (2018). Nasalization by *Nasalis larvatus* : Larger noses audiovisually advertise conspecifics in proboscis monkeys. *Science Advances*, 4(2), eaaq0250.
- Köhler, W. (1929). *Gestalt Psychology*. New York: Liveright.
- Köhler, W. (1947). *Gestalt Psychology (2nd. Ed.)* (2ème). New York: Liveright.

- Kordsmeyer, T. L., Hunt, J., Puts, D. A., Ostner, J., & Penke, L. (2018). The relative importance of intra- and intersexual selection on human male sexually dimorphic traits. *Evolution and Human Behavior, 39*(4), 424-436.
- Krakauer, A. H., Tyrrell, M., Lehmann, K., Losin, N., Goller, F., & Patricelli, G. L. (2009). Vocal and anatomical evidence for two-voiced sound production in the greater sage-grouse *Centrocercus urophasianus*. *Journal of Experimental Biology, 212*(22), 3719-3727.
- Kreiman, J., & Gerratt, B. R. (2005). Perception of aperiodicity in pathological voice. *The Journal of the Acoustical Society of America, 117*(4), 2201-2211.
- Krom, G. de. (1995). Some Spectral Correlates of Pathological Breathy and Rough Voice Quality for Different Types of Vowel Fragments. *Journal of Speech, Language, and Hearing Research, 38*(4), 794-811.
- Kuhl, P. K., Andruski, J. E., Christovich, I. A., Christovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., ... Lacerda, F. (1997). Cross-Language Analysis of Phonetic Units in Language Addressed to Infants. *Science, 277*(5326), 684-686.
- Lameira, A. R., Hardus, M. E., Kowalsky, B., de Vries, H., Spruijt, B. M., Sterck, E. H. M., ... Wich, S. A. (2013). Orangutan (*Pongo* spp.) whistling and implications for the emergence of an open-ended call repertoire: A replication and extension. *The Journal of the Acoustical Society of America, 134*(3), 2326-2335.
- Lameira, A. R., Hardus, M. E., Mielke, A., Wich, S. A., & Shumaker, R. W. (2016). Vocal fold control beyond the species-specific repertoire in an orang-utan. *Scientific Reports, 6*(1).
- Laver, J., & Trudgill, P. (1979). Phonetic and linguistic markers in speech. *Cambridge University Press*, 1-32.
- Law Smith, M. J., Perrett, D. I., Jones, B. C., Cornwell, R. E., Moore, F. R., Feinberg, D. R., ... Hillier, S. G. (2006). Facial appearance is a cue to oestrogen levels in women. *Proceedings of the Royal Society B: Biological Sciences, 273*(1583), 135-140.
- Leaderbrand, K., Dekam, J., Morey, A., & Tuma, L. (2008). The Effects of Voice Pitch on Perceptions of Attractiveness: Do You Sound Hot or Not? *Winona State University Psychology Student Journal, 6*.
- Leirer, V. O., Hamilton, D. L., & Carpenter, S. (1982). Common first names as cues for inferences about personality. *Personality and Social Psychology Bulletin, 8*(4), 712-718.
- Leongómez, J. D., Binter, J., Kubicová, L., Stolařová, P., Klapilová, K., Havlíček, J., & Roberts, S. C. (2014). Vocal modulation during courtship increases proceptivity even in naive listeners. *Evolution and Human Behavior, 35*(6), 489-496.
- Leongómez, J. D., Mileva, V. R., Little, A. C., & Roberts, S. C. (2017). Perceived differences in social status between speaker and listener affect the speaker's vocal characteristics. *PLOS ONE, 12*(6), e0179407.
- Levickij, V. V. (2013). Phonetic Symbolism in Natural Languages. *Glottology International Journal of Theoretical Linguistics, 4*(1), 20.
- Levero, F., Mathevon, N., Pisanski, K., Gustafsson, E., & Reby, D. (2018). The pitch of babies' cries predicts their voice pitch at age five. *Biology Letters, 14*(7), 15.
- Lieberman, D. E., McCarthy, R. C., Hiiemae, K. M., & Palmer, J. B. (2001). Ontogeny of postnatal hyoid and larynx descent in humans. *Archives of Oral Biology, 46*(2), 117-128.
- Lieberman, P., Harris, K. S., Wolff, P., & Russell, L. H. (1971). Newborn infant cry and nonhuman primate vocalization. *Journal of Speech and Hearing Research, 14*(4), 718-727.
- Marcinkowska, U. M., Galbarczyk, A., & Jasienska, G. (2018). La donna è mobile? Lack of cyclical shifts in facial symmetry, and facial and body masculinity preferences—A hormone based study. *Psychoneuroendocrinology, 88*, 47-53.
- Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: Sound-shape correspondences in toddlers and adults. *Developmental Science, 9*(3), 316-322.
- Mayew, W. J., Parsons, C. A., & Venkatachalam, M. (2013). Voice pitch and the labor market success of male chief executive officers. *Evolution and Human Behavior, 34*(4), 243-248.
- McGlone, R. E., & Hollien, H. (1963). Vocal Pitch Characteristics of Aged Women. *Journal of Speech and Hearing Research, 6*(2), 164-170.
- Mehrabian, A. (2001). Characteristics attributed to individuals on the basis of their first names. *Genetic, Social, and General Psychology Monographs, 127*(1), 59.
- Meynadier. (2013). Éléments de phonétique acoustique. Accessible à <https://doi.org/10.13140/rg.2.1.2153.8800>
- Miller, N., Maruyama, G., Beaver, R. J., & Valone, K. (1976). Speed of Speech and Persuasion. *Journal of Personality and Social Psychology, 34*(4), 615.
- Mitani, J. C., & Stuht, J. (1998). The evolution of nonhuman primate loud calls: Acoustic adaptation for long-distance transmission. *Primates, 39*(2), 171-182.
- Mitchell, K. R., Mercer, C. H., Prah, P., Clifton, S., Tanton, C., Wellings, K., & Copas, A. (2019). Why Do Men Report More Opposite-Sex Sexual Partners Than Women? Analysis of the Gender Discrepancy in a British National Probability Survey. *The Journal of Sex Research, 56*(1), 1-8.
- Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist, 111*(981), 855-869.
- Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences, 109*(41), 16474-16479.
- Newman, S.-R., Butler, J., Hammond, E. H., & Gray, S. D. (2000). Preliminary report on hormone receptors in the human vocal fold. *Journal of Voice, 14*(1), 72-81.
- Nolan, F. (2003). Intonational equivalence: An experimental evaluation of pitch scales. *Proceedings of the 15th International Congress of Phonetic Sciences, 771*, 4.
- Nowak, J., Pawłowski, B., Borkowska, B., Augustyniak, D., & Drulis-Kawa, Z. (2018). No evidence for the immunocompetence handicap hypothesis in male humans. *Scientific Reports, 8*(1).
- O'Connor, J. J. M., Fraccaro, P. J., Pisanski, K., Tigue, C. C., & Feinberg, D. R. (2013). Men's Preferences for Women's Femininity in Dynamic Cross-Modal Stimuli. *PLoS ONE, 8*(7), e69531.

- Oguchi, T., & Kikuchi, H. (1997). Voice and interpersonal attraction. *Japanese Psychological Research*, 39(1), 56-61.
- Ohala, J.J. (1996). Ethological theory and the expression of emotion in the voice. *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP '96*, 3, 1812-1815.
- Ohala, J.J. (1984). An Ethological Perspective on Common Cross-Language Utilization of F₀ of Voice. *Phonetica*, 41(1), 1-16.
- Ordin, M., & Mennen, I. (2017). Cross-Linguistic Differences in Bilinguals' Fundamental Frequency Ranges. *Journal of Speech, Language, and Hearing Research*, 60(6), 1493-1506.
- Owren, M. J., Amoss, R. T., & Rendall, D. (2011). Two organizing principles of vocal production: Implications for nonhuman and human primates. *American Journal of Primatology*, 73(6), 530-544.
- Pelham, B. W., Mirenberg, M. C., & Jones, J. T. (2002). Why Susie sells seashells by the seashore: Implicit egotism and major life decisions. *Journal of Personality and Social Psychology*, 82(4), 469-487.
- Pemberton, C., McCormack, P., & Russell, A. (1998). Have women's voices lowered across time? A cross sectional study of Australian women's voices. *Journal of Voice*, 12(2), 208-213.
- Penke, L., & Asendorpf, J. B. (2008). Beyond global sociosexual orientations: A more differentiated look at sociosexuality and its effects on courtship and romantic relationships. *Journal of Personality and Social Psychology*, 95(5), 1113-1135.
- Pépiot, E. (2013). *Voix de femmes, voix d'hommes: Différences acoustiques, identification du genre par la voix et implications psycholinguistiques chez les locuteurs anglophones et francophones*. Accesible à <https://tel.archives-ouvertes.fr/tel-00821462>
- Perfors, A. (2004). The effect of sound symbolism on perception of facial attractiveness. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 26(26), 2.
- Pettijohn, T. F., & Jungeberg, B. J. (2004). Playboy Playmate Curves: Changes in Facial and Body Feature Preferences Across Social and Economic Conditions. *Personality and Social Psychology Bulletin*, 30(9), 1186-1197.
- Pilcher, J. (2016). Names, Bodies and Identities. *Sociology*, 50(4), 764-779.
- Pipitone, R. N., & Gallup, G. G. (2008). Women's voice attractiveness varies across the menstrual cycle. *Evolution and Human Behavior*, 29(4), 268-274.
- Pisanski, K., & Rendall, D. (2011). The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. *The Journal of the Acoustical Society of America*, 129(4), 2201-2212.
- Pisanski, K., & Feinberg, D. R. (2013). Cross-Cultural Variation in Mate Preferences for Averageness, Symmetry, Body Size, and Masculinity. *Cross-Cultural Research*, 47(2), 162-197.
- Pisanski, K., Fraccaro, P. J., Tigue, C. C., O'Connor, J. J. M., Röder, S., Andrews, P. W., ... Feinberg, D. R. (2014). Vocal indicators of body size in men and women: A meta-analysis. *Animal Behaviour*, 95, 89-99.
- Pisanski, K., Hahn, A. C., Fisher, C. I., DeBruine, L. M., Feinberg, D. R., & Jones, B. C. (2014). Changes in salivary estradiol predict changes in women's preferences for vocal masculinity. *Hormones and Behavior*, 66(3), 493-497.
- Pisanski, K., Jones, B. C., Fink, B., O'Connor, J. J. M., DeBruine, L. M., Röder, S., & Feinberg, D. R. (2016). Voice parameters predict sex-specific body morphology in men and women. *Animal Behaviour*, 112, 13-22.
- Pisanski, K., Oleszkiewicz, A., & Sorokowska, A. (2016). Can blind persons accurately assess body size from the voice? *Biology Letters*, 12(4), 20160063.
- Pisanski, K., Mora, E. C., Pisanski, A., Reby, D., Sorokowski, P., Frackowiak, T., & Feinberg, D. R. (2016). Volitional exaggeration of body size through fundamental and formant frequency modulation in humans. *Scientific Reports*, 6(1).
- Pisanski, K., Cartei, V., McGettigan, C., Raine, J., & Reby, D. (2016). Voice Modulation: A Window into the Origins of Human Vocal Control? *Trends in Cognitive Sciences*, 20(4), 304-318.
- Pisanski, K., Bhardwaj, K., & Reby, D. (2018). Women's voice pitch lowers after pregnancy. *Evolution and Human Behavior*, 39(4), 457-463.
- Pisanski, K., Oleszkiewicz, A., Plachetka, J., Gmiterek, M., & Reby, D. (2018). Voice pitch modulation in human mate choice. *Proceedings of the Royal Society B: Biological Sciences*, 285(1893), 20181634.
- Pitcher, B. J., Mesoudi, A., & McElligott, A. G. (2013). Sex-Biased Sound Symbolism in English-Language First Names. *PLoS ONE*, 8(6), e64825.
- Pitcher, B. J., Briefer, E. F., & McElligott, A. G. (2015). Intrasexual selection drives sensitivity to pitch, formants and duration in the competitive calls of fallow bucks. *BMC Evolutionary Biology*, 15(1).
- Puts, D.A. (2005). Mating context and menstrual phase affect women's preferences for male voice pitch. *Evolution and Human Behavior*, 26(5), 388-397.
- Puts, D.A., Gaulin, S. J. C., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior*, 27(4), 283-296.
- Puts, D.A., Hodges, C. R., Cárdenas, R. A., & Gaulin, S. J. C. (2007). Men's voices as dominance signals: Vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior*, 28(5), 340-344.
- Puts, D.A. (2010). Beauty and the beast: Mechanisms of sexual selection in humans. *Evolution and Human Behavior*, 31(3), 157-175.
- Puts, D.A., Barndt, J. L., Welling, L. L. M., Dawood, K., & Burriss, R. P. (2011). Intrasexual competition among women: Vocal femininity affects perceptions of attractiveness and flirtatiousness. *Personality and Individual Differences*, 50(1), 111-115.
- Puts, D.A., Apicella, C. L., & Cardenas, R. A. (2012). Masculine voices signal men's threat potential in forager and

- industrial societies. *Proceedings of the Royal Society B: Biological Sciences*, 279(1728), 601-609.
- Puts, D.A., Bailey, D. H., Cárdenas, R. A., Burriss, R. P., Welling, L. L. M., Wheatley, J. R., & Dawood, K. (2013). Women's attractiveness changes with estradiol and progesterone across the ovulatory cycle. *Hormones and Behavior*, 63(1), 13-19.
- Puts, D.A., Doll, L. M., & Hill, A. K. (2014). Sexual Selection on Human Voices. In V. A. Weekes-Shackelford & T. K. Shackelford (Éd.), *Evolutionary Perspectives on Human Sexual Psychology and Behavior* (p. 69-86).
- Puts, D.A., Bailey, D. H., & Reno, P. L. (2015). Contest Competition in Men. In *The Handbook of Evolutionary Psychology* (p. 385-402). John Wiley & Sons, Ltd.
- Puts, D.A., Hill, A. K., Bailey, D. H., Walker, R. S., Rendall, D., Wheatley, J. R., ... Ramos-Fernandez, G. (2016). Sexual selection on male vocal fundamental frequency in humans and other anthropoids. *Proceedings of the Royal Society B: Biological Sciences*, 283(1829), 20152830.
- Puts, D.A., & Aung, T. (2019). Does Men's Voice Pitch Signal Formidability? A Reply to Feinberg et al. *Trends in Ecology & Evolution*, 34(3), 189-190.
- Quené, H. (2007). On the just noticeable difference for tempo in speech. *Journal of Phonetics*, 35(3), 353-362.
- Rabinov, C. R., Kreiman, J., Gerratt, B. R., & Bielamowicz, S. (1995). Comparing Reliability of Perceptual Ratings of Roughness and Acoustic Measures of Jitter. *Journal of Speech Language and Hearing Research*, 38(1), 26.
- Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia—A Window Into Perception, Thought and Language. *Journal of Consciousness Studies*, 8(12), 3-34.
- Re, D. E., O'Connor, J. J. M., Bennett, P. J., & Feinberg, D. R. (2012). Preferences for Very Low and Very High Voice Pitch in Humans. *PLoS ONE*, 7(3), e32719.
- Reby, D., & McComb, K. (2003). Anatomical constraints generate honesty : Acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour*, 65(3), 519-530.
- Rendall, D., Vokey, J. R., & Nemeth, C. (2007). Lifting the curtain on the Wizard of Oz : Biased voice-based impressions of speaker size. *Journal of Experimental Psychology: Human Perception and Performance*, 33(5), 1208-1219.
- Renoult, J. P., Bovet, J., & Raymond, M. (2016). Beauty is in the efficient coding of the beholder. *Royal Society Open Science*, 3(3), 160027.
- Renoult, J. P., & Mendelson, T. C. (2019). Processing bias : Extending sensory drive to include efficacy and efficiency in information processing. *Proceedings of the Royal Society B: Biological Sciences*, 286(1900), 20190165.
- Révis, J. (2017). *La voix et soi : Ce que notre voix dit de nous*. Solal.
- Rhodes, G., Yoshikawa, S., Clark, A., Lee, K., McKay, R., & Akamatsu, S. (2001). Attractiveness of Facial Averageness and Symmetry in Non-Western Cultures : In Search of Biologically Based Standards of Beauty. *Perception*, 30(5), 611-625.
- Richmond, E. J., & Rogol, A. D. (2007). Male pubertal development and the role of androgen therapy. *Nature Clinical Practice Endocrinology & Metabolism*, 3(4), 338-344.
- Rosch, E. (1999). Principles of Categorization. In *Concepts : Core Readings* (p. 22). Cambridge, Ma: Mit Press.
- Rose, P. (1991). How effective are long term mean and standard deviation as normalisation parameters for tonal fundamental frequency? *Speech Communication*, 10(3), 229-247.
- Rosenfield, K. A., Sorokowska, A., Sorokowski, P., & Puts, D. A. (2019). Sexual selection for low male voice pitch among Amazonian forager-horticulturists. *Evolution and Human Behavior*.
- Ryan, M. J. (1990). Sexual selection, sensory systems and sensory exploitation. *Oxford surveys in evolutionary biology*, 7, 157-195.
- Sataloff, R. T., Caputo Rosen, D., Hawkshaw, M., & Spiegel, J. R. (1997). The aging adult voice. *Journal of Voice*, 11(2), 156-160.
- Saxton, T. K., Mackey, L. L., McCarty, K., & Neave, N. (2016). A lover or a fighter? Opposing sexual selection pressures on men's vocal pitch and facial hair. *Behavioral Ecology*, 27(2), 512-519.
- Scherer, K. R. (1978). Personality inference from voice quality : The loud voice of extroversion. *European Journal of Social Psychology*, 8(4), 467-487.
- Schneider, B., Cohen, E., Stani, J., Kolbus, A., Rudas, M., Horvat, R., & van Trotsenburg, M. (2007). Towards the Expression of Sex Hormone Receptors in the Human Vocal Fold. *Journal of Voice*, 21(4), 502-507.
- Šebesta, P., Kleisner, K., Tureček, P., Kočnar, T., Akoko, R. M., Třebický, V., & Havlíček, J. (2017). Voices of Africa : Acoustic predictors of human male vocal attractiveness. *Animal Behaviour*, 127, 205-211.
- Sell, A., Bryant, G. A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., ... Gurven, M. (2010). Adaptations in humans for assessing physical strength from the voice. *Proceedings of the Royal Society B: Biological Sciences*, 277(1699), 3509-3518.
- Shih, S. S., Ackerman, J., Hermalin, N., Inkelas, S., & Kavitskaya, D. (2018). Pokémonikers : A study of sound symbolism and Pokémon names. *Proceedings of the Linguistic Society of America*, 3(1), 42.
- Shinohara, K., & Kawahara, S. (2013). The sound symbolic nature of Japanese maid names. *Proceedings of the 13th Annual Meeting of the Japanese Cognitive Linguistics Association*, 13(183-193), 11.
- Shirazi, T. N., Puts, D. A., & Escasa-Dorne, M. J. (2018). Filipino Women's Preferences for Male Voice Pitch : Intra-Individual, Life History, and Hormonal Predictors. *Adaptive Human Behavior and Physiology*, 4(2), 188-206.
- Sidhu, D. M., & Pexman, P. M. (2015). What's in a Name? Sound Symbolism and Gender in First Names. *PLoS ONE*, 10(5), e0126809.
- Sidhu, D. M., Pexman, P. M., & Saint-Aubin, J. (2016). From the Bob/Kirk effect to the Benoit/Éric effect : Testing the mechanism of name sound symbolism in two languages. *Acta Psychologica*, 169, 88-99.
- Sidhu, D. M., & Pexman, P. M. (2018). Five mechanisms of sound symbolic association. *Psychonomic Bulletin & Review*, 25(5), 1619-1643.

- Sidhu, D. M., & Pexman, P. M. (2019). The Sound Symbolism of Names. *Current Directions in Psychological Science*, 096372141985013.
- Siegmán, A. W., & Boyle, S. (1993). Voices of Fear and Anxiety and Sadness and Depression : The Effects of Speech Rate and Loudness on Fear and Anxiety and Sadness and Depression. *Journal of Abnormal Psychology*, 102(3), 430.
- Simpson, A. P. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass*, 3(2), 621-640.
- Skrinda, I., Krama, T., Kecko, S., Moore, F. R., Kaasik, A., Meija, L., ... Krams, I. (2014). Body height, immunity, facial and vocal attractiveness in young men. *Naturwissenschaften*, 101(12), 1017-1025.
- Slepian, M. L., & Galinsky, A. D. (2016). The voiced pronunciation of initial phonemes predicts the gender of names. *Journal of Personality and Social Psychology*, 110(4), 509-527.
- Smith, B. L., Brown, B. L., Strong, W. J., & Rencher, A. C. (1975). Effects of Speech Rate on Personality Perception. *Language and Speech*, 18(2), 145-152.
- Smith, D. R. R., Patterson, R. D., Turner, R., Kawahara, H., & Irino, T. (2008). *The processing and perception of size information in speech sounds*. *The Journal of the Acoustical Society of America*, 117(1), 305-318.
- Smith, K. M., Olkhov, Y. M., Puts, D. A., & Apicella, C. L. (2017). Hadza Men With Lower Voice Pitch Have a Better Hunting Reputation. *Evolutionary Psychology*, 15(4), 147470491774046.
- Sobin, C., & Alpert, M. (1999). Emotion in Speech : The Acoustic Attributes of Fear, Anger, Sadness, and Joy. *Journal of Psycholinguistic Research*, 28(4), 347-365.
- Sorensen, D., & Horii, Y. (1982). Cigarette smoking and voice fundamental frequency. *Journal of Communication Disorders*, 15(2), 135-144.
- Sorokowski, P., Puts, D., Johnson, J., Żółkiewicz, O., Oleszkiewicz, A., Sorokowska, A., ... Pisanski, K. (2019). Voice of Authority : Professionals Lower Their Vocal Frequencies When Giving Expert Advice. *Journal of Nonverbal Behavior*, 43(2), 257-269.
- Starnberger, I., Preininger, D., & Hödl, W. (2014). The anuran vocal sac : A tool for multimodal signalling. *Animal Behaviour*, 97, 281-288.
- Stearns, S. C. (1992). *The evolution of life histories* (Repr). Oxford: Oxford Univ. Press.
- Steinpreis, R. E., Anders, K. A., & Ritzke, D. (1999). The Impact of Gender on the Review of the Curricula Vitae of Job Applicants and Tenure Candidates : A National Empirical Study. *Sex Roles*, 41(7-8), 509-528.
- Street, R. L., Brady, R. M., & Lee, R. (1984). Evaluative responses to communicators : The effects of speech rate, sex, and interaction context. *Western Journal of Speech Communication*, 48(1), 14-27.
- Street, R. L., Brady, R. M., & Putman, W. B. (1983). The Influence of Speech Rate Stereotypes and Rate Similarity or Listeners' Evaluations of Speakers. *Journal of Language and Social Psychology*, 2(1), 37-56.
- Sugiyama, L. S. (2015). Physical Attractiveness in Adaptationist Perspective. In D. M. Buss (Éd.), *The Handbook of Evolutionary Psychology* (p. 292-343). John Wiley & Sons, Ltd.
- Svantesson, J.-O. (2017). Sound symbolism : The role of word sound in meaning: Sound symbolism. *Wiley Interdisciplinary Reviews: Cognitive Science*, 8(5), e1441.
- Tartter, V. C. (1980). Happy talk : Perceptual and acoustic effects of smiling on speech. *Perception & Psychophysics*, 27(1),
- Taylor, A. M., & Reby, D. (2010). The contribution of source-filter theory to mammal vocal communication research : Advances in vocal communication research. *Journal of Zoology*, 280(3), 221-236.
- Teixeira, J. P., Oliveira, C., & Lopes, C. (2013). Vocal Acoustic Analysis – Jitter, Shimmer and HNR Parameters. *Procedia Technology*, 9, 1112-1122.
- Thornhill, R., & Gangestad, S. W. (2008). *The evolutionary biology of human female sexuality*. Oxford ; New York: Oxford University Press.
- Tigue, C. C., Borak, D. J., O'Connor, J. J. M., Schandl, C., & Feinberg, D. R. (2012). Voice pitch influences voting behavior. *Evolution and Human Behavior*, 33(3), 210-216.
- Titze, I. R. (1994). *Principles of voice production*. Englewood Cliffs, N.J: Prentice Hall.
- Titze, I.R., Riede, T., & Mau, T. (2016). Predicting Achievable Fundamental Frequency Ranges in Vocalization Across Species. *PLOS Computational Biology*, 12(6), e1004907.
- Torre, P., & Barlow, J. A. (2009). Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders*, 42(5), 324-333.
- Trask, R. L. (2012). *A Dictionary of Phonetics and Phonology*. Accessible à <http://public.eblib.com/choice/publicfullrecord.aspx?p=199923>
- Traunmüller, H., & Eriksson, A. (1995). The frequency range of the voice fundamental in the speech of male and female adults. *Unpublished Manuscript*, 11.
- Trivers, R. (1972). Parental investment and sexual selection. In *Sexual Selection & the Descent of Man* (p. 136-179). New York: Aldine de Gruyter.
- Twenge, J. M. (1997). Changes in masculine and feminine traits over time : A meta-analysis. *Sex Roles*, 36(5-6), 305-325.
- Twenge, J. M. (2001). Changes in women's assertiveness in response to status and roles : A cross-temporal meta-analysis, 1931-1993. *Journal of Personality and Social Psychology*, 81(1), 133-145.
- Ullakonoja, R. (2007). Comparison of pitch range in Finnish (L1) and Russian (L2). *Proceedings of the 16th International Congress of Phonetic Sciences*, 6-10.
- Vaissière, J. (2015). *La phonétique*. Paris: Presses universitaires de France.
- Valentova, J. V., Varella, M. A. C., Havlíček, J., & Kleisner, K. (2017). Positive association between vocal and facial attractiveness in women but not in men : A cross-cultural study. *Behavioural Processes*, 135, 95-100.

- van Bezooijen, R. (1995). Sociocultural Aspects of Pitch Differences between Japanese and Dutch Women. *Language and Speech, 38*(3), 253-265.
- van Borsel, J., Janssens, J., & De Bodt, M. (2009). Breathiness as a Feminine Voice Characteristic : A Perceptual Approach. *Journal of Voice, 23*(3), 291-294.
- Vermeulen, A., Goemaere, S., & Kaufman, J. M. (1999). Testosterone, body composition and aging. *Journal of endocrinological investigation, 25*(5), 110-116.
- Voelter, Ch., Kleinsasser, N., Joa, P., Nowack, I., Martínez, R., Hagen, R., & Voelker, H. U. (2008). Detection of hormone receptors in the human vocal fold. *European Archives of Oto-Rhino-Laryngology, 265*(10), 1239-1244.
- Voracek, M. (2014). No effects of androgen receptor gene CAG and GGC repeat polymorphisms on digit ratio (2D:4D) : A comprehensive meta-analysis and critical evaluation of research. *Evolution and Human Behavior, 35*(5), 430-437.
- Vukovic, J., Feinberg, D. R., Jones, B. C., DeBruine, L. M., Welling, L. L. M., Little, A. C., & Smith, F. G. (2008). Self-rated attractiveness predicts individual differences in women's preferences for masculine men's voices. *Personality and Individual Differences, 45*(6), 451-456.
- Vukovic, Jovana, Feinberg, D. R., DeBruine, L., Smith, F. G., & Jones, B. C. (2010). Women's voice pitch is negatively correlated with health risk factors. *Journal of Evolutionary Psychology, 8*(3), 217-225.
- Waser, P. M., & Brown, C. H. (1986). Habitat acoustics and primate communication. *American Journal of Primatology, 10*(2), 135-154.
- Watkins, C. D., Fraccaro, P. J., Smith, F. G., Vukovic, J., Feinberg, D. R., DeBruine, L. M., & Jones, B. C. (2010). Taller men are less sensitive to cues of dominance in other men. *Behavioral Ecology, 21*(5), 943-947.
- Weiss, B., & Burkhardt, F. (2010). Voice Attributes Affecting Likability Perception. *Eleventh Annual Conference of the International Speech Communication Association, 4*.
- Wendahl, R. W. (1966). Some Parameters of Auditory Roughness. *Folia Phoniatica et Logopaedica, 18*(1), 26-32.
- Wheatley, J. R., Apicella, C. A., Burriss, R. P., Cárdenas, R. A., Bailey, D. H., Welling, L. L. M., & Puts, D. A. (2014). Women's faces and voices are cues to reproductive potential in industrial and forager societies. *Evolution and Human Behavior, 35*(4), 264-271.
- Wich, S., & Nunn, C. (2002). Do male « long-distance calls » function in mate defense? A comparative study of long-distance calls in primates. *Behavioral Ecology and Sociobiology, 52*(6), 474-484.
- Wilson, M. L., Miller, C. M., & Crouse, K. N. (2017). Humans as a model species for sexual selection research. *Proceedings of the Royal Society B: Biological Sciences, 284*(1866), 20171320.
- Xu, Y., Lee, A., Wu, W.-L., Liu, X., & Birkholz, P. (2013). Human Vocal Attractiveness as Signaled by Body Size Projection. *PLoS ONE, 8*(4), e62397.
- Yamazawa, H., & Hollien, H. (1992). Speaking fundamental frequency patterns of Japanese women. *Phonetica, 49*(2), 128-140.
- Young, R. K., Kennedy, A. H., Newhouse, A., Browne, P., & Thiessen, D. (1993). The Effects of Names on Perception of Intelligence, Popularity, and Competence. *Journal of Applied Social Psychology, 23*(21), 1770-1788.
- Zahavi, A. (1975). Mate selection—A selection for a handicap. *Journal of Theoretical Biology, 53*(1), 205-214.
- Zimmerer, F., Jügler, J., Andreeva, B., Möbius, B., & Trouvain, J. (2014). Too cautious to vary more ? A comparison of pitch variation in native and non-native productions of French and German speakers. *Proceedings to the 7th Speech Prosody Conference, 5*.
- Zuckerman, M., & Driver, R. E. (1988). What sounds beautiful is good: The vocal attractiveness stereotype. *Journal of Nonverbal Behavior, 13*(2), 67-82.
- Zwebner, Y., Sellier, A.-L., Rosenfeld, N., Goldenberg, J., & Mayo, R. (2017). We look like our names : The manifestation of name stereotypes in facial appearance. *Journal of Personality and Social Psychology, 112*(4), 527-554.

Remerciements

Nous remarquerons que la section « *Remerciements* » est traditionnellement située au début de la thèse. Dans mon cas, elle apparaît ici, en toute logique, comme le point final d'une course, d'un long chemin traversé tant par le lecteur que par moi-même.

Alors, pour la personne qui, par nécessité, ou un jour d'ennui, s'est saisi de ce manuscrit, je la remercie d'avoir parcouru de bout en bout les méandres de mon travail. Vous pouvez maintenant souffler. Et pour la personne qui m'a connu pendant cette période de ma vie, ces quelques lignes vous sont adressées.

Je m'imagine la thèse comme une nébuleuse pièce de théâtre, dont les différents actes ne sont pas ce que l'on appelle la « *première, la deuxième et la troisième année de doctorat* ». Ils correspondent en fait à quelque chose de plus subtil, de plus personnel. Si nous pensons en connaître les contours avant de s'employer à la besogne, le scénario n'est en réalité qu'improvisation. Certaines personnes entrent en scène, d'autres la quitte, mais toutes restent précieuses.

Les personnages :

A Alice, Benoit et Kévin, qui ont composé la musique.

A Driss, Rémi, Héloïse et Pierrot, qui ont crapahutés en ma compagnie.

A Yohann et Sylvain, dont les correspondances éloignées furent indispensables.

A Léa, dont la douceur et la tendresse n'ont pas d'équivalent.

A Pauline, dont le charme et l'intelligence me fait sourire.

A Alex, dont la bonne humeur est plus que contagieuse.

A Arnaud, dont la bienveillance (et la patience ?) semble inépuisable.

A Sarah, dont la gentillesse est inappréciable.

A Guangli, dont la pensée me semble étrangement familière.

Et à celles et ceux que j'ai oublié... ou qui m'ont oublié.

Les spectateurs :

A mon père, ma mère, ma sœur et mes grands-parents, dont la présence ne m'est que trop précieuse.

La régie :

A mon équipe de recherche, pour avoir accepté de supporter mon cynisme.

A Valérie, le soleil.

A Michel, la lumière.

Et à Melissa, mon ESS.

Évolution de la voix humaine : le rôle de la sélection sexuelle

Résumé : Il a été suggéré que la voix grave des hommes résulterait de l'action de la compétition intrasexuelle pour signaler aux compétiteurs la dominance, la menace et la masculinité, tandis que la voix relativement aiguë des femmes serait le produit de la compétition intersexuelle pour signaler la fertilité et la féminité. En effet, au-delà du message linguistique, la voix humaine révèle de précieuses informations biologiques et sociales sur la qualité et la condition des locuteurs telles que le sexe, l'âge, la dimension corporelle, la personnalité et possiblement le statut social. Ces indices prennent toute leur importance lorsqu'il s'agit d'évaluer des compétiteurs et d'éventuels partenaires sexuels. Au cours de cette thèse, nous avons ainsi étudié le rôle fonctionnel de la voix humaine sous l'angle de la sélection sexuelle. Premièrement, nos travaux suggèrent que les préférences vocales ne sont pas universelles et qu'elles dépendent de l'environnement culturel en question, puisque plusieurs de nos résultats dans une population de locuteurs francophones montrent que les hommes sont attirés par des voix relativement graves chez les femmes, contrairement à ce qui est majoritairement observée dans les populations anglophones. De même, la plupart des études se sont focalisées sur la hauteur et le timbre, mais nos résultats suggèrent que la qualité phénotypique peut être exprimée par d'autres éléments de la qualité vocale tels que la raucité, le souffle et divers éléments prosodiques. Deuxièmement, les interprétations évolutives jusque-là évoquées dans la littérature pour expliquer ces préférences restent insatisfaisantes. En effet, nos résultats montrent d'une part que la voix des hommes n'est pas corrélée au taux de testostérone, remettant en question l'idée d'un signal « *honnête* » de l'immunocompétence et, d'autre part, que la modulation vocale, correspondant à un pattern dynamique de la voix en contexte interactionnel, souligne l'importance d'étudier la voix dans des situations écologiquement valides. Enfin, nous avons montré via le principe du symbolisme phonétique que le dimorphisme sexuel de la voix humaine se traduit également au niveau de la composition sonore des prénoms et de leur attribution en fonction du sexe. Pour conclure, notre travail offre de nouvelles pistes de réflexion et établit la sélection sexuelle comme un paradigme de choix pour étudier la voix humaine.

Mots clés : sélection sexuelle, choix de partenaire, compétition, parole, voix, hauteur, timbre, qualité vocale, code-fréquence, symbolisme phonétique.

Evolution of the human voice : the role of sexual selection

Abstract : It has been suggested that the deep voices of men have been selected through intrasexual competition to signal dominance, threat and masculinity to competitors, whereas the high voices of women have been selected through intersexual competition to signal fertility and femininity. Indeed, beyond the linguistic message, the human voice conveys valuable biological and social information about the quality and condition of the speakers such as sex, age, body configuration, personality and possibly social status. These cues are crucial when assessing competitors and potential sexual partners. In this thesis, we studied the functional role of the human voice through the lens of sexual selection. Firstly, our work suggests that vocal preferences are not universal and depend on the culture under study, since several of our results in a population of French speakers show that men are attracted by relatively low-pitched voices in women, as opposed to what has been observed in English speaking populations. Likewise, most studies have focused on pitch and timbre, but our results suggest that phenotypic quality may be expressed through other elements of voice quality such as roughness, breathiness, and various prosodic elements. Secondly, the evolutionary interpretations hitherto evoked in the literature to explain these preferences remain unsatisfactory. Indeed, our results show that, on one hand, men's voices are not correlated with testosterone levels, suggesting that they are not a "*honest*" signal of immunocompetence and, on the other hand, that vocal modulation, which corresponds to a dynamic modification of voice quality in an interactional context, emphasizes the importance of studying voice in ecologically valid situations. Finally, we have shown through the principle of sound symbolism that the sexual dimorphism of the human voice is also reflected in the sound composition of first names and their attribution according to sex. To conclude, the present work offers several new lines of research and establishes sexual selection as a key evolutionary paradigm to study the human voice.

Keywords : sexual selection, mate choice, competition, speech, voice, pitch, timbre, vocal quality, frequency-code, sound symbolism.