



Quel est mon âge d'après ma voix ? Effets de la variété régionale et de la génération.

Nicolas Audibert¹, Cécile Fougeron¹, Fany Barbier^{1,2},
Léa Croze^{1,2}, Camille Lavoine^{1,2}, Hélène Rance^{1,2}

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

(2) Département Universitaire d'Enseignement et de Formation en Orthophonie,
Sorbonne Université UPMC, 91 Boulevard de l'hôpital, 75013 Paris, France
{nicolas.audibert, cecile.fougeron}@sorbonne-nouvelle.fr

RESUME

L'âge d'un locuteur peut souvent être estimé à partir de sa voix seule. Nous testons 112 locuteurs francophones âgés de 50 à 89 ans, répartis en 4 variétés régionales (français, belges, suisses, québécois), 2 sexes et 4 décennies. 13 auditeurs français jeunes (22-31) et 13 âgés (70-95) ont estimé l'âge du locuteur comme appartenant à l'une des 4 classes d'âge. L'âge perçu apparaît comme un indicateur assez fiable de l'âge chronologique d'un locuteur, mais avec une tendance à surestimer l'âge des [50-59] ans et sous-estimer l'âge des locuteurs de 70 ans ou plus. L'estimation dépend de l'âge de l'auditeur : les auditeurs âgés surestiment davantage l'âge des locuteurs de la génération précédente [50-69]. L'origine partagée ou non entre auditeur et locuteur affecte aussi les réponses : les auditeurs français estiment plus âgés les locuteurs d'une autre variété régionale.

ABSTRACT

Guess my age from my voice. Effect of regional variety and generation.

The age of a speaker can often be estimated from his/her voice alone. We test 112 speakers of French aged 50 to 89, distributed into 4 regional varieties (French, Belgian, Swiss, Quebecois), 2 sexes and 4 decades. 13 young (22-31) and 13 old (70-95) French listeners estimated the speaker's age as belonging to one of the 4 age classes. Perceived age appears as a fairly reliable cue to the chronological age of a speaker, but with a tendency to overestimate the age of the speakers in the [50-59] group and underestimate the age of speakers from 70 y.o. Estimations depend on listener's age: older listeners overestimate more the age of speakers of the younger generation ([50-69]). The sharing of origin between speaker and listener also affects answers: French listeners rate the speakers of another regional variety as older.

MOTS-CLES : Age, vieillissement, perception, variation régionale, effet de l'auditeur.

KEYWORDS: Aging, perception, dialectal variation, listener effect.

1 Introduction

Dans les études concernant les effets de l'âge sur la parole de locuteurs adultes, une question récurrente porte sur la définition de la notion même de vieillissement. A partir de quel âge est-on vieux ? Sur quels critères définir les groupes d'âge comparés dans une étude transversale sur le

vieillessement ? Doit-on se référer à l'âge chronologique du locuteur, à l'image renvoyée/perçue, à ses performances cognitives, à des mesures/marqueurs physiologiques, à sa condition physique, etc. ? Concernant ce dernier aspect par exemple, Ramig (1983) a comparé des locuteurs de 3 groupes d'âge chronologique (25-35), (45-55), (65-75), chacun formés de 8 locuteurs en bonne condition et 8 locuteurs en mauvaise condition physique. Que les pauses soient comptabilisées ou non, le débit de parole des locuteurs âgés (groupe 65-75) est plus lent que celui des locuteurs jeunes, quelle que soit la condition physique. Toutefois, ces différences sont plus marquées dans un sens comme dans l'autre chez les locuteurs en mauvaise condition physique. On notera que dans cette étude, le groupe considéré comme « âgé » va de 65 à 75 ans, alors que dans d'autres, comme chez Fletcher et al (2015), les plus jeunes locuteurs vont de 65 à 69 ans et les plus âgés de 85 à 89 ans.

Dans cette littérature sur le vieillissement et ses effets sur la parole, force est de constater que les études sont très peu comparables dans leur définition des « locuteurs jeunes » par rapport aux « locuteurs âgés ». Certains se basent sur une définition « occidentale » (cf Pierce et al. 2013) de la personne âgée, avec une césure liée au changement d'activité vers 65 ans, d'autres prennent en compte des changements physiologiques (ex. changement hormonaux chez la femme à partir de 50 ans, réduction des fibres nerveuses à partir de 60), des changements neurologiques affectant le contrôle moteur ou les fonctions cognitives. Outre ces critères hétérogènes, la variabilité dans la définition des groupes d'âge est accrue par la nécessité dans de nombreuses études de constituer des groupes de taille homogène en fonction des locuteurs disponibles. Quels que soient les critères de groupement, on sait qu'une forte variabilité inter-individuelle est à attendre d'un découpage en âge chronologique et que l'hétérogénéité des groupes augmente pour les locuteurs âgés. En effet, l'âge accentue les différences entre les individus au niveau biologique (Woodruff & Birren 1975), mais aussi au niveau de la prise de médication ou de l'isolement social qui peuvent aussi affecter la parole. Afin de pallier cette hétérogénéité dans les groupes, plusieurs études préfèrent se référer à l'âge perçu du locuteur qu'à son âge chronologique. Pour autant, comme toutes mesures issues d'un traitement perceptif subjectif, mesurer un âge perçu n'est pas exempt de problèmes.

En effet, même si tout un chacun peut prétendre avoir déjà estimé l'âge d'un locuteur à partir de sa voix, au téléphone par exemple, la littérature montre que cette estimation peut être influencée par plusieurs facteurs. Plusieurs études ont montré que les auditeurs pouvaient estimer l'âge d'un individu à partir de sa voix de façon relativement précise. De bonnes corrélations entre âge perçu et âge chronologique ont été trouvées dans la littérature et dans plusieurs langues (corrélations comprises entre $r=.66$ et $r=.91$ pour l'anglais américain, l'allemand, l'italien et le japonais, voir Hunter & Ferguson, 2016 pour une revue). Pourtant, malgré ces corrélations élevées, les études montrent une tendance à la sous-estimation de l'âge perçu pour les locuteurs les plus âgés (Huntley et al., 1987, Kido & Kasuya, 2004, Hunter & Ferguson 2016). La précision de l'estimation semble également dépendre de l'âge de l'auditeur, ou plutôt du décalage de génération entre l'auditeur et le locuteur. Pour Hollien & Tolhurst (1978), les auditeurs n'arrivent à évaluer l'âge que de locuteurs qui leurs sont proches en âge. Pour d'autres, ce n'est pas une question de « familiarité avec la génération », mais une capacité d'estimation qui est moins bonne chez les auditeurs âgés par rapport au plus jeunes (Kreiman & Paçun, 1985 ; Lindville & Korabic 1986 ; Goy, Pichora-Fuller & van Lieshout, 2016). Enfin, il semblerait que le traitement des indices de l'âge dans la voix et la parole d'un individu ne relève pas uniquement d'un traitement acoustique. En effet, d'après Nagao (2006), il est plus facile d'estimer l'âge d'un locuteur de sa propre langue que dans une autre langue, tandis que Braun et Cerrato (1999) ne trouvent pas de bénéfice à partager la même langue que le locuteur.

Dans cette étude, nous chercherons à mieux comprendre les processus en jeu dans la définition d'un âge perçu, avec comme objectif à plus long terme de pouvoir analyser les caractéristiques

acoustiques de la parole de personnes âgées en fonction de leur âge chronologique vs. perçu. Pour l’heure, nous chercherons à savoir comment l’estimation de l’âge du locuteur varie en fonction de son âge chronologique et en fonction de caractéristiques partagées ou non entre le locuteur et l’auditeur qui le juge : la génération et la variété régionale du français.

2 Méthode

2.1 Locuteurs et prétraitement des stimuli

112 locuteurs francophones âgés de 50 à 89 ans ont été sélectionnés dans la base MonPaGe_HA (Fougeron et al., 2018). Les locuteurs se répartissent en quatre variétés régionales : 32 français (Île-de-France), 32 belges (Mons), 24 suisses (Genève), 24 québécois (Montréal). Au sein de chaque groupe régional, la distribution est équilibrée entre les deux sexes et quatre classes d’âge (50-59, 60-69, 70-79, 80-89), dont les caractéristiques sont données dans la Table 1. La production en lecture de la phrase entièrement voisée « Anne-Marie et moi allons à la mer » (16 phonèmes, 10 syllabes) a été choisie car elle présentait un bon compromis entre le besoin d’avoir suffisamment de matière pour un jugement perceptif et les contraintes de durée de l’expérience. Les enregistrements ont été normalisés en intensité, et recoupés pour éliminer les éventuels bruits extérieurs audibles présents au début et à la fin des enregistrements. Les parties coupées ont été remplacées par un silence de 200 ms, avec un lissage sur 5 ms pour éviter les artefacts de lecture audio.

Groupe	Sexe	Français	Belges	Suisses	Québécois
50-59 ans	F	53 (1.2)	53.5 (3)	53.7 (2.1)	52 (1.6)
	H	51 (1.2)	52.8 (1.1)	52.3 (1.2)	54 (2.9)
60-69 ans	F	64.3 (2.8)	64.5 (1.7)	64.7 (3.1)	63.7 (1.7)
	H	64 (2.4)	64.5 (1.8)	64.7 (0.9)	63.3 (2.9)
70-79 ans	F	76.8 (2.3)	73.3 (3.1)	75.3 (1.7)	71.7 (1.2)
	H	75 (2.1)	77 (1.4)	73 (3.6)	74 (2.2)
80-89 ans	F	82.5 (1.8)	85.5 (2.5)	83 (2.4)	83.3 (2.1)
	H	84 (2.7)	83.5 (1.1)	85.7 (3.3)	81 (0.8)

TABLE 1 : Moyenne (écart-type) de l’âge des 112 locuteurs dans chaque sous-groupe. Chaque case correspond à 4 locuteurs pour français et belges, 3 pour suisses et québécois.

2.2 Test de perception et auditeurs

La tâche principale du test de perception était d’identifier l’âge du locuteur comme appartenant à l’une des quatre décades possibles : 50-59 ans ; 60-69 ; 70-79 ; 80-89. Les choix de réponses proposés aux auditeurs sont donc distribués de façon équilibrée relativement à l’âge chronologique des locuteurs, ce qui correspond à l’approche majoritairement adoptée dans la littérature pour une évaluation catégorielle. Il était également demandé d’évaluer le sexe du locuteur, toutefois les résultats correspondants ne sont pas traités ici. Afin d’évaluer la consistance intra-auditeur, les productions de 16 locuteurs (8 français, 8 belges) ont été incluses deux fois dans le test, donnant un total de 128 stimuli à traiter dans la version complète du test. En raison de leur fatigabilité plus importante, une version réduite du test sans les 48 stimuli suisses et québécois a été mise en place à

destination d'auditeurs âgés, pour un total de 80 stimuli présentés. Un ordre aléatoire différent était proposé à chaque auditeur, la randomisation étant effectuée avec la contrainte de ne pas présenter deux fois consécutives les stimuli inclus en double pour l'évaluation de la consistance intra-auditeur. Un questionnaire a été soumis à l'ensemble des auditeurs avec des questions portant sur la ou les professions ayant été exercées, les langues parlées et les lieux de vie afin de contrôler le degré d'exposition à différents accents. Une évaluation de la qualité de l'audition pour chacune des deux oreilles a également été incluse, ainsi qu'une question sur le port éventuel de verres correcteurs.

La version complète du test, d'une durée totale d'environ 35 minutes en incluant la réponse au questionnaire, a été passée par *13 femmes étudiantes en 4ème année d'orthophonie âgées de 22 à 31 ans (25.1 en moyenne)*. Pour ce groupe d'auditeurs, une question spécifique portant sur le degré d'exposition aux troubles de la voix ou de la parole a été ajoutée et fera l'objet d'une analyse ultérieure des différences inter-auditeurs. La version réduite du test a été passée par *13 personnes (11 femmes, 2 hommes) âgées de 70 à 95 ans (86.6 en moyenne)* recrutées au sein d'un EPHAD de banlieue parisienne. La durée totale était d'environ 50 minutes en incluant la réponse au questionnaire et la passation du test d'évaluation des fonctions cognitives MMSE (Mini Mental State Examination). Ces auditeurs âgés ont été sélectionnés avec l'aide du médecin gériatre assurant leur suivi afin de n'inclure que des personnes considérées non-dépendantes et ne souffrant pas de déficits cognitifs avérés, sans troubles visuels ni auditifs majeurs.

3 Résultats

La performance des auditeurs pour la réalisation de la tâche de catégorisation de l'âge des locuteurs est évaluée par les mesures suivantes, globalement ou par sous-groupe de locuteurs ou d'auditeurs :

1. Taux de réponses « correctes », c'est-à-dire pour lesquelles la classe d'âge choisie par l'auditeur est celle correspondant à l'âge chronologique du locuteur ;
2. Corrélation de Spearman entre l'âge chronologique du locuteur et la classe d'âge choisie par l'auditeur, considérée comme une variable ordinale ;
3. Nombre de décennies d'écart entre la classe d'âge choisie par l'auditeur et celle correspondant à l'âge chronologique du locuteur, converti en nombre d'années pour plus de clarté.

3.1 Performances globales : une tâche complexe

Tous locuteurs et auditeurs confondus, on observe que le taux de réponses « correctes » n'est que de 38%, et la corrélation entre âge du locuteur et classe d'âge choisie de $r=.45$. Bien que ce résultat soit significativement meilleur que le hasard ($\chi^2(1)=98.08$, $p<2.10^{16}$), permettant ainsi de rejeter l'hypothèse d'une absence d'effet de l'âge sur les jugements, l'âge perçu des locuteurs ne correspond donc pas exactement à leur classe d'âge chronologique. L'analyse de l'écart entre classe d'âge choisie et classe d'âge de référence permet toutefois de nuancer ce tableau : les confusions concernent en effet surtout les décennies adjacentes, avec un écart moyen en valeur absolue de 8.3 ans. Seules 18% de l'ensemble des réponses sont données avec plus d'une classe d'âge d'écart.

La tâche de catégorisation de l'âge des locuteurs s'avère complexe, comme l'indique l'analyse de la fiabilité intra-auditeur à partir des réponses aux 16 stimuli inclus en double : globalement, ces stimuli ne sont évalués à l'identique lors des deux présentations que dans 58% des cas. Là encore, l'analyse de l'amplitude de l'écart entre les deux présentations permet de nuancer : l'écart moyen en valeur absolue n'est que de 4.8 ans, avec seulement 6% des réponses pour lesquelles on observe plus d'une classe d'âge d'écart entre les deux présentations.

3.2 Effet de l'âge des auditeurs sur l'estimation de l'âge des français et belges

Les performances des auditeurs jeunes (ci-après AJ) et âgés (ci-après AA) sont comparées sur l'ensemble des stimuli communs aux deux versions du test (soit 64 locuteurs français et belges). Les performances sont globalement différentes entre auditeurs jeunes et les auditeurs âgés. En effet le groupe AJ obtient un taux de réponses « correctes » de 41%, un écart moyen de 7.6 ans et une corrélation entre âge chronologique du locuteur et catégorie de réponse de $r=.55$, tandis que le groupe AA obtient un taux de réponses « correctes » de 32%, un écart moyen de 9.8 ans et une corrélation de $r=.29$. Ainsi pour les auditeurs jeunes, l'âge perçu des locuteurs se rapproche plus de leur âge chronologique. On peut en revanche noter que la consistance intra-auditeur moyenne est meilleure pour AA que pour AJ, tant en termes de taux de réponses « correctes » (63% vs. 53%) que d'années d'écart (4.3 vs. 5.3).

L'effet de différents facteur sur l'écart entre classe d'âge choisie et classe d'âge du locuteur est évalué au moyen d'un modèle linéaire mixte, à l'aide du package de R lme4 (Bates et al., 2015). L'utilisation d'un modèle paramétrique supposé s'appliquer à une variable continue pourrait être problématique avec cette variable qui ne peut prendre que 7 valeurs entières comprises entre -3 et +3. Toutefois plusieurs études (par ex. Norman, 2010) ont montré que des variables ordinales telles que des échelles de Likert pouvaient être analysés au moyen de modèles paramétriques sans pour autant accroître le risque d'erreur de type I ou II (respectivement rejet ou acceptation à tort de H_0).

Le groupe d'âge des auditeurs, celui des locuteurs et la variété régionale des locuteurs ont été inclus dans le modèle en tant qu'effets fixes, l'interaction entre âge des auditeurs et âge des locuteurs étant également prise en compte. Les variations liées aux différences inter-locuteurs et inter-auditeurs ont été modélisées par une structure aléatoire incluant des pentes aléatoires pour les différences entre locuteurs en fonction du groupe d'âge des auditeurs, ainsi que pour les différences entre auditeurs en fonction du sexe et du groupe d'âge des locuteurs. Bien qu'on ne puisse supposer a priori que tous les auditeurs traitent de façon similaire les différences entre variétés régionales des locuteurs, la non-convergence du modèle avec la pente aléatoire correspondante nous a conduit à ne modéliser cet aspect de la variabilité inter-auditeur que par une ordonnée à l'origine (*intercept*) aléatoire. Le modèle retenu, pour lequel l'inspection des résidus ne révèle pas de déviation majeure aux conditions de normalité et d'homoscédasticité, est donc (loc désignant le locuteur et aud l'auditeur) :

$$\text{ecartAge} \sim \text{groupeAgeAud} * \text{groupeAgeLoc} + \text{varieteRegionaleLoc} + \\ (1 + \text{groupeAgeAud} | \text{loc}) + (1 + \text{sexeLoc} | \text{aud}) + (1 + \text{groupeAgeLoc} | \text{aud})$$

Les effets principaux et l'interaction sont illustrés par la Figure 1. La significativité de l'effet des facteurs fixes a été testée en comparant le modèle sans interaction au même modèle sans le facteur considéré dans un test de vraisemblance, et l'effet de l'interaction en comparant le modèle avec vs. sans interaction. Ces comparaisons révèlent un effet significatif de l'ensemble des facteurs fixes ($\chi^2(16)=73.60$, $p=7.10^{16}$ pour le groupe d'âge des locuteurs, $\chi^2(1)=12.63$, $p=.0003$ pour la variété régionale), le groupe d'âge des auditeurs excepté ($\chi^2(1)=2.45$, $p=.117$). En revanche, l'interaction entre le groupe d'âge des auditeurs et celui des locuteurs est significative ($\chi^2(3)=14.07$, $p=0.003$).

Pour cette interaction, la significativité du contraste entre AJ et AA a été estimée pour chaque groupe d'âge de locuteurs. Tandis que cette différence est significative pour les groupes d'âge 50-59 ($t(45.15)=2.76$, $p=.008$) et 60-69 ($t(42.36)=2.17$, $p=.036$), elle ne l'est ni pour le groupe d'âge 70-79 ($t(41.28)=0.21$, $p=.083$) ni pour le groupe 80-89 ($t(38.31)=-1.34$, $p=.189$). Ces contrastes nous indiquent donc que les auditeurs âgés surestiment plus l'âge des locuteurs les plus jeunes que ne le

font les auditeurs jeunes, tandis que l'évaluation de l'âge des locuteurs les plus âgés est similaire entre groupes d'auditeurs. Quant à eux, les contrastes entre groupes d'âge de locuteurs au sein de chaque groupe d'âge d'auditeur sont tous significatifs ($p < 0.037$ pour toutes les comparaisons), à l'exception du contraste entre locuteurs de 70-79 ans vs. 80-89 ans évalués par le groupe d'auditeurs jeunes ($t(68.79) = 2.13$, $p = .155$).

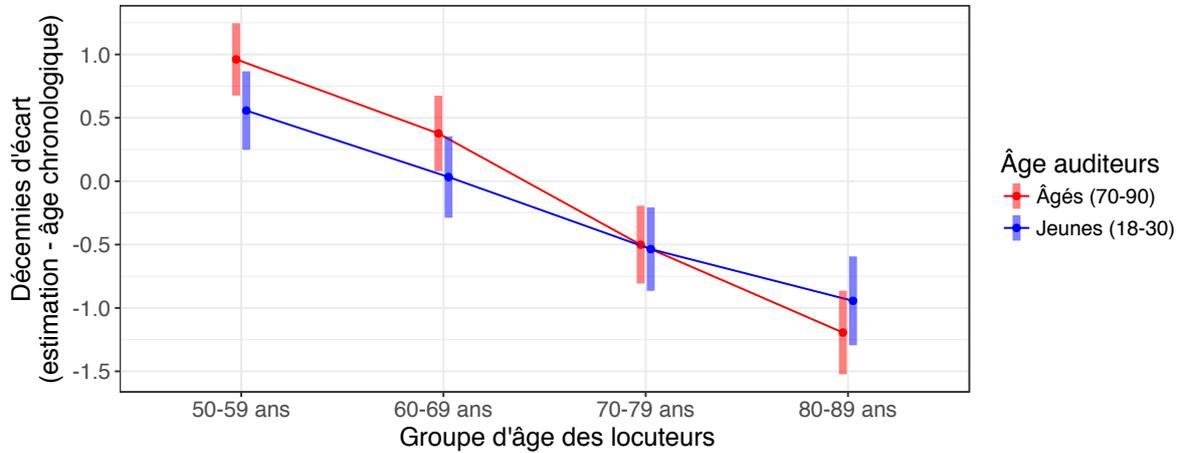


FIGURE 2: Effet du groupe d'âge des locuteurs, du groupe d'âge des auditeurs et de l'interaction, valeurs prédites par le modèle. Les barres verticales représentent les intervalles de confiance.

3.3 Effet de la variété régionale des locuteurs sur le jugement des auditeurs jeunes

Sur l'ensemble des stimuli évalués par les auditeurs jeunes, c'est pour les locuteurs suisses que les performances sont les meilleures (47% de réponses « correctes », écart moyen de 6.6 ans en valeur absolue et corrélation de $r = .63$). En revanche les niveaux de performance dans les trois autres groupes de locuteurs sont très proches les uns des autres (taux de réponses « correctes » entre 39% et 40%, écart moyen entre 7.5 et 7.8 ans), hormis pour les corrélations entre âge du locuteur et catégorie de réponse en raison de différences dans la direction des décalages : $r = .58$ pour les français, $r = .52$ pour les belges, et $r = .49$ pour les québécois. La consistance intra-auditeur, évaluée sur les français et belges, est quant à elle plus élevée pour les locuteurs français (56% de réponses « correctes », 4.9 ans d'écart moyen) que belges (51% de réponses « correctes », 5.7 ans d'écart).

De même que précédemment, l'effet de la variante régionale compte tenu des groupes d'âge des locuteurs a été analysé au moyen d'un modèle linéaire mixte (effets illustrés par la Figure 3) avec l'écart entre classe d'âge choisie et classe d'âge du locuteur comme variable dépendante : $\text{ecartAge} \sim \text{groupeAgeLoc} * \text{varieteRegionaleLoc} + (1 + \text{sexeLoc} | \text{aud}) + (1 + \text{groupeAgeLoc} | \text{aud})$

L'inspection visuelle des résidus ne révèle pas de déviation majeure aux conditions de normalité et d'homoscédasticité. Les tests de vraisemblance révèlent un effet significatif de la variété régionale ($\chi^2(3) = 62.32$, $p = 2.10^{-13}$) et du groupe d'âge des locuteurs ($\chi^2(3) = 48.24$, $p = 2.10^{-10}$), mais pas de l'interaction entre ces deux facteurs ($\chi^2(9) = 12.35$, $p = .194$). Les contrastes entre variétés régionales nous indiquent que l'âge des locuteurs français est estimé comme significativement moins élevé que celui des locuteurs d'autres variétés régionales ($p < .0001$ pour les trois comparaisons, $p > 0.55$ pour les comparaisons entre autres variétés du français). De plus, les mêmes contrastes entre variétés régionales au sein de chaque groupe d'âge révèlent que cette différence entre locuteurs français et autres variétés régionales se neutralise pour les locuteurs de 80 à 89 ans (tous $p > .209$).

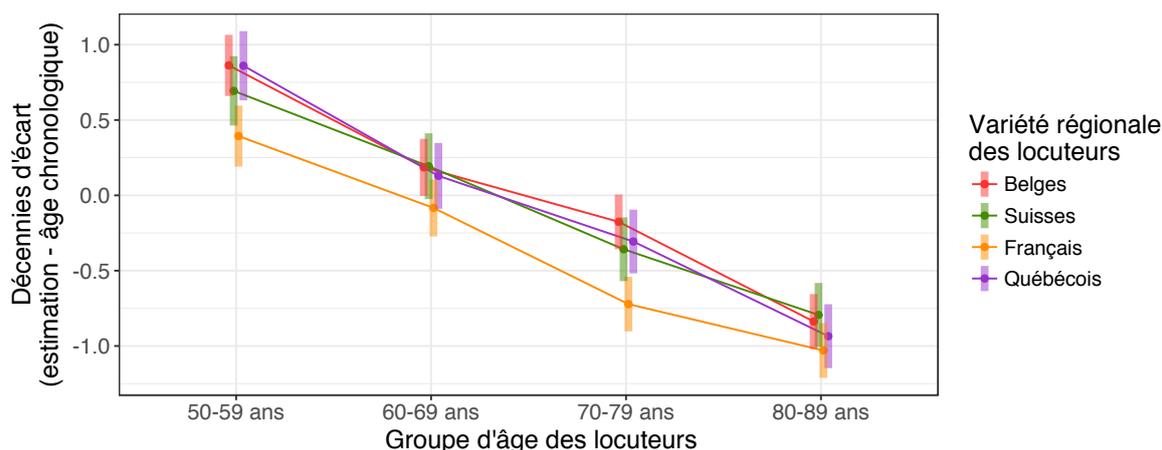


FIGURE 3: Effet du groupe d'âge, de la variété régionale des locuteurs et de l'interaction, valeurs prédites par le modèle. Les barres verticales représentent les intervalles de confiance.

4 Discussion et conclusion

Le premier résultat qui ressort de cette étude montre que l'âge perçu de nos locuteurs couvrant 4 décennies successives (plutôt que des groupes très éloignés en âge) correspond relativement bien à leur âge chronologique, à environ 8 ans près. Toutefois, on observe une tendance à surestimer l'âge du premier groupe ([50-59] ans) et à sous-estimer l'âge des locuteurs des deux derniers groupes. En d'autres termes, c'est dans le groupe [60-69] ans que les locuteurs « sonnent » le plus proche de leur âge chronologique, alors que les [50-59] paraissent plus vieux qu'ils ne le sont et les plus de 70 ans, plus jeunes qu'ils ne le sont. Ce résultat est toutefois à pondérer par le fait que dans ce type de tâche les auditeurs ont tendance à centraliser les réponses vers le milieu de l'échelle et que les faibles taux d'accord intra-juge pour les doubles jugements en test-retest montrent que la tâche demandée aux auditeurs est relativement complexe. Il serait intéressant de comparer sur les mêmes locuteurs, les résultats de cette tâche de catégorisation en quatre classes d'âge à ceux d'une tâche, que nous n'avons pas retenue car nous semblait a priori plus complexe, dans laquelle on demande à l'auditeur une estimation de l'âge exact du locuteur. Pour autant, la tendance à la surestimation de l'âge des locuteurs plus jeunes et la sous-estimation des locuteurs les plus âgés que nous observons se retrouve aussi dans la littérature. Par exemple, Hunter & Ferguson (2017) observent que les jugements de l'âge d'un même locuteur enregistré longitudinalement sur 50 ans tendent à la surestimation de son âge dans la cinquantaine et une sous-estimation de son âge après 70 ans.

Le second résultat important concerne la différence dans les jugements en fonction de l'âge des auditeurs. L'âge perçu d'un locuteur est plus proche de son âge chronologique lorsque cet âge perçu est estimé par des auditeurs jeunes par rapport à des auditeurs âgés. L'analyse des interactions montre par ailleurs que c'est particulièrement sur les locuteurs de moins de 70 ans (50-59 et 60-69) que l'effet de l'âge de l'auditeur est important. Contrairement à ce qu'on aurait pu attendre, ce n'est pas pour l'estimation de l'âge des locuteurs les plus âgés que les juges diffèrent : au-delà de 70 ans les locuteurs sont perçus comme plus jeunes –en d'autres termes leur âge est sous-estimé– par les auditeurs jeunes et âgés. Au contraire, ce sont les locuteurs de 50 à 69 ans qui voient leurs âges davantage surestimés (et sont donc perçus plus âgés que leur âge chronologique) lorsqu'ils sont jugés par des auditeurs âgés que par des auditeurs jeunes. Ces résultats trouvent écho dans la littérature où une différence de performance entre auditeurs âgés et auditeurs jeunes a été relevée (e.g. Lindville & Korabic 1986; Goy, Pichora-Fuller & van Lieshout, 2016). Cette différence ne semble pas liée à une difficulté particulière des auditeurs âgés pour cette tâche comme le suggèrent

Kreiman & Papçun (1985). En effet dans notre expérience, les auditeurs âgés semblent même globalement meilleurs en termes d'accord intra-juge que les auditeurs jeunes. Huntley, Hollien, & Shipp (1987) expliquent les différences entre groupe de juges, non pas par une différence d'âge en soi (dans leur étude les personnes âgées et les adolescents montrent des performances différentes de celles des adultes d'âge moyen), mais par une différence de familiarisation et d'expérience avec certaines voix, provoquant un bénéfice dans le jugement de ses pairs. Dans notre étude, cela signifierait que les auditeurs âgés qui ont plus de 70 ans surestimerait davantage l'âge des locuteurs n'appartenant pas à la même génération ([50-69]). Dans ce cas que dire des auditeurs jeunes (22-31 ans) qui n'appartiennent à aucune des générations représentées parmi les locuteurs à tester ? Il semblerait que les locuteurs les moins jeunes au sein de ce groupes de jeunes aient de meilleures performances que les plus jeunes, comme le montre la corrélation dans le groupe AJ entre l'âge de l'auditeur et les performances ($r=.67$ pour le taux de bonnes réponses, $r=-.68$ pour le nombre de décennies d'écart). Pour autant, l'explication des estimations globalement plus justes/proches de l'âge chronologique de la part du groupe d'auditeurs jeunes pourrait tout simplement résider dans le fait que ce groupe, constitué d'orthophonistes en fin de formation, est probablement plus « expert » en écoute de voix à des fins d'évaluation que le groupe de locuteurs âgés. La comparaison entre le bénéfice acquis par l'expérience et la familiarisation par rapport à celui acquis par l'expertise dans les capacités analytiques d'écoute fera l'objet d'une étude future comparant les évaluations d'auditeurs experts naïfs à celles d'auditeurs naïfs appariés en âge.

Enfin, nous avons pu également montrer que l'origine partagée ou non entre auditeur et locuteur affecte aussi les réponses. Dans cette étude, tous les auditeurs sont français et estiment plus âgés les locuteurs d'une autre variété régionale que la leur. Ici encore, il serait possible d'expliquer ce résultat en termes de familiarisation avec les voix à juger, à l'image des meilleures performances obtenues pour les auditeurs connaissant le locuteur évalué (Hunter & Ferguson ,2017). Toutefois, la tendance à surestimer l'âge des locuteurs présentant un éventuel accent régional est aussi à mettre en rapport avec les résultats de Stölten & Engstrand (2003) qui ont montré une corrélation entre la perception de l'âge et la perception de la force d'un accent dialectal en suédois, cet accent étant perçu comme plus fort chez les sujets âgés et inversement. Drager (2010) a également montré que la perception d'un auditeur, en l'occurrence l'identification des voyelles, dépend des caractéristiques sociales (origine, âge) qu'il attribue au locuteur. Afin d'approfondir ces hypothèses, il sera nécessaire d'examiner les réponses d'auditeurs belges, suisses et québécois sur les mêmes données.

Remerciements

Ce travail est soutenu en partie par l'ANR VoxCrim (ANR-17-CE39-0016) et le Labex EFL (ANR-10-LABX-0083).

Références

- BATES, D., MAECHLER, M., BOLKER, B., WALKER S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.
- BRAUN, A., CERRATO, L. (1999). Estimating speaker age across languages. Actes de *ICPhS 1999*, 1369-1372.
- CERRATO, L., FALCONE, M., PAOLONI, A. (2000). Subjective age estimation of telephonic voices. *Speech Communication*, 31(2-3), 107-112.

- DRAGER, K. (2011). Speaker age and vowel perception. *Language and Speech*, 54(1), 99-121.
- FLETCHER, A. R., MCAULIFFE, M. J., LANSFORD, K. L., LISS, J. M. (2015). The relationship between speech segment duration and vowel centralization in a group of older speakers. *The Journal of the Acoustical Society of America*, 138(4), 2132-2139.
- FOUGERON C., DELVAUX V., MÉNARD L., LAGANARO M. (2018) The MonPaGe_HA Database for the Documentation of Spoken French Throughout Adulthood, Actes de *LREC 2018*.
- GOY H, PICHORA-FULLER K.M., VAN LIESHOUT P. (2016) Effects of age on speech and voice quality ratings. *The Journal of the Acoustical Society of America*, 139, 1648-1659.
- KREIMAN, J., PAPÇUN, G. (1985). Voice discrimination by two listener populations. *The Journal of the Acoustical Society of America*, 77, S9.
- HOLLIEN, H., TOLHURST, G. (1978). The aging voice. In *Transcripts of the 7th symposium care of the professional voice, II: Life span changes in the human voice*, 67-73. New York: Voice Foundation.
- HUNTER, E. J., FERGUSON, S. H., NEWMAN, C. A. (2016). Listener estimations of talker age: A meta-analysis of the literature. *Logopedics Phoniatrics Vocology*, 41(3), 101-105.
- HUNTLEY, R., HOLLIEN, H., SHIPP, T. (1987). Influences of listener characteristics on perceived age estimations. *Journal of Voice*, 1(1), 49-52.
- LINVILLE, S. E., KORABIC, E. W. (1986). Elderly listeners' estimates of vocal age in adult females. *The Journal of the Acoustical Society of America*, 80(2), 692-694.
- NAGAO, K. (2006). *Cross-language study of age perception*. Doctoral dissertation, Indiana University, Bloomington.
- NORMAN, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in health sciences education*, 15(5), 625-632.
- PIERCE, J. E., COTTON, S., PERRY, A. (2013). Alternating and sequential motion rates in older adults. *International journal of language & communication disorders*, 48(3), 257-264.
- PTACEK, P. H., SANDER, E. K. (1966). Age recognition from voice. *Journal of Speech, Language, and Hearing Research*, 9(2), 273-277.
- RAMIG, L. A. (1983). Effects of physiological aging on speaking and reading rates. *Journal of communication disorders*, 16(3), 217-226.
- SHIPP, T., HOLLIEN, H. (1969). Perception of the aging male voice. *Journal of Speech, Language, and Hearing Research*, 12(4), 703-710.
- STÖLTEN, K., ENGSTRAND, O. (2003). Effects of perceived age on perceived dialect strength: A listening test using manipulations of speaking rate and F0. *Phonum*, 9, 29-32.
- WOODRUFF, D. S., BIRREN, J. E. (1975). *Aging: Scientific perspectives and social issues*. New York : D. van Nostrand.