

# Creating Healthy Character Voices: Strategies for Older Sounding Voices

Alex Owen-Hill

27/07/2024

Word Count: 2,503

*Research Question: What specific vocal techniques can vocal performers use to make their voice sound older than their natural voice?*

## **1. Introduction**

Character voices are those that involve the vocal performer changing their spoken voice away from their usual patterns of speech to create a particular perceived effect (Amin, Marziliano, and German, 2012). They can do this by varying either source properties, filter properties (Tokuda, 2021), or both.

This paper explores the existing literature on spoken character voices, focusing on those "vocal characteristics that result in speakers sounding old" (Linville, 1996, p.190). This includes identifying which properties affect the perceived sound of the voice and how changing these away from the performer's usual speech patterns can create a voice that sounds older than the performer's natural voice.

From the recommendations found in the literature, I also propose a practical framework that can be taught to performers to create and sustain older-sounding character voices that differ significantly from their natural voice. This framework will provide some of the theoretical underpinning to an upcoming practical course on character voice creation for comedy performers.

## **2. Character information included in the voice**

A person's voice can strongly affect how they are perceived by others. Even with brief utterances, such as saying 'Hello,' listeners show a high consistency in their ratings of a speaker's perceived personality (McAleer, Todorov, and Belin, 2014) .

Such perceived information (McAleer, Todorov, and Belin, 2014; Story, Titze, and Hoffman, 2001) can include the person's:

- Age
- Sex or gender
- Physique

- Health
- Socio-linguistic factors (e.g., dialect, accent, familial tendencies)
- Emotional and psychological state
- Idiosyncratic speech patterns
- Personality factors (e.g. trustworthiness, competence, confidence)

This means that, when creating a new character voice, a skilled vocal performer can impart a significant amount of information about that character simply by varying the acoustic properties of their voice (Amin, Marziliano, and German, 2012).

In this paper, I focus on voices that are perceived as old by the listener. Age is a useful property to focus on as the voice contains a lot of age-related information; listeners can estimate a speaker's age from a voice sample almost as accurately as from looking at a photograph of the speaker (Krauss, Freyberg, and Morsella, 2002).

### **3. Acoustic and physical properties influencing the perceived age of a character voice**

A potential challenge when studying character voices is that the underlying vocal mechanisms differ between individuals (Zhang, 2021). To minimize this variability, some studies investigate voice impersonators performing various distinct voices (Amin, Marziliano, and German, 2012; Amin et al., 2014; Zetterholm, 2006), allowing researchers to isolate the key acoustic properties of character voices while controlling for individual differences in vocal anatomy and mechanics.

This section examines vocal four properties often used in such research (Amin, Marziliano, and German, 2012) and their impact on the character's perceived age: fundamental frequency (F0), speech rate, formant frequencies (F1 and F2), and vocal quality.

#### **3.1. Fundamental frequency (F0)**

The fundamental frequency of a voice is the mean frequency of the vocal fold oscillation over a particular utterance and correlates with the perceived pitch (Amin et al., 2014). A voice with a high

F0 might be described perceptually as high pitched or squeaky while a low F0 voice might be described as low pitched or deep. It is primarily controlled by adjusting the tension of the vocal folds; when the vocal folds are lengthened, they become more tense and vibrate more quickly, increasing F0 (Zhang, 2016). The cricothyroid and thyroarytenoid muscles are primarily responsible for this action (Movahhedi et al., 2021).

A prevailing view is that younger voices have a higher F0 and older voices have a lower F0 (Waller and Eriksson, 2016). Studies of both professional voice actors (Amin, Marziliano, and German, 2012; Amin et al., 2014) and non-actors found people reduced the pitch of their voice when they tried to sound older regardless of the speaker's age (Waller and Eriksson, 2016).

In reality, this is a simplification. Because of developmental changes throughout a person's lifetime, vocal pitch does change with age, but the effect is not linear (Linville, 1996). It drops markedly in the transition between childhood and young adulthood; after this, the effect varies depending on the sex of the speaker (Linville, 1996).

In male voices, F0 follows a U-shaped trend where it lowers until around 40-50 years old, then gradually rises again (Linville, 1996). From its highest at young adulthood, it falls by up to 10 Hz by middle age. It then increases with age and can raise by up to 35 Hz (Torre and Barlow, 2009). In musical terms, a 45-year-old voice with an F0 of 100 Hz (G2) could increase to 135 Hz (C#3) by the age of 85.

In female voices, F0 remains fairly steady in adulthood, then decreases suddenly after the menopause (Eichhorn et al., 2018), when it can drop by 10-35 Hz (Torre and Barlow, 2009). In musical terms, a voice with an F0 of 220 Hz (A3) before this change could drop to 185 Hz (F#3).

### **3.2. Articulation and speech rate**

The articulation rate is a rough measure of the speed at which an individual articulates words (Jacewicz et al., 2009). It is often measured as the mean syllables per second or words per minute and does not include pauses (Amin, Marziliano, and German, 2012; Amin et al., 2014). The speech rate is a similar measure that includes pauses.

Regarding perceived age, various studies have shown that older adults tend to speak slower than younger adults (Bóna, 2014; Jacewicz et al., 2009). For example, speakers between 20-30 years

old were estimated to be 8 years older than their actual age by slowing their speech rate by 10% (Harnsberger et al., 2008). However, speech rate is influenced by "a whole range of social and demographic variables such as pertaining to age, gender or geographic region of origin" (Jacewicz, Fox, and Wei, 2010, p.839).

When creating character voices, it can be useful to break this rough measure into finer-grained properties. For instance, performers can reduce their speech rate by lengthening their articulation (reducing articulation rate) or introducing pauses (Bosker et al., 2013; Bóna, 2014; Jacewicz et al., 2009). There are also different types of pauses that can be added, including silent pauses, filled pauses, false starts, repetitions, and abnormally lengthened syllables (Duez, 1982).

### **3.3. Formant frequencies (F1 and F2)**

Formant frequencies are the resonant frequencies of the vocal tract that shape the perceived spoken sounds (Story, Titze, and Hoffman, 2001). The first two formants (F1 and F2) are usually enough information to determine the vowel being spoken; they are known as the "vowel formants" and the dynamic relationship between them will influence the speaker's phonetic quality (Story, Titze, and Hoffman, 2001).

Regarding perceived age, the effect of vowel formants is more subtle as they barely change over the course of a person's life (Eichhorn et al., 2018). However, varying vowel formants away from the speaker's natural speech patterns is a fundamental aspect of creating a distinct character voice (Amin, Marziliano, and German, 2012).

The exact changes a vocal performer should make to affect the perceived age of a character will depend on specific traits of that character (e.g. accent, socioeconomic class). For example, research into the New Zealand English accent found that older speakers maintain a distinction between the sounds of diphthongs /ɪə/ (as in NEAR) and /eə/ (as in SQUARE), while for younger speakers these have merged into a single sound /eə/ (Drager, 2011; Hay, Warren, and Drager, 2006). This is known as a vowel shift and has been shown to vary depending on the age of the speaker (Drager, 2011).

To identify the most suitable vowel sounds for their character's age, the performer can use an accent resource, such as the International Dialects of English Archive (Persley, 2013) and select speakers close to this age, accent, and cultural background.

### 3.4. Vocal quality measures

Vocal quality is a very important aspect of creating character voices and covers a range of approaches in the literature (Maryn et al., 2009).

The configuration of the vocal tract will affect both the vowel formant frequencies and the vocal quality (Story, Titze, and Hoffman, 2001). Various systems exist in the literature to describe the configuration of the vocal tract, including Estill (Benson, 2017), Complete Vocal Technique (CVT) (Aaen et al., 2024), and resonant voice therapies like Laver-Messac (Yiu, Lo, and Barrett, 2017). Examples of changes to the vocal tract in these systems include raising or lowering the larynx, changing lip position, lowering the velum, and altering the tongue placement (Benson, 2017).

Vocal quality measures that are most often used to determine a speaker's age are those related to the person's health and fitness (Linville, 1996). For example, two commonly researched vocal perturbation measures are jitter (the short-term variability of pitch) and shimmer (the short-term variability of loudness) (Rojas et al., 2020). Both jitter and shimmer are significantly higher in elderly speakers than they are in young to middle-aged speakers (Dehqan et al., 2013). Various factors affect this change, including diminished laryngeal neuromotor control, muscle atrophy, and a looser, thinner lamina propria with age (Rojas et al., 2020). However, the impact of age on these vocal stability measures reduces when a person improves their vocal health and control, such as through regular singing (Lortie et al., 2017).

Other vocal quality measures affected by age are creak, also known as vocal fry (a low frequency irregular phonation) (Grama, Eiswirth, and Buchstaller, 2023), breathiness (the sound of excess breath caused by incomplete glottal closure) (Vipperla, Renals, and Frankel, 2010), and roughness (voices that sound grating, raspy, or strained) (Eddins, Kopf, and Shrivastav, 2015).

From a practical perspective, performers can apply various vocal techniques to affect jitter and shimmer. Some vocal effects, their method of production, and the impact on jitter and shimmer are (Aaen et al., 2024):

- Distortion — Vibration of the ventricular folds combined with medial-lateral and anterior-posterior constriction of the ventricular folds — Jitter increases by around 5-18%. Shimmer increases by around 2 dB (Aaen et al., 2024).

- Grunt — Vibrations of the entire supraglottic structure with large amplitude vibration of vocal folds — Jitter increases by around 20-40%. Shimmer increases by around 2-2.5 dB (Aaen et al., [2024](#)).
- Growl — Vibration of the arytenoid cartilages against, or adjacent to, the epiglottis — Jitter increases by around 0.5-3%. Shimmer increases by around 1.5-3 dB (Aaen et al., [2024](#)).
- Rattle — Vibration of the arytenoid cartilages against or adjacent to one another — Jitter increases by around 2-7%. Shimmer increases by around 2.5-3% (Aaen et al., [2024](#)).
- Creak or vocal fry — A train of discrete singlet, doublet or triplet low-frequency, glottal pulses during a single vibratory cycle — Jitter increases by around 7-14%. Shimmer increases by around 1 dB (Blomgren et al., [1998](#)).

The amount of such effects that the performers should add to their speech will depend on the age, health, and other properties of the character, but the effect is often quite subtle; for example, speakers of between 80-89 years old have been found to have only 0.64% more jitter and 0.1 dB more shimmer than speakers between 60-69 years old (Rojas et al., [2020](#)). For this reason, effects like subtle rattle or creak (vocal fry) would be more suitable than, say, grunt (Aaen et al., [2024](#); Blomgren et al., [1998](#)).

It is also possible to affect the vocal quality by altering the thickness and opening of the vocal folds. For example, to achieve the breathy quality associated with some older voices, the thyroarytenoid muscles need to be barely adducted to keep the vocal folds from closing completely (Zhang, [2016](#)).

## 4. A blueprint for the older-sounding voice

From the above discussion, we can create a blueprint for vocal performers looking to create older-sounding character voices.

Here is a three-step process that performers can follow to create a character voice:

### 4.1. Step 1: Adjust fundamental frequency to suit desired age

The fundamental frequency ( $F_0$ ) is a strong cue to help listeners detect a speaker's age (Waller and Eriksson, [2016](#)) so is a good starting place when creating a voice.

First, the performer should determine the fundamental frequency of their own voice. Then modify their pitch to the new frequency.

Vocal performers can use the following information to adjust their voice to suit their character's desired age (Torre and Barlow, 2009):

- For male voices of speakers aged 20-30. Lower the pitch by up to 10 Hz to emulate a middle aged voice. From there, to emulate a voice aged 60 to 85 years, raise the pitch by up to 35 Hz.
- For male voices of speakers aged 40-50. Raise the pitch by up to 35 Hz to emulate a voice aged 60 to 85 years.
- For female voices of pre-menopausal speakers. Lower the pitch by 10-35 Hz to emulate a voice aged over 60 years.

#### **4.2. Step 2: Modify speech and articulation rate**

The speech rate is the next focus when creating a character voice. It incorporates articulation rate and pauses.

Common changes to make a voice sound older include decreasing speech rate and introducing pauses (Jacewicz, Fox, and Wei, 2010).

A general trend throughout a person's lifetime is that speech rate increases non-linearly from childhood until it reaches a peak at around 40-50-years-old, when it begins to fall again. The performer can use this information to compare their natural speaking rate to the desired age of their character.

Older adults tend to pause more, have longer pause duration, greater hesitancy, and less precise articulation. The performer can introduce such pauses to their character voice to create the impression of increased age (Linville, 1996).

#### **4.3. Step 3: Adjust vocal qualities**

Performers can modify a range of vocal quality measures to their voice to create the perception of age (Linville, 1996; Rojas et al., 2020).

Common changes that occur with age and examples of how performers can achieve them include:

- **Breathiness** — A breathy quality is often associated with older voices when laryngeal muscle atrophy causes vocal fold bowing, causing air to escape during phonation (Rojas et al., 2020). This quality can be added by using a more relaxed, under-energised vocal fold attack when phonating while keeping the vocal folds slightly apart.
- **Hoarseness and harshness** — Older voices often sound hoarse as the same muscle atrophy can lead to difficulties controlling intensity (Rojas et al., 2020). This effect can be mimicked by creating a rougher, uneven phonation by adding extreme vocal effects like creak or rattle (Aaen et al., 2024).
- **Jitter** — Aging voices also often display more jitter than younger voices (Vipperla, Renals, and Frankel, 2010). To introduce more jitter, the speaker can fluctuate the pitch of their voice, such as producing jumps in pitch by switching in and out of falsetto (Echternach et al., 2011).
- **Shimmer** — Similarly, older voices have more shimmer (Vipperla, Renals, and Frankel, 2010). Shimmer can be added by adding effects that change the intensity of the vocal signal, such as closing the false vocal folds in a controlled way or adding vocal fry.

## 5. Conclusions

Creating an older-sounding voice for a character can be a challenging and nuanced task for vocal performers (Waller and Eriksson, 2016).

Through understanding the acoustic changes that occur in the aging voice, such as alterations to fundamental frequency, speech rate, and vocal qualities, performers can more reliably create believable older-sounding voices (Amin, Marziliano, and German, 2012).

In this paper, I have outlined the main acoustic factors that the performer can affect voluntarily in their delivery, as well as the physical properties of the vocal mechanism that affect these.

I have also provided a blueprint for vocal performers to follow when creating an older-sounding voice. By following a systematic approach and prioritising vocal health, performers can effectively create unique and convincing character voices that accurately reflect the intended age of their characters.

# Bibliography

Aaen, M., McGlashan, J., Christoph, N., and Sadolin, C., 2024. Extreme vocal effects distortion, growl, grunt, rattle, and creaking as measured by electroglottography and acoustics in 32 healthy professional singers. *Journal of voice* [Online], 38 (3), 795.e21–795.e35. Available from: <https://doi.org/10.1016/j.jvoice.2021.11.010>.

Amin, T.B., Marziliano, P., and German, J.S., 2012. Nine voices, one artist: linguistic and acoustic analysis. *Proceedings - IEEE international conference on multimedia and expo* [Online], pp.450–454. Available from: <https://doi.org/10.1109/ICME.2012.142>.

Amin, T.B., Marziliano, P., German, J.S., and Member, S., 2014. Glottal and vocal tract characteristics of voice impersonators. *IEEE transactions on multimedia* [Online], 16 (3), pp.668–678. Available from: <https://doi.org/10.1109/TMM.2014.2300071>.

Benson, E.A., 2017. The estill voice model: theory and translation. *Voice and speech review* [Online], 11 (1), pp.112–114. Available from: <https://doi.org/10.1080/23268263.2017.1370810>.

Blomgren, M., Chen, Y., Ng, M.L., and Gilbert, H.R., 1998. Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. *The journal of the acoustical society of america* [Online], 103 (5), pp.2649–2658. Available from: <https://doi.org/10.1121/1.422785>.

Bosker, H.R., Pinget, A.F., Quené, H., Sanders, T., and Jong, N.H. de, 2013. What makes speech sound fluent? the contributions of pauses, speed and repairs. *Language testing* [Online], 30 (2), pp.159–175. Available from: <https://doi.org/10.1177/0265532212455394>.

Bóna, J., 2014. Temporal characteristics of speech: the effect of age and speech style. *The journal of the acoustical society of america* [Online], 136 (2), EL116–EL121. Available from: <https://doi.org/10.1121/1.4885482>.

Dehqan, A., Scherer, R.C., Dashti, G., Ansari-Moghaddam, A., and Fanaie, S., 2013. The effects of aging on acoustic parameters of voice. *Folia phoniatrica et logopaedica* [Online], 64 (6), pp.265–270. Available from: <https://doi.org/10.1159/000343998>.

Drager, K., 2011. Speaker age and vowel perception. *Language and speech* [Online], 54 (1), pp.99–121. Available from: <https://doi.org/10.1177/0023830910388017>.

Duez, D., 1982. Silent and non-silent pauses in three speech styles. *Language and speech* [Online], 25 (1), pp.11–28. Available from: <https://doi.org/10.1177/002383098202500102>.

Echternach, M., Sundberg, J., Zander, M.F., and Richter, B., 2011. Perturbation measurements in untrained male voices' transitions from modal to falsetto register. *Journal of voice* [Online], 25 (6), pp.663–669. Available from: <https://doi.org/10.1016/j.jvoice.2010.01.013>.

Eddins, D.A., Kopf, L.M., and Shrivastav, R., 2015. The psychophysics of roughness applied to dysphonic voice. *The journal of the acoustical society of america* [Online], 138 (6), pp.3820–3825. Available from: <https://doi.org/10.1121/1.4937753>.

Eichhorn, J.T., Kent, R.D., Austin, D., and Vorperian, H.K., 2018. Effects of aging on vocal fundamental frequency and vowel formants in men and women. *Journal of voice* [Online], 32 (5), 644.e1–644.e9. Available from: <https://doi.org/10.1016/j.jvoice.2017.08.003>.

Grama, J., Eiswirth, M.E., and Buchstaller, I., 2023. Tracking creak from early to late adulthood: a panel study from the north east of england. *Proceedings of the 20th international congress of phonetic science*, pp.2049–2053.

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* [Online], 22 (1), pp.58–69. Available from: <https://doi.org/10.1016/j.jvoice.2006.07.004>.

Hay, J., Warren, P., and Drager, K., 2006. Factors influencing speech perception in the context of a merger-in-progress. *Journal of phonetics* [Online], 34 (4), pp.458–484. Available from: <https://doi.org/10.1016/j.wocn.2005.10.001>.

Jacewicz, E., Fox, R.A., O'Neill, C., and Salmons, J., 2009. Articulation rate across dialect, age, and gender. *Language variation and change* [Online], 21 (2), pp.233–256. Available from: <https://doi.org/10.1017/S0954394509990093>.

Jacewicz, E., Fox, R.A., and Wei, L., 2010. Between-speaker and within-speaker variation in speech tempo of american english. *The journal of the acoustical society of america* [Online], 128 (2), pp.839–850. Available from: <https://doi.org/10.1121/1.3459842>.

Krauss, R.M., Freyberg, R., and Morsella, E., 2002. Inferring speakers' physical attributes from their voices. *Journal of experimental social psychology* [Online], 38 (6), pp.618–625. Available from: [https://doi.org/10.1016/S0022-1031\(02\)00510-3](https://doi.org/10.1016/S0022-1031(02)00510-3).

Linville, S.E., 1996. The sound of senescence. *Journal of voice* [Online], 10 (2), pp.190–200. Available from: [https://doi.org/10.1016/S0892-1997\(96\)80046-4](https://doi.org/10.1016/S0892-1997(96)80046-4).

Lortie, C.L., Rivard, J., Thibeault, M., and Tremblay, P., 2017. The moderating effect of frequent singing on voice aging. *Journal of voice* [Online], 31 (1), 112.e1–112.e12. Available from: <https://doi.org/10.1016/j.jvoice.2016.02.015>.

Maryn, Y., Roy, N., Bodt, M.D., Cauwenberge, P.V., and Corthals, P., 2009. Acoustic measurement of overall voice quality: a meta-analysis. *The journal of the acoustical society of america* [Online], 126 (5), pp.2619–2634. Available from: <https://doi.org/10.1121/1.3224706>.

McAleer, P., Todorov, A., and Belin, P., 2014. How do you say 'hello'? personality impressions from brief novel voices. *Plos one* [Online], 9 (3). Available from: <https://doi.org/10.1371/journal.pone.0090779>.

- Movahhedi, M., Geng, B., Xue, Q., and Zheng, X., 2021. Effects of cricothyroid and thyroarytenoid interaction on voice control: muscle activity, vocal fold biomechanics, flow, and acoustics. *The journal of the acoustical society of america* [Online], 150 (1), pp.29–42. Available from: <https://doi.org/10.1121/10.0005275>.
- Persley, N.H., 2013. *An innovative idea: a review of the international dialects of english archive*. Available from: <https://doi.org/10.1017/S0266078413000229>.
- Rojas, S., Kefalianos, E., Vogel, A., Redenlab, D., and Australia, V., 2020. How does our voice change as we age? a systematic review and meta-analysis of acoustic and perceptual voice data from healthy adults over 50 years of age. *Journal of speech, language, and hearing research* • [Online], 63, pp.533–551. Available from: <https://doi.org/10.23641/asha>.
- Story, B.H., Titze, I.R., and Hoffman, E.A., 2001. The relationship of vocal tract shape to three voice qualities. *The journal of the acoustical society of america* [Online], 109 (4), pp.1651–1667. Available from: <https://doi.org/10.1121/1.1352085>.
- Tokuda, I., 2021. The source–filter theory of speech. In: Oxford University Press. Available from: <https://doi.org/10.1093/acrefore/9780199384655.013.894>.
- Torre, P. and Barlow, J.A., 2009. Age-related changes in acoustic characteristics of adult speech. *Journal of communication disorders* [Online], 42 (5), pp.324–333. Available from: <https://doi.org/10.1016/j.jcomdis.2009.03.001>.
- Vipperla, R., Renals, S., and Frankel, J., 2010. Ageing voices: the effect of changes in voice parameters on asr performance. *Eurasip journal on audio, speech, and music processing* [Online], 2010. Available from: <https://doi.org/10.1155/2010/525783>.
- Waller, S.S. and Eriksson, M., 2016. Vocal age disguise: the role of fundamental frequency and speech rate and its perceived effects. *Frontiers in psychology* [Online], 7 (NOV). Available from: <https://doi.org/10.3389/fpsyg.2016.01814>.
- Yiu, E.M.-L., Lo, M.C., and Barrett, E.A., 2017. A systematic review of resonant voice therapy. *International journal of speech-language pathology* [Online], 19 (1), pp.17–29. Available from: <https://doi.org/10.1080/17549507.2016.1226953>.
- Zetterholm, E., 2006. Same speaker–different voices. a study of one impersonator and some of his different imitations. *Proceedings of the 11th australian international conference on speech science and technology*. Australasian Speech Science and Technology Association, pp.70–75.
- Zhang, Z., 2016. Cause-effect relationship between vocal fold physiology and voice production in a three-dimensional phonation model. *The journal of the acoustical society of america* [Online], 139 (4), pp.1493–1507. Available from: <https://doi.org/10.1121/1.4944754>.
- Zhang, Z., 2021. Contribution of laryngeal size to differences between male and female voice production. *The journal of the acoustical society of america* [Online], 150 (6), pp.4511–4521. Available from: <https://doi.org/10.1121/10.0009033>.