

Sweet voice: The role of voice quality in a Japanese feminine style

REBECCA L. STARR

*Department of English Language and Literature, Faculty of Arts and
Social Sciences, National University of Singapore Blk AS5,
7 Arts Link, Singapore 117570
rstarr@nus.edu.sg*

ABSTRACT

‘Sweet voice’, a distinctive Japanese vocal style, illustrates the role played by voice quality as a marker of authenticity in the construction of linguistic styles. The acoustic properties and sociopragmatic functions of sweet voice, as performed by professional voice actresses, are analyzed using data from anime programs, paraphernalia, and fan discourse. Sweet voice is shown to be connected to a traditional notion of Japanese femininity, and licenses the positive use of grammatical features of Japanese Women’s Language. The mature, traditional image conveyed by sweet voice contrasts with the youthful cuteness of burikko and related vocal styles, illustrating that multiple notions of femininity operate within Japanese popular culture. The interplay of voice quality and grammatical features suggests that perceptions of conscious control at different levels of language play a crucial role in social meaning. (Voice quality, Japanese, language and gender, style, authenticity)*

INTRODUCTION

As observed by Lise Skov and Brian Moeran in *Women, media, and consumption in Japan*, Japanese public life is saturated with female voices.

We hear women everywhere. So audible that we hardly listen to their carefully enunciated voices greeting us at the end of almost every telephone line, and on every visit to every commercial building in every city in Japan.... Female are the voices of instruction, warning mothers and children to mind their hands and feet on the escalator, reminding passengers not to forget their belongings on the train, recommending consumers to purchase this detergent, that wine, those contact lenses, cleansing creams or leopard-spotted leotards. Anonymous images, anonymous voices, yet ever present. (Skov & Moeran 1995b:1–2)

The ubiquity of these voices, however, is not their only distinctive quality. The voices sound extraordinarily similar, as if one determined woman were following you around the city of Tokyo with a megaphone. But there is something else, as well: the voices sound oddly alien, and completely unlike announcement voices in other countries. Visitors to Japan often remark that they sound ‘high’ (Miller 2004:151). What they are attempting to describe has, in fact, little to do with

pitch—rather, these voices are all produced using a distinctive voice quality, resulting in a vocal style I refer to here as ‘sweet voice’.¹

Voice quality remains a relatively unstudied area in sociolinguistics, partly due to the methodological challenges inherent in describing phenomena resulting from the many possible complex configurations of the vocal tract (Poděšva 2007). An examination of the acoustic and sociopragmatic functioning of sweet voice may provide us with insights into how voice quality contributes to the creation of style. Specifically, this study seeks to clarify the role of sweet voice in the performance of certain types of femininity in Japanese. This research follows the model of recent work in Japanese language and gender that investigates specific communities and linguistic practices rather than making sweeping statements about how all Japanese women speak (e.g. Okamoto & Shibamoto Smith 2004, 2008; Inoue 2006; Gagné 2008).

The voices of Japanese women have been persistent targets of popular interest and generalization, particularly in the area of pitch, which is closely intertwined with voice quality both in physical and perceptual terms. The current common wisdom on Japanese women’s pitch was codified in 1995 by New York Times columnist Nicholas Kristof, who declared that certain enlightened Japanese women were rebelling against squeaky-high pitch norms in a piece entitled, ‘Japan’s feminine falsetto falls right out of favor’ (Kristof 1995). In the Kristofian view, traditional Japanese femininity requires high pitch, and women who do not use high pitch are therefore challenging the status quo. This view appeared once again on the front page of the New York Times in Hiroko Tabuchi’s (2013) article about the All-Japan Phone Answering Competition, titled ‘Japan’s top voice: high, polite and on the phone’. She reports that ‘some experts explicitly tell women to speak in a higher voice than usual to sound feminine and energetic’ (A1), and portrays women who reject high pitch as modern and untraditional (Tabuchi 2013). In the following analysis, I suggest that the characterization of traditional feminine voices as ‘high’ is problematic in a few respects: most crucially, the confusion of voice quality and pitch has led to a broader conflation of distinct Japanese feminine styles that index different social meanings.

While the suggestion has been made that breathy phonation is perceived as feminine (Ohara 2004), the sociolinguistic literature has not yet rigorously addressed the role of voice quality in the construction of feminine Japanese styles. Previous phonetic studies focusing on automatic speech processing, however, have investigated sociopragmatic functions of phonation types. Ito (2003, 2004) and Campbell & Mokhtari (2003) find that breathiness correlates with polite, formal speech to unfamiliar addressees; Sadanobu (2004) identifies several pragmatic functions of pressed voice including emphasis and admiration. Campbell & Mokhtari (2003) also establish that pitch and voice quality vary independently in their analysis of one female Japanese speaker, with situations involving the most breathiness not always correlating with high pitch. Some recent phonetic work suggests that voice quality is beginning to play a phonemically contrastive role in Japanese: Kong, Yoneyama, & Beckman (2014) find that women who use phonation rather

than voice onset time as a cue for the voiced/voiceless distinction are more likely to use higher pitch and are perceived as more feminine. This research raises questions about the relationship between phonation, pitch, and gendered styles that must be addressed from a sociolinguistic as well as phonetic perspective.

Sweet voice is an example of a stylized, professional voice. In other words, sweet voice is largely the province of professional announcers and voice actors, and is rarely if ever produced by ‘ordinary’ women. Addressed later in the discussion of the voice’s physiological properties, the absence of sweet voice in nonprofessional speech is not merely a matter of convention, but of physical limitation; the sweet voice is difficult to produce, and only those who have trained extensively are able to consistently replicate it. Native Japanese listeners exposed to sweet voice are immediately able to recognize it as a professional voice, often describing it as *anime no koe*, ‘voice from Japanese animation’. As a result, issues of exaggeration or fakeness that apply to other vocal styles that are used by real women, such as the cute *burikko* vocal style described in Miller (2004), do not apply here—the sweet voice is always ‘fake’ in some sense. Nonetheless, I argue here that sweet voice plays a crucial role in establishing authenticity within the frame of the fictional world in which it occurs. I follow Bucholtz (2003) in treating authenticity and the notion of the authentic speaker as emerging from sociolinguistic ideologies and practices; under this view, communities develop ideologies of authenticity that privilege certain features of language as authentic. In the case of sweet voice, its indexation of authenticity results from an ideology in which voice quality is perceived to be beyond conscious control relative to other aspects of speech.

This study approaches sweet voice first from an acoustic perspective and then from a qualitative perspective, looking at how sweet voice functions in the construction of fictional characters and in fan perception.

SWEET VOICE IN ANIME

Why anime?

The sweet voice can be heard in a vast array of contexts—from video games, to television commercials, to public announcements. This analysis, however, focuses on the role of sweet voice in anime (Japanese animated films and television programs). Anime is a particularly suitable format in which to examine the sweet voice style for a number of reasons. From a practical perspective, it is difficult to acquire a large amount of high-quality data from other sources, such as public announcements. Anime has the advantage of not only providing a lot of speech data, but also data from many different speakers, and data from the same speakers using different styles as they portray different characters. Moreover, many of the professional voice actresses, or *seiyuu*, who perform in anime are also involved in producing sweet voice in other contexts; Ouhara Sayaka, for example, performs one of the voices included in the present anime study, does announcements for multiple

train lines, and works as a radio DJ (Haikyou 2008). In terms of gaining insight into the sweet voice, anime is ideal because it provides two levels of context through which we may examine the sociopragmatic functions of this style: the function of the sweet voice within the program itself, and the evaluation of sweet voice and sweet voice characters by fans of the program. While the body of anime programs containing sweet voice encompasses a wide variety of genres and levels of sophistication, the role of sweet voice within these programs nonetheless follows certain recognizable patterns that reveal multiple underlying ideologies relating to gender, Japanese society, and the nature of the human voice and language. This analysis is usefully supplemented by a study of how these programs are perceived by audience members; anime fans are prolific analysers, debaters, and classifiers of characters and genres, creating a superstructure of meta-content and interpretation that can extend or subvert the structure of the original program. This fan-created superstructure can feed back into the creation of new anime programs, so that structures and ideologies of fan culture are integrated into the culture of the programs themselves. Thus, examining both the interior world of the program and the exterior world of fan evaluation is crucial in understanding the structures and styles found in anime.

Seiyuu

Central to this examination of sweet voice are the women who produce it, the seiyuu. The voice acting industry in Japan is highly developed, and seiyuu can achieve a level of prominence comparable to 'live' actors or pop idols.² Seiyuu are involved in voiceover work of all types, including dubbing for foreign films and the adult entertainment industry, but the heart of the profession lies in performing voices for anime; it is their association with particular anime characters that leads to a seiyuu's popularity. While seiyuu can follow multiple paths to entering the profession, they are generally the products of two years at a specialized training school followed by another few years of apprenticeship at their talent agency.³ This extensive, centralized training may account for how the sweet voice has emerged as a distinctive and uniform style.

Table 1 lists the ten seiyuu whose vocal performances are analyzed below.⁴ The fact that seiyuu make use of different vocal styles to portray a variety of characters allows two avenues of acoustic analysis: examining the sweet voice performances as a whole, and then contrasting the sweet voice styles with nonsweet voice styles produced by the same speakers. Both the sweet and nonsweet voice characters are listed for the four seiyuu whose inter-character styles are contrasted. Notice that many of these seiyuu are employed by the same two talent agencies, Ken and Aoni Productions. It is therefore perhaps not so surprising that, although these are different women, the sweet voices they produce are at times startlingly similar.

The four nonsweet voice performances are arguably not as uniform as the sweet voice performances, but they were selected to be representative of relatively neutral,

TABLE 1. *Seiyuu and characters included in voice quality study. (Title and date of anime in parentheses; for English titles see Appendix A).*

Seiyuu	Birth year ^a	Talent agency	Sweet voice character		Nonsweet voice character	
			Name	Description	Name	Description
Nabatame Hitomi (NH)	1976	Ken	Jinguuji Kanade (Gokujou seitokai, 2005)	Student council president	Ise Nanao (Bleach, 004)	Lieutenant
Noda Junko (NJ)	1971	Aoni	Kusunoki Yurara (Petopeto-san, 2005)	Older girl	Tashigi (One piece, 1999) Nausicaä (Kaze no tani no Nausicaä, 1984)	Marine officer Princess
Shimamoto Sumi (SS)	1954	N/A	Kusakabe Yasuko (Tonari no Totoro, 1988) Fujino Shizuru (Mai HiME, 2004; Mai Otome, 2005)	Mother Student council president	Leila Balt (Uchuu no Stellvia, 2003)	Princess Flight instructor
Shindou Naomi (SN)	1972	Aoni	Isabella (Paradise kiss, 2005)	Fashion design student		
Chiharu Suzuka (CS)	1958	Ken	Mother (Binbou shimai monogatari, 2006)	Mother		
Hisakawa Aya (HA)	1968	Aoni	Carmen 99 (Gun X sword, 2005)	Bounty hunter		
Inoue Kikuko (IK)	1964	Office Anemone	Harada Rika (Hachimitsu to kuroubaa, 2005)	Interior designer		
Ouhara Sayaka (OS)	1975	Haikyuu	Yagami Sachiko (Death note, 2006)	Mother		
Satou Ai (SA)	1955	Ken	Kamiyama Misa (Kamisama kazoku, 2006)	Mother		
Touma Yumi (TY)	1966	Aoni		Older sister		

^aYears of birth for the seiyuu were obtained from the ‘Anime news network’ database (animenewsnetwork.com) and cross-checked with their professional websites and other sources when available. While they may not be entirely accurate, I include them here to give some indication of the age range of the speakers.

nonextreme vocal styles. The nonsweet performances all involve voicing female characters; because young male characters are commonly voiced by women in anime, there were several available examples of male character performances by these seiyuu, but they were excluded from this analysis.

Acoustic properties of sweet voice

From a perceptual perspective, the sweet voice style is dramatically distinctive. It has a light timbre and the tense resonance of a singing voice; the speaker sounds as though she is smiling. While speakers shift between different levels of breathiness, a unifying voice quality is retained throughout the performance.⁵ The first analysis presented here investigates whether some acoustic correlates of this style may be identified.

Acoustic measures of voice quality. Although voice quality is sometimes taken as synonymous with phonation, in its broader definition voice quality results from a combination of oral vocal tract settings, including settings of the lips, tongue body, and so on, as well as the laryngeal and glottal settings that comprise phonation (Esling 2012). The sweet voice style certainly involves particular configurations of multiple aspects of the vocal tract beyond the laryngeal source (based on videos of seiyuu performances, for example, this style consistently involves lip spreading, resulting in a ‘smiling’ quality). Nevertheless, what is perceptually and articulatorily most salient about sweet voice is its phonation.⁶ Thus, the acoustic analysis of sweet voice given here focuses on quantitative analysis of phonation, following the practice of previous acoustically based studies of Japanese voice quality (Ito 2003, 2004; Ishi 2004; Kong et al. 2014).

Phonation refers to the process by which the airstream is shaped by the configuration of the vocal folds and other parts of the larynx. Creaky phonation, for example, is produced through the adduction and thickening of the vocal folds, resulting in irregular, slow vibration (Laver 1980). Although changes in laryngeal settings are reflected in the resulting acoustic signal, interpreting the complex movements of the vocal apparatus via acoustic data has proven surprisingly difficult; many different measures have been employed in the phonetics and speech pathology literature, yielding varying degrees of correlation with human perception (Maryn, Roy, de Bodt, van Cauwenberge, & Corthals 2009). These measures often suffer from some degree of ambiguity in discriminating among several possible articulatory sources of a particular acoustic pattern, such as breathiness versus nasalization (Ishi 2004). Another issue in the acoustic measurement of phonation is that the interpretation of differences in the acoustic signal are often reduced to a simplified framework in which common phonation types—creaky, modal, breathy—are laid out in a continuum based on the size of the aperture through which the airstream passes (Gordon & Ladefoged 2001). In reality, laryngeal configuration involves more variables than simply vocal-fold aperture size.

In spite of these drawbacks, acoustic measurements can provide some indications of the physiological settings underlying this style. The most common measures used in studies of phonation are H1-H2, H1-A1, and H1-A3 (Stevens & Hanson 1995; Iseli, Shue, & Alwan 2007; Bishop & Keating 2012; Keating 2014). ‘H’ in this terminology refers to harmonic and ‘A’ refers to formant amplitude; numbers indicate where they occur in the signal (see Figure 1).

H1-H2, meaning the magnitude of the first harmonic relative to that of the second harmonic, correlates with the open quotient, the percentage of time during the glottal cycle that the glottis is open (Swerts & Veldhuis 2001; Kreiman, Shue, Chen, Iseli, Gerratt, Neubauer, & Alwan 2012). H1-A1 is a measure of the bandwidth of the first formant and H1-A3 correlates with spectral tilt, indicating the rate at which upper harmonics lose amplitude. For all three of these measures, higher values are consistent with breathiness, which is characterized by lowered tension in the vocal folds and a wider glottal aperture, resulting in a broader bandwidth for F1 and less energy in the upper harmonics (Gordon & Ladefoged 2001). Higher values of these measures can also result from other vocal settings, however, such as falsetto (a type of phonation in which the vocal folds are stretched and vibrate at a higher frequency), which has proven quite difficult to distinguish acoustically from other phonation types (Podesva 2007; Keating 2014).

As sweet voice is a professional vocal style, previous studies of professionally trained voices, including singers and radio broadcasters, are particularly relevant. Two characteristics of trained voices are most salient here: enhanced periodicity and enhanced energy in the upper harmonics. Periodicity refers to the regularity of the vibration produced by the vocal folds: low periodicity correlates with

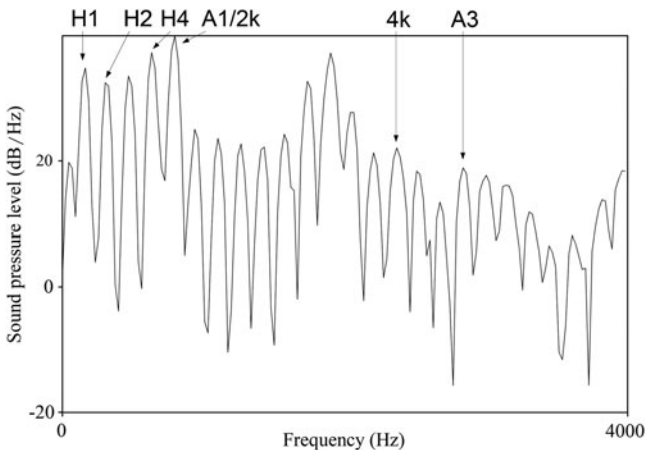


FIGURE 1. Sample spectral slice indicating H1, H2, A1, A3, 2k, and 4k.

dysphonic voices and nonmodal phonation (Hillenbrand, Cleveland, & Erickson 1994; Heman-Ackah, Heuer, Michael, Ostrowski, Horman, Baroody, Hillenbrand, & Sataloff 2003). Trained singers are thought to achieve more regular periodicity by controlling the extrinsic muscles that maintain the larynx in a constant position (Sundberg & Askenfelt 1983). Periodicity has been successfully measured using Cepstral Peak Prominence (CPP). CPP measures cepstral peak magnitude relative to the overall amplitude of the cepstral signal, calculated using a cepstrum created by Fourier transformation of a Fourier transformation spectrum (Garrett 2013). CPP correlates highly with perception of dysphonic voices, and is considered a more reliable measure of breathiness than other common measures such as relative amplitude of the first harmonic or spectral tilt (Hillenbrand et al. 1994; Heman-Ackah et al. 2003; Samlan & Story 2011; Garret 2013; Brinca, Batista, Tavares, Goncalves, & Moreno 2014). Voices with higher CPP are evaluated more positively in several respects: most notably, higher values correlate with perceived sexual attractiveness in both genders (Balasubramaniam, Bhat, Srivastava, & Eldose 2012). Another measure related to periodicity is harmonic-to-noise ratio (HNR), also calculated using the cepstrum. HNR at various frequency ranges has been used as a measure of roughness, breathiness, and falsetto versus modal phonation (de Krom 1993; Shue, Chen, & Alwan 2010; Keating 2014). Both CPP and HNR measures are employed in this analysis.

The second relevant feature of trained voices is upper harmonic energy: specifically, a phenomenon known as the Actor's Formant (AF). The AF is a spoken analogue of the Singer's Formant, both of which consist of a peak in spectral energy in the 3 kHz region (e.g. Leino 1993; Sundberg 1987, 2001). The AF is thought to correlate with what is perceived as a ringing or resonant voice quality, perhaps resulting from a narrowing of the laryngeal opening (Master, de Biase, Brasilia, & Laukkanen 2008; Lin, Jayakody, & Looi 2009). In the realm of voice artists, Warhurst, McCabe, Yiu, Heard, & Madill (2013) found that male commercial radio broadcasters had a more prominent AF than public radio broadcasters or nonbroadcaster controls. The AF has not been consistently found in female voices, however (Master, de Biase, & Madureira 2012). AF is most commonly quantified by measuring the relative magnitude of the highest peak in the 3-4 kHz region with the highest peak lower in the spectrum; here, this measure is operationalized as 2k-4k, meaning the highest peak around 4 kHz subtracted from the highest peak around 2 kHz.

Methodology. This analysis makes use of the six measures described above: H1-H2, H1-A1, H1-A3, 2k-4k, CPP, and HNR. The data were drawn from the ten sweet voice performances listed in Table 1, with the addition of the four nonsweet voice performances specified in the table to allow for an additional intra-speaker analysis. An average of approximately one minute of data per speaker was collected and then analyzed using the software Praat (Boersma & Weenink 2012) for all measures but CPP and HNR, which were automatically extracted using the software VoiceSauce (Shue 2009). Because these measures

have been shown to vary significantly by vowel context, only tokens of /a/ vowels were analyzed (/a/ is often preferred in phonation studies due to the height of F1, which minimizes the influence of F1 on the first two harmonics) (Iseli et al. 2007). Praat measures were extracted automatically, taking measurements at 25%, 50%, and 75% of vowel duration.⁷ CPP and HNR were automatically extracted by VoiceSauce using the default settings. HNR in VoiceSauce is calculated in four measures with increasing ranges: from 0–500 Hz, 0–1500 Hz, 0–2500 Hz, and 0–3500 Hz. As all HNR measures produced similar trends, only the broadest range (HNR35) is presented.

The sweet and nonsweet voice performances analyzed here include tokens that are more perceptibly breathy than others. It is of interest to study the individual acoustic characteristics of these different phonation types to get a clearer picture of how they combine to form an overall style. Therefore, tokens were manually classified by the author's perception into creaky, modal, and breathy categories (ambiguous tokens were classified as modal). As I explore in the intra-speaker analysis, 'modal' may not be an entirely precise description of the nonbreathy phonation used in sweet voice; for now, we use it as a convenient label for the default phonation type in this style.

Inter-speaker analysis. The first analysis seeks to acoustically characterize the ten sweet voice performances. Perceptual categorization by the author of the phonation type of each token finds that most speakers employ a mix of modal and breathy phonation, with very few creaky tokens (Figure 2). As creaky tokens are so rare, subsequent analysis contrasts only modal and breathy tokens.

Only two speakers (SS & TY) use more than 50% breathy phonation. The breathiness in these roles appears to serve two rather different stylistic purposes,

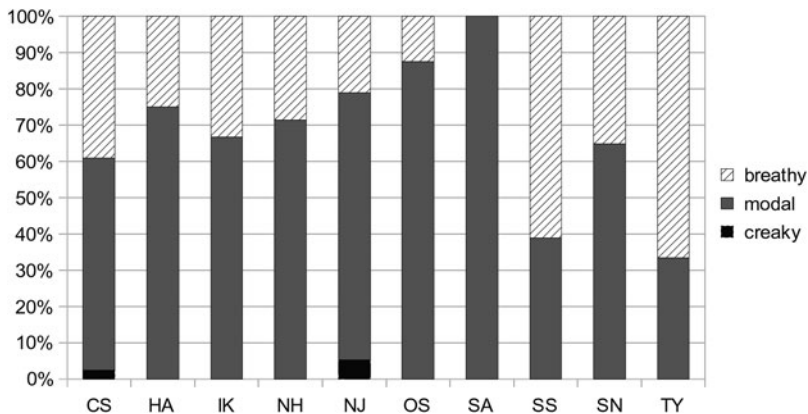


FIGURE 2. Proportion of perceptually creaky, modal, and breathy tokens for ten sweet voice performances.

as Shimamoto is portraying a mother interacting with her young children while Touma is portraying an alluring young woman flirting with boys. The role of sweet voice in constructing character types is discussed below (FUNCTION OF SWEET VOICE IN ANIME).

Mean values for each acoustic measure for all sweet voice performances are given in Table 2. In terms of the distinction between modal and breathy phonation, while not all speakers show the expected trend for each measure, overall these results are consistent with previous work: H1-H2, H1-A1, H1-A3, and 2k-4k tend to be higher for breathy phonation, while CPP and HNR are lower. Linear mixed-model regressions performed in Rbrul (Johnson 2009) identify phonation as a highly significant factor in predicting each measure, controlling for F0, time point, and speaker (see Appendix B). Although the H1-based measures fail to reflect expected trends for three speakers (NH, NJ, OS), the CPP and HNR measures are consistent across the board; thus, as in previous work, periodicity is found to be a more reliable correlate of perceived breathiness (e.g. Hillenbrand et al. 1994).

The mean fundamental frequency (F0) of the ten sweet voice performances is briefly examined here.⁸ As indicated in Table 2, the relationship between perceived phonation type and mean F0 is not particularly robust across speakers. Linear mixed-model analysis, however, identifies a small but significant positive correlation between breathiness and higher F0 when controlling for speaker (see Appendix B). This finding is somewhat surprising because low F0 is often associated with breathiness, a pattern that does not hold here (Honorof & Whalen 2004).

Rather than only comparing the pitch of these sweet voice characters to the pitch of ordinary Japanese women, it seems most relevant to first interpret them in relation to the other voice types that commonly appear in anime. In Teshigawara's (2003) study of the voices of anime heroes and villains, she finds two groups of pitch targets for female heroic characters: 320 Hz and 400 Hz. As shown in Figure 3, the pitch of sweet voice characters is generally lower than both of these targets. Removing the exceptionally high Touma Yumi and Satou Ai, who average at 307.92 Hz, the mean pitch of the sweet voice characters is 262.7 Hz. This finding is consistent with the intuition that, while sweet voice can occur at a range of pitches, it is most often distinctively low by female anime voice standards. As discussed in the following section on the function of sweet voice within the anime setting, sweet voice characters are not heroines but rather are used to portray older-sister types; this lower pitch target helps distinguish them from main characters and establishes their relative maturity.

In comparison to the speech of ordinary Japanese women, sweet voice is relatively high: Masataka (1992) finds that Japanese mothers have an average pitch of 177 Hz when speaking to adults. When addressing children, however, their pitch increases gradually from 175 Hz to 252 Hz as they produce more utterances trying to elicit a child response. By contrast, Ohara (1999) finds a much higher average pitch for adult women, with her population of five female university students

TABLE 2. *Acoustic measures by phonation type, ten sweet voice performances.*

Spkr	N	F0		H1-H2		H1-A1		H1-A3		CPP		HNR35		2k-4k	
		modal	brthy	modal	brthy	modal	brthy	modal	brthy	modal	brthy	modal	brthy	modal	brthy
CS	123	231.1	233.8	5.26	6.97	0.34	1.22	17.71	15.51	19.77	16.12	22.47	18.69	17.60	16.38
HA	12	243.1	321.6	-1.19	5.98	-3.25	0.00	8.11	23.81	21.15	16.03	25.63	17.10	12.28	20.37
IK	72	252.8	281.3	4.21	8.40	-1.83	2.57	14.37	18.78	21.26	16.77	25.36	18.49	14.43	14.79
NH	42	286.1	221.7	2.04	-7.93	-1.04	-2.88	9.87	-1.16	20.18	15.92	21.30	18.54	13.97	12.89
NJ	57	286.6	233.2	6.18	4.34	0.79	0.73	16.44	7.58	17.82	17.32	24.79	19.48	17.28	12.58
OS	48	241.4	207.3	5.25	-5.22	4.9	-0.88	21.01	6.70	20.28	15.57	30.22	22.45	18.66	16.35
SA	30	305.8	NA	-6.97	NA	-12.27	NA	-14.22	NA	20.98	NA	39.00	NA	3.426	NA
SS	108	277.3	305.3	-1.51	0.86	-8.66	-3.35	-0.88	4.49	21.96	17.77	31.71	24.89	4.28	8.18
SN	171	257.4	260.3	0.69	6.96	-3.86	3.98	8.06	18.01	22.74	16.37	27.50	18.99	10.99	16.50
TY	9	335.5	274.8	-5.4	7.27	-10.11	1.00	-1.8	10.81	22.37	16.13	25.21	17.98	6.26	10.78

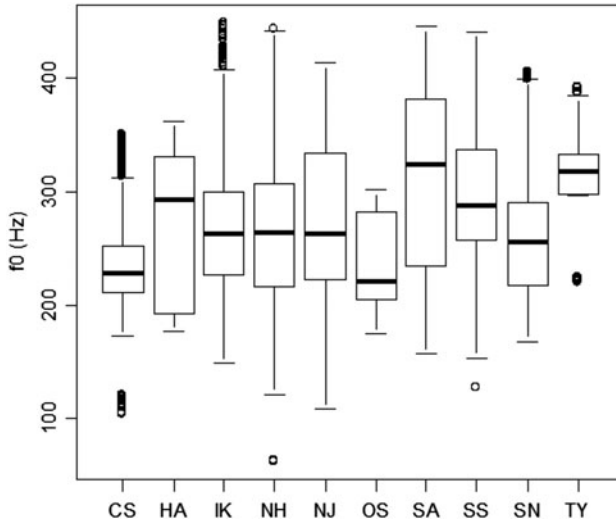


FIGURE 3. Boxplot indicating range of f_0 (Hz) for ten sweet voice characters (labeled by seiyuu).

using an average F_0 of 261.6 Hz when addressing a professor and 253.2 Hz when addressing friends. Taken together, these studies situate the pitch of sweet voice in the higher end of the normal female range in Japanese. Nonetheless, in a media speech context, and particularly in anime, the sweet voice is relatively low.

Intra-speaker analysis. As indicated in Table 1, contrasting nonsweet styles were collected for four of the seiyuu studied. From a qualitative perceptual standpoint, the sweet and nonsweet styles were extremely different from each other, to the extent that it was difficult to detect that the same speaker was producing both of them. Examining changes within individual speakers' performance of sweet and nonsweet vocal styles may help address the question of what makes the sweet voice so distinctive, while controlling for possible physiological and stylistic differences between speakers.

In the perceptual analysis, both the nonsweet and sweet performances consist of a mix of modal and breathy tokens, with only one speaker producing a significantly different proportion of breathy and modal (Shimamoto Sumi, $\chi^2(1) = 30.404$, $p < .0001$) (Figure 4). As was the case for the sweet voice performances examined above, the nonsweet performances contain only a handful of creaky tokens (from the same seiyuu, Noda Junko). The difference in styles cannot, therefore, be accounted for by a change in the proportion of phonation types.

Table 3 lists the mean acoustic measures calculated for the four seiyuu in their two different styles. Mixed-effects linear regression modeling was carried out in Rbrul for each of the acoustic measures among the four sweet and nonsweet

SWEET VOICE

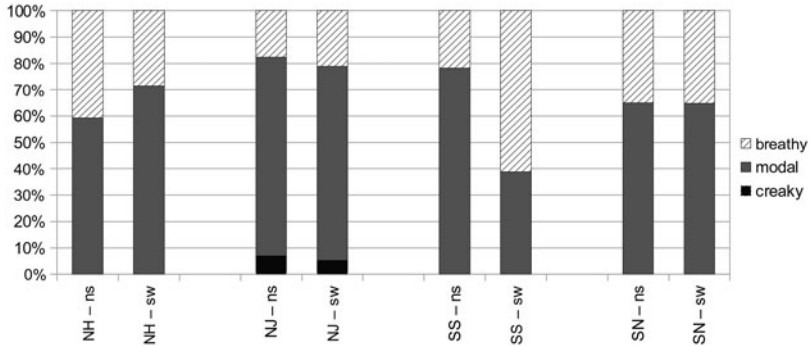


FIGURE 4. Proportion of perceptually creaky, modal, and breathy tokens for nonsweet vs. sweet performances.

performances, controlling for speaker, phonation type, F0, and time point (see Appendix B). Sweet versus nonsweet style was a significant factor in predicting all measures except for CPP. For H1-H2, H1-A1, and H1-A3, sweet voice style correlated with higher values. As shown in Table 3, this effect can be seen among both the modal and breathy tokens. In contrast, for 2k-4k, sweet voice correlates with lower values: this trend is consistent for all but one of the cells. This lower 2k-4k value indicates that the sweet voice style has greater energy in the upper harmonics: specifically, it provides evidence for the Actor’s Formant. This is consistent with the perception that sweet voice is more resonant. Similarly, the HNR35 value is significantly higher for sweet voice than nonsweet voice. This measure is less consistent than 2k-4k, however, with two speakers showing higher HNR for sweet voice and two showing lower HNR. Finally, CPP shows no consistent difference between the two styles, with two speakers (SS and SN) displaying remarkably similar CPP values in both styles for each phonation type.

The mixed model indicates a small but significant positive correlation between sweet voice and higher pitch (see Appendix B). As shown in Figure 5, however, it is evident that this is not a robust phenomenon, as only two of the speakers demonstrate this F0 pattern.

In sum, while not all performances showed the same trend, the intra-speaker acoustic analysis has found several significant differences between nonsweet and sweet voice performances: higher values for H1-H2, H1-A1, H1-A3, HNR35, and F0, and lower values for 2k-4k. These trends are not wholly consistent with a one-dimensional phonation scale explanation: for example, for increased breathiness we would expect higher values on the H1-related measures, but not lower 2k-4k or higher HNR. These results are also inconsistent with falsetto, which has a lower HNR than modal phonation (Keating 2014).

In light of these acoustic findings, let us return to the overall perception of these nonsweet versus sweet performances. The sweet voice performances have lighter

TABLE 3. *Acoustic measures by phonation type, four nonsweet vs. sweet voice performances.*

Spk.	Sty.	N	F0		H1-H2		H1-A1		H1-A3		CPP		HNR35		2k-4k	
			modal	brthy	modal	brthy	modal	brthy	modal	brthy	modal	brthy	modal	brthy	modal	brthy
NH	nsw	81	227.4	176.8	0.6	111.44	-2.64	9.07	15.75	19.63	18.76	13.95	27.09	22.78	22.06	17.54
	sw	42	286.1	221.7	2.04	-7.93	-1.04	-2.88	9.87	-1.16	20.18	15.92	21.30	18.54	13.97	12.89
NJ	nsw	510	305.0	311.5	-10.5	-0.94	-13.8	-4.62	-0.14	8.87	23.15	15.72	32.80	25.55	11.36	13.52
	sw	57	286.6	233.2	6.18	4.34	0.79	0.73	16.44	7.58	17.82	17.32	24.79	19.48	17.28	12.58
SS	nsw	96	299.5	314.3	-13.2	-3.43	-18.7	-5.65	-8.42	3.14	21.03	17.76	11.64	9.32	7.65	11.63
	sw	108	277.3	305.3	-1.51	0.86	-8.66	-3.35	-0.88	4.49	21.96	17.77	31.71	24.89	4.28	8.18
SN	nsw	60	250.9	166.5	-3.25	2.06	-5.3	2.8	10.32	21.24	21.21	16.36	20.57	14.85	17.32	21.77
	sw	171	257.4	260.3	0.69	6.96	-3.86	3.98	8.06	18.01	22.74	16.37	27.50	18.99	10.99	16.50

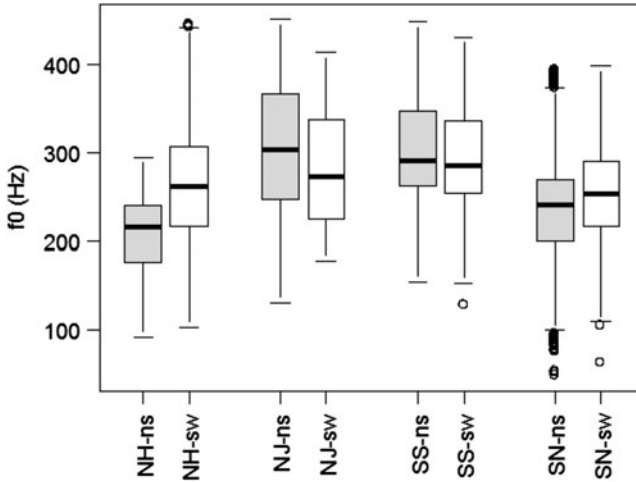


FIGURE 5. Boxplots of mean F0 for four seiyuu nonsweet (gray) vs. sweet (white) vocal styles.

timbres; as noted earlier, to a layperson, they may be perceived as ‘higher’ voices although they are occurring at the same absolute F0. Most notably, the performances from the same seiyuu sound so different that they are easily mistaken for two different individuals. In order to achieve this shift, the seiyuu must be making dramatic changes in their vocal apparatus. Moreover, we can speculate that sweet voice involves an unusual vocal setting, as this style is not normally encountered in everyday speech and only appears among trained professionals.

The most obvious example of a phonation type less commonly used in speech is falsetto; in the production of falsetto voice, the vocal folds are stretched so that they become thinner and vibrate more quickly, giving the voice a lighter timbre. A falsetto explanation for sweet voice is appealing in some ways, but it is not consistent with the higher HNR or lower 2k-4k observed in sweet voice—more importantly, sweet voice simply does not sound like falsetto voice. Sweet voice has a more tense sound, with a resonance that is reminiscent of a trained female singer in the middle of her range. In fact, there is a term used in the study of singing voices for a register lower than falsetto: ‘head voice’ or ‘head register’. In head register, the vocal ligaments are elongated as in falsetto voice, but the vocalis muscle is tensed; the resultant sound is more resonant than falsetto (Hirano, Vennard, & Ohala 1970; Miller 2000). Professional singers generally practice developing more resonance in their head register as a way to avoid a perceptible gap between modal and falsetto as their pitch rises (Cryar 2014). Thus, resonance in head voice is a crucial feature that distinguishes trained and untrained voices. Björkner (2006) observed a higher open quotient in head register compared to chest (modal) voice in female singers; this finding is consistent with the higher H1-H2 acoustic measure found in the present analysis.

The category of head register is not generally used in the study of spoken phonation, where it is apparently grouped in with falsetto. Certainly, speaking in head register is not a common phenomenon. Nonetheless, the notion that one can be trained to speak in head register, as distinct from falsetto and chest register, is widely accepted in the vocal coaching field (Jahn 2011).⁹ The fact that producing a resonant head voice generally requires vocal training may explain the observation that sweet voice is quite difficult to replicate and is not used by ordinary Japanese women. Indeed, the head register theory can account for several of our observations about sweet voice. Acoustically, the higher H1-H2 and H1-A3 values for the sweet voice performances are consistent with head voice. The perception of the sweet voice style as lighter is also explained by head voice, in which the vocal folds are thinned and tensed, resulting in a lighter timbre. Previous research suggests that listeners use voice quality cues to estimate the location of an unfamiliar speaker's F0 within that speaker's pitch range: for example, pitches with higher H1-A3 values are rated as being lower in the speaker's range (Honorof & Whalen 2004; Bishop & Keating 2012). This suggests that what listeners are hearing when they perceive that the sweet voice is 'high' is not in fact that speakers are producing a high F0, but rather that the F0 is perceived as being relatively low in the speaker's overall pitch range due to its higher spectral tilt.

Summary of acoustic study results. This acoustic examination of the sweet voice style sought to characterize the physiological basis of the sweet voice by examining the performances of ten seiyuu. Sweet voice performances consisted predominantly of a mix of perceptually modal and breathy tokens, corresponding to significant differences in H1-H2, H1-A1, H1-A3, CPP, HNR, and 2k-4k measures. Mean F0 of sweet voice performances was lower than that of female anime heroines but comparable to the higher end of ordinary Japanese female voices. An intra-speaker analysis of four seiyuu performing nonsweet and sweet vocal styles found higher values for sweet voice for the measures H1-H2, H1-A1, H1-A3, and HNR. Lower values were found for 2k-4k. Mean F0 was higher for sweet voice for two of the four seiyuu. The use of head register is proposed as an explanation for the acoustic and perceptual characteristics of sweet voice style.

Given the difficulty in studying voice quality through acoustic data alone, it may be profitable in the future to expand this analysis by studying the articulatory process of sweet voice production via ultrasound, EGG, or other similar methodologies.

Function of sweet voice in anime

Sweet voice characters. After developing a sense of the physical properties of sweet voice, it is time to move on to examining how the sweet voice functions within the anime setting. Again, this analysis primarily makes use of the ten

sweet voice performances listed in [Table 1](#), as well as a handful of additional performances from similar programs viewed in the course of this research.

Female characters with sweet voices appear in various genres of anime, including both male-oriented (shounen) and female-oriented (shoujo) programs, targeting multiple age ranges. In spite of this, certain commonalities exist among the characters who have sweet voices. Four general principles concerning sweet voice character traits can be outlined as follows.

- i. Sweet voice characters are primarily (relatively) older women in positions of traditional female authority. The roles of the ten sweet voice characters in [Table 1](#) include three mothers, two student council presidents, one older sister, two older girls, and one interior designer. Other sweet voice character occupations include teacher, school nurse, and goddess. The one nontraditional occupation in the ten performances is Carmen 99 (*Gun X sword*, 2005), a kind-hearted bounty hunter.
- ii. Sweet voice characters are extraordinarily beautiful. Due to the stylized nature of the representation of the human form in anime, it is often difficult to determine by looking at a character whether they are meant to be attractive. Thus, voice quality provides a valuable cue to the audience in determining how a character's appearance is perceived by others. Characters with sweet voices are often beautiful to the degree that they elicit remarks of admiration, for example, "I've never seen such a beautiful creature before!" (*Petopeto-san*). Even mothers with sweet voice, who one might imagine are excluded from the heterosexual marketplace, are explicitly stated to be beautiful; in *Tonari no Totoro*, for example, the mother is asked by her young daughter if she will have hair as beautiful as hers one day. The use of beautiful here is crucially different from cute—sweet voice characters are not cute, but rather possess a mature, classical beauty.
- iii. Sweet voice characters are neither villainesses nor heroines. As illustrated in principle i, sweet-voiced women are generally positive supporting characters who take care of the protagonist (either male or female) in some fashion. These characters rarely undertake independent action, instead reacting to events around them, most often with calm resignation.
- iv. Sweet voice characters are sexualized yet ladylike. With the exception of the characters in young children's films, the sweet voice characters have their sexuality highlighted in some respect, though their behavior is generally far from lewd. These characters also tend to dress conservatively, wearing long skirts. Of the ten sweet voice characters in [Table 1](#), three are explicitly or implicitly nonheterosexual.

Readers familiar with anime tropes will recognize that the character qualities described above are not representative of the average female anime character. A more typical anime girl is well-represented by one of the nonsweet voice examples from [Table 1](#), Nausicaä from the Miyazaki film *Kaze no tani no Nausicaä*, a spunky, intelligent young princess who saves her community and brings peace to the world. The sweet voice characters, with their calm maturity, stand out in a medium dominated by frenetic youth and the aesthetic of cuteness.

Interaction of Japanese Women's Language and sweet voice. In addition to the distinctive behaviors and character traits associated with the sweet voice, the use of sweet voice in anime is accompanied by a set of linguistic features known as Japanese Women's Language (JWL). Contrary to what its name suggests, scholars agree that JWL is not commonly used today by most Japanese women, but rather only by a subset, and only in certain contexts (Okamoto & Shibamoto Smith 2004:3). Nevertheless, JWL continues to be associated with women's speech in Japan, and carries pragmatic implications of supposedly traditional Japanese feminine features including softness and nonassertiveness (Inoue 2006:12).

A selection of previously studied JWL features that occur in the sweet voice data are listed in Table 4. The presence of each feature in the ten sweet voice performances and four nonsweet performances is listed on the right.

While these data are limited by the brevity and content of the analyzed passages, and by the lack of a full-scale variationist analysis, some clear trends are nonetheless present. The most frequently observed JWL features in the sweet voice performances studied are sentence-final rising *ne* and *wa* (features such as interjection *ara* are more heavily bound to context and thus more rare). The fact that *ne* and *atashi* each occurred once in the nonsweet clips may indicate that they are less marked than other JWL features.

As described in Okamoto & Shibamoto Smith (2008), two discourses are relevant in the discussion of JWL: norms dictating that women should use 'polite, gentle, and refined' speech and norms regarding which linguistic features index

TABLE 4. Selection of JWL features with number of observations in sweet and nonsweet performances.

FEATURE	EXAMPLE	%	
		SWEET	NONSWEET
sentence-final <i>ne</i> with rising pitch (Enyo 2005)	<i>aikawarazu ne</i> 'the same as always' (Gun X word, 2005)	7/10	1/4
sentence-final particle <i>wa</i> (Takahara 1991; Okamoto & Shibamoto Smith 2008)	<i>chikaku ni aru hazu desu wa</i> 'should be nearby' (Mai Otome, 2004)	6/10	0
first-person pronoun <i>atashi</i> and <i>atakushi</i> (Shibamoto Smith 2003)	<i>atashi no kodomo no koro to sokkuri</i> 'exactly the same as when I was a child' (Tonari no Totoro, 1988)	4/10	1/4
tag-question <i>kashira</i> (Okamoto & Shibamoto Smith 2008)	<i>Petopeto-san datta kashira?</i> 'I wonder, aren't you Petopeto-san?' (Petopeto-San, 2005)	4/10	0
interjection <i>ara</i> (Okamoto & Shibamoto Smith 2008)	<i>ara ara, mou shoukai wo kuremashita ga</i> 'ah, well then, I will introduce myself' (Gun X sword, 2005)	3/10	0
sentence-final particles <i>no yo</i> (Okamoto & Shibamoto Smith 2008)	<i>kekko takai no yo</i> 'quite expensive' (Paradise kiss, 2005)	2/10	0

these traits (Okamoto & Shibamoto Smith 2008:88). Within this framework, the authors observe that various linguistic features may be identified as contributing to a feminine style: for example, while many linguistic norms for JWL relate to using features from standard Japanese, viewed as more refined than dialects, feminine styles can also be indexed using certain regional dialects. Specifically, the authors observe cases in which polite forms of the Kyoto dialect are used to construct a feminine style on television dramas (Okamoto & Shibamoto Smith 2008:102). Indeed, in the sweet voice corpus there is one example of a speaker who uses the Kyoto dialect (Mai HiME, 2004). This finding underscores the point that the features listed above are simply a subset of possible grammatical features that may be used to construct JWL.

The use of JWL in anime is particularly significant in that it introduces these features to young viewers who do not necessarily hear JWL from speakers in their everyday lives. In her account of her own experiences with JWL, for example, Inoue (2006:8) identifies particular television characters, including a teacher and an older-sister character, as the earliest speakers of JWL she ever encountered. Thus, JWL's function in anime is not merely interesting as a reflection of societal patterns, but may be crucial in the creation and perpetuation of second-order indices that connect linguistic features with particular feminine traits in Japanese society.

Given that sweet voice characters consistently use JWL, we might ask whether sweet voice carries any social significance independently of JWL. Is it simply a part of the women's language package? In fact, an asymmetry exists between the use of sweet voice and JWL in anime: while sweet voice does not occur without JWL, JWL can occur without sweet voice. Moreover, when JWL does occur without sweet voice, and in particular when certain marked features occur, such as sentence-final *wa*, the pragmatic function of the JWL is often quite different than it is in the case of the sweet voice characters—it implies a negative character evaluation. To put it another way, it is only when JWL co-occurs with sweet voice that it is portrayed positively.

Characters who use JWL but lack a sweet voice often fall into the category of the spoiled *oujousama*, or princess. The spoiled princess is a comedic villain, a girl with money to burn on expensive feminine clothing and a fancy education, who seeks to portray herself as proper and refined, but is inevitably revealed to have violent, unladylike tendencies under the surface.¹⁰ Her trademark is her theatrical high-pitched evil laugh, accompanied by a stylized gesture of her hand poised to one side of her mouth. Through her expensive upbringing, the spoiled princess has acquired features of JWL; she flaunts her socioeconomic privilege by making use of extremely formal, feminine language at all times. While she can use grammatical features to put on a front of false femininity, she lacks a sweet voice quality—her voice is quite ordinary and modal. No matter how she disguises her personality, her unpleasant internal character prevents her from having a sweet voice.

The nonsweet voice of the spoiled-princess type reveals an ideology of linguistic authenticity in which speakers are able to manipulate their own use of grammatical features, but cannot control their voice quality. This is corroborated by the observation that characters are generally unable to turn sweet voice on and off. Within the setting of anime, sweet voice cannot be faked; one must be an authentically sweet person to have a sweet voice.¹¹ This system has a level of irony to it, since viewers are aware that sweet voice is always, in some sense, a fake voice, performed by seiyuu who are able to produce a multitude of different vocal styles. The spoiled-princess type also reveals a negative view of JWL; in anime, sounding like a traditional Japanese woman is not necessarily desirable.

If we consider the broader linguistic context of anime, it is not particularly surprising that JWL features would become a resource for indexing negative character qualities. Mainstream anime celebrates youthful values and rejects, to some extent, traditional Japanese values and the world of adult responsibilities. More specifically, the aesthetic of cuteness that pervades anime, with its Western orientation and romanticization of the childlike, can be seen as a rejection of traditional Japanese female roles (Kinsella 1995:249). In a marketplace in which the greatest cultural capital is held by the modern and the young, JWL, as an index of traditional femininity, becomes the way that ‘other people’ talk. Even in the case of sweet voice characters, JWL still marks the other, although these other people are mothers and sisters who are portrayed positively. In the case of the spoiled princess, JWL indexes the inauthentic other, a rich girl who misguidedly attempts to buy her way into a refinement that can only be naturally possessed.

A primary function of sweet voice within the anime setting, then, is to provide authenticity to feminine gender performance. We might also argue that the sweet voice, as an index of physical beauty, fills a role that is played by women’s language in more mainstream Japanese cultural traditions. Washi (2004:87) cites Kotaka Otsuma, an early twentieth-century advocate of women’s higher education, in her discussion of the promotion of women’s language as a means of enhancing physical appearance.

Say there is a lady. Say she possesses the beauty of a flower, looking as one imagines Yang Guifei or Cleopatra to have been in days of yore... What would it be like if that lady then opened her mouth and spoke in *beranmee kotoba* [a lower-class dialect]? Alas! her lovely countenance and arching brow would lose their luster; her very garb would appear a sham. Thus decorous language is more beautiful still than fair looks. (Otsuma 1929:1–2)

As this applies to anime, when a lady opens her mouth and speaks without a sweet voice, it is the fair trappings of grammatical features that can appear false.

Perception of sweet voice among fans

In the previous section, I observed some commonalities and distinctions among various female anime characters and, based on this information, proposed two relevant character classes: sweet-voice characters and spoiled-princess characters.

By engaging in this practice, I have joined the company of thousands of anime viewers who interpret and re-interpret the characters, settings, and events of anime programs, creating new categories and discourses through which anime is consumed. The present section explores how fans interpret the significance of the sweet voice. While there is little explicit discussion of the sweet voice in itself, we can gain insight into how the sweet voice is interpreted through fans' perception of the characters who use sweet voice, and of the seiyuu who produce them.¹²

Character-focused fan perception. Although the female characters in anime do not bear a close resemblance to live human beings, they nonetheless inspire feelings among certain fans.¹³ As a result, much discussion has been devoted to what exactly makes different types of characters attractive.

One of the outcomes of this discourse is the notion of moe. Moe has a few meanings, but here we use it to refer to the erotic appeal of an anime character. Underlying the discourse of moe is the assumption that the emotional attraction that one feels toward an anime character is distinct from the appeal of a real human being, and therefore worthy of its own system of classification and discussion. On the website Heisei Democracy, a blog covering Japanese popular 'two-dimensional' culture (i.e. anime and manga), an author named Shingo has published several ambitious articles related to moe, one of them titled 'A discourse on moe, Part I: Introduction and a moetic taxonomy'. Within his extensive moetic taxonomy, we find a description of a class of characters called the 'Yamato Nadeshiko type'.

The gentle yet strong, nurturing yet uncompromising Yamato Nadeshiko is one of the ideal forms of the Japanese woman that has endured since antiquity. This is a more mature character than the others, who is made beautiful by the weight of a few additional years of experience resting on her shoulders. The natural inclination of the moetic observer is toward regression to a childlike state, succumbing [sic] to the temptation of ministrations at her capable hands. Famous Yamato Nadeshiko: Belldandy (Ah! Megami- sama)... and any other character voiced by Inoue Kikuko. (Shingo 2005)

In this excerpt, Shingo refers specifically to voice as a defining feature of this character type, supporting our notion that voice quality is a significant contributor to the identity of these sweet voice characters. In fact, he brings up one of the seiyuu from Table 1, Inoue Kikuko (the most famous of the seiyuu included in this study), and declares her voice to be a sufficient condition for any character to qualify as a Yamato Nadeshiko type.

What exactly does he mean by Yamato Nadeshiko? Yamato Nadeshiko, a metaphor for the ideal Japanese woman literally referring to a species of flower (the fringed pink), is a potent term with deep historical roots that remains in use today (Bullock 1994; Sugihara & Katsurada 1999). The notion of Yamato Nadeshiko had its heyday in twentieth-century imperialist Japan, in which, in contrast with previous eras, Japanese women were touted as the anchors of the home and 'mothers of the nation', rather than viewed simply as inferior to men (Skov & Moeran 1995a). The chronicler of Japan, Lafcadio Hearn, elegantly summarized the Yamato Nadeshiko ideal while describing the appeal of the Japanese woman to a Western audience:

A being working only for others, thinking only for others, happy only in making pleasure for others, —a being incapable of unkindness, incapable of selfishness, incapable of acting contrary to her own inherited sense of right,—and in spite of this softness and gentleness ready, at any moment, to lay down her life, to sacrifice everything at the call of duty; such was the character of the Japanese woman. (Hearn 1905:396–97)

Given that 100 years separate the above description and the anime involved in the present study, it is almost alarming how closely the sweet voice characters adhere to Hearn's notion of the character of the Japanese woman. While these remarkable women have never existed in the real Japan, in the Japan of the imagination they are alive and well.

Returning to Shingo's description of the Yamato Nadeshiko type, another significant point he makes is that this class of character is the most mature of the moe categories. Moe characters are normally extremely young girls, and the Yamato Nadeshiko is the single exception, made 'more beautiful' by her maturity. What qualities have enticed these fans of youthful cuteness to move to a higher age bracket? While Shingo's description avoids explicit discussion of sexuality, the appeal here is clear; regardless of how ladylike the Yamato Nadeshiko may appear, her maturity implies sexual experience. The sexual appeal of the older woman in Japan was noted by Clammer (1995:210) in his study of men's magazines, in which he observed a contrast between cute younger girls and the 'true sensuality' embodied by more mature women. This opposition of cuteness and mature beauty is one of the central distinctions that forms the structure of classification of sexually attractive women negotiated in the discourse of moe.

With the arrival of the Yamato Nadeshiko label, our sweet voice character, once a free-standing curiosity with an unusual voice, now finds herself interpolated into an elaborate structure in which she is contrasted with her younger, cuter peers, and connected to ancient female archetypes of the ideal woman. This system of female classification is in one sense unique to anime fan subculture, but in other ways part of a wide-reaching discourse implicated in numerous aspects of Japanese life. The phenomenon of categorizing the sexual appeal of women is hardly exclusive to anime fans, but moetic discourse is part of a distinctive subculture that engages in extensive quantification and standardization of the consumption of female characters, to the point of developing notions such as the zettai ryouiki, 'absolute area', meaning the distance of bare skin between the hem of the skirt and the top of the socks, for which the ideal ratio has been calculated to be 4:1:2.5 with a margin of error of $\pm 25\%$ (Dark Mirage 2008). This penchant for detailed classification is reminiscent of another Japanese subculture, that of advertising agencies, which in the 1980s yielded a plethora of classifications for types of female consumers, such as the 'yenjoy gal', the young working woman who likes to spend money during her leisure time (Skov & Moeran 1995b:37). The identification of the Yamato Nadeshiko type, then, can be interpreted both in relation to the array of contrasting anime character types and to

other subtypes of women who have been classified, measured, accounted for, and consumed in ancient and contemporary Japan.

Seiyuu-focused fan perception. A slightly different class of fan discussion surrounding the sweet voice focuses on the voice actress performing the vocal style, rather than on the character for whom the voice is performed. While not all successful seiyuu become big stars, some of them are able to parlay their success in performing particular roles into personal fame, moving from voice acting into recording music and making personal appearances. One seiyuu who has managed this feat particularly well is Inoue Kikuko, who has carved a successful niche for herself based on her sweet voice performances. Accounts of Inoue's voice written by fans commonly include both perceptual descriptors (e.g. 'soft', 'sweet') and descriptors which reference typical sweet voice roles (e.g. 'motherly', 'angelic').

Often called the voice of a Goddess, Inoue Kikuko is a very popular seiyuu mostly recognised for her very soft and gentle speaking voice. Kikuko gained recognition from her role as Tendo Kasumi in *Ranma 1/2*. It was due to this role that she earned her nick name as 'Onee-chan' (big sister). Her role as Belladandy in *Ah! My Goddess* is most known among seiyuu fans. Applying her tender angelic voice to great advantage, she was able to portray the gentleness of a pure goddess so fittingly. (Seiyuu Directory 2002)

If you hear a sweet, motherly voice in anime, Kikuko Inoue is probably behind it as one of the premier voice actresses in Japan. (tv.com 2008)

Inoue-san continues to be hired a lot because she has a voice quality that is ideal for gentle mother and teacher roles.¹⁴ (2channel 2013)

Some fan descriptions of Inoue's sweet voice performances draw a connection between Inoue's natural speaking voice and her professional sweet voice. While fans are aware that Inoue performs both sweet and nonsweet voice roles, they consider the sweet voice performances to be closer to her true character.

Kikuko's natural speaking voice is gentle itself, so voice acting gentle roles is not of particular challenge to her, this is the very reason I find her ability to voice act characters such as Panthar to be a display of her true skills. (CC 2002)

It is no accident that fans draw this association between Inoue and the roles she performs; Inoue has deliberately created a public persona for herself that evokes the qualities of the sweet voice characters she has portrayed. As noted in the Seiyuu Directory quote above, Inoue has adopted the nickname *onee-chan*, a familiar term for older sister. She also, with much tongue-in-cheek fanfare, claims to be only seventeen years old, although she has been a professional seiyuu for longer than that (Manbow 2008). Like her characters, Inoue wears her hair long and dresses conservatively, preferring long skirts and long sleeves. The voice she uses for public events and interviews is not the sweet voice, but it is extremely breathy; we might call it a real-world approximation of the sweet voice. Although in some respects Inoue's public persona is reminiscent of the cute Japanese pop idols described in Kinsella (1995) and Miller (2004), in the sense that she is creating an artificial character of a specific age and style in a calculated appeal to fans, the

aesthetic of this ‘sweet voice idol’ is entirely different from that of a cute pop idol. Inoue is crafting a fantasy of a traditional, conservative Japanese young woman, one who wants nothing more than to settle down and take care of you. As shown in Shingo (2005), while this fantasy is not exactly cutting edge, it has its own distinctive appeal for fans.

Kikuko-san has a unique sex appeal.¹⁵ (Iriyama 2011)

It seems I am likely to become a mature woman chaser because of Inoue Kikuko.¹⁶ (Katyosukatyosu 2011)

For her fans, Inoue represents the Yamato Nadeshiko ideal, and central to that ideal is her voice:

If I was able to design the perfect woman, for her voice it would be Kikuko Inoue’s voice. (Tronman63 2007)

Summary of fan perception. This examination of character-focused and seiyuu-focused fan discourse has demonstrated the prominent role of sweet voice in fan perceptions of Yamato Nadeshiko characters. These characters are older than the typically attractive female character type in anime, and their appeal lies in their sexual maturity and their embodiment of the ideal traditional Japanese woman who will care for you and yet be subservient to you. Fans of particular seiyuu show awareness both of the perceptual qualities of the seiyuu’s voice and of the nature of the characters they portray. The example of the seiyuu Inoue Kikuko illustrates how voice actresses can exploit sweet voice performances to create successful idol personae marketing a fantasy of the perfect woman.

A final look at anime: The male sweet voice

I have argued that the sweet voice in anime is closely intertwined with ideologies of traditional feminine beauty and the notion of the ideal Japanese woman. This does not, however, preclude the sweet voice from crossing gender lines. While it is far more rare than the female sweet voice, a male version of sweet voice does appear in certain settings. One notable use of male sweet voice in anime is in the series *Death note* (2006), in which the main character, Yagami Light (voiced by actor Mamoru Miyano), has a sweet voice. While most anime featuring male protagonists tell the tales of average young men who are marked for greatness, *Death note* tells the story of a boy who is far from average, and the use of sweet voice contributes crucially to viewers’ understanding of his character. From the moment Light’s voice is heard, the novelty of a male protagonist with a sweet voice suggests to viewers that something is not as it seems; Light is presented as the hero of the story, and yet, viewers are immediately made wary of him. Light’s sweet voice reflects several aspects of his character. First, as in the female sweet voice, the style communicates his physical attractiveness and charisma to viewers, letting them know that he is perceived by other characters as extraordinarily beautiful.

His voice also guides expectations about Light's sexuality; while the female sweet voice amplifies the traditionally feminine characteristics of a female character, a mismatch exists between Light's gender and his vocal style, suggesting that Light may not be heterosexual. The association between the Yamato Nadeshiko perfect woman and the sweet voice is also exploited here, for Light, as his name suggests, is indeed perfect in most respects, his fatal flaw being his intolerance of imperfection. In fact, Light turns out to be a sociopath and a gifted serial killer, albeit one with a beautiful voice.

Fans of Mamoru Miyano's voicing of Light point to features similar to those of the female version of sweet voice.

I think you know what I think of Light's voice. :DD But I'll say it anyway. MIYANO MAMORU
OMGH [sic] YOUR VOICE IS GORGEOUS. It's so smooth and sounds like it's coming near you.²¹

"Sounds like it's coming near you" may refer to the breathy quality prominent in Light's voice, which produces an effect reminiscent of an intimate whisper in your ear.

The popularity of Light's voice among fans brings up another intriguing issue: what happens to a sweet voice when it is dubbed into another language? In the case of English, at least, dubbers are faced with a dilemma; on the one hand, the sweet voice or anything resembling it would sound exaggerated and bizarre in English, but on the other hand, dubbing with a less stylized voice often results in a loss of characterization and emotional intensity. Dubbers generally opt for the latter of these two options, resulting in English versions with relatively lackluster vocal performances that compare poorly with the originals. (This is one of the reasons why non-Japanese-speaking fans prefer subtitles to dubbing.) Thus, for now, sweet voice remains a Japanese phenomenon, but one that is consumed by both Japanese and non-Japanese-speaking viewers.

CULTURAL IMPLICATIONS OF SWEET VOICE

Although some aspects of modern Japanese popular culture reject traditional notions and sensibilities, framing the Japanese cultural marketplace as a struggle between the opposing forces of the traditional and the popular, and by implication between the Eastern and the Western, would be a grave oversimplification. As we have seen in the case of sweet voice, traditional cultural models may be reinvented within the scope of popular culture and integrated into new systems of social meaning; Lafcadio Hearn's selfless Japanese woman may find herself transformed into an elegant transsexual attending a fashion design high school. Thus, the ancient is not replaced by the modern, but becomes modern in itself.

Crucially, while traditional conceptions of womanhood such as the Yamato Nadeshiko may appear to have endured unchanged from antiquity until today, changes in the systems of difference surrounding these conceptions have altered their significance. The Yamato Nadeshiko of the early Showa period, for instance, was a relatively positive model of women as powerful homemakers, in

comparison to the older notion of women as being ruled in the home as in all other spheres (Skov & Moeran 1995a:22). Today, the Yamato Nadeshiko appears subservient and unidimensional next to her more modern counterparts, but also more mature and alluring. In particular, we must understand the function of sweet voice in light of the culture of cuteness that has dominated Japanese popular culture to a greater or lesser extent since the 1980s (Kinsella 1995:220). Cuteness presents a model of femininity that is naïve and virtually asexual, rejecting the responsibilities of adulthood in favor of a romanticized and stylized artificial childhood. The cute voice, in turn, sounds high and immature, and the cute speaking style rejects Japanese women's language in favor of invented, childlike forms (Kinsella 1995:225). Presented in contrast with cuteness, then, the sweet voice and Yamato Nadeshiko character represent dramatically different values and aesthetic qualities. This is not meant to suggest that a competition of feminine ideals is taking place within Japanese popular culture in which one ideology must triumph over the other, but simply that these contrasting notions highlight each other's differences and appeal to the consumer, or the 'moetic observer', in different respects. Furthermore, the two femininities described here are joined by other contrasting notions, with some arising from popular media and others created by the consumers themselves. In sum, a multitude of femininities co-exist within the sphere of Japanese popular culture, and many of these manifest themselves in distinctive vocal styles.

This discussion has come a long way from Kristof's (1995) monolithic headline, 'Japan's feminine falsetto falls right out of favor'. As for the sweet voice, it shows no signs of falling out of favor, although the same cannot necessarily be said of Japanese Women's Language; if the media is the last bastion of JWL, and JWL is showing signs of losing cultural capital in the media, it is possible that, in a few decades, even the most spoiled of princesses will abandon her sentence-final particles.

CONCLUSION

This study has explored the role of voice quality in the creation of style. Voice quality has a unique role to play in style, because its status as a feature considered to be beyond the realm of conscious control makes it a source of authenticity. Thus, voice quality can license particular linguistic behaviors, such as the use of Japanese Women's Language, that would otherwise be seen as inauthentic. In the specific case of the sweet voice, a Japanese vocal style used by professional voice actresses rather than ordinary women, voice quality is a prominent factor in the evaluation of the characters and performers who make use of this style. Sweet voice contrasts with other feminine Japanese styles both in terms of its acoustic properties and due to its association with traditional notions of the ideal woman. Differences between sweet voice and other vocal styles remind us that there is more than one way of performing femininity in Japanese.

NOTES

*This article was initiated as a class project with my colleague Rebecca Dayle Greene. Thanks to participants at the NWAV and AAA conferences and the Stanford Style Retreat for their valuable feedback on previous versions. Thanks to Sakiko Kajino, Jane Stuart-Smith, Jenny Cheshire, and an anonymous reviewer for their comments. Special thanks to my Japanese informants and to the anime fan community whose tireless enthusiasm and fansubbing efforts made this project possible.

¹I adopt the use of 'sweet voice', *amai koe* in Japanese, to refer to this style as there does not appear to be a more established term in use. *Amai koe* is used in a general sense to describe a sweet or seductive-sounding speaking or singing voice; an anonymous reviewer points out that it is also sometimes used to particularly refer to a sexually excited voice (as heard in adult videos), which overlaps with the style under discussion here. The descriptors *kagayaku*, 'sparkling', and *kiyoraka*, 'pure', appear most often on voice-actor profiles to describe those who frequently employ this style, although the terms are not used consistently.

²One way in which the fame of *seiyuu* is comparable to that of conventional actors is that their pictures, personal information, and performances are catalogued and critiqued on numerous websites. A simple Google search indicates that there is greater online interest in particular *seiyuu* than in even the most famous American voice actors; Hayashibara Megumi, the voice of Hello Kitty, has 1,690,000 hits (using only her Japanese name), while Mel Blanc, the voice of Looney tunes, has 508,000.

³Many *seiyuu* talent agencies have associated training schools that serve as recruitment centers as well as rich sources of income. Ken Productions, for example, has an associated training school, School Duo, in which *seiyuu* hopefuls can train for one year for the chance to audition to make it to 'junior' level. According to their curriculum description, students receive training in vocal performance and pronunciation, and in specific genres including Western film dubbing, narration, anime, radio drama, etc. (Kenpro Groups 2008).

⁴Because there is no preexisting register of *seiyuu* who perform sweet voice, I selected the first ten *seiyuu* I observed using this vocal style. The roles selected for analysis were chosen based on the most recent performance by that *seiyuu* providing sufficient linguistic data, at the time of data collection.

⁵Interested readers may listen to a synthesized version of sweet voice in the Japanese audio of the Google Translate web service (translate.google.com). Other examples may be found in the films and programs referenced in the article.

⁶By 'articulatorily salient', I refer to my own replications of the sweet voice, in which it is evident that the vocal folds are tensed as in singing in the head voice register. In future work, articulatory studies of performers using tools such as ultrasound may provide more detailed insight into vocal tract settings in the production of sweet voice.

⁷Praat script identifies formants using LPC and harmonics via spectral slices created with a 3 ms Hamming window. Modified from 'Nasality measurement' script originally written by Rebecca Scarborough.

⁸F0 was calculated in VoiceSauce using the Snack algorithm.

⁹One of the suggestions I came across on a vocal coaching video on how to produce head register was the following: "I would like to throw in an idea, for any fans of Japanese Anime. Many characters in Anime speak far higher than in our Western culture and impersonating them may also help to improve head voice (and maybe your Japanese too)." (freethoughtmusic 2013, online: <http://www.youtube.com/watch?v=5SFVIKWAZYI>)

¹⁰Examples of spoiled-princess characters include Mishima Sayoko (*Aa! Megami-sama*, 2005), Saotome Miki (*Kidou shinsengumi moeyo Ken*, 2005), Miyuki Mayu (*Ai yori aoshi*, 2002), Kiryuu Nanami (*Shoujo kakumei utena*, 1997), and Kodachi Kuno (*Ranma ½*, 1989).

¹¹The lone exception in these data to the principle that characters cannot manipulate sweet voice is *Kamisama kazoku* (2006), in which a goddess is able to physically transform herself and turn on sweet voice to beguile her younger brother's friends.

¹²The present analysis of fan comments and discourse is based on searches for mentions of anime vocal styles and *seiyuu* who perform sweet voice styles on informational websites (e.g. Anime News

Network), blogs and message boards (e.g. 2channel, Hatena), and comments on seiyuu performances on YouTube.

¹³As in the case of anime itself, there are many different types of anime consumers and fans. The prototypical anime fans are the otaku, primarily Japanese, male, adolescent and adult fans who are often considered to be sexually obsessed with anime (Hu 2010). Aside from this group, however, there are many anime fans around the world who participate in fandom in various ways, from sharing their own subtitles, to creating art, to dressing up and attending anime conventions. Indeed, anime fandom has increasingly become globalized as anime has grown in popularity abroad while its domestic market has decreased (Bryce, Barber, Kelly, Kunwar, & Plumb 2010). The present analysis is therefore not restricted to Japanese fans.

¹⁴Translated from original Japanese: “井上さんは結構起用され続けてる。優しい母親や先生役にぴったりの声質だからなあ。”

¹⁵Translated from original Japanese: “喜久子さんって 独特の色気があるよね。”

¹⁶Translated from original Japanese: “井上喜久子のせいで熟女好きになってしまいそうだ。”

REFERENCES

- 2channel (2013). 皆口裕子(46)井上喜久子(17 + 31)林原めぐみ(46) [‘Minaguchi Yuko (46), Inoue Kikuko (17 + 31), Megumi Hayashibara(46)’]. Online: <http://www.logchannel.com/bbs/ogame/1368626634/>; accessed May 26, 2014.
- Balasubramaniam, Radish Kumar; Jayashree S. Bhat; Manav Srivastava; & Aimy Eldose (2012). Cepstral analysis of sexually appealing voice. *Journal of Voice* 26(4):412–15.
- Bishop, Jason, & Patricia Keating (2012). Perception of pitch location within a speaker’s range: Fundamental frequency, voice quality, and speaker sex. *Journal of the Acoustical Society of America* 132 (2):1100–12.
- Björkner, Eva (2006). Why so different? Aspects of voice characteristics in operatic and musical-theatre singing. Stockholm: KTH School of Computer Science and Communication dissertation.
- Boersma, Paul, & David Weenink (2012). Praat: Doing phonetics by computer. [Computer program] Version 5.3.14 (28 April 2012).
- Brinca, Lilia F.; Ana Paula Batista; Ana Ines Tavares; Ilidio Goncalves; & Maria L. Moreno (2014). Use of cepstral analyses for differentiating normal from dysphonic voices: A comparative study of connected speech versus sustained vowel in European Portuguese female speakers. *Journal of Voice* 28(3):282–86.
- Bryce, Mio; Christie Barber; James Kelly; Siris Kunwar; & Amy Plumb (2010). Manga and anime: Fluidity and hybridity in global imagery. *Electronic Journal of Contemporary Japanese Studies*. Article 1.
- Bucholtz, Mary (2003). Sociolinguistic nostalgia and the authentication of identity. *Journal of Sociolinguistics* 7(3):398–416.
- Bullock, Ben (ed.) (1994). What Is Yamato Nadeshiko? Online: <http://www.sljfaq.org/afaq/yamatonadeshiko.html>; accessed August 1, 2008.
- Campbell, Nick, & Parham Mokhtari (2003). Voice quality: The 4th prosodic dimension. *Proceedings of the 15th International Congress of Phonetic Sciences*, 2417–20.
- CC. 2002. Seiyuu directory. Online: http://www.geocities.com/seiyuu_directory/kikuko.html; accessed August 15, 2008.
- Clammer, John (1995). Consuming bodies: Constructing and representing the female body in contemporary Japanese print media. In Skov & Moeran 1995a, 197–219.
- Cryar, Morgan (2014). Head voice and falsetto. Become a singing master. Online: <http://www.become-a-singing-master.com/head-voice-and-falsetto.html>; accessed June 1, 2014.
- Dark Mirage (2008). Zettai Ryouiki. Online: <http://tvtropes.org/pmwiki/pmwiki.php/Main/ZettaiRyouiki>; accessed August 15, 2008.

- de Krom, Guus (1993). A cepstrum-based technique for determining a harmonic-to-noise ratio in speech signals. *Journal of Speech and Hearing Research* 36:254–66.
- Enyo, Yumiko (2005). Gender performance and intonation in a Japanese sentence-final particle yo ne. Paper presented at NWAV 2005, New York.
- Esling, John H. (2012). Voice quality. In Carole A. Chappelle (ed.), *The encyclopedia of applied linguistics*. Blackwell. doi: 10.1002/9781405198431.wbeal1271.
- Gagné, Isaac (2008). Urban princesses: Performance and women's language in Japan's gothic/lolita subculture. *Journal of Linguistic Anthropology* 18(1):130–50.
- Garrett, Rachel (2013). Cepstral- and spectral-based acoustic measures of normal voices. Theses and Dissertations. University of Wisconsin-Milwaukee. Paper 217.
- Gordon, Matthew, & Peter Ladefoged (2001). Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29:383–406.
- Haikyou (2008). Ouhara Sayaka Profile. Online: <http://haikyo.or.jp/PROFILE/woman/11570.html>; accessed September 1, 2008.
- Hearn, Lafcadio (1905). *Japan: An attempt at interpretation*. New York: Macmillan.
- Heman-Ackah, Yolanda D.; Reinhardt J. Heuer; Deirdre D. Michael; Rosemary Ostrowski; Michelle Horman; Margaret Baroody; James Hillenbrand; & Robert T. Sataloff (2003). Cepstral Peak Prominence: A more reliable measure of dysphonia. *Annals of Otolaryngology, Rhinology and Laryngology* 112(4):324–33.
- Hillenbrand, James; Ronald Cleveland; & Robert Erickson (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech and Hearing Research* 37:769–78.
- Hirano, Minoru; William Vennard; & John Ohala (1970). Regulation of register, pitch, and intensity of voice: An electromyographic investigation of intrinsic laryngeal muscles. *Folia Phoniatrica* 22:1–20.
- Honorof, Douglas N., & D. H. Whalen (2004). Perception of pitch location within a speaker's F0 range. *Journal of the Acoustical Society of America* 117(4):2193–200.
- Hu, Tze-Yue G. (2010). *Frames of anime: Culture and image-building*. Hong Kong: Hong Kong University Press.
- Inoue, Miyako (2006). *Vicarious language: Gender and linguistic modernity in Japan*. Berkeley: University of California Press.
- Iriyama, Sho (2011). Comment on 水樹奈々、福圓美里、こやまきみこ、井上喜久子 釘宮理恵 [‘Nana Mizuki, Misato Fukuen, Kimiko Koyama, Inoue Kikuko, Kugimiya Rie’]. Online: http://www.youtube.com/watch?v=vcpVGk_f5eM; accessed May 26, 2014.
- Iseli, Markus; Yen-Liang Shue; & Abeer Alwan (2007). Age, sex, and vowel dependencies of acoustic measures related to the voice source. *The Journal of the Acoustical Society of America* 121(4):2283–95.
- Ishi, Carlos Toshinori (2004). A new acoustic measure for aspiration noise detection. *Proceedings of INTERSPEECH* 2004, 941–94.
- Ito, Mika (2003). The contribution of voice quality to politeness in Japanese. *Proceedings of VOQUAL '03*, Geneva, 157–62.
- (2004). Politeness and voice quality: The alternative method to measure aspiration noise. *Proceedings of Speech Prosody* 2004, 213–16.
- Jahn, Anthony (2011). Talk better, sing better. Voice Council Magazine, August 19, 2011. Online: <http://voicecouncil.com/talk-better-sing-better>; accessed June 1, 2014.
- Johnson, Daniel Ezra (2009). Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistic Compass* 3(1):359–83.
- Katysukatyosu (2011). Comment on 水樹奈々、福圓美里、こやまきみこ、井上喜久子 釘宮理恵 [‘Nana Mizuki, Misato Fukuen, Kimiko Koyama, Inoue Kikuko, Kugimiya Rie’]. Online: http://www.youtube.com/watch?v=vcpVGk_f5eM; accessed May 26, 2014.
- Keating, Patricia A. (2014). Acoustic measures of falsetto voice. Paper presented at the annual meeting of the Acoustical Society of America. Providence, RI.
- Kenpro Groups (2008). School duo. Online: <http://www.school-duo.com/>; accessed September 1, 2008.
- Kinsella, Sharon (1995). Cuties in Japan. In Skov & Moeran 1995a, 220–54.

- Kong, Eun Jong; Yoneyama Kiyoko; & Mary Beckman (2014). Effects of a sound change in progress on gender-marking cues in Japanese. Poster presented at LabPhon 2014, Tokyo, Japan.
- Kreiman, Jody; Yen-Liang Shue; Gang Chen; Markus Iselt; Bruce R. Gerratt; Jürgen Neubauer; & Åber Alwan (2012). Variability in the relationships among voice quality, harmonic amplitudes, open quotient, and glottal area waveform shape in sustained phonation. *The Journal of the Acoustical Society of America* 132(4):2625–32.
- Kristof, Nicholas (1995). Japan's feminine falsetto falls right out of favor. *New York Times*, December 13 1995.
- Laver, John (1980). *The phonetic description of voice quality*. Cambridge: Cambridge University Press.
- Leino, T. (1993). Long-term average spectrum study on speaking voice quality in male actors. Proceedings of the Stockholm Music Acoustics Conference, 206–10. Stockholm, Sweden.
- Lim, Emily; Dona Jayakody; & Valerie Looi (2009). The singing power ratio and timbre-related acoustic analysis of singing vowels and musical instruments. Voice Foundation's 38th Annual Symposium: Care of the Professional Voice, 3–7 Jun 2009, Philadelphia, PA.
- Mambow (2008). Inoue Kikuko's official site. Online: <http://www.mambow.com>; accessed August 1, 2008.
- Martyn, Youri; Nelson Roy; Marc de Bodt; Paul van Cauwenberge; & Paul Corthals (2009). Acoustic measurement of overall voice quality: A meta-analysis. *Journal of the Acoustical Society of America* 126(5):2619–34.
- Masataka, Nobuo (1992). Pitch characteristics of Japanese maternal speech to infants. *Journal of Child Language* 19(2):13–23.
- Master, Suelly; Noemi de Biaz; Maria Chari Biaz; & Anne-Maria Laukkanen (2008). Acoustic and perceptual analysis of Brazilian male actors' and nonactors' voices: Long-term average spectrum and the actor's formant? *Journal of Voice* 22(2):146–54.
- ; Noemi de Biaz; & Sandra Madureira (2012). What about the 'actor's formant' in actresses' voices? *Journal of Voice* 26(3):117–22.
- Miller, Laura (2004). You are doing burikko! Censoring/screening artificers of cute femininity in Japanese. In Shigeo Okamoto & Janet Shibamoto Smith (eds.), *Japanese language, gender, and ideology: Cultural models and real people*, 148–65. New York: Oxford University Press.
- Miller, Richard (2000). *Training soprano voices*. New York: Oxford.
- Ohara, Yumiko (1999). Performing gender through voice pitch: A cross-cultural analysis of Japanese and American English. In Ursula Pasero & Friedrike Braun (eds.), *Wahrnehmung und Herstellung von Geschlecht* ['Perceiving and performing gender'], 105–66. Wiesbaden: VS Verlag für Sozialwissenschaften.
- (2004). Prosody and gender in workplace interaction. In Shigeo Okamoto & Janet Shibamoto Smith (eds.), *Japanese language, gender, and ideology: Cultural models and real people*, 222–40. New York: Oxford University Press.
- Okamoto, Shigeo, & Janet Shibamoto Smith (2004). Introduction. In Shigeo Okamoto & Janet Shibamoto Smith (eds.), *Japanese language, gender, and ideology: Cultural models and real people*, 3–20. New York: Oxford University Press.
- , & Janet Shibamoto Smith (2008). Constructing linguistic femininity in contemporary Japan: Scholarly and popular representations. *Gender and Language* 2(1):87–112.
- Osuma, Kotaka (1929). 日常常識儀作法 ['Etiquette and manners as daily knowledge']. Tokyo: Okamura Shoten.
- Podsva, Robert (2007). Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics* 14(4):478–504.
- Sadanobu, Toshiyuki (2004). A natural history of Japanese pressed voice. *Journal of the Phonetic Society of Japan* 8(1):29–44.
- Samlan, Robin A., & Brad H. Story (2011). Relation of structural and vibratory kinematics of the vocal folds to two acoustic measures of breathy voice based on computational modeling. *Journal of Speech, Language, and Hearing Research* 54:1267–83.

- Seiyuu Directory (2002). Inoue Kikuko. Online: http://www.geocities.com/seiyuu_directory/kikuko.html; accessed August 1, 2008.
- Shibamoto Smith, Janet (2003). Gendered structures in Japanese. In Marlis Hellinger & Hadumod Bußmann (eds.), *Gender across languages: The linguistic representation of women and men*, 201–26. Philadelphia, PA: John Benjamins.
- Shingo (2005). A discourse on moe, part I: Introduction and a moetic taxonomy. Heisei Democracy. Online: <http://heiseidemocracy.com/2005/02/28/62/>; accessed October 1, 2006.
- Shue, Yen-Liang (2009). Voicesauce: A program for voice analysis. Online: <http://www.seas.ucla.edu/spapl/voicesauce/>.
- ; Gang Chen; & Abeer Alwan (2010). On the interdependencies between voice quality, glottal gaps, and voice-source related acoustic measures. Proceedings of *INTERSPEECH* 2010, 34–37.
- Skov, Lise, & Brian Moeran (1995a). *Women, media, and consumption in Japan*. Honolulu: University of Hawai'i Press.
- , & Brian Moeran (1995b). Introduction: Hiding in the light: from Oshin to Yoshimoto Banana. In Skov & Moeran 1995a, 1–74.
- Stevens, Kenneth, & Helen M. Hanson (1995). Classification of glottal vibration from acoustic measurements. In Osamu Fujimura & Minoru Hirano (eds.), *Vocal fold physiology: Voice quality control*, 147–70. San Diego, CA: Singular Publishing Group.
- Sugihara, Yoko, & Emiko Katsurada (1999). Masculinity and femininity in Japanese culture: A pilot study. *Sex Roles* 40(7–8):635–46.
- Sundberg, Johan (1987). *The science of the singing voice*. Dekalb: Northern Illinois University Press.
- (2001). Level and centre frequency of the singer's formant. *Journal of Voice* 15(2):176–86.
- , & Anders Askenfelt (1983). Larynx height and voice source: A relationship? In Diane M. Bless & James H. Abbs (eds.), *Vocal fold physiology: Contemporary research and clinical issues*, 307–16. San Diego, CA: Collegehill.
- Superflavor (2007). Superflavor LiveJournal Jan 20, 2007. Online: <http://superflavor.livejournal.com/>; accessed August 1, 2008.
- Stewarts, Marc, & Raymond Veldhuis (2001). The effect of speech melody on voice quality. *Speech Communication* 33(4):297–303.
- Tabuchi, Hiroko (2013). Japan's top voice: High, polite, and on the phone. *New York Times*, December 14, 2013. A1.
- Takahara, Kumiko (1991). Female speech patterns in Japanese. *International Journal of the Sociology of Language* 92:61–85.
- Teshigawara, Mihoko (2003). *Voices in Japanese animation: A phonetic study of vocal stereotypes of heroes and villains in Japanese culture*. Victoria: University of Victoria dissertation.
- Tronman63 (2007). Kikuko Inoue 17yo & Ami Koshimizu short clip. Online: http://www.youtube.com/watch?v=XssN_nFzLRs; accessed August 15, 2008.
- TV.com (2008). Kikuko Inoue bio. Online: <http://www.tv.com/kikuko-inoue/person/116271/biography.html>; accessed August 15, 2008.
- Warhurst, Samantha; Patricia McCabe; Edwin Yiu; Robert Heard; & Catherine Madill (2013). Acoustic characteristics of male commercial and public radio broadcast voices. *Journal of Voice* 27(5):655.e1–655.e7.
- Washi, Rumi (2004). 'Japanese female speech' and language policy in the World War II era. In Shikego Okamoto & Janet Shibamoto Smith (eds.), *Japanese language, gender, and ideology: Cultural models and real people*, 76–91. New York: Oxford University Press.

APPENDIX A: ANIME TELEVISION PROGRAMS
AND FILMS REFERENCED

- Aa! Megami-sama* ['Ah! My goddess'] (2005). Produced by AIC.
- Ai yori aoshi* ['Bluer than indigo'] (2002). Produced by J. C. Staff.
- Binbou shimai monogatari* ['Story of poor sisters'] (2006). Produced by Toei Animation.
- Bleach* (2004). Produced by Studio Pierrot.
- Death note* (2006). Produced by Madhouse.
- Gokujou seitokai* ['Best student council'] (2005). Produced by J. C. Staff.
- Gun X sword* (2005). Produced by AIC.
- Hachimitsu to kuroubaa* ['Honey and clover'] (2005). Produced by J. C. Staff.
- Kamisama kazoku* ['God family'] (2006). Produced by Toei Animation.
- Kaze no tani no Nausicaä* ['Nausicaä of the valley of the wind'] (1984). Produced by Takahata Isao.
- Kidou shinsengumi moeyo Ken* ['Moeyo Ken'] (2005). Produced by Trinet Entertainment.
- Mai HiME* ['My HiME'] (2004). Produced by Sunrise.
- Mai Otome* ['My Otome'] (2005). Produced by Sunrise.
- One piece* (1999). Produced by Toei Animation.
- Paradise kiss* (2005). Produced by Madhouse.
- Petopeto-san* (2005). Produced by Enterbrain Inc., Genco, and Media Factory.
- Ranma ½* (1989). Produced by Kitty Films and Studio Deen.
- Shoujo kakumei utena* ['Revolutionary girl utena'] (1997). Produced by J. C. Staff.
- Tonari no Totoro*. ['My neighbor Totoro'] (1988). Produced by Studio Ghibli.
- Uchuu no Stellvia* ['Stellvia of the universe'] (2003). Produced by XEBEC.

SWEET VOICE

APPENDIX B: MIXED EFFECTS MODELS OF ACOUSTIC MEASURES IN RBRUL

INTER-SPEAKER ANALYSIS:

Log F0

Factor	Coef.	p
speaker [rand]		
phon (breathy)	.002	.00206

REML dev REML AIC df intercept grand mean
 -38599.65 -38591.65 4 2.424 2.414

CPP

Factor	Coef.	p
speaker [rand]		
phon (breathy)	-2.265	<.0001
timepoint (start)	-1.296	<.0001
log F0	-1.098	<.0001

REML dev REML AIC df intercept grand mean
 113704.5 113718.5 7 21.211 19.441

H1-H2

Factor	Coef.	p
speaker [rand]		
log F0	16.76	<.0001
phon (breathy)	1.399	.000262

REML dev REML AIC df intercept grand mean
 4819.729 4829.729 5 -37.977 2.85

HNR35

Factor	Coef.	p
speaker [rand]		
phon (breathy)	-3.165	<.0001
timepoint (start)	-1.918	<.0001
log F0	-4.562	<.0001

REML dev REML AIC df intercept grand mean
 143409.8 143423.8 7 35.078 24.678

H1-A1

Factor	Coef.	p
speaker [rand]		
phon (breathy)	2.139	<.0001
log F0	15.248	<.0001

REML dev REML AIC df intercept grand mean
 4628.982 4638.982 5 -38.126 -1.445

2k-4k

Factor	Coef.	p
speaker [rand]		
phon (breathy)	0.894	.000834

REML dev REML AIC df intercept grand mean
 4358.973 4366.973 4 13.02 13.094

H1-A3

Factor	Coef.	p
speaker [rand]		
log F0	29.495	<.0001
phon (breathy)	1.722	<.0001

REML dev REML AIC df intercept grand mean
 4910.533 4920.533 5 -61.678 10.748

INTRA-SPEAKER ANALYSIS:

Log F0

Factor	Coef.	p
speaker [rand]		
phon (breathy)	-.016	<.0001
style (sweet)	.014	<.0001

REML dev REML AIC df intercept grand mean
 -34900.21 -34890.21 5 2.415 2.422

CPP

Factor	Coef.	p
speaker [rand]		
phon (breathy)	-2.358	<.0001
timepoint (start)	-1.546	<.0001
log F0	2.912	<.0001

REML dev REML AIC df intercept grand mean
 118709.6 118723.6 7 11.829 19.823

H1-H2

Factor	Coef.	p
speaker [rand]		
phon (breathy)	3.467	<.0001
style (sweet)	3.39	<.0001

REML dev REML AIC df intercept grand mean
 8433.869 8443.869 5 -0.302 -3.748

HNR35

Factor	Coef.	p
speaker [rand]		
phon (breathy)	-2.345	<.0001
style (sweet)	1.97	<.0001
timepoint (start)	-1.614	<.0001
log F0	-3.628	<.0001

REML dev REML AIC df intercept grand mean
 241680.5 241696.5 8 32.083 24.36

H1-A1

Factor	Coef.	p
speaker [rand]		
phon (breathy)	4.483	<.0001
style (sweet)	2.841	<.0001
log F0	12.737	<.0001

REML dev REML AIC df intercept grand mean
 8380.147 8392.147 6 -34.344 -7.36

2k-4k

Factor	Coef.	p
speaker [rand]		
style (sweet)	-1.089	<.0001
phon (breathy)	0.984	<.0001
log F0	-6.634	<.0001

REML dev REML AIC df intercept grand mean
 7136.577 7148.577 6 29.267 12.475

H1-A3

Factor	Coef.	p
speaker [rand]		
phon (breathy)	4.034	<.0001
log F0	13.395	<.0001
style (sweet)	1.521	.00323

REML dev REML AIC df intercept grand mean
 8738.356 8750.356 6 -24.046 5.274

(Received 1 November 2013; revision received 9 June 2014; accepted 23 June 2014; final revision received 5 August 2014)