VOICES IN JAPANESE ANIMATION: HOW PEOPLE PERCEIVE THE VOICES OF GOOD GUYS AND BAD GUYS

Mihoko Teshigawara

Department of Linguistics
University of Victoria, B.C., Canada

1. INTRODUCTION

Japanese *anime*, an animation medium which depicts the world as inhabited by good and bad characters, is wildly popular in Japan and other parts of the world. The few scholarly studies that have considered this medium are still in the development stage (Lent, 2001). The present study examines the voices of characters in Japanese *anime*, focusing on the articulatory and perceptual characteristics of the voices of male and female heroes and villains.

Vocal stereotyping plays an important role in animation: voices need to reflect the physical attributes and personality traits of characters and the vocal stereotypes that consumers, filmmakers, and voice actors share. Previous studies on vocal stereotypes (Zuckerman & Miyake, 1993) reveal that people infer similar personality traits from voices. However, few studies have investigated the acoustic correlates of personality in speech (Aronovitch, 1976; Zuckerman & Miyake, 1993), and few have investigated the auditory correlates identified by phoneticians and the correspondence between auditory correlates and laypersons' perceptions (Biemans, 1998; van Bezooijen, 1988). This study will attempt to fill the gap between these two types of study, using speech samples from Japanese *anime*.

Prior to the present paper, the author identified auditory characteristics of the voices of 88 characters (44 heroes, 42 villains, and two supporting roles) from 20 animated cartoons, using Laver's (1994, 2000) framework for voice quality description. The following summarizes the auditory characteristics identified across categories (for more details, see Teshigawara, 2003, in press):

1. Heroes' voices exhibited an absence of pharyngeal constriction and the presence of breathy voice.
2. The majority of villains' voices exhibited pharyngeal constriction and harsh voice caused by tense laryngeal tension settings; however, pharyngeal expansion accompanied by lowered larynx was observed in a majority of female and some male villains.

In order to investigate whether the identified auditory characteristics contribute to people's perceptions of good and bad characters, Japanese laypersons' perceptions of selected speech samples were examined in an experimental setting. Speech samples were selected based on the auditory analysis results regarding the perceived pharyngeal states and phonatory settings from a separate study (Teshigawara, 2003, in press). In addition to prototypical heroes and villains, heroes that exhibited pharyngeal constriction and harsh voice and villains without these characteristics were included in order to specifically examine the roles of these auditory characteristics in vocal stereotyping. Participants were asked to rate their impressions of cartoon characters using trait items in the following four categories: physical traits, personality traits, emotional states, and vocal characteristics. It was hypothesized that participants would attribute less favorable physical traits, personality traits, emotional states, and vocal characteristics to speakers who exhibited pharyngeal constriction/expansion no matter which roles they played in the original cartoons. Statistical analyses were performed in order to examine the relationship between the auditory characteristics of the voices and participants' trait ratings of the speakers.

2. METHOD

2.1 Stimuli

In light of the auditory characteristics summarized above, the 88 character voices were divided into two groups, representative and non-representative: representative meaning that characters exhibited auditory characteristics appropriate to their role, and non-representative meaning that characters exhibited auditory characteristics opposite to or simply atypical of their role. Within these two groups, characters were examined according to role, gender and age (adult vs. child). For example, villains showing either pharyngeal constriction or

expansion were categorized into the representative villain group, while those showing neither trait fell into the non-representative villain group. There were 16 possible groups: hero or villain (2) × gender (2) × age (2) × representativeness (2). However, since there was only one child villain (male) in the corpus, this classification system yielded only 13 groups. Two speakers were chosen for each of 12 groups, with the exception of the child male villain group, which had only one speaker. In addition, two supporting roles (one male child and one female adult) exhibiting the auditory characteristics of villains' voices, that is, pharyngeal constriction and harsh voice, were added in order to see whether they would be rated similarly to heroes or villains. Therefore, the voices of 27 speakers in total were chosen as the basis for experimental stimuli.

Noise-free speech samples of these 27 speakers had been stored on a personal computer for auditory and acoustic analyses for a separate study (Teshigawara, 2003, in press). They had been recorded from VHS tapes of the cartoons. In order to elicit listeners' responses to the voices independent of verbal content, it was necessary to mask the contents of the speech samples. Of the five content-masking techniques investigated by Scherer, Feldstein, Bond, and Rosenthal (1985), random splicing was chosen since it retains voice quality information (the focus of this study), while the other techniques mask it. First, in order to create stimuli representative of each speaker, speech portions produced with a voice quality setting deviating from the speaker's normal setting were removed, with the exception of characters who were consistently angry or shouting. Intensities were standardized across speakers so that the maximum intensity was between 70 and 72 dB. Following previous research using the random-splicing technique (Friend & Farrar, 1994; Scherer, et al., 1985; van Bezooijen & Boves, 1986), after removing pauses, the digitized speech samples were divided into 250-ms segments. The first and last 3 ms of each segment were linearly attenuated to zero amplitude in order to avoid the introduction of transients (Friend & Farrar, 1994). In order to create a 5-s stimulus for each speaker, 20 250-ms segments were prepared and rearranged so that segments could not occur in the same relative order in the spliced stimulus as in the original.

In order to counterbalance the effects of ordering, two stimulus orders (A and B) were used: in A, the 27 speakers were randomly ordered disregarding the speaker groups, while B was the reverse of A. For each speaker, the speaker number was announced followed by the 5-s stimulus; after 1 s of silence, the same segment was repeated, followed by 70 s of silence. This gave participants a total of 81 s to rate each speaker, which, according to previous studies (Kido & Kasuya, 2001; van Bezooijen, 1995), is considered sufficient to rate the 21 trait items selected in this experiment. Participants were given a practice session in which they rated an additional three speakers before rating the 27 target speakers.

2.2 Procedures and Participants

Twenty-one trait items were initially chosen to be used in the questionnaire for the rating session. English translations are given for the items as follows: gender (female or male); age group (0–10; 11–18; 19–35; 36–60; over 61); physical characteristics ("big," "good-looking"); personality traits, 11 in total, of which three were chosen for their pertinence to heroes of Japanese *anime* in particular ("selfless," "loyal," "devoted") (Levi, 1998), three were thought to be universal characteristics of heroes ("brave," "intelligent," "strong"), and five represented each of the five factors in the NEO Personality Inventory ("sociable," "calm," "curious," "conscientious," "sympathetic") (McCrae & Costa, 1987); emotional states ("positive emotion"); and vocal characteristics ("high-pitched," "loud," "relaxed," "pleasant," "attractive"). The 19 adjectives were rated on 7-point scales, from 1 (*not at all true*) to 7 (*extremely true*). In the questionnaire, all 19 adjective items were grouped according to the category and were never mixed disregarding categories. The remaining two multiple-choice items, gender and age were considered to be categories on their own, yielding six categories all together. In order to counterbalance ordering effects of trait categories and items within categories, two orders of the six trait categories – (I) gender, age, physical characteristics, personality, emotion, vocal characteristics; (II) age, gender, vocal characteristics, emotion, personality, physical characteristics – and, where applicable, two *item orders* within trait categories (a, b) were prepared. These were systematically combined yielding four questionnaire types (i.e., Ia, Ib, IIa, IIb).

Thirty-two participants (15 males, 17 females; average age 22.8 years old) were recruited from Nagoya University, Japan and the vicinity. In total, eight experimental conditions were yielded, combining the two stimulus orders (A and B) and the four questionnaire types (Ia, Ib, IIa, IIb). Four participants were assigned to each experimental condition, with the exception of the two groups that used Questionnaire Ia, in which five participants listened to stimulus order A and three listened to B.

Experimental sessions were run in groups of up to seven in a soundproof room in the School of Letters building at Nagoya University. Using a CD player, the experimenter played a CD containing instructions recorded by the author, a practice session and the 27 target stimuli. The same instructions were given in the questionnaire booklet as well. Participants were told that they would hear two 5-s content-masked cartoon speech excerpts for each speaker, and they were asked to rate impressions of the speakers' traits on 7-point scales and choose appropriate groups for gender and age. After the experiment, participants completed a questionnaire including demographic information about themselves and their exposure to *anime*. Each session lasted less than one hour.

3. RESULTS AND DISCUSSION

3.1 Reliability

In order to examine the consistency of participants' trait ratings, Cronbach's alpha was calculated for each of the 21 trait items across participants. SPSS version 11.5 was used for the statistical analyses performed in this paper. As shown in Table 1, the Cronbach's alphas were very high, ranging between .90 and .99 for all but two items (.87 for "sociable" and .80 for "positive emotion"). Intraclass correlation, which measures the reliability between any two pairs of judges (participants in the case of the present study) as opposed to the aggregate reliability of all the judges measured by Cronbach's alpha (Rosenthal & Rosnow, as cited in Hecht & LaFrance, 1995), was also calculated for each item across participants. As was the case with Hecht and LaFrance (1995), the intraclass reliabilities were lower (r = .11 to .82) than Cronbach's alphas, which indicates that there was considerably more variability in the judgments of trait items at the individual level.

Table 1
Reliability of Ratings

Trait item	Cronbach's alpha	Intraclass correlation
Gender	0.99	0.82
Age	0.98	0.63
Physical characteristics		
Big	0.97	0.52
Good-looking	0.98	0.56
Personality traits		
Brave	0.90	0.22
Selfless	0.94	0.33
Loyal	0.94	0.33
Devoted	0.93	0.30
Intelligent	0.94	0.31
Strong	0.92	0.26
Sociable	0.87	0.17
Calm	0.93	0.29
Curious	0.94	0.32
Conscientious	0.93	0.31
Sympathetic	0.95	0.39
Emotional state		
Positive	0.80	0.11
Vocal characteristics		
High-pitched	0.98	0.59
Loud	0.92	0.27
Relaxed	0.92	0.27
Pleasant	0.95	0.35
Attractive	0.95	0.35

Note. Figures are based on all 32 participants.

In addition, since an anomaly was found in a close inspection of the raw data, bivariate correlations were calculated for each pair of participants and for each participant relative to the average ratings of all participants, separately for each item. It was found that ratings of one of the participants in the condition group that had stimulus order A and questionnaire Ia had significant correlations with the average ratings for only six out of 21 items (p.

< .05). According to the experimenter, this particular participant complained of fatigue and boredom after the experimental session; therefore, it was decided that results for this participant would be removed from the rest of the analysis.

Due to the asymmetry of the stimuli in child villain groups, the following analyses are confined to the adult hero and villain groups, which have full contrasts in gender and representativeness–16 speakers, two from each of the eight groups, i.e., hero or villain (2) × gender (2) × representativeness (2). (Supporting roles are not discussed in this paper due to space limitations.) In addition, among the 21 trait items, five items (gender, age, "big" from physical traits, and "high-pitched" and "loud" from vocal characteristics) were removed from the present analysis because it appeared that factors other than pharyngeal states were decisive in influencing participants' impressions of these factors. (See Teshigawara, 2003, for an analysis of the omitted speakers and items.)

Table 2
Means of Selected 16 Items by Speaker Group

	Hero				Villain			
	Male		Female		Male		Female	
Items	Rep	Non-rep	Rep	Non-rep	Rep	Non-rep	Rep	Non-rep
Physical characteristic								
Good-looking	5.45	3.98	5.21	5.02	3.27	4.26	2.77	4.74
Personality traits								
Brave	5.06	4.77	4.13	4.45	5.24	4.48	3.74	3.65
Selfless	4.55	3.85	4.81	5.08	4.13	3.89	3.24	4.53
Loyal	5.03	4.23	4.90	4.95	4.37	4.31	3.31	4.27
Devoted	4.71	4.60	5.44	5.32	4.82	4.08	3.94	5.10
Intelligent	4.76	3.87	4.65	4.65	4.23	5.00	3.77	3.90
Strong	5.06	4.77	3.66	4.39	5.73	4.40	4.69	3.73
Sociable	4.29	4.19	4.05	4.65	3.48	3.44	3.23	4.69
Calm	4.48	3.45	3.90	3.61	3.27	4.40	2.66	3.71
Curious	4.11	4.44	4.10	3.85	3.18	3.02	3.42	4.23
Conscientious	4.90	3.98	4.98	5.21	4.98	4.95	3.63	4.66
Sympathetic	4.94	4.05	4.85	5.06	3.73	3.69	3.06	4.52
Emotional state								
Positive	4.13	3.82	3.39	3.00	2.87	3.76	2.60	3.45
Vocal characteristics								
Relaxed	4.53	3.85	3.08	2.89	2.90	4.03	2.79	3.13
Pleasant	4.77	3.81	4.08	3.81	3.18	3.73	2.55	3.68
Attractive	4.84	3.98	4.53	4.29	3.61	3.87	2.74	4.08

Note. Rep stands for representative groups and Non-rep for non-representative groups. $n = 2$ for each group.

3.3 Analyses of Variance: Heroes Versus Villains

In order to examine whether participants responded to stimuli according to the auditory characteristics of the voices, that is, in reference to differences in pharyngeal state, a series of three-factor repeated measures ANOVAs with three between-subjects factors for control purposes was carried out for each item. The three within-subjects factors were role (hero or villain), gender (male or female), and representativeness (representative or non-representative). The three between-subjects factors were the two stimulus orders and the two category and item orders in the questionnaire. For the purpose of Type I error protection, the Bonferroni correction was used; therefore, an alpha level .003 (i.e., .05 divided by 16) was adopted instead of the standard .05 level. Table 3 summarizes the significant main and interaction effects that emerged in the analyses. It also includes partial eta squared (η^2) as an estimated effect size for each main and interaction effect.

As can be seen in Table 3, the main effect of role was significant for 12 of the 16 items. Therefore, it can be

said that the voices of heroes were generally perceived as having more favorable physical and personality traits, emotional states, and vocal characteristics than those of villains. Main effects for the factors of gender and representativeness did not emerge in as many items as for role (six main effects for gender, and three for representativeness). This result seems reasonable, given that it was not hypothesized that pharyngeal states would differentiate the two genders. In addition, representativeness conveys different pharyngeal states depending on role; for instance, non-representative characters include both heroes exhibiting laryngeal sphincteering and villains exhibiting an open airway. In other words, representativeness is meaningful only in relation to role.

A number of interactions were found between any combination of two of the three factors and among the three. Except for "strong" and "curious," the same patterns were observed in the direction of interactions across items. In the interaction between role and gender, females were rated significantly higher when they were heroes than villains, which was the expected tendency; however, males' ratings were generally consistent across roles or significantly higher for villains than heroes (Figure 1). Six items showed this tendency and they are all personality trait items: "selfless," "loyal," "devoted," "intelligent," and "conscientious." (For "curious," the direction of the interaction was reversed; male heroes were rated significantly higher than male villains, while female heroes were rated significantly lower than female villains.) This tendency was not predicted; rather, it was expected that both male and female villains would receive low scores.

Table 3

Results From Analyses of Variance of Participants' Trait Ratings for Adult Heroes and Villains

Source	df	Trait items: Personality: Good-looking	Brave	Selfless	Loyal	Devoted	Intelligent	Strong
Role (H/V)	1							
F		81.09**	7.33	44.63**	32.93**	26.99**	5.56	1.87
η^2		.78	.24	.66	.59	.54	.20	.08
MSE	23	.99	.85	.52	.95	.62	.71	.91
Gender (G)	1							
F		3.94	68.46**	9.13	1.71	10.78	4.60	56.99**
η^2		.15	.75	.28	.07	.32	0.17	.71
MSE	23	.61	.74	.64	.55	.91	.52	.85
Representativeness (R)	1							
F		12.71*	4.36	2.60	0.10	0.47	0.01	23.66**
η^2		.36	.16	.10	.00	.02	.00	.51
MSE	23	.53	.63	.62	.75	.35	1.00	.58
H/V*G	1							
F		3.55	5.76	17.87**	19.12**	15.06*	29.87**	0.00
η^2		.13	.20	.44	.45	.40	.57	.00
MSE	23	.58	.89	.63	.60	.45	.60	.68
H/V*R	1							
F		87.58**	3.83	16.24*	22.28**	4.87	20.38**	32.88**
η^2		.79	.14	.41	.49	.18	.47	.59
MSE	23	.58	.70	.52	.49	.38	.62	.85
G*R	1							
F		48.39**	12.98*	27.56**	19.11**	36.27**	0.21	12.99*
η^2		.68	.36	.55	.45	.61	.01	.36
MSE	23	.39	.54	.89	.73	.40	.70	.63
H/V*G*R	1							
F		0.62	0.02	2.29	0.29	33.79**	16.60**	2.77
η^2		.03	.00	.09	.01	.60	.42	.11
MSE	23	.49	.54	.57	.60	.43	.57	.58
Others[a]						R*QO2		
F						15.08*		
η^2						.40		

Table 3 (Continued)

Source	df	Trait items: Personality: Sociable	Calm	Curious	Conscientious	Sympathetic	Emotion: Positive	Vocal characteristics: Relaxed	Pleasant	Attractive
H/V	1									
F		16.38*	15.44*	28.04**	4.38	95.69**	30.04**	14.75*	67.39**	46.21**
η^2		.42	.40	.55	.16	.81	.57	.39	.75	.67
MSE	23	1.27	.52	.99	.70	.60	.34	.55	.61	.90
G	1									
F		10.31	18.06**	3.84	0.77	6.00	42.65**	72.83**	24.19**	2.14
η^2		.31	.44	.14	.03	.21	.65	.76	.51	.09
MSE	23	.56	.57	.72	.46	.86	.40	.60	.29	.72
R	1									
F		27.67**	5.28	5.29	0.99	3.29	6.32	1.78	1.54	1.25
η^2		.55	.19	.19	.04	.13	.22	.07	.06	.05
MSE	23	.51	.56	.34	.46	.68	.76	.82	.52	.90
H/V*G	1									
F		3.31	4.58	15.70*	83.49**	3.01	10.52	9.02	0.00	3.15
η^2		.13	.17	.41	.78	.12	.31	.28	.00	.12
MSE	23	.77	.58	.99	.40	.66	.35	.82	.70	.53
H/V*R	1									
F		4.32	79.01**	2.15	15.66*	29.11**	35.28**	34.79**	57.54**	48.65**
η^2		.16	.78	.09	.41	.56	.61	.60	.71	.68
MSE	23	.75	.59	.64	.74	.57	.35	.58	.55	.55
G*R	1									
F		40.25**	3.19	1.55	31.76**	46.10**	0.12	0.84	11.02	22.22**
η^2		.64	.12	.06	.58	.67	.01	.04	.32	.49
MSE	23	.48	.63	.51	.61	.58	.38	.42	.57	.50
H/V*G*R	1									
F		4.20	5.45	16.56**	0.05	1.02	0.05	11.13	0.08	1.31
η^2		.15	.19	.42	.00	.04	.00	.33	.00	.05
MSE	23	.62	.46	.56	.54	.72	.45	.50	.45	.65
Others[a]										
			H/V* SO					G*SO* QO1* QO2		
F			23.68**					15.61*		
η^2			.51					.40		

Note. [a] "Others" lists significant interactions with between-subjects factors (stimulus, category, and item orders; SO, QO1, QO2, respectively).
* $p < .003$. ** $p < .001$.

In the interaction between role and representativeness, heroes were rated significantly higher when they were representative than non-representative, while villains were rated significantly lower when they were representative than non-representative (Figure 2). This trend was observed in 12 of the 16 items, which include "good-looking," "selfless," "loyal," "intelligent," "calm," "positive emotion," "relaxed," and so forth. (The direction of the interaction in "strong" was reversed; therefore, speakers with pharyngeal constriction/expansion were rated as significantly stronger than those without.) Thus, participants attributed less favorable physical traits, personality traits, emotional states, and vocal characteristics to speakers who exhibited non-neutral pharyngeal states (i.e., pharyngeal constriction or more than a slight degree of pharyngeal expansion) regardless of the roles they played in the original cartoons. This pattern corresponds with the hypothesis set out in Section 1, and reveals that the classification of auditory characteristics into representative and non-representative based on the pharyngeal states identified in Teshigawara (2003, in press) was valid.

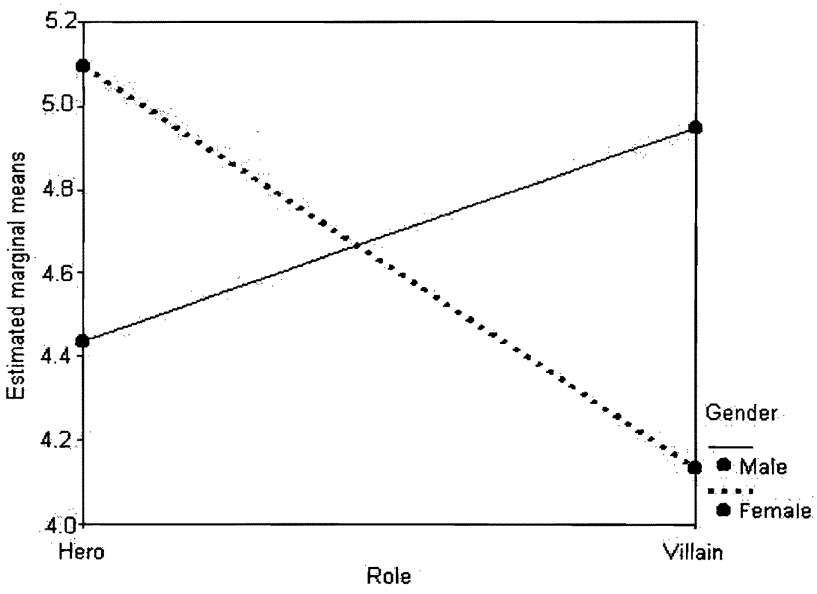

Figure 1. Interaction between role and gender for the item "conscientious."

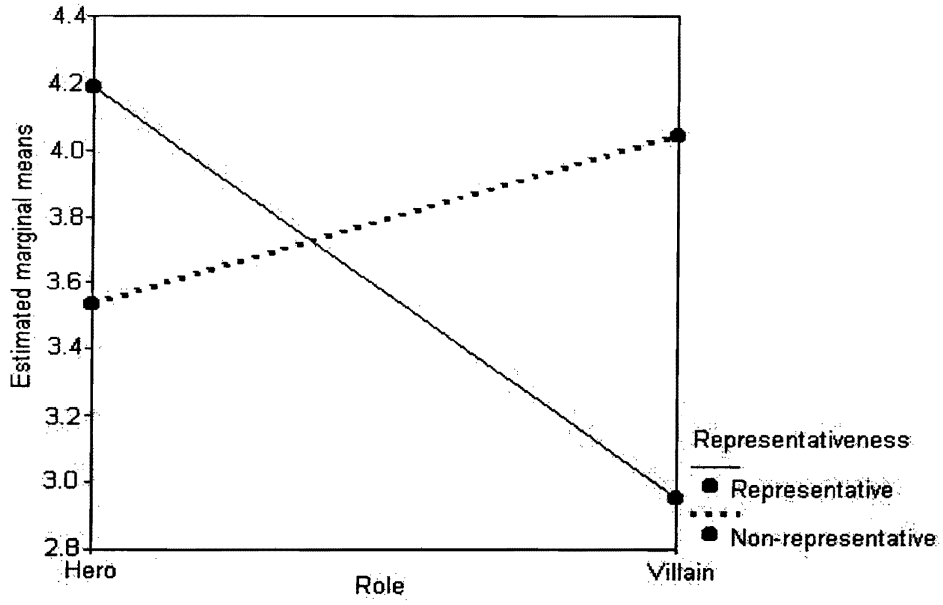

Figure 2. Interaction between role and representativeness for the item "calm."

The interaction between gender and representativeness emerged in ten items. In this pattern, males were rated significantly higher when they were representative than non-representative, while females were rated significantly lower when they were representative than non-representative (Figure 3). The items where this pattern emerged included "good-looking," "selfless," "sociable," and "sympathetic." This pattern was not predicted. These two kinds of interaction involving gender appear to have been caused by the relatively high scores of representative male villains. Relative to average ratings, representative male villains are much closer to heroes or non-representative villains than representative female villains are to the other groups except for the item "strong." It would be interesting to see whether gender would play such a significant role in a study using more naturalistic

samples as perceptual stimuli.

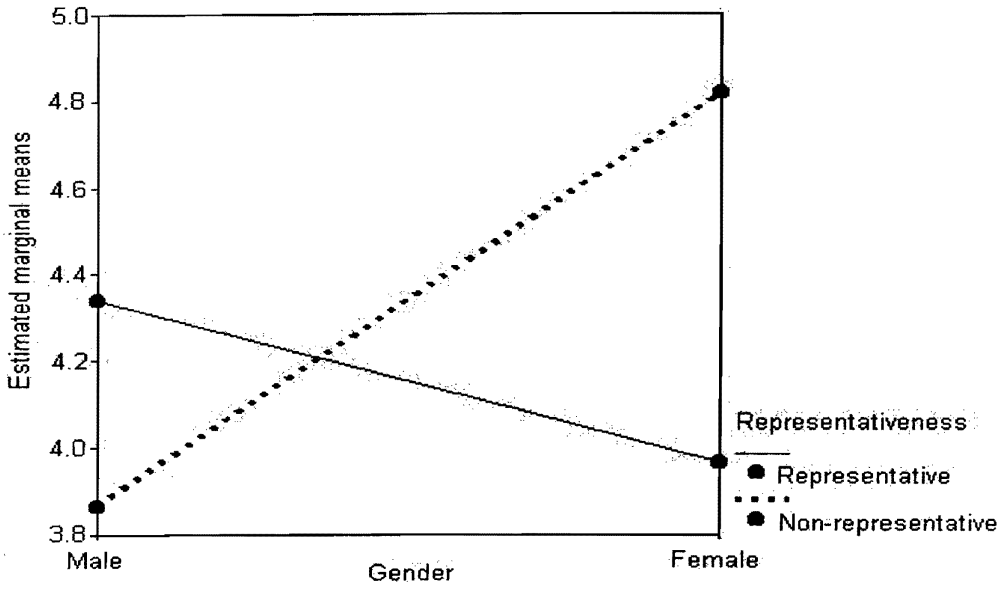

Figure 3. Interaction between gender and representativeness for the item "sympathetic."

Three-way interactions were found for three items: "devoted," "intelligent," and "curious." Of the three, "devoted" and "curious" share the same direction of interaction: male representative villains were rated higher than non-representative villains for these items, while for females, the same interaction direction was found as in the regular two-way interaction between role and representativeness. For the item "intelligent," the regular role-representativeness interaction was seen for males, whereas for females, a clear difference between representative and non-representative emerged for villains, but not for heroes.

In addition, three significant but unexpected interactions between one of the three factors and one or more control factors (i.e., stimulus order, category order, and item order) emerged: an interaction between representativeness and item order for "devoted" ($p = .001$); an interaction between role and stimulus order for "calm" ($p < .001$); and a four-way interaction among gender and all three control factors for "relaxed" ($p = .001$). It is not easy to interpret these results. For instance, the item "devoted" appeared in the following two orders in the questionnaire: (a) loyal, devoted, intelligent; (b) curious, devoted, intelligent. In (a), "devoted" appeared fourth among all the personality traits, while in (b), it appeared fifth. According to the ANOVA results, these two groups of participants rated representative characters more or less similarly while they rated non-representative characters significantly differently, with the ratings of the item order (a) group being higher than those of the group using item order (b). Interpreting the interaction between role and stimulus order in "calm" is more straightforward. The stimulus order A group rated heroes and villains similarly, while the stimulus order B group rated heroes significantly higher than villains. It is apparent that the control factors, possibly in combination with other factors that have not been assessed in the present analyses, may have played a role in these interactions. In order to determine whether the control conditions used in this study affect listeners' trait ratings, it would be necessary to replicate the experiment with a much larger number of participants in each condition group. In addition, more careful controlling may be necessary in future research.

4. CONCLUSION

This paper attempted to correlate auditory characteristics and laypersons' perceptions of the voices of heroes and villains in Japanese *anime*. Based on the auditory characteristics of the voices of heroes and villains identified in Teshigawara (2003, in press), 27 characters were selected as target stimuli from Japanese cartoons. Thirty-two participants listened to content-masked speech excerpts of the 27 target speakers and rated their

impressions of physical and personality traits, emotional states and vocal characteristics on 7-point scales. A series of three-factor repeated measures ANOVAs was carried out for each of the selected 16 items in order to examine whether participants responded to the stimuli according to the differences in pharyngeal state among the characters. In addition to significant main effects for the factor role, a number of significant interaction effects between any combination of two of the three factors and among the three also emerged. The interaction between role and representativeness emerged in the majority of the interactions observed, which suggests that participants attributed less favorable physical traits, personality traits, emotional states, and vocal characteristics to speakers who exhibited non-neutral pharyngeal states (i.e., pharyngeal constriction or more than a slight degree of pharyngeal expansion) regardless of the roles they played in the original cartoons. This pattern reveals that the classification of auditory characteristics into representative and non-representative based on the pharyngeal states identified in Teshigawara (2003, in press) was valid. Main effects for gender, and interaction effects between role and gender and between gender and representativeness also emerged for a number of items, suggesting that gender also played an important role in the perception of the characters. In order to more fully investigate the correspondence between auditory correlates and laypersons' perceptions of the voices of heroes and villains in Japanese *anime*, it is necessary to examine statistical correlations between the two, an analysis which is underway in Teshigawara (2003).

REFERENCES

Aronovitch, C. D. (1976). The voice of personality: Stereotyped judgments and their relation to voice quality and sex of speaker. *The Journal of Social Psychology, 99*, 207–220.

Biemans, M. (1998). Production and perception of gendered voice quality. *Symposium 'Perceiving and Performing Gender,' Zentrum für interdisziplinäre Frauenforschung*, Kiel - Germany, 1998, 63–72.

Friend, M., & Farrar, M. J. (1994). A comparison of content-masking procedures for obtaining judgments of discrete affective states. *Journal of the Acoustical Society of America, 96*, 1283–1290.

Hecht, M. A., & LaFrance, M. (1995). How (fast) can I help you? Tone of voice and telephone operator efficiency in interactions. *Journal of Applied Social Psychology, 25*, 2086–2098.

Kido, H., & Kasuya, H. (2001). Tsujo hatsuwa no seishitsu ni kanrenshita nichijo hyogengo: Chosyu hyoka ni yoru chushutsu [Everyday expressions associated with voice quality of normal utterance: Extraction by perceptual evaluation]. *Nippon Onkyo Gakkai shi* [Journal of the Acoustical Society of Japan], *57*, 337–344.

Laver, J. (1994). *Principles of phonetics.* Cambridge, UK: Cambridge University Press.

Laver, J. (2000). Phonetic evaluation of voice quality. In R. D. Kent & M. J. Ball (Eds.), *Voice quality measurement* (pp. 37–48). San Diego, CA: Singular.

Lent, J. A. (Ed.) (2001). *Animation in Asia and Pacific.* Bloomington, IN: Indiana University Press.

Levi, A. (1998). The new American hero: Made in Japan. In M. L. Kittelson (Ed.), *The soul of popular culture: Looking at contemporary heroes, myths, and monsters* (pp. 68-83). Chicago: Open Court.

McCrae, R. R., & Costa, P. T., Jr., (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*, 81–90.

Scherer, K. R., Feldstein, S., Bond, R. N., & Rosenthal, R. (1985). Vocal cues to deception: A comparative channel approach. *Journal of Psycholinguistic Research, 14*, 409–425.

Teshigawara, M. (2003). Voices in Japanese animation: A phonetic study of vocal stereotypes of good guys and bad guys in Japanese culture. Manuscript of doctoral dissertation in preparation, University of Victoria, Victoria, British Columbia, Canada.

Teshigawara, M. (in press). Voices in Japanese animation. In *Proceedings of the Western Conference on Linguistics 2002*, November 1–3, 2002, University of British Columbia, Vancouver, British Columbia, Canada.

van Bezooijen, R. (1988). The relative importance of pronunciation, prosody, and voice quality for the attribution of social status and personality characteristics. In R. van Hout & U. Knops (Eds.), *Language attitudes in the Dutch language area* (pp. 85–103). Dordrecht: Foris.

van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech, 38*, 253–265.

van Bezooijen, R., & Boves, L. (1986). The effects of low-pass filtering and random splicing on the perception of speech. *Journal of Psycholinguistic Research, 15*, 403–417.

Yarmey, A. D. (1993). Stereotypes and recognition memory for faces and voices of good guys and bad guys. *Applied Cognitive Psychology, 7*, 419–431.

Zuckerman, M., & Miyake, K. (1993). The attractive voice: What makes it so? *Journal of Nonverbal Behavior, 17*, 119–135.