Aggressiveness of the Growl-Like Timbre: Acoustic Characteristics, Musical Implications, and Biomechanical Mechanisms

CHEN-GIA TSAI, LI-CHING WANG, SHWU-FEN WANG, YIO-WHA SHAU, AND TZU-YU HSIAO
National Taiwan University, Taipei, Taiwan

WOLFGANG AUHAGEN
Martin-Luther-Universität Halle-Wittenberg, Germany

The term growl typically refers to low-pitched, rough sounds uttered by animals. Humans occasionally use growl-like voices to express excessive emotions. Acoustically characterized by loud dynamics and low values of the harmonic-to-noise ratio, growl-like sounds usually express anger and excitement associated with aggression. We propose a biomechanical model relating the aggressive characteristic of the growl-like timbre to the motor mechanisms underlying growl production in humans, highlighting how an abdominal muscle contraction enhances spine stability, which plays a critical role in physical attacks. This model was supported by the experimental data of activation of the deepest abdominal muscle during resting, singing, and growling. We found a significant positive correlation between the abdominal muscle activity associated with producing voice and the perceived aggressiveness intensity of voice. The cognition of growl-like sounds is discussed from the perspectives of biomechanics, evolutionary biology, and cognitive science.

Received September 22, 2008, accepted November 12, 2009.

Key words: growl, aggressiveness, Beijing opera, heavy metal, spine stability

It generally is believed that musical timbres are determined by their spectral and temporal features (e.g., Risset & Wessel, 1999). Some attempts to relate the emotional expression of timbre to its acoustic characteristics have been made using psychoacoustic models. For example, auditory roughness is an important parameter inducing unpleasant qualities in the sound. Since its introduction by Helmholtz (1885/1954), roughness has been considered to be due to rapid beatings in auditory channels, or critical bands. Using stimuli with simultaneous pure-tone pairs, Plomp and Levelt (1965) plotted an idealized curve of the relationship between consonance and critical bandwidth. This curve serves as the basis of some roughness estimation models. Applications of these models to music research range from non-tonal orchestral timbres (Pressnitzer, Mcadams, Winsberg, & Fineberg, 2000) to the sounds of the Middle Eastern mijwiz, the Indian tambura, and Bosnian ganga songs (Vassilakis, 2005).

Another approach to examining timbre perception is to reveal the physiological states and body movements associated with producing musical sounds. Communicative processes in musical activities comprise both the performers’ encoding of specific emotion and intention in the music and the listeners’ decoding of them (Juslin, 2001). As the mechanisms underlying listeners’ capacity to decode music are not well understood, it may be helpful to consider this issue within the broader framework of action understanding. Understanding complex movements and action goals is critical during social interactions. In the rapidly-progressing field of social cognitive neuroscience, investigations of action understanding have polarized into two camps (Keysers & Gazzola, 2007): one emphasizes the mechanisms of consciously thinking about the states of others (Amadio & Frith, 2006); the other emphasizes the so-called shared circuits that are involved in one’s own actions, sensations, and emotions, and in perceiving those of others (Gallese, Keysers, & Rizzolatti, 2004). The latter approach is closely related to the concept of motor simulation or reenactment. Recent data of brain imaging suggested that the neural network subserving vocal sound production may be involved in the cognition of pleasant music (Koelsch, Fritz, von Cramon, Muller, & Friederici, 2006). What makes this type of motor representation possible is a mechanism of as-if neural simulation: while listening to a pleasant tune, the vocal motor system is activated as if the listener hums this tune. Through listeners’ covert reenacting, performer’s emotion and intention can be recognized. In contrast to the psychoacoustic models...
Acoustic Characteristics and Production
Mechanisms of Growl-Like Sounds

While there has been much scientific research on the
operatic (Western classical) singing style, non-operatic
styles rarely have been analyzed. The pioneer study of
growl-like singing techniques by Borch and colleagues
(Borch, Sundberg, Lindestad, & Thalen, 2004) focused
on the so-called distorted voice used in singing rock
music, demonstrating that the distorted voice is pro-
duced with simultaneous vibrations of the vocal folds
and supraglottic structures. During normal phonation
the supraglottic structures do not vibrate; when pro-
ducing distorted voice they are adducted, involved in
vibration under high subglottal pressure. The vibrating
supraglottic mucosa intermittently obstructs the free
passage of the glottal air pulses and imposes amplitude
modulation upon the voice produced by the vibrating
vocal folds. The growl is sometimes described as the
voiced aryepiglottic trill (Esling, 1996), implying that
the vibration of the aryepiglottic folds, a portion of
the supraglottic structures, causes amplitude modulation
of the voice. In addition to the aryepiglottic folds, pro-
ducing a distorted voice is associated with the vibration
of other supraglottic structures such as the ventricular
folds and the anterior part of the mucosa covering the
arytenoid structures (Borch et al., 2004). Unlike the
vocal folds, elasticity of these supraglottic structures is
not controlled by skeletal muscles and cannot be ad-
justed at will. The supraglottic mucosa always vibrates
deeper than the vocal folds when driven by the glottal
flow. Borch et al. (2004) showed that the supraglottic
mucosa of a rock singer vibrated at 160 Hz while the
evocal folds vibrated at 160 Hz × 3 = 480 Hz. This
vibratory pattern shows interesting similarity with that
observed during throat singing (Lindestad, Södersten,
Merker, & Granqvist, 2001). In the throat singing, the
vibrations of the vocal folds and the supraglottic mu-
cosa always seem harmonically related. In the distorted
singing, however, coupling of these two oscillators also
can cause aperiodic laryngeal vibration that generates
broadband noise (Borch et al., 2004).

Growl vocalization is a crucial defining quality of some
singing styles in Chinese opera, especially in the singing
style of the jing role. The jing role is famous for his col-
olorfully painted face as well as an extraverted, aggressive,
and sometimes childish personality. Figure 1 displays
an example of the use of normal and growl vocalization
in the jing’s singing style. The initial voice was relatively
clean, with mild vibrato, as shown by the wavy lines in
the spectrogram. From 1.23 s on this vocalization was

Based on low-level auditory processing, the approach of
reenactment pays more attention to the motor system
involved in producing musical sounds (Molnar-Szakacs
& Overy, 2006).

The present paper accounted for the acoustic charac-
teristics and emotional expression of the growl-like
timbre, as well as the motor mechanisms underlying
growl production. The growl (or roar) typically refers
to low-pitched, rough sounds uttered by animals.
Humans occasionally use growl-like voices to express
excessive emotions. Although noise-rich growl-like
sounds generally are avoided in Western classical music,
they can be found in singing styles of ethnic and pop
musics. In Chinese opera, the extravert, impulsive
personality of the jing role is reflected by his growl voice
(Tsai, 2005). In Japanese theatre, a Noh percussionist’s
voice, kakegoe, may growl at the beginning of phonation
(Sakakibara, Fukus, Imagawa, & Tayama, 2004). Vocalists
of heavy metal and hard rock frequently use the growl
for expressing anger and violence.

In addition to a loud dynamic level, the growl-like
timbre acoustically is characterized by the presence of
inharmonic components that are distributed among
adjacent harmonics in a spectrum. According to psy-
choacoustic models of roughness, these inharmonic
components add a rough quality to a growl-like sound
because they interfere within critical bands. It remains
to be answered, however, why the growl-like timbre
usually communicates anger and aggressiveness to lis-
teners. While the unpleasant quality of the growl-like
timbre can be explained by psychoacoustic models of
roughness, we argue that its aggressive characteristic is
closely related to the muscle recruitment for growl
production. One goal of this paper was to outline a
causal link between the motor mechanisms underlying
growl production and the preparation of aggres-
sive behaviors.

The rest of this paper is organized as follows. First, we
describe the acoustic characteristics and production
mechanisms of the growl-like timbre. Second, we review
the possible musical meanings and the psychological
correlates of the growl-like timbre. Third, we conduct an
experiment to measure the activity of the deepest ab-
dominal muscle during resting, singing, and growling—a
muscle playing a central role in preparing high-loading
tasks such as physical attacks. This experiment also esti-
imated the harmonic-to-noise ratio (HNR) and the per-
ceived intensity of aggressiveness of these singing and
growling voices. Finally, we discuss the experimental
results from the perspectives of biomechanics, evolution-
ary biology, and cognitive science.
distorted, as can be noted by the sudden presence of broadband noise in the spectrogram. The growl-like timbre is manifested around 1.6 s, and the envelope of its waveform shows regular amplitude modulation at approximately 89 Hz. According to Daniel and Weber’s (1997) roughness model, amplitude modulations at 75-150 Hz of a sound give rise to the sensation of roughness. As the vibration frequency of the supraglottic structures during human growl usually lies within this frequency range, the roughness of growl voices tends to be higher than in normal voices.

Figure 2 compares the spectra and the HNR values of the clean voice (marked by * in Figure 1) and the growl voice (marked by ▼ in Figure 1). The former is characterized by the predominance of harmonics, whereas the latter contains significant sidebands around some harmonics (marked by numbers in Figure 2). The HNR value of the clean voice is 19.8 dB, suggesting a significant energy difference between the harmonic and inharmonic spectral components. On the contrary, the HNR value of the growl voice is 3.6 dB, suggesting a small energy difference between the harmonic and inharmonic spectral components. In the frequency domain, the rough quality of the growl voice is attributed to significant spectral energy distributed among adjacent harmonics. In the time domain, this quality of the growl voice is attributed to its amplitude modulation. Our finding is consistent with a previous investigation on vocal expression in humans and monkeys (Leinonen, Laakso, Carlson, & Linnankoski, 2003), which demonstrated the presence of spectral noise and low-frequency (< 100 Hz) amplitude modulation of their angry voices.

Some musical instruments also produce growl-like sounds. Wind instruments such as the trumpet and the saxophone produce noise-rich sounds through simultaneously playing a note and singing into an instrument. This playing technique is called growling, occasionally used in rock, jazz, and blues. When the lips (or reed) and the vocal folds co-oscillate with nonlinear coupling, the musical sound will show aperiodicity with the appearance of sidebands, broadband noise, or subharmonics in the spectrum (Gibiat & Castellengo, 2000). Subharmonics are the spectral components at integral submultiples of the fundamental frequency to which they are related (e.g., Neubauer, Edgerton, & Herzl, 2004). It should be noted that growing on the trumpet and the saxophone is sometimes used with soft dynamics, and the sound may express sadness or a laid-back feeling rather than aggressiveness.

Distorted tones of the electric guitar, frequently used in hard rock, punk rock, and heavy metal, are obtained through various forms of electric clipping, i.e., truncation of the part of an input signal exceeding certain voltage limits. The distortion effect in the electric guitar changes
the sound of a signal by additional high-frequency harmonics or combination tones at the various sums and differences of the frequency components of the input signal. Besides the distortion effect, hitting a string harder on the electric guitar also generates loud, noise-rich angry tones (Gabrielson & Juslin, 1996).

The principal bowed-string instrument of Beijing opera, the jinghu, also produces growl-like sounds. Figure 3 presents the spectrogram and time course of HNR of a sustained jinghu tone. This tone was played with a special left-hand technique to generate the effect of alternation between the clean timbre and the growl-like timbre. The latter is known as the artificial noise, tiger tone, or blossoming tone by jinghu musicians (Tsai, 2006). The time course of HNR on top of the spectrogram shows coherent decreases with the presence of broadband noise.

Previous studies on the vocal and musical expressions of emotion have suggested many acoustic cues associated with anger, such as fast tempo, high sound level, much variability in sound level, much high-frequency energy, high pitch level, much pitch variability, rising pitch contour, fast tone attacks, and microstructural irregularity (Juslin & Laukka, 2003). It is striking to find that most are also cues associated with happiness. While these arousal-related acoustic cues are not sensitive to emotional valence, the HNR might be a key indicator for distinguishing between anger and happiness.

**Aggressiveness Expressed by Musical Sounds**

The growl-like timbre is characterized by a loud dynamic and the presence of inharmonic components distributing between adjacent harmonics. With regard to psychoacoustics, these inharmonic components induce a rough sound quality because they increase interferences within critical bands. Note that a rough voice is not necessarily growl-like. An example is the singing voice of the Russian lament, whose sound quality is characterized by roughness induced by subharmonics (Mazo, Ericson, & Harvey, 1995). Whereas soft rough voices tend to convey sadness, loud rough voices usually give the impression of anger and aggression.

Among a limited number of studies on the noise-rich growl-like timbre, most are concerned with rock music and heavy metal music. Sangild (2004) addressed three distinct uses of noise: expressionist, introverted, and minimal. The expressionist gesture, commonly found in rock music, is defined as “the gesture of rush, rage, excitement, aggression, pain, fear, anger, emotional chaos and ecstasy.” Walser (1993) drew parallels between distorted guitar tones with human shouts, indicating that “distortion functions as a sign of extreme power and intense expression by overflowing its channels and materializing the exceptional effort that produces it” (p. 42). Although distorted tones of the electric guitar are open to multiple meanings in rock music, Walser (1995) stressed that “it is only at a particular historical moment that distortion begins to be perceived in terms of power rather than failure, intentional transgression rather than accidental overload.” (p. 42).

Thus, the growl-like timbre in rock music may be associated with an elevated arousal level and the muscle recruitment for high-loading tasks such as transgression. Further insights into the meanings of growl-like musical sounds were provided by an interdisciplinary study on the singing style of the jing role in Beijing opera (Tsai, 2005). Since growl production is considered detrimental to vocal health in humans, there has been much clinical research on the voice disorders possibly due to the excessive use of the growl. An example is the condition of vocal nodules, also known as screamer’s nodules. Previous investigations on vocal nodules suggested extravert, socially dominant, aggressive, and impulsive tendencies, as well as emotional coping strategies and immature behavior in people who often produce growl-like voices (Baker, 2002; Green, 1989; McHugh-Munier, Scherer, Lehmann, & Scherer, 1997; Moses, 1954; Roy, Bless, & Heisey, 2000). These clinical findings strikingly match the personality traits of the jing role in Beijing opera (Tsai, 2005). Moreover,
Tsai (2005) analyzed two major subtypes of the jīng role: tong-tsue and jīa-tzi. The former seldom growls and his anger is usually implicitly expressed on the stage (cold anger), whereas the latter frequently uses the growl and vents his anger openly. The distinction of tong-tsue and jīa-tzi is reminiscent of the distinction between anger and aggression made in the psychology literature: anger is an internal feeling state typically associated with an increased tendency to hurt others; aggression refers to the actual act of hurting others (for a recent review, see Wilkowski & Robinson, 2008). It has been proposed that individuals with superior effortful control abilities may be able to override their automatic tendencies toward aggression in a state of anger (e.g., Eisenberg, Smith, Sadosky, & Spinrad, 2004). Effortful control has been specifically defined as the ability to suppress a dominant response to perform a subdominant response (Rothbart, 1989). In this regard, growl production of the jīa-tzi role may be closely linked to the preparation or execution of aggressive acts, as effortful control fails to suppress outward manifestations of anger.

Taken together, the growl-like timbre can be viewed as a sign of extreme power (Walser, 1993) associated with rage, excitement, and sometimes ecstasy (Sangild, 2004). The elevated level of autonomic arousal represented by a growl may be the physiological basis of aggression. On the other hand, the ‘extreme power’ for the execution of aggressive acts—or more specifically, the contribution of growl-related muscle recruitment to aggressive acts—has been poorly understood. To clarify how physiological changes and emotional states influence vocal production, Johnstone and Scherer (2000) pointed out that anger yields increased tension in the laryngeal musculature coupled with increased subglottal air pressure, and this changes the production of sound at the glottis and therefore changes the timbre of the voice. This description, however, provides little information about the functions of larynx adduction and subglottal pressure elevation in aggressive acts. In the following, we argue that larynx adduction and subglottal pressure elevation in aggressive acts—in the laryngeal musculature coupled with increased tension—may be the physiological basis of growl voices produced by the human voice.

**Experimental Study**

Several biomechanical mechanisms are involved in the production of human growl, and the present experiment focused on abdominal mechanisms. We hypothesized that growling may originally function as a preparatory strategy for aggressive acts. During limb movements associated with a physical attack, reactive forces are imposed on the body that are equal in magnitude but opposite in direction to those producing the movement (Bouisset & Zattara, 1981). Because the structure of the spine is inherently unstable, muscle activation is essential for maintaining trunk posture when sudden loads are applied. To prepare a physical attack, we may activate abdominal muscles to stabilize the spine with adduction of the vocal folds and the supraglottic structures. A byproduct of this action may be a growl voice generated by the vibrating larynx under high subglottal pressure. This pressure is achieved by abdominal bracing, i.e., consciously activating all the abdominal musculature. Because of this association of sound with action, the aggressive characteristic of growl voices is almost universally recognized.

Spine stability plays an important role in high-loading acts, and several muscles contribute to it. Our experiment focused on the activity of the deepest abdominal muscle: transversus abdominis (TrA). An initial series of studies on the contribution of abdominal muscles to rapid limb movements showed that the TrA was consistently the first muscle activated (Hodges & Richardson, 1997). This finding suggests a key contribution of TrA to the preparation of the spine for perturbation during physical attacks. Unfortunately, because of its anatomical position, direct measurements of TrA activity usually have employed invasive electromyographic techniques involving fine wire or needle electrodes (Critchley, 2002). Recently, a good correlation between the electromyographic activity and the thickness change in TrA measured by ultrasound scanning has been demonstrated (Hodges et al., 2003). Therefore, we used ultrasound to estimate TrA activation during resting, growling, and singing. TrA thickness and activity were expected to reach their maximum during growling.

The major aim of the present experiment was to examine the correlations between TrA activity, the growl-like timbre, and its aggressive characteristic. We measured the HNR values and the perceived intensity of aggressiveness of recorded voices, expecting that the relative change in TrA thickness and HNR would be negatively correlated, and the relative change in TrA thickness and aggressiveness intensity would be positively correlated.

**Method**

The experimental study of human growl consisted of two parts. Part I aimed to detect TrA activity during growling, singing, and resting. Part II aimed to measure
the perceived aggressiveness intensity of the voices recorded in Part I.

**Part I**

**Participants**

Eight male students of National Taiwan University, ranging from 20 to 24 years of age ($M = 22.3, SD = 1.4$) participated in Part I of the experiment. Because greater values of TrA thickness of men than women have been reported (e.g., Teyhen et al., 2008), only male participants were included in the present study for intersubject consistency. No participants were singers or actors. Potential participants were excluded if they had voice problems, a current complaint of low back pain, or a history of surgery to the lumbar spine. Height and body mass data were collected on all participants. All participants were paid for their participation and gave written informed consent per approval by the Ethics Committee of National Taiwan University Hospital, Taipei, Taiwan.

**Procedure**

Participants were provided a review of the experimental procedure and voice samples of singing and growling prior to ultrasound examination. During ultrasound examination, participants were asked to lie on a bed and then to perform two types of vocalization: growling and singing. This posture was chosen to diminish the low-level activation of TrA associated with the standing posture (Cresswell, Oddsson, & Thorstensson, 1994; Hodges & Richardson, 1997). In the growling condition, participants deeply inhaled and then growled the vowel /a/ for 0.5-1.0 s. In the singing condition, participants deeply inhaled and then sang the vowel /a/ for 0.5-1.0 s, with a steady fundamental frequency between 100 Hz and 150 Hz. Participants first repeated the growling task until data from three vocalizations with a sound pressure level of 80-85 dB were successfully collected. Then, the singing task was repeated, until data from three vocalizations with a sound pressure level of 80-85 dB were collected. Sound pressure level was measured with a precision sound level meter (B&K 2235; Bruehl&Kjaer, Naerum, Denmark) placed about 10 cm from the lips, where the instant measurement could be seen by the participant. The reason for restricting the range of sound pressure level was to avoid the confounding effects of voice loudness.

Ultrasound images of the antero-lateral abdominal wall were obtained using a real-time ultrasound apparatus with a 10 MHz linear array transducer (HDI 5000; ATL Ultrasound, Bothell, WA, USA). Gel was interposed between the transducer and the skin. The image gain was adjusted to produce the best contrast between the tissues. Initially, the transducer was placed transversely 10 cm right of the umbilicus, adjacent and perpendicular to the abdominal wall. In the B-mode of ultrasound, a cross-sectional image with a depth of a few cm was seen from the entire length of the transducer. Then, the position of the transducer was adjusted horizontally to best visualize the mid-section of TrA, where its boundary was approximately parallel to the skin (Figure 4). During ultrasound examination, singing and growling voices were picked up by a microphone (AT9440, Audio-Technica, Tokyo, Japan). These voices were digitized (44,100 sampling rate, 16 bit mono) and underwent HNR computation using the Praat speech signal analysis software (Praat, Institute of Phonetic Sciences), whose HNR analysis is based on auto-correlation and the obtained value is a logarithmic ratio (see Boersma, 1993, for details).

To help control for the influence of respiration on TrA thickness, ultrasound images were collected both at the end of deep inhalation (resting condition) and again during vocalization (growling and singing conditions). These images were processed off-line to extract TrA thickness. For each participant, three measurements of TrA thickness in the resting condition were used to establish interrater reliability for this measurement technique. We computed the intra-condition correlation coefficient (ICC) for the percent accuracy scores. The relative change in TrA thickness was determined by dividing the thickness measurement in the two vocalization conditions by the average thickness measured in the resting condition (contracted-rest ratio).
Part II

PARTICIPANTS
Twenty eight students of National Taiwan University ranging from 19 to 25 years of age (M = 20.1, SD = 1.7) participated in Part II of the experiment.

PROCEDURE
Participants were asked to rate the aggressiveness intensity of all voices recorded in Part I (ranging from 1 = none, 2 = low, 3 = medium, 4 = medium-high, to 5 = high) with two anchor points, which were a non-aggressive voice and a highly aggressive voice selected by the experimenters from among these voices. Each run of aggressiveness-rating contained 2 (vocal types) × 3 (voice samples per type) × 8 (participants) = 48 pseudo-randomized trials. A web-based interface was built to generate these trials for a blind test on the aggressiveness intensity of singing and growling voices.

Because the data were not normally distributed, we used the non-parametric Wilcoxon paired signed rank test to compare the TrA thickness ratio, HNR, and perceived aggressiveness intensity between the singing and growling conditions (Table 2). This result suggests that our participants understood the instructions properly and were able to produce growl-like voices with a hoarse quality and an aggressive expression.

Scatterplots for the Spearman correlation analyses are shown in Figure 5. There was a significant negative correlation between the TrA thickness ratio and HNR, \( \rho(46) = -0.77, p < .0005 \), and a significant positive correlation between the TrA thickness ratio and aggressiveness intensity, \( \rho(46) = 0.74, p < .0005 \). It also can be noted that all values of HNR for the singing condition (represented by the hollow circles in Figure 5) were higher than 10 dB, whereas all values of HNR for the growling condition (represented by the filled circles in Figure 5) were lower than 10 dB. This result is consistent with an acoustic study on dog barks, which reported low values of HNR (< 10 dB) for hoarse barks (Riede, Herzl, Hammerschmidt, Brunnberg, & Tembrock, 2001).

Discussion
The major aim of the present study was to outline a causal link between the aggressive characteristic of the growl-like timbre and the abdominal mechanism underlying growl production. In our two-part experiment we measured the relative changes in TrA thickness during singing and growling, as well as the HNR values and perceived aggressiveness intensities of these voices. The finding of a significant positive correlation between TrA activity and the aggressiveness intensity (Figure 5) supports our hypothesis that the aggressive characteristic of the growl-like timbre may be related to the contribution of TrA contraction to trunk stability, which plays a critical role in physical attacks.

<table>
<thead>
<tr>
<th>TABLE 1. Descriptive Statistics from Part I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean ± SD</td>
</tr>
<tr>
<td>Height (cm) 173.56 ± 3.93</td>
</tr>
<tr>
<td>Body mass (kg) 70.12 ± 13.96</td>
</tr>
<tr>
<td>Body mass index (kg/m²) 23.22 ± 3.85</td>
</tr>
<tr>
<td>TrA thickness (mm) 2.94 ± 1.31</td>
</tr>
<tr>
<td>Note: The values of TrA thickness were measured at the end of deep inhalation (resting condition).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean ± SD</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Singing</td>
</tr>
<tr>
<td>Growling</td>
</tr>
<tr>
<td>z</td>
</tr>
<tr>
<td>TrA thickness ratio 1.54 ± 0.34 3.11 ± 1.17 -4.26 &lt; .0005 0</td>
</tr>
<tr>
<td>HNR (dB) 22.49 ± 5.50 0.99 ± 3.96 -4.29 &lt; .0005 0</td>
</tr>
<tr>
<td>Aggressiveness intensity 1.51 ± 0.40 4.26 ± 0.44 -4.29 &lt; .0005 1</td>
</tr>
</tbody>
</table>
Growl Production and Spine Stability

Inharmonic spectral components of a growl voice induce a rough quality because they interact within critical bands, but psychoacoustic models only provide part of the explanation of the aggressive expression of growl voices. Here, we propose a model to clarify the relationship between aggression and growl production. Physical aggressive behaviors, including attacking others and being attacked, require increased control of the spine. Because the structure of the spine is inherently unstable, muscle activation is essential for intervertebral control when the limbs are being moved. During limb movements the configuration of the body is altered and reactive forces are imposed on the body that are equal in magnitude but opposite in direction to those producing the movement (Bouisset & Zattara, 1981). Both attacking and being attacked involve a challenge to trunk posture, and the central nervous system prepares for this by recruiting trunk muscles to stabilize the spine. Among the abdominal and erector spine muscles, the TrA has been found to be the first muscle activated during the preparatory process occurring before limb movements (Hodges & Richardson, 1997).

To our knowledge, there are two mechanisms involved in the contribution of TrA activity to spine stability: increasing muscle stiffness and increasing intra-abdominal pressure (Cholewicki, Juluru, Radebold, Panjabi, & McGill, 1999). It is well established that skeletal muscle stiffness is proportional to the muscle force (Cholewicki & McGill, 1996). Therefore, contracting the muscles surrounding the lumbar spine can increase its stiffness and stability (Cholewicki & McGill, 1996). In addition, a contraction of TrA results in a reduction of abdominal circumference with an increase in intra-abdominal pressure because of the horizontal fiber orientation of the TrA. Cresswell and colleagues (1994) measured the intra-abdominal pressure and TrA activity during loading and movement tasks, identifying a close relationship between them. Their results indicated that the TrA was always the first muscle active in both unexpected and self-loading conditions. Recently, the contribution of elevated intra-abdominal pressure to spine stability was separated from increased TrA stiffness (Hodges, Eriksson, Shirley, & Gandevia, 2005), and the former may be regarded as an ancillary mechanism to the latter. Moreover, adduction of the vocal fold as well as the supraglottic structures prevent the movement of abdominal contents, and thereby help maintain a higher intra-abdominal pressure.

The growl may originally function as a preparatory strategy of spinal control. Pioneering work by Negus (1949) addressed the importance of closing the glottis for building up intra-thoracic pressure and thus stabilizing the body for violent physical activity. When preparing to attack, we adduct the vocal folds and supraglottic structures, activating the abdominal muscles to increase intra-abdominal pressure. A by-product...
of this action is a growl voice generated by the vibrating larynx under high subglottal pressure. Indeed, producing a distorted voice was associated with higher values of subglottal pressure compared to those typically observed in classically trained singers (Borch et al., 2004). Elevated subglottal pressure during growl production is achieved by a rapid upward displacement of the thoracic diaphragm due to elevated intra-abdominal pressure (Figure 6).

Some musicians believe that the death growl is distorted by use of ‘stomach muscles’ (e.g., see http://www.ultimatemetal.com/forum/musicians-discussion/254974-vocals.html), and the present study demonstrates increased TrA activity associated with growl production. Figures 5 shows that the TrA thickness ratio negatively correlates with the HNR value of voice and positively correlates with the aggressiveness intensity of voice. TrA contraction has been found to contribute to respiration by increasing expiratory air flow rate (Agostoni & Campbell, 1970). Since significant vibration of the supraglottic mucosa is necessary for growl production, TrA is intensively activated to produce sufficiently high subglottal pressure, generating a strong glottal flow for inducing significant vibration of the supraglottic mucosa.

Evolutionary Perspective

The term growl typically refers to low-pitched, rough sounds uttered by animals. Therefore, it is of interest to explore the phylogeny and adaptive functions of this behavior. Within the framework of “motivation-structural principles of vocal communication,” Morton (1977) noted that “birds and mammals use harsh, relatively low-frequency sounds when hostile” (p. 855). This vocal behavior may evolve independently in mammals and birds, with the adaptive significance of advertising the signaller’s body size. To compete for resources such as food, mates, or territory, numerous animal species produce aggressive calls to repel competitors. Because body size often determines fighting ability and reproductive success, it is important to assess the reliability of acoustic variables that may serve as indicators of body size. Fitch (1997) suggested a reliable indicator of body size: formant dispersion, which is the averaged difference between successive formant frequencies and was found to be closely tied to vocal tract length. If the vocal tract approximates a simple tube, the frequency difference between successive formants emphasized by the resonance of the vocal tract is inversely correlated with its length. Because vocal tract length tends to be constrained by body size, formant dispersion and body size should be negatively correlated, as can be observed in rhesus macaques (Fitch, 1997) and domestic dogs (Riede & Fitch, 1999). An ideal way to accentuate formant information may be a noise-rich call or roar (Fitch, Neubauer, & Herzel, 2002). On the contrary, a high-pitched, clear call cannot convey details of the vocal tract resonance because its spectral components are widely dispersed, and estimating the vocal tract length by the spectral envelope becomes very difficult. Therefore, low-pitched noise-rich calls of mammals and birds often serve as advertisements of their body size.

Two kinds of information in a growl should be distinguished: the signaller’s fighting ability, and the strength of intention to fight. If the formant dispersion of a growl faithfully reflects the signaller’s body size, this noise-rich sound conveys the information about fighting ability. In addition, the loudness of a growl reflects the signaller’s TrA strength, which is critical to fighting ability. For example, Taekwondo competitors occasionally growl before they attack. Their growl may be a preparatory process before limb movements (attacking the rival) and for unpredictable high-loading conditions (attacked by the rival) through the recruitment of TrA for stabilizing the spine. The growl prior to a Taekwondo match not only initiates this TrA activation, but may also threaten their rivals by “advertising” the strength of their TrA. Besides humans, enhancing spine stability by growling also may play a role in animals’ fighting behavior. An in vivo porcine study showed that elevated intra-abdominal pressure, contraction of the diaphragm, and TrA provided a mechanical contribution
to the control of spinal intervertebral stiffness (Hodges et al., 2003). Future investigations are needed to examine the activation of abdominal muscles of humans and animals during fighting.

As the growl functions as a preparatory strategy for physical attacks, this vocalization reveals the intention to fight. This is reminiscent of Tinbergen’s *intention movements* for attack, which was defined as the incomplete or preparatory phases for attack (Tinbergen, 1952). When an animal growls in the face of a stronger rival, the sound may reveal a shorter vocal tract and a lower fight ability than the rival, but its high sound level still conveys the intention of engaging in a serious fight. The stronger individual has to assess the potential cost of fighting and would retreat if the benefit of gaining the resource is not significantly higher than the cost of a serious fight. Signalling the intention to fight is important in the case that small individuals value the resource more highly than larger individuals (Dugatkin & Ohlsen, 1990). In addition to a resource value asymmetry between the contestants, recent studies suggest several possible answers to the question “why are small animals aggressive?” (Just & Morris, 2003; Morrell, Lindström, & Ruxton, 2005) and may lend support to our view that the loud growl of small individuals emphasizes the strong intention to fight.

The production and perception issues are equally important in addressing the growl’s functions. We suggest that producing a loud growl is closely related with TrA activation and spinal control, and a low value of voice HNR is critical to signalling the vocal tract length and body size. It should be noted that the Valsalva manoeuvre (forced expiration against a closed glottis) could maintain a higher intra-abdominal pressure than a growling larynx. Indeed, the Valsalva manoeuvre is widely performed by aircraft pilots, weight lifters, and other athletes for stabilizing their trunk. A fully closed glottis, however, does not produce any sound and is of no significance in acoustic communication.

**Musical Expression and Motor Simulation**

Growl-like sounds are sometimes intentionally played on musical instruments such as the electric guitar and the *jinghu*, and the terminology of such sounds reveals a close relationship between vocal expression and instrumental music. Proposals about the relationship between vocal expression and music have a long history (for a recent review, see Juslin & Laukka, 2003). Indeed, growl-like sounds produced by musical instruments tend to represent aggressiveness, apparently in the same manner as growl voices.

Recent findings of mirror neurons in neuroscience may shed new light on the relationship between vocal and musical expressions. Mirror neurons discharge when an individual executes an action, or when he/she sees or hears other individuals perform the same action (e.g., Keysers & Gazzola, 2007). The mirror neuron system is involved in one’s own actions, sensations, and emotions, and in perceiving those of others. Molnar-Szakacs and Overy (2006) proposed a model of the possible involvement of the human mirror neuron system in representing meaning and affective responses to music. This model is supported by the neuroscientific finding that both listening to pleasant music (Koelsch et al., 2006) and music imagery (Halpern & Zatorre, 1999) evoke coactivation of auditory and motor regions in the human brain. According to the simulation mechanism implemented in the mirror neuron system, listening to growl-like sounds produced by the *jinghu* or the electric guitar may induce a covert rehearsal of growling in listeners, and thereby the listeners can recognize the emotional content of these sounds. Alternatively, music decoding may be mediated by an inferential interpretive system rather than the mirror neuron system. Future investigations would examine whether the motor-related neural circuits are involved in the cognition of growl-like sounds.

**Conclusion**

There are two major acoustic characteristics of the growl-like timbre: a high sound level and a low value of HNR. From a perceptual viewpoint, the former is related to an elevated arousal level, and the latter is critical to conveying information concerning body size. The present study proposes a model that relates aggressiveness of the growl-like timbre to the motor mechanisms underlying growl production in humans. This model has been supported by the experimental finding of a significant positive correlation between TrA activity and the perceived aggressiveness intensity. While the present study offers a tentative model for future experiments and developments, it should be pointed out that reliability and deception in animal growls have been found to vary across species (Fitch & Hauser, 2003; Fitch & Reby, 2001; Searcy & Nowicki, 2005). More importantly, there are numerous recordings of blues singers or in the playing style of cornetists, trumpeters, and trombonists of the swing period (Kernfeld, 2002) that may not implicate an association of the hoarse sound quality and aggressiveness. We believe that the musical uses and meanings of the growl-like timbre are more or less shaped by culture.
Author Note

We wish to thank Ian Cross, Lien-Siang Chou, and three anonymous reviewers for helpful suggestions on earlier versions of this manuscript, Hsiang-Yu Liang and Feng-Pu Yang for assistance in the experimental study, and Hsiu-Chuan Hung for advice in statistics.

This work was supported by Grant NSC 97-2410-H-002-165 and NSC 96-2314-B-002-081-MY3, Taiwan.

Correspondence concerning this article should be addressed to Chen-Gia Tsai. Ph.D., Graduate Institute of Musicology, National Taiwan University, No.1, Sec. 4, Roosevelt Road, Taipei, Taiwan 106. E-MAIL: tsaichen-gia@ntu.edu.tw

References


