HEAVY METAL VOCALS: A TERMINOLOGY COMPENDIUM

Susanna Mesiä and Paolo Ribaldini
* Sibelius Academy, Finland
** University of Helsinki, Finland

Abstract

The aim of this study is to provide a compendium of vocal terminology, which can be referred to when addressing vocals in the academic research about Heavy Metal. The purpose is reached through the implementation of the vocabulary commonly employed in contemporary Finnish pop/jazz vocal pedagogy. Such vocabulary is kept as simple and understandable as possible, in order for scholars untrained in singing to effectively employ it in their studies. In order to facilitate the task of understanding how the different vocal characteristics works, famous performances from the HM repertoire are brought as examples.

General characteristics of Heavy Metal music

Previous studies report many different opinions about when exactly HM came into existence (Lilja, 2009, pp. 29-47). Walser (1993) states that the namesake albums Led Zeppelin II, Deep Purple In Rock, and Black Sabbath, all released between 1969 and 1970 by English bands, marked a fundamental milestone in defining the primary features of Heavy Metal (pp. 1-16).

The main musical characteristics include a) loud volumes and distortion; b) strong riffs usually played on electric guitar, sometimes on keyboards; c) frequent use, according to each singer’s style, of high-pitched strong vocals; d) great display of virtuosity derived by the aesthetics of Western art music practice (Walser, 1993, pp. 53-107); e) predominance of other modes than major and minor, and preference for modal harmony rather than tonic-dominant (i.e. ‘classical’) structures (for a definition of “mode”, see Tagg, 2009, pp. 45-48); f) extensive use of power-chords. A power-chord is an amplified and distorted chord – usually performed on guitar or keyboard - formed by the chord root, the perfect fourth or fifth, and sometimes also the perfect octave (Walser, 1993, pp. 2-3; Lilja, 2009, pp. 102-104).

Conventionally, HM is considered to be split into three main sub-genres from the 1980s: traditional, pop, and extreme metal. In spite of the many fundamental differences between traditional and pop metal in their instrumental and extra-musical characteristics, both these sub-genres are vocally based on the same technique, so they are here brought together as one. Unlike the two sub-genres mentioned above, extreme metal employs different vocal tools, of which the main one is grunting. When it comes to the vocals, an earlier research by Zangger Borch and Sundberg (2011) identifies power, loudness and high energy as fundamental characteristics of rock (and therefore HM) vocals, which clearly fit into the general features of the genre described above.
How previous literature on Heavy Metal deals with singing

Of all the instruments used in HM music, vocals are the less addressed by previous academic studies, sometimes even through a technically imprecise terminology. Nevertheless, when criticizing what has been said in such works, one should be aware that vocals were out of their main purposes. The criticisms presented in this paper don’t mean to compromise the positive value in the focus of those researches, being it music-analytical, historical, sociological, and so on. This study is meant to be a complementary support, which integrates, rather than opposes, what has been said so far in the HM academic community.

At best, earlier literature shortly describes in general terms ‘how it sounds’, but leaves completely unaddressed the question of ‘how it is produced’. Most often, the short description of HM vocals raise many questions, which have been unanswered in the field’s academic studies so far. The focus goes on guitars, bass, keyboards, and sometimes drums. The former three instruments are usually dealt with in analytical contexts, e.g. in the abovementioned studies by Walser and Lilja. In addition to that, Cope (2010) concerned himself with the influence of drum patterns in shaping the change from traditional to extreme HM (pp. 96-103). Walser, for example, states that vocals somehow try to imitate the sonic configuration of distorted guitars:

The vocal sounds of heavy metal are similar, in some ways, to the guitar sounds. Quite often, vocalists distort their voices, for many of the same reasons that guitar players distort theirs. Heavy metal vocalists project brightness and power by overdriving their voices (or by seeming to), and they also sing long sustained notes to suggest intensity and power; sometimes heavy vibrato is used for further intensification (Rob Halford of Judas Priest is a prominent example). The tough solo voice, the norm of vocal delivery, is occasionally supplemented by a chorus of backup voices, most often during the chorus section of the song. These additional voices serve to enlarge the statements of the solo vocalist, enacting the approval or participation of the larger social world, or at least a segment of it. (Walser 1993, p. 45)

Voice distortion is, especially in traditional HM, frequently encountered, but not so often. Ozzy Osbourne and Ian Gillan, for example, seldom use any distortion at all, and the same goes for Rob Halford’s performances in (at least) his first decade with Judas Priest. It is true that the traditional HM singing focuses on brightness and power, and that it may present vibrato and distortion, although it is not clear what ‘overdriving’ means in this context, or why singers would try to imitate the guitar sound.

In addition to what Walser already presented, Lilja says:

Vocal style is generally high pitched tenor, frequently employing falsetto. Still, there are exceptions to the rule; [...] Deep Purple’s Ian Gillan and Led Zeppelin’s Robert Plant clearly derive much of their vocabulary from blues and rock ‘n’ roll. For instance, Gillan’s singing in “Speed King” (In Rock 1970) significantly resembles the gospel-influenced vocal style of Little Richard; [...] Both Gillan and Plant use a lot of vocal embellishment that is common in gospel-derived styles (cf. e.g. Ray Charles, Sam Cooke) and various forms of the blues (e.g. Robert Johnson, Billie Holiday). (Lilja 2009, p. 36)
Here Lilja talks about tenor voice as a style, as if it was possible to choose whether to use it or not. Properly said, tenor is a singing voice type mostly based on the physiological structure of the singer, therefore it is not a matter of choice. A voice type is identified by range, qualities and characteristics of the voice. Tenor is the male voice type with the highest range. On the other hand, it is correct to say that traditional HM vocalists make extensive use of the falsetto register, which does not rely so much on the range but rather on a specific technique.

Other material employs more sociological and anthropological perspectives. In this respect, Weinstein’s *Heavy Metal* (2000) is reputed a cornerstone work about the sub-culture of the genre. Vocals are considered in respect of the social and symbolic role of the singer, but hardly from a musical point of view.

The heavy metal vocalist is an individual and is not submerged in a vocal group. But the singer is also embedded in the band; [...] The heavy metal code for the singer is distinctive. One major requirement is the explicit display of emotionality, which contrasts with the punk vocal principle of the flat, unemotional voice. But not all methods of emotional display are permitted. The plaintive, nasal whine of country music and the falsetto of doo-wop are rejected. The singing is openmouthed, neither gritted nor crooned. The range of emotions is wide, including pain, defiance, anger, and excitement. As in other features of the genre, softness, irony, and subtlety are excluded. [...] The heavy metal singer’s voice must also sound very powerful. It is amplified not merely by electronic devices, but by a robust set of lungs and vocal chords [sic]. Special sounds, especially screams, serve to emphasize the power and emotionality of the voice. Led Zeppelin’s Robert Plant and Judas Priest’s Rob Halford are well known for their wails and yowls. [...] Other singers use an operatic voice, although it cannot be pure toned. There must be a blues-tinged toughness in the voice. Ronnie James Dio’s voice [...] is an example of this gritty-operatic option. (Weinstein 2000, pp. 26-27)

Weinstein seems here to ignore, in the first place, that punk music delivers a huge load of anger and frustration, and the voice itself doesn’t escape this task: it is then quite inappropriate to address the punk vocal style as flat and unemotional. Perhaps, it could be said that the vocal melodies are not very clear and varied, and the types of voice are often much closer to yelling rather than melodizing. Apparently, Lilja and Weinstein refer to different ideas when employing the term ‘falsetto’. The former seems to recall the register of falsetto, i.e. the high register of a male voice, whereas the latter probably conceives it as a delicate and breathy sound, which is indeed typical of the doo-wop style. Traditional HM vocalists don’t necessarily exclude either gritting, crooning, or singing their lines open-mouthedly: Ian Gillan, for example, croons the lower parts in Deep Purple’s ‘Child In Time’ (1970), but opens his mouth on the top tone E, and powerfully hits the top A at the end of the vocal climax in the same song. Other singers privilege one particular setting over the others, but it would be daring to say that one singing manner is completely left out. Similarly, excluding “softness, irony, and subtlety”, as Weinstein does, seems to fit the fastest and most aggressive HM repertoire, but doesn’t properly consider the many power-ballads and more quiet parts of songs, e.g. ‘Victim Of Changes’ by Judas Priest (1976), which present a complex and varied structure. Whereas common sense suggests that the physical characteristics of the HM singer’s vocal apparatus must match the power of the sound, it’s not necessary to be equipped with either extraordinarily strong lungs, or thick vocal folds. Tenors
usually have the smallest vocal folds among male voices, and the amount of breath inhaled when singing pop/jazz music is only a minor part of the maximum capacity. Operatic voices belong to the classical music tradition and, aside from rare exceptions (e.g. former Nightwish vocalist Tarja Turunen), are basically left out of HM canons. Many characteristics of the other instrumental parts indeed owe a lot to the Western art music practices, but vocals don’t borrow much from this repertoire. The definition of ‘pure toned’ might be misleading, since in popular music the standard of pureness and cleanness of the vocal parts is much more blurred than in the classical tradition, and the individual singer’s timbre may be characterized, for example, by breathy or raspy sound, slightly off-tune ornaments imitating blue notes, and so on. Finally, whereas it is true that blues singing is one of the major inspirations of HM vocalists, not all of them borrow from the same musical tradition. Ronnie Dio himself, though very much influenced by blues in his Elf and Rainbow years, hardly uses blues tones or swung rhythms in his later career with Black Sabbath and with his self-named band. It seems quite contradictory, then, to describe Dio’s voice as ‘operatic-gritty’, when the two types of sound present opposite characteristics, as shown later in this paper.

**Physiology of singing**

We are aware that singing is a complicated activity that can not be explained in detail in such a short presentation. Therefore, we discuss the voice production from an angle that could be beneficial to HM researchers. To be able to provide researchers terminology with which singing may be aurally analyzed, we need to discuss the physiology of singing. Understanding how different sounds are produced will help the researchers to identify different singing phenomena.

**Figure 1: Larynx and vocal tract.**

To make an audible sound, the human voice needs different components: a sound source that vibrates (that are two true vocal folds in the larynx), an air-
flow that makes them vibrate (that is the exhalation from the lungs through trachea) and resonators that reinforce the quiet vibration of the vocal folds (that are all spaces in the vocal tract above the vocal folds). By making changes in the closure of vocal folds, the airflow and the shape of the vocal tract a singer is able to produce sounds with different pitch and quality. In this paper, these different aspects are discussed in three categories; firstly the vocal fold level that affects for example the register which the singer is using, secondly the vocal tract level in which all resonating sounds and effects are produced, and thirdly other vocal phenomena that usually are a combination of the first two.

**Relating vocal registers**

Defining registers in the human voice has been controversial. Ingo R. Titze (2000), one of the most famous researchers of vocology, defines registers as “perceptually distinct regions of vocal quality that can be maintained over some ranges of pitch and loudness” (p. 282). In other words, the voice is produced similarly by the vocal folds, and sounds alike while singing in one register. It is generally agreed that registers are associated with the glottal voice source, such that each register is produced with a specific set of vocal fold vibration characteristics. Logopedics, phonetics and vocology, sciences which do not approach singing but vocal habilitation, identify four different registers in human voice. These are 1) vocal fry (pulse register), 2) modal, 3) falsetto and 4) whistle. These registers may be aurally demerged from each other.

Figure 2: The registers.

In *vocal fry* vocal folds are tightly against each other and only a small part in front of the glottis opens while the vocal folds vibrate (Laukkanen & Leino, 2001, p. 49). Vocal fry may sometimes be heard in normal speech, especially in the end of a phrase or when a person is speaking lazily or without any support from the body. Vocal fry produces a ‘popping’ sound, in which each vibration of the vocal folds may be heard separately. Such sound may be heard in the
beginning of David Coverdale’s vocals in ‘Here I Go Again’ by Whitesnake [0:11] in both the studio version from 1982 and the one from 1987.

The term modal identifies the register in which speech is usually produced. In singing research, though, the modal register itself is often divided into two, the lower speech register and the higher head register. In speech register, vocal folds have a thick depth of contact, and therefore produce a full-bodied rich sound. Examples are Eric Adams at the beginning of ‘Heart Of Steel’ by Manowar (1988) [0:33] and Rob Halford in the initial verse of ‘Turbo Lover’ by Judas Priest (1986) [0:25]. When singing in head register, the vocal folds close only from the upper part, producing less overtones and therefore a thinner sound, as exemplified by Ian Gillan hitting high pitches in ‘Child In Time’ by Deep Purple [2:37], or again by Halford at the beginning of ‘Savage’ by Judas Priest (1978) at the beginning of the vocals. Sometimes, in order to prevent audible changes in the sound, the two registers are mixed, making them sound as only one wide, smooth and coherent register. Still, the term mixed refers to the mixing technique of speech and head registers, and is therefore not a register itself. A good example is Ronnie James Dio’s performance in ‘Gates Of Babylon’ by Rainbow (1978) [0:58].

Falsetto register is defined in various ways depending on the literature. According to Thurman, Welch, Theimer, and Klitzke (2004) it refers to a voice quality that adult males can produce within the female pitch range and is female-like in quality (p.26). Estill Voice Training (2009b) re-defines it as a ‘voice quality’, characterized by the vocal folds vibrating completely or partially separated (p. 21). According to the latter interpretation, falsetto may also be produced by female voices. In lower ranges, falsetto, if defined as sound produced with open or partly open vocal folds, sounds breathy, but in higher ranges may be difficult to separate from the head register. A good female example is respectively Amy Lee in the beginning of Evanescence’s ‘My Immortal’ (2003) [0:11] and male example Ian Gillan in ‘Child In Time’ by Deep Purple [1:53].

The whistle register is commonly considered to be a register that can only be produced in very high frequencies, mainly by children and women. On the other hand, Thurman, Welch, Theimer, and Klitzke (2004) consider it to be the parallel register of the female voice to falsetto register of the male voice (p.26). There are some examples of possible male whistle register use in HM, such as Nicola Sedda in ‘Colors’ (2014) [2:47]. This register is hypothesized to be produced by forming a small opening between the vocal folds to enable a whistling sound (Laukkanen & Leino 2001, p. 50, author’s translation).

The term grunt (also called growl in some research) describes the vocal technique used in most extreme sub-genres, such as black, grind and death metal. Songs such as ‘Mourning Palace’ by Dimmu Borgir (1997) [0:13] and ‘Nemesis’ by Arch Enemy (2005) [0:09] are suitable examples. The sound has some similarities with ‘undertone’ singing used in ethnic voice techniques, like Tuvan throat-singing or Tibetan Buddhist monks’ chanting. Eckers & al. (2009) suggest that, in death metal singing, “... at least two forms of supraglottal laryngeal constriction are used to obtain the typical death metal sound which differs in the use of vocal folds and ventricular folds as well as in supraglottal adduction effort” (p. 1747). In simpler words, this means that the sound is mostly produced through the vibration of other parts of the larynx than the true vocal folds. The research results are incoherent, probably because the term ‘growl’ may be understood differently. According to Sadolin (2012, pp. 192, 196), the difference between a growl and a grunt is that growl is produced with true vo-
cal folds (producing a sound with a clearly audible fundamental pitch) and added with an effect, and grunt may be done without major participation of the vocal folds, while the whole vocal tract is vibrating with low frequencies (producing a sound with no clear fundamental pitch). Another research by Eckers and others (2009) shows that the analyzed singers achieved the grunt sound in slightly different ways from each other. Therefore, it is not possible to summarize one and only grunt technique at the moment.

Moving from any register to another with a sudden and aurally clear change of the sound, is called vocal break. Emphasizing vocal breaks is common in yodeling and many contemporary popular music genres. If the singer wishes to make the break between speech and head registers - or vice versa - unheard, the two registers must be mixed (see above).

**Related to sound and vocal effects**

In popular music there is no ideal sound for the human voice. Zangger Borch and Sundberg (2011) suggest that voice usage may differ substantially between different singing styles (p. 532). The aesthetic value of singing is often considered through the musical tradition of the genre, and as expression of the singer’s personality. Singers therefore often make great changes in their voice production looking for a distinct personal sound to separate them from other singers. These differences are usually aurally detectable, even by a non-trained ear.

In rock genres it is quite common to add a timbral ornament conventionally called distortion as an expressivity tinge to the singing voice. There is a variety of distorted sounds, which are all produced with a different technique. Zangger Borch, Sundberg, Lindestadt, and Thalén (2004) name these sounds ‘dist tones’. They are produced with a vibration of the supraglottic mucosa (the resonating part above the vocal folds), that may be ‘aperiodic or periodic’ (irregular or regular). Periodic vibrations may, again, often be found in ethnic voice production such as throat singing, whereas the distortion in rock is often aperiodic. Often the personal distorted sound of a singer is the combination of the vibration of the false vocal folds, the epiglottis, the aryepiglottic fold and other supraglottic mucosa. In the same study, ‘dist’ singing is deemed to be very taxing to the voice, and voice rest is recommended after a prolonged use of such techniques, for example in a demanding HM concert. (pp. 152-153). Nevertheless, no unmistakable scientific data about the topic exists yet, the vocal technique of distortion may considerably vary from singer to singer, and some singers can be particularly resistant. Sadolin (2012) also argues that intentional distortion can be done without compromising vocal health (p. 179). A suitable example of male distortion is Ronnie James Dio in ‘Don’t Talk To Strangers’ [1:06] and female Noora Louhimo in ‘Raven’ by Battle Beast [0:27].

According to Sundberg (1994), vibrato presents differing acoustical and/or physiological characteristics in different music traditions and cultures. It is conventionally defined as a regular pulsating change in the tone, and/or an undulation of the fundamental frequency; furthermore, it may be intentional or unintentional, depending on the control and the awareness of the singer. The kind of vibrato that changes the pitch is called laryngeal vibrato, in which two elements are considered: the vibrato rate, which specifies the number of undulations per second (how ‘fast’ the vibrato is), and the extent, which describes how far the vibrato pitch departs up and down from the original (how ‘wide’ it is). (Sundberg, 1994, p.45). An example of such vibrato is Ian Gillan’s
performance in 'Highway Star' (live in 1972) by Deep Purple [2:08]. The other way to produce vibrato, which Sundberg doesn’t discuss, is the \textit{hammer} vibrato. It is assumed that this vibrato is produced at the vocal folds level by quickly changing their closure. The hammer vibrato does not change the pitch, but may be identified as a pulsation (Sadolin, 2012, p. 209). Such vibrato was used for example by Edith Piaf, and the sound thus produced resembles a goat.

There is no joint understanding of the term \textit{twang}, but in most literature it is defined as a metallic, sometimes piercing effect, added to the phonation by lowering the epiglottis, thus narrowing the vocal tract. The amount of twang may be adjusted by the singer, from small narrowing causing brightness to almost full closure which causes \textit{distortion} (see above). Twang may aurally be confused with nasality, but the sound is not produced using nasal resonance. A good example is the final part of 'Victim Of Changes' by Judas Priest [from 6:35]. Jo Estill’s (2009b) definition of \textit{nasalized twang} refers to a nasal sound accomplished by opening the back of the soft palate during phonation (p. 42). Such clear use of nasality is found in country music, especially sung by female singers, but is not really common in HM, so examples are hard to find as such.

\textbf{Other vocal phenomena}

There are different views on what is the optimal \textit{height of the larynx} while singing. It is commonly acknowledged that lifting or lowering the larynx from its middle position, accomplished by external muscles of the larynx, alters the sound. As the vocal tract is lengthened when the larynx is lowered, this causes a darker sound. Raised larynx on the contrary brightens the sound color of the voice. This difference is clearly heard when comparing the classical singing tradition (with a lower larynx) and popular music singing (higher larynx).

There has been several research done in recent decades on vocal technique used in popular music and musical theatre singing, in which the singer seems to reach high pitches with a loud chest register sound. This phenomenon is conventionally called \textit{belting}. The physiology behind the term is still controversial, as in a survey conducted in the USA vocal teachers with classical background maintain that belting is a style of forced declamation, in which the chest voice mechanism is taken upward, beyond its normal frequency limits, without mixing or changing voice qualities. Some teachers assumed, on the basis of their aural perception and physiologic production, that belting is not chest voice, but a quality unique unto itself (Spivey, 2008). According to Estill studies, the latter is the case. The sound may resemble speech register in higher range but the production of the voice is different from speech (McDonald Klimek, Obert, & Steinhauer, 2009a, p. 66). This technique can be exemplified by Bruce Dickinson in 'Tears of the Dragon' [1:19].

\textbf{Figure 3:} Recommended terminology.
The vocal technique used in classical singing has a distinct sound created by, for example, lowering the larynx and expanding the vocal tract. Also, the legato lines and vowel-based pronunciation differ from popular music singing. It has not been commonly used within HM music. Nightwish was the first band to be successful in combining HM and classical singing. Tarja Turunen’s use of voice in the higher parts of ‘Sleeping Sun’ by Nightwish (1999) [3:30] exemplifies this classical technique.

Conclusions

As a conclusion, we suggest what is shown in Picture 3 as a possible way to address vocals in HM singing. Hopefully, this provides the academic field with a stable but simple and basic framework, describing vocals from the points of view of both production and perception.

In sketching out their fundamental characteristics, it must be remembered that they seldom appear in their ‘pure’ configuration: most of the time, they are combined together, therefore producing particular, sometimes very personal and recognizable vocal sounds.

Vocal registers happen at the level of vocal folds. Effects are caused by the intervention of other parts in the vocal tract. Other vocal phenomena belong to neither of the two previous groups, and are characterized by their own specific configurations of the vocal fold and vocal tract levels or activity of the external muscles of the larynx.

If we consider the most successful and well-known traditional HM repertoire, it is clear that some features are primary to this genre. Among the registers, the modal (either speech or head, with an almost certain probability of mixing between the two) produces the most powerful and energetic sound, whereas falsetto is used more rarely due to its sometimes breathy and softer sound. Distortion, vibrato, and twang are very common, and are often combined with the desired register to create the trademark sound of a successful HM singer.

From the perspective of HM studies, a particularly important consequence is the clear difference between distortion and grunt: although they both produce very aggressive timbres, the former is an effect and still leaves audible the fundamental pitch of the tone, whereas the latter may be considered as separate from registers, and doesn’t have a fundamental pitch, therefore producing sounds which are not really ‘tuned’. This is why male and female voice can hardly be distinguished one from another when grunting, a phenomenon which has favored the rise of many extreme metal bands led by female vocalists (e.g. Angela Gossow in Arch Enemy).

Bibliography


**Discography**


Nicola Sedda (2014). ‘Colors’. Colors EP.


