Manual for Voice Map Method Manual

Contents

This is a short, practical introduction to the Voice Map Method, VMM — a device for enhancing communication between composer and singer.

The method was created and tested as part of a doctoral project. For the theoretical background, see the project report: miika.info/VMM

Some information in this manual is provided via links to short videos. The full list of videos can be found here: miika.info/VMM/videomanual

1 What is the VMM?

The VMM is an assessment system that helps a classically trained opera singer and composer better communicate with each other during the process of creating new vocal music. The needs of composers with little experience of vocal music are especially considered. The VMM enhances communication during the creative process, especially at the beginning stages.

The voice is measured in the first step, the Voice Map Analysis (VMA). VMA software is a Max/MSP standalone patch. This measurement requires work from the singer, comparable to a very thorough vocal warm-up in front of a microphone. The full range of the singer's voice is studied, and most pitches are sung more than once. The output, charted on a graph (a.k.a. the Voice Map), illustrates certain aspects of the voice: the different areas, range, dynamic possibilities, and singer's formant. Thereafter, the singer and composer discuss these results, taking the List of Good Questions as their starting point. Two hours or less is normally required to complete the VMM process.

VMM process:

- Voice Map Analysis
    - Preparation
    - Naming the areas
    - Soundcheck
    - Measurement
    - Interpretation
- List of Good Questions

1.1 Who should use the VMM?

The VMM has been created to aid communication between singers and composers. Only singers who understand their vocal registers and have a good control of their voice can undertake the VMA, which is a fundamental part of the VMM. In the context of the VMM, the registers are called 'areas' (refer to the report, section 1.3: Areas, for further details). All singers participating in the testing phase were classically trained.

Through the systematic use of this method, the composer is able to gain an understanding of the functionality of the voice in general, and above all, of how the voice of a specific singer functions.

1.2 How can the VMM help the user?

Each area of the singer's voice is evaluated systematically using the VMA, whereby the composer hears each tone sung as softly as possible and as loudly as possible. Dynamic values of each tone are automatically shown in graph form, making it easier to discuss specific points, such as the extremes of the singer's range, or pitches that might be produced as part of more than one area. Although the VMM identifies the extremes of the voice, it doesn't encourage exploiting them.

The VM also informs the composer of the so-called singer's formant (i.e., ring, twang, *squillo*), an indicator of the voice's ability to be heard through the sound of the orchestra.

These communications help to engender an understanding of what the composer requires from the singer as well as how to compose efficiently for their voice. To demonstrate the practical use of the VMA, some theoretical, experience-based examples are provided here below.

1.3 Areas

The word 'area' can be used as a synonym of 'register' but at the same time, it can be understood more loosely as something the singer feels in their voice.

The range of the voice is divided into different areas. This division is based on the vocal mechanism – the physical sensation of singing and the timbre of the voice. When the tones are similar in both, they belong to one and the same area.

Before the singer enters into the analysis situation, they should write down the names and approximate ranges of their vocal areas. The exact pitches will then be checked during the analysis. The vocal areas will likely overlap.

> **Example 1: Being aware of the areas**
> The composer has a concept of a vocal piece consisting of two contradicting sections. In the first section, the voice changes timbre quickly, whereas in the second section, glissando lines with a very homogenous timbre are sung. Via the VMM, the natural areas of the singer's voice are identified, and from the discussion that follows, the composer can find out how the singer feels about using those areas. In the final composition, the timbral changes comprise a mixture of areas and some extended techniques. Most glissandi are sung in the range of one area. These decisions allow the singer to use the physical aspects of their voice efficiently and concentrate on the few places that come less naturally.
>
> When another singer wants to perform the piece, the piece is transposed and the new singer's voice functions similarly in this new range.

2 The VMA Software and Accessories

This section introduces the technical equipment and processes involved in the VMA.

Video 1: Technical set-up and preparation: miika.info/VMM/videomanual1

2.1 Equipment

Zoom H4n Pro or Zoom H6 recorder

Testing was done on a Zoom device, which is deemed by many composers and other musicians to provide adequate quality. Thus, all the calibrations and instructions given in this manual relate to this device, but there are no technical reasons that prevent the use of similar recorders or microphones. Before beginning the testing, make sure that the device is updated and fully functional.

Mini USB to USB cable, 2 m or longer

The device needs to be connected to the computer so that it can be used as the audio interface.

Microphone stand

This should be at least 1.8 m tall, and you should be able to fix the Zoom to it.

Computer, with 300 MB of free space

Typically, audio files use less than 100 MB. Max/MSP software is not needed since the VMA software functions as standalone.

[Testing for this project was done with a 2015 MacBook Pro.]

Headphones for the singer

Closed-back headphones are preferable. They are used to enable the singer to hear the reference tones, so average audio quality is sufficient. The cable should be long enough to allow the singer to be positioned 50 cm from the microphone.

A silent space

It is essential that the testing space is reserved for private use during this analysis session, preferably for two hours. Noise coming from outside may disturb the singer

and also affect the program. The rehearsal rooms at universities often prove too noisy, but a standard soundproof room will do. The composer and the computer used for the analysis can be placed in the same space as the singer.

Optional:

Headphones and a splitter plug-in for the composer

With headphones, the composer can hear the reference tones exactly as the singer hears them. This may in some instances (such as when there is a strong vibrato) help the composer connect the sung tones to absolute pitches, but it can also hinder the listening to the actual timbre of the voice.

MIDI keyboard and USB cable

The reference tones can be played by clicking on the screen of the computer running the VMA software, but also by way of an external keyboard.

2.2 Voice Map Analysis software

The functionalities and components of the VMA software can be seen in figure 1.

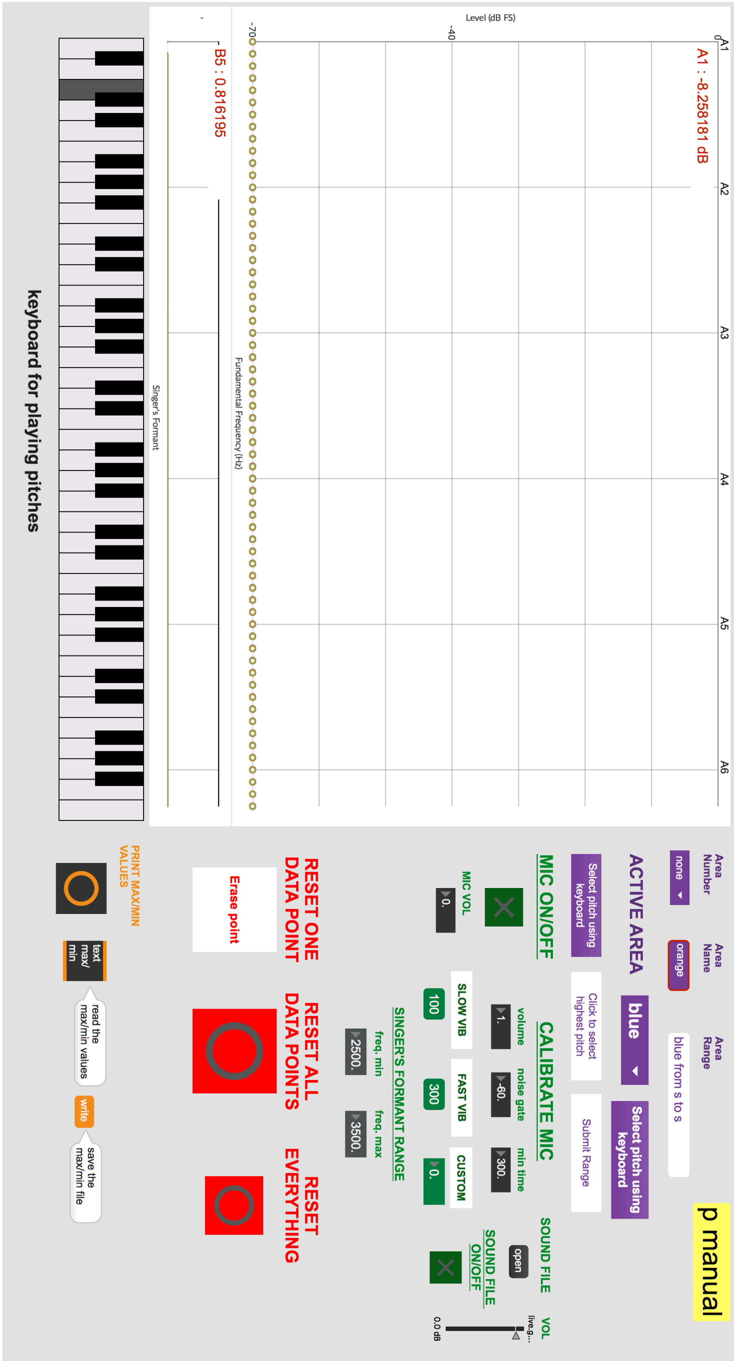

Figure 1. Components of the VMA software

2.3 Preparations

The voice measurement process should be clutter-free, in order for the singer to be able to concentrate on singing and the composer on listening. Therefore, a well-planned technical setup is crucial.

- Open the VMA software, a standalone patch for Max/MSP
- Connect the Zoom audio device to the computer via the cable; Zoom should start automatically
- In Zoom, select **Audio I/F**, **48 kHz** and then **Connect**; now you can use the audio device as an audio interface
- *For H4N*: On the side of the Zoom device is the **Rec level** button — use it to set the recording level at level 20 (practical in most analysis situations)
- *For H6:* On the Zoom device are volume knobs. Choose the one for the right stereo pair and set the recording level at 4 (practical in most analysis situations)
- In Mac OSX, go to **System preferences**, choose **Sound**, **Voice in**, and from the list, choose **H4** or **H6**; now the computer is using the audio device for sound input
- Open the VMA patch, a Max/MSP window opens; choose **Options** and **Audio status** then choose **Input device**, **H4** or **H6;** in the same window, click **Audio on/off**: Max/MSP uses the audio device for audio input, and the default **Audio** setting is **on**.

2.3. 1 Input of areas in the program

Video 2: Setting up the areas: miika.info/VMM/videomanual2

The areas of the singer's voice need to be set manually. This can be done by either the composer or the singer.

- In the menu, select **Choose Area to Add**. Start by choosing **Area 1**
- Type the name of this area into the corresponding field, choose **Area name**
- From the **Active Area**, choose the area you just named. You may have to click twice.
- Select **Click to define area range**. Two new options appear
- Select **Click to select lowest pitch**, then click on the lowest pitch of this area using the keyboard located at the bottom left of the screen. For reference, each A and the respective octave are indicated in the graph with scientific pitch notation, A4 is 440 Hz
- Select **Click to select highest pitch**, then click on the highest pitch of this area using the same keyboard
- Click on **Submit Range**. The range should now be indicated in the field **Area Range**
- Repeat for all the other areas

2.4 Soundcheck

The singer is positioned 50 cm from the audio device and sings the allocated tones as loudly as possible. As they sing these tones, the composer needs to keep an eye on the Current Vol., which should not be higher than -5 dB. Adjust the volume as you wish, but if it is set lower than this, the VMA might not be fully comparable with another singer's values. Note: adjustments to the volume should not be done mid-analysis.

3 Measurement

Video 3: Measurement: miika.info/VMM/videomanual3

It is advisable for the composer to run the application. Another person can also do this, in which case the composer can just listen to the voice and make notes.

Before starting, the singer should do a light warm-up to be sure that the VMA relays a representative image of their voice, and also to check that they can sing at the extremes of their voice in a healthy way.

As part of the analysis, the program automatically saves all the sounds. These may be helpful for the composer to reference afterwards. Before doing any actual measurements, the singer and composer should agree as to how the sound files can be used, whether they want these files to be saved permanently, how they are saved after the analysis, and for how long.

These AIFF files are designated with **area name** and **file number**. For example, the third recording of the area 'chest' would be chest3.aiff. A new file and file number is always created when the microphone is turned on, by pressing **MIC ON/OFF**.

The composer has the possibility to listen to the audio files after the VMA has been completed. Sound files can also be used to create a new Voice Map. To do so, open the first file by pressing **OPEN SF**, then **SOUND FILE ON/OFF**. Then go through all the sound files in the same way. For sound files, **SF VOL** is used instead of **MIC VOL**.

When all the areas are analysed, the Voice Map is ready and can be discussed. The Voice Map can be manually saved as a screenshot. The data can also be saved as a list: click **Read Max/Min** then **write** to save it as a text file. The software does not save these files automatically.

3.1 Singing

The singer should take a comfortable position facing the audio device, such that the distance between the device and their mouth is 50 cm. As it can be difficult to maintain this distance, it should be checked between measurement cycles.

The Voice Map shows how softly and how loudly each tone of each area can be sung. This is usually done by singing the scales, as indicated in 3.1.1. Using scales has proven to be a more natural way for the singer to proceed; furthermore, the data obtained is more reliable when the tones are sung numerous times.

The singer should sing vowel /a/ as per the International Phonetic Alphabet. This is an open front, unrounded vowel sound. At the extremes of the full ambitus, the vowel may change somewhat.

Each tone should be sung for roughly two seconds. The required length of the tone can be adjusted in the software, but shorter values often yield less satisfactory results and should only be used if really needed, such as for extremely high tones.

Typical order of a cycle:

- Singer and composer agree on which area to analyse and whether it will first be sung softly or loudly
- Area is chosen (**Active Area**)
- Singer indicates they are ready to start
- Composer gives the reference tones (a scale of three to five tones) using the on-screen keyboard or a physical keyboard, playing each tone for about two seconds
- Composer switches on microphone (**MIC ON/OFF**)
- Singer sings the tones at a similar tempo and with moderate legato

- Reference tones (within the area) and singing are repeated
- When appropriate, one of these options may be taken:
    - Singer signals they want to make a correction or repeat something
    - Singer or composer signals that they need to discuss how to proceed
    - Composer signals they have finished the area and may proceed to another
- Composer switches off microphone

In this manner, all the tones of the area are analysed as softly and as loudly as possible (but not necessarily in this order).

The composer observes the screen and can ask for any of the tones to be reanalysed. Sometimes, because of the vibrato or technical issues, the software may leave holes in the graph, which clearly do not represent the auditive reality.

Do not feel discouraged if the first few cycles aren't successful and need to be repeated, which could happen if the tones were sung too briefly or with too much legato between tones, for instance.

The singer should use their ordinary singing voice. Making the voice sound very breathy and noisy would not generate representative results and therefore should not be used in the VMA.

The program is provided with a setup for slow, fast, and adjustable vibrato (**SLOW VIB**, **FAST VIB**, and **CUSTOM**). Once one of those is activated, the software expects the tone to vibrate between tones at the given speed. A slow vibrato (100 pulses per second) is typical for a large and dramatic voice. A fast vibrato (300 pulses per second) is more characteristic of somewhat lighter voices. Moreover,

for many lighter voices that are trained in contemporary music, a full analysis could be completed without any vibrato adjustment. If the composer and singer are uncertain about which option to choose, they might try testing this with different tones to see which set can identify a test tone most reliably. The speed of the vibrato may also be tailored using the **CUSTOM** button, if the given options do not yield satisfactory results. Where the vibratos are rather wide, the software sometimes misinterprets the tone, reading it as the next chromatic tone. In most cases, this does not disturb the analysis, especially when it is done using scales.

3.1.1 Scales

In the testing phase, the most satisfying results usually come when the area is gone through using a sequence of two ascending and two descending whole tones, repeating this phrase a half-step higher, as in figure 2.

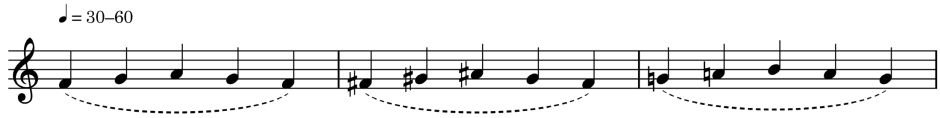

Figure 2. Using whole tones to go through the area

Sometimes the sequence can be transposed chromatically downwards instead of upwards.

At the extremes of the range, especially for the very high tones, it is often easier to sing chromatic scales upwards, as in figure 3.

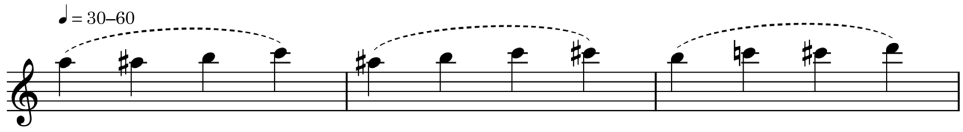

Figure 3. Using chromatic steps to go through the extremes of the area

If some of the tones are audible but the point does not appear on the graph, this

phrase should simply be sung again. Usually, going through the area using diatonic scales as seen in figures 2 and 3, yields good values across the full range, and the missing tones in the graph can then be sung individually.

While singing a sequence or a scale, the singer may use some legato, but if it goes too close to a *portamento*, the program will not be able to analyse the individual tones.

If the singer notices that the last tone they sang was not aesthetically solid, or if they notice that they sang a tone in a different area, the analysis should be briefly paused and the points on the graph referring to these tones should be erased (see section 3.2). The composer can also remark on differences in timbre.

3.1.2 Whole voice

Different voices function differently, so singers' preferences vary in terms of how to proceed within their vocal range. The following suggestions are based on the experiences of testing.

The area in the middle, or middle to low, is analysed first, thereafter continuing to the lower area(s), and from there ascending all the way to the highest area or areas, going through all of them .

For singing each tone of one specific area, two systems are equally popular: 1) going from low to high, and 2) starting in the middle of the area, going down and then all the way up.

Regarding the loudness of the voice, it is usually preferable to first sing as softly as possible and then as loudly as possible. But for the highest areas, it may be easier to sing loudly and then softer, but this sometimes yields unsatisfactory results.. If the singer finds that the very highest notes are only obtainable when

sung very loudly or the lowest notes extremely softly, then no minimum or maximum values need to be analysed.

The singer should always decide the order, as long as all the areas are gone through systematically.

Example 2 [hypothetical]: Analysing the four areas of the voice
(NB: The voice may also have more or fewer areas than four)

- Soundcheck
- Second lowest area
    - In the first cycle, the singer sings the tones too quickly and the software cannot recognise the tone
    - In the second cycle, the singer uses a very airy sound that doesn't represent their actual voice
- In the third cycle, the singer begins softly from the middle of the area, using a sequence of whole tones, first descending and then ascending. The softest tones can be produced so softly that they fall under the noise gate. This would be equivalent to the singer singing *niente.*
    - The highest tone feels different to the singer; it was already part of the higher register and gets erased from the graph
    - The loudest tones are sung ascending and then descending
- Lowest area
    - The softest tones are analysed from the middle of the area — first ascending and only then descending to the lower extreme

    - As the loudest tones are analysed, the dynamic is almost identical to the softest tones; this phenomenon is common to all voices and it is quite natural that the contours meet at the lowest pitch, since it can only be produced very softly
    - As the voice gets louder, the vibrato also widens; the software interprets some of the tones to be a halftone too high, but since all the tones are gone through with the whole tone sequences, these tones also eventually receive the right data

- Second highest area

    - The soft tones cannot be produced as softly as before
    - During the process, the software does not give any value to one tone in the middle of the area, although it is audibly present and part of the sung sequence
    - The cycle is restarted and the composer only gives the singer this missing tone to sing, which usually solves the problem
    - In the higher range, the singer needs to adjust the vowel somewhat, it is not precisely an /a/ anymore; this is for anatomical reasons only and does not adversely affect the analysis

- Highest area

    - The singer can produce the tones of this extreme area only in forte dynamics
    - Only loudly sung notes are analysed and the singer and composer discuss the limited dynamic range

3.2 Editing the graph

Sometimes a wrong tone or dynamic value needs to be deleted from the graph. This kind of error could be the result of background noise or the singer producing a note that is clearly in a different area, or other such causes. In this situation, pause the analysis cycle. Click **Erase point**, and select the pitch with the unwanted data using the keyboard. After this, you can choose to **Erase maximum value**, **Erase minimum value**, **Erase formant**, or **Erase all**. Then the analysis cycle can proceed.

4 Interpretation of the Voice Map

Video 4: Interpretation of the Voice Map: miika.info/VMM/videomanual4

The Voice Map is divided into two graphs, separated by a horizontal line. Together these graphs contain all the VMA data. The upper graph displays the full dynamic of the voice. The lower graph displays the singer's formant.[57] The x-axis indicates the musical pitch, and the y-axis indicates the dynamic (in dB) or the amount of singer's formant. The on-screen keyboard serves as a reference for reading the pitch.

All data points that belong to the same area are the same colour. On the upper graph, each tone has two dB values: the maximum level is indicated with a filled-in circle, and the minimum with a hollow circle. The tones outside the singer's range have a default value of -70 dB, which is seen on the x-axis (bottom of the graph).

The singer's formant appears as a solid line on the lower graph, where the same colour coding applies.

[57] Technically, the maximum dynamic from the range of the singer's formant.

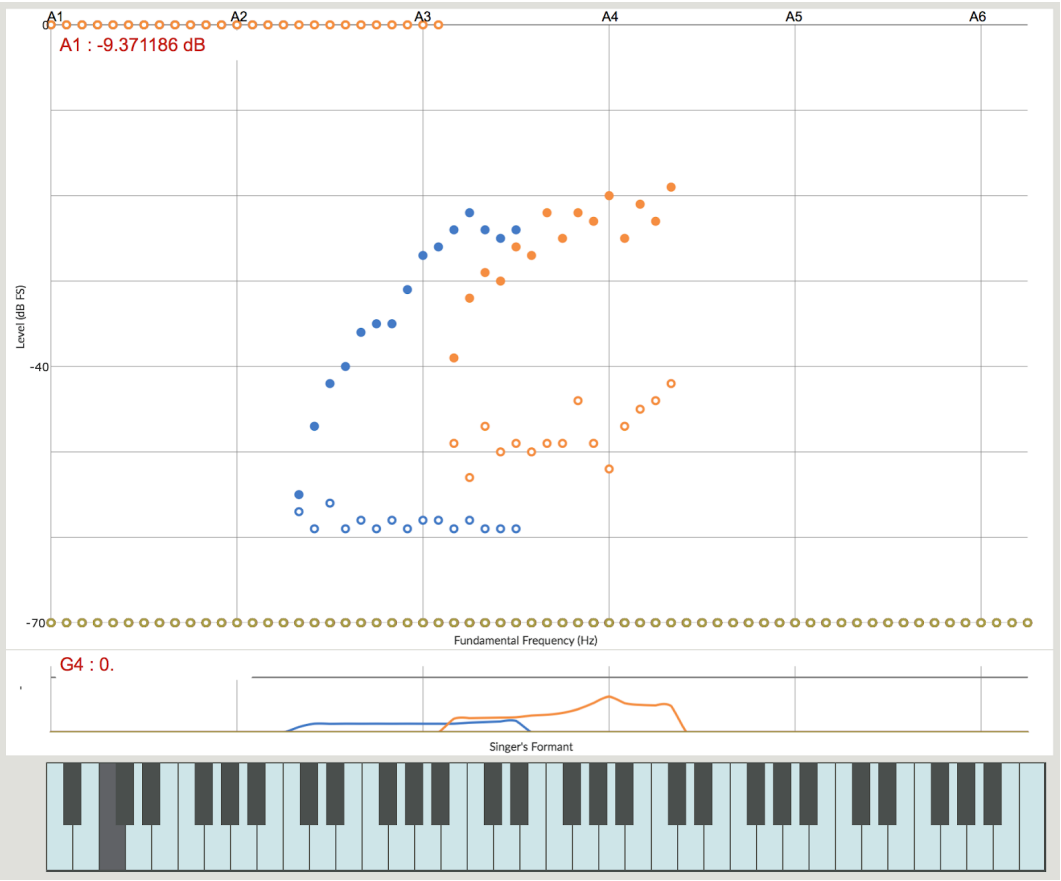

Figure 4 . Theoretical example of the Voice Map with two areas

These values give a good basic understanding of the voice, providing common ground for further discussion.

The composer should pay attention to the following aspects:

- For each area
    - Connecting each area with the memory of the timbre that was heard
    - General shapes of the contours

    - Comparison between the various maximum values
    - Addressing the minimum values similarly
    - -70 dB, the theoretical minimum value, indicates that the tone was not analysed
    - -60 dB is roughly the threshold of the background noise; tones with this dynamic can be considered niente

- Connection between the areas

    - Overlapping
    - Width of the areas

- For the singer's formant

    - Tones with moderately loud dynamics and a loud singer's formant can typically be heard through a larger orchestra than would be the case for tones with a soft singer's formant

Although the values are theoretically exact, the measurements relate to human beings and a person's voice alters slightly from day to day. However, the graph gives a very good idea of how the voice works, and changes of a few decibels are not significant.

Figure 4 is a theoretical example of a graph with two areas: blue and orange. The following information can be gleaned from it.

> **Example 3 [hypothetical]: Interpreting the graph in figure 4**
> The blue area ranges from C#3 to D#4. The minimum values of these tones are very low, close to the threshold of software analysis. The tones are going to be soft enough for practical purposes. For this area, the very first tones can only be sung softly, but from F3 onwards, the dynamic possibilities are rather open. The singer's formant does not make the voice more audible in the blue area.

The orange area has a range of B3 to C#5. This area needs a bit more energy, so the singer cannot sing it quite as softly as the blue area but still rather softly. Importantly, the minimum values get louder for tones A#4 to C#5. C#5, when sung as softly as possible (orange area), is approximately as loud as D#3 sung as loudly as possible (blue area). The maximum values of the orange area start from modest, at the bottom, and rise somewhat at the top of the area. The maximum is somewhat louder than the loudest maximum values of the blue area. Here, the singer's formant starts to have an effect. The dynamic maximum of A4 is only a little louder than C4, but A4 has a much more apparent singer's formant. This means that if a singer should sing a long tone accompanied by a chamber orchestra, the A4 would probably be audible, and C4 possibly not.

The tones that are part of more than one area are particularly interesting. In this case, B3 to D#4. This so-called passaggio is usually not very comfortable for singers and they need to learn how to mask the transition. From the composer's point of view, this can also be an opportunity: if these tones are reached from below, they are probably sung within the blue area and have its timbre. If the tones are reached from above, they can have the timbre of the orange area. In this case, the dynamic range of the blue area is much broader in the passaggio. If the composer wants to create a melody that ends with a C4 being sung loudly, the melody should preferably reach it from below rather than ascend to it.

The composer should always hold a discussion with the singer before composing for special parts of their voice, such as how to achieve very high and low tones, or tones that are part of more than one area. It should also be remembered that the VMA provides information on the limits of the voice in a closed analysis situation, and each analysis produces the tones only briefly and possibly just once. Usually the analysis situation informs the composer of how tiresome the tones are to produce. This should nevertheless be discussed in further detail during the next step, the List of Good Questions.

5 List of Good Questions

The VMA and the Voice Map serve as a starting point for communication. They provide auditory and visual information that help the singer and composer identify certain phenomena and talk about it. The questions are based on info obtained from literature and on the experience of the participants. The questions are directed at the singer, but should be seen as themes that invite open dialogue. Questions 1-8 stem from the Voice Map, continuing to the more general aspects of the voice, and eventually to the more holistic artistry of the singer.

Questions to address immediately after the analysis:

1. Can you see yourself in the Voice Map? What aspects of your voice does it accurately represent and what is not shown? Was the way your voice sounds today representative of what it typically sounds like?
2. The Voice Map emphasises the extremes of the voice. What pitches and dynamics do you consider to be:
   a) the main part of your voice (*tessitura*)
   b) the extremes of your voice that are usable occasionally and briefly
   c) very unreliable and only usable on special occasions, such as in recordings
3. The analysis was done with the vowel /a/. How does the text change at the extremes of your voice? What part of your range is optimal for text production and what part makes the text understandable? What kind of vowels would help to produce the tones at the extremes of your voice?
4. Are there some parts or points (pitch and dynamic) that are especially difficult or tiring for your voice, for example the *passaggi*?
5. Are there some parts of your voice or certain vocal effects (experimental singing techniques, laughter, tongue click, etc.) that you find especially interesting and would like to develop further?

6. What kinds of musical textures and performative styles are you especially used to and find easy to learn and produce? These might entail very long legato lines, fast coloraturas or jumps, passages with demanding intervals such as microtones, or extreme physicality and performance skills on stage
7. What musical textures and performative styles are least comfortable or demand a lot of time and energy to achieve?
8. How would you like to continue the artistic process? [This question is for the singer and the composer]

Questions to address during the artistic process:

It is highly recommended that the composer and singer meet after the analysis situation and before the first performance. Before the second of these meetings, the composer should re-read their notes about the voice and preferably listen to the recordings created by the VMA software, as these might aid the composition process.

The remaining questions relate to a situation whereby the composer already has some material that can be tried out with the singer and there is still time to make changes.

9. How does it feel to sing the material? Is it tiring for the voice? Does it need to have more pauses? How do you find the length of the piece?
10. Are there any specific sections where the text, dynamic, pitch, and phrasing could be improved to become more idiomatic or better suited to the composer's musical or dramatic idea? (Of course, the dramatic concept sometimes requires the music to be unidiomatic)
11. How shall we proceed from here?

This manual is for use in applying the full VMM, but the List of Good Questions alone

could also prove a useful as a starting point for communication, after which another such system would be required if the composer is to obtain a good understanding of the timbre and demands of different parts of the singer's voice.

6 Safe use of the program

To ensure the safety and comfort of the singers, it is advisable to follow these instructions.

6.1 Vocal health

The VMA requires the singer to sing at the extremes of their dynamic and pitch ranges. The analysis should not be done if the singer feels that their voice is not completely healthy. If the normal functionality of the voice changes, the analysis should immediately be halted. The VMM has been tested with classically trained opera singers, so to utilise it with other people, such as but not limited to untrained singers, may cause unforeseeable problems.

The VMM is devised to enhance the process of creating vocal music that fits the singer. For many reasons, the results may still be unsatisfactory and the singer and the composer should keep on communicating during the composition process and before the performance. With all vocal music, including music composed using the VMM, the singer should be aware of how it could affect and strain their voice.

During the testing phase, none of the singers mentioned that the VMA had caused excessive stress to them or to their voice. Many of the singers mentioned that they discovered new and inspiring aspects of their voice. According to the feedback on compositions created during the testing phase (*Voice Box*, *NOS*, and

Voice is Voices), they are very idiomatic for the singers who premiered them. This is also the goal of the process.

6.2 Data storage

The program saves all the audio that was used to make the analysis. It is saved in the form of audio files in a folder automatically created at the beginning of the VMA process. Before the analysis, the composer and singer must decide:

- whether they both agree to save the analysis files
- how the recordings, numeric data and screenshots are saved
- whether the recordings are destroyed afterwards, and if so, when?

7 Legal disclaimer

Legal disclaimer: Miika Hyytiäinen and the University of Arts, Helsinki are not responsible for any damage caused by the VMA or for the performance of a composition created using the VMA. Due to the limited and very specific test group, the VMA should not be used for medical diagnosis or as a tool for creating a Voice Range Profile.

Miika Hyytiäinen and the University of Arts, Helsinki are not responsible for inadequate use or storage of any of the data that is collected by the VMA software.