

Gestures in Vocal Performance and the Experience of the Listener: A Case Study of Extra-Semantic Meaning-Making in the Singing of Olavi Virta

Author(s): Marko Aho

Source: *Popular Music*, Jan., 2009, Vol. 28, No. 1 (Jan., 2009), pp. 33-51

Published by: Cambridge University Press

Stable URL: <https://www.jstor.org/stable/40212425>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



Cambridge University Press is collaborating with JSTOR to digitize, preserve and extend access to *Popular Music*

JSTOR

Gestures in vocal performance and the experience of the listener: a case study of extra-semantic meaning-making in the singing of Olavi Virta

MARKO AHO

Folk Music Institute, Kaustinen, Finland
E-mail: marko.aho@kaustinen.fi

Abstract

At the level of micro-intonation, singing provides an object for intimate corporeal identification for the listener. In fact, the listener can, by way of vocal identification, achieve a direct corporeal relationship with the vocal performance. Similarly, the researcher can use his/her own purposeful identification and imitation as a means of observation. 'Once in a century' is a description often attached to the legendary singer Olavi Virta (1915–1972), who had a significant impact on Finnish popular music during the 1950s and is considered a cultural icon by many Finns. In this article, the vocal performance on the novelty recording 'Luonasi jos oisin' ('If I were there with you')/Isn't it Romantic?' is studied in order to map out its gestural content. Olavi Virta was in possession of a rich arsenal of vocal gestures. The gestures shared a relationship with the semantic content of the lyrics – not always mimetic but often contrapuntal, polarising, or estranging.

Introduction

The singing voice is one essential aspect of the *performance* of music. In analysing performance, the traditional musicological methods – relying on metaphors to depict the organic *form* of music – are of little help. The awkwardness of trying to force an art form which is essentially both audible and temporal into a frozen visual representation (a standard procedure of music analysis) has perplexed researchers of popular music. This is because in popular music, most often the importance lies not in *what* is being performed, but in *how* it is being performed. Some writers in popular music studies have sought inspiration from the history of the aesthetics of music, turning their attention to thoughts prior to the autonomic and organisist view of music which has dominated since the age of romanticism. These writers have instead championed the idea, deriving from the classical era, of music as rhetoric (e.g. Brackett 1995, p. 25; Frith 1996, p. 259). Here, music is understood to be oration, aiming to affect the listeners. Rhetoric is a means used to make the listeners react, take interest in and identify with what they are hearing. What is important is the concrete surface, not the abstract substructure. The most far-reaching implication of such a view is that

analytical attention should also be targeted on the musical experience of the listener.

The singing voice, in some sense the most intimate¹ means of musical expression, provides us with a fruitful context for studying the experience of listening to music. During the past decade, most of what has been written about singing in the new musicological and popular music studies contexts has been anchored to the corporeal listener (e.g. Poizat 1992; Kostenbaum 1993; Dunn and Jones 1994; Potter 1998). In most of these cases, writers have sought alternatives to the conventional way of approaching singing purely as a physiological, anatomical and acoustic phenomenon.² In this article, the central epistemological question is: how do we gain knowledge of expression in singing outside the semantics of the lyrics? A solution is found through an approach which relies on listening and mimicking as means of analysis. This article aims to gain insight into the modes of extra-semantic meaning-making one may encounter in crooning-style pop singing. An analysis based on listening and mimicking might be considered as an obligatory preliminary step towards more rigorous procedures of measuring and observing. Although visual representations are not the object of analysis here, visualisations will be used for illustrative purposes. In addition to the illustrations, the music itself is also made available to the reader (see information concerning this at the end of the article).

'Once in a century' – Olavi Virta

If one had to name the '*essential singer*' in the history of Finnish popular music, without doubt it would be Olavi Virta (1905–1972). This singer of Finnish twentieth-century schlager-music remains an icon today, fifty years on. Newcomers to the scene are still compared to him, and, without exception, unfavourably. Although the 'Finnish Carlos Gardel', as he has sometimes been referred to, is first and foremost remembered as the singer of the Finnish version of the Argentinean Tango, he also recorded waltzes, American show tunes and cover versions of popular Italian tunes during his illustrious career. His fans have cherished the idea that if there ever was a Finnish popular-music singer who was of genuine world class and could match the performances of Gardel, Frank Sinatra and Edith Piaf, it would have been Olavi Virta.

In 1998, the Finnish Broadcast Company, in its popular weekly radio show, *Iskelmäradio* ('Schlager-radio'), dedicated a special programme to Olavi Virta ('Olavi Virta's path to the tango-kingship of Finland'). During the broadcast, the host Maarit Niiniluoto posed this question to listeners: Why do you admire Olavi Virta? Some examples of the answers sent to the programme are reproduced below:³

His sonorous, beautiful, masculine, homely, affective and technically skilful voice affected my emotions.

I get goose bumps when I listen to him: the quality of the voice is so lush, that you feel as if you were sinking in it.

A charismatic legend; an interpreter of big emotions; unforgettable and the only one that makes the tango touch you deep down in your heart.

Something very peculiar in his voice makes the listeners feel every time that the song is just for them personally.

Olavi Virta had a soft and pleasant voice and [there was] a great emotional charge in his songs.

As an interpreter of tango, a beautiful and trembling voice, which no one has been able to emulate.

A pure, distinct, pleasant, awesome, fantastic voice.⁴

If we take these testimonies seriously, then the obvious place to search for the secret of Olavi Virta's lasting success is in his *singing voice*. In fact, most of what has been written about him points to the same thing – 'it's in his voice!': 'masculine', 'warm', 'sensuous', 'brilliant', 'coming straight at you' – these are the attributes that time and time again the media⁵ and his fans attach to his voice.

What exactly do his fans mean in these descriptions? Their accounts seem to suggest that in his singing Virta interprets the ideas and the atmosphere of the lyrics in an extraordinarily animated way. Furthermore, a recurring observation by listeners is that they get the sensation that he is singing to them personally. What could it be in Virta's art that brings out these intimate feelings? Could we approach his singing voice and its effects on the listener in more detail? If the singing voice is the explanation for the allure of Virta, then *what exactly is it* in his voice that makes it so special?

A great interpreter of songs – but what is interpretation?

Why do telemarketers have that smile when they talk, even though no-one is there to see? Why do dogs understand the commanding tone in a voice even though the semantic meaning is, in most cases, incomprehensible? Why do some dogs fail to obey if the voice sounds insecure? Inter-human, and indeed, human-to-animal communication is partly based on the ability to convey and interpret information through the movement and postures of the body and the face. Similarly, our voice communicates feelings, attitudes and gestures, whether we want it to or not.⁶ Speech carries within it the semantic meaning of the words; it also gives out signals so that the receiver can detect different kinds of tacit messages. The emotional state has a significant effect on the vocal delivery of the speaker. Subtle variations in, for example, prosody, length and accentuation of the syllables, and breathing, all convey information outside the semantics of the utterances. This communication is primordial. Although culture regulates all human expressiveness, research has suggested that the vocal utterances of the basic emotions (sorrow, fear, joy, hate) are not only trans-cultural but, to some extent, trans-species⁷ (Laukkanen and Leino 1999, p. 98).

All the parameters of speech are present in singing. If the abstract structure of music, the *composition*, determines the macro-intonation of a musical performance, then the *interpretation* of that composition takes place at the micro-intonation level, limited by no more than the boundaries given in the composition and the general features of a given musical style. At the level of micro-intonation, singing provides the listener with a remarkable object for intimate identification. Indeed, the listener can, by way of vocal identification and in some cases by concretely singing along, achieve a direct corporeal relationship with the vocal performance (e.g. Frith 1996; Cubitt 2000). Similarly, the researcher can also use his/her own purposeful identification and imitation as an instrument for observation. The researcher can assign his/her own body as an 'experimental tool' in order to obtain unmediated insight into the poorly verbalisable corporeal dimensions of music reception. Tarvainen classifies *emphatic listening* as a manner of analysis, where bodily identification is fundamental:

The researcher can listen to the audible research material with her or his whole body. This means that he or she does not strain her- or himself to listen to the features that have been decided beforehand, but instead lets her or his own body find a contact surface from the material. It is often said that singing and music in general can affect people. To acknowledge this motion in oneself, and to use this experiential knowledge purposefully in analysis, is the essential (of emphatic listening).⁸ (Tarvainen 2004)

The method of analysis employed in this study has been guided by this concept of emphatic listening (see also Tarvainen 2006a and 2006b). Accordingly, this study will focus on what is most often termed a 'gesture'. Ekman (1977, p. 39) acknowledges, after extensive research on kinesthetic communication, that there is no perfect common expression for all the phenomena in non-linguistic communication. Nevertheless, the concept of gesture might be the most suitable to indicate all non-linguistic messages sent by a human with his/her body. A gesture can be a movement, posture of the body and/or a facial expression made either in association with words or independently. According to Mead (Coker 1972, pp. 10–11), gesture is the basis of and precursor to language, and acts as a stimulus in adapting the behaviour of one organism to match the behaviour of another. A gesture is always accompanied by an emotional stance. The nature of the reception of a gesture is non-reflective and immediate; if a gesture is preceded by a conscious effort to generate it, it is already language which, in turn, is tied to a social consciousness of attitudes of 'others' towards attitudes and gestures of 'mine'.

According to his fans, Virta is a great 'interpreter of songs'. The hypothesis of this study is that he either gestures within the micro-intonation of the performance – appositely so – whatever the denotations and connotations the macro-intonation of the song (in the singer's case it is chiefly the melody of the vocal part) might carry, or the semantic content of the lyrics. Here, the focus is on the latter, a less problematic choice than the former as one would need to examine all possible denotations and connotations of the melody and harmony of the song to begin with. This article, therefore, poses the question: does Virta express through his gestures the semantic content of the song lyrics?

The analysis involved two phases. First, the author listened to the vocal performance of the song phrase by phrase and noted all gesture-related features of each phrase. Second, the author mimicked these features – the validity was checked by fellow researchers informally (the facial photographs below are of the author). To enhance the reliability of the analysis, the reports of six test-subjects, who were given the task of making notes on all the gestures they heard in the vocal performance on the original recording, were analysed.⁹ The interpretations of the gestures were formulated with precision in mind and in an attempt to minimise the unavoidable subjective nature of these descriptions. The analyses aimed to discover if *something significant was taking place at the micro-level*. At this stage, only secondary attention was given to the consistency of what was reported by the participants in the study.

The study focuses on a novelty recording which Virta made in Stockholm in 1955, a mid-tempo ballad called 'Luonasi jos oisin ('If I were with You')/Isn't it Romantic' in Stockholm. This song, by Richard Rodgers and Lorenz Hart, originally featured in the motion picture *Love Me Tonight* (1932). The Finnish lyrics are faithful to the mood of the original English lyrics, and were written by Sauvo Puhtila, the most successful lyricist in the Finnish popular-music industry during the 1950s. The recording was not intended for commercial use, but as a gift for his wife Irene, to whom he dedicates the song during the introduction. It was only published in 1992 on a CD compilation of all Virta's known recordings. He sings the song first in Finnish, then in English, on top of a plain piano accompaniment by the conductor of the recording session, Georg de Godzinsky, Virta's compatriot. This recording of Virta's singing, close-microphoned and with a quiet piano backing, provides us with the opportunity to become absorbed in his micro-level means of expression. First, however, let us familiarise ourselves with Virta's singing at a general level.

'Full-tension singing'

Peter von Bagh, a renowned Finnish popular culture writer, probably best summarises it when he said that while some might secretly imagine themselves singing like Reijo Taipale (another Finnish veteran star), no one even dreams of singing like Virta as such technical command is communicated in his voice. Nevertheless, there is no fundamental *stylistic* difference between Virta and other male singers of the 1950s and early 1960s. In Finland, genuine operatic voices in the fashion of Mario Lanza (e.g. Mauno Kuusisto or Gerorg Ots¹⁰) and numerous lesser talents who adapted the masculine baritone-sound during the tango-craze of the early 1960s, and used a full-bodied vocal delivery. This style has survived at the margins of Finnish pop to this day, largely thanks to the annual televised tango-singing contest which offers the possibility of a major break-through for young aspiring Finnish Schlager-singers. Technical ability, however, is only one aspect of Virta's singing. His apparent technical flexibility allows him to easily master all styles of the 1950's dance-music genre.

In the recording 'Luonasi jos oisin/Isn't It Romantic?', Virta utilises the *crooning* style of singing, the American popular music male voice standard from Bing Crosby onwards. The advent of the microphone in the 1930s enabled much wider dynamics in singing than previously. Consonants could be more pronounced as unvoiced sounds were clearly recorded. In this particular performance, there is more variation in micro-intonation, dynamics, phrasing and placement of resonance than in most of Virta's singing on record. He varied his vocal delivery within certain limits for different dance rhythms and sub-styles of pop songs, so that he would, for example, sing tangos with a rather heavier mode of vocal delivery than he uses for this delicate ballad.

The performance is very dynamic, containing sudden explosive fortes at the beginning of phrases, which are followed by fast and wide decrescendos, spiced up with a 'quivering' sort of vibrato. In this song, the highest tone is D, which is nowhere near Virta's top register. Virta's upper register was genuinely that of a tenor; loud and rich in the upper partials where the top notes of most pop singers are distinctively 'thin'. This 'full-tension singing', as Virta himself called it, points to the fact that he sang with a considerable sound pressure and combined a well-supported chest voice with some brilliance-giving head resonance. Again, this is a feature of Western art singing, which Virta does not assign to every note, but which is present in the performance under study. Virta does not use the full operatic voice at any point, but always keeps it tilted towards the crooning style. On occasion, Virta thins down his voice. One mannerism of the swing period of post-World War Two Finland, as elsewhere, was to thin down the voice quite pointedly in the upper register: 'The soft whining . . . was apparently well suited for the sentiment of the period, when the male voice was preferably a bit syrupy' (Jalkanen and Kurkela 2004, p. 365). Virta, however, abandoned this way of singing at the beginning of the 1950s. One reason for this might well have been that he joined the highly successful singing quartet *Kipparikvartetti*, in which the other members were classically trained; another could be his numerous roles in musicals. Whatever the reason, the turning point was around 1952:

The crooning-humming is gone, and Virta's sound has clearly a wider palette of nuances than before. Long notes come with a natural-sounding, at times a very wide vibrato, and the *fortes* sound brave, but however remain smooth. Virta still remained a microphone-singer in the sense that one cannot find art-singing mannerisms in his delivery. Even the performances of some classic tangos and operetta-pieces present a different kind of grandiloquence than we have been used to in stage-singing. For softer passages Virta still produced intimate nuances with a skilful

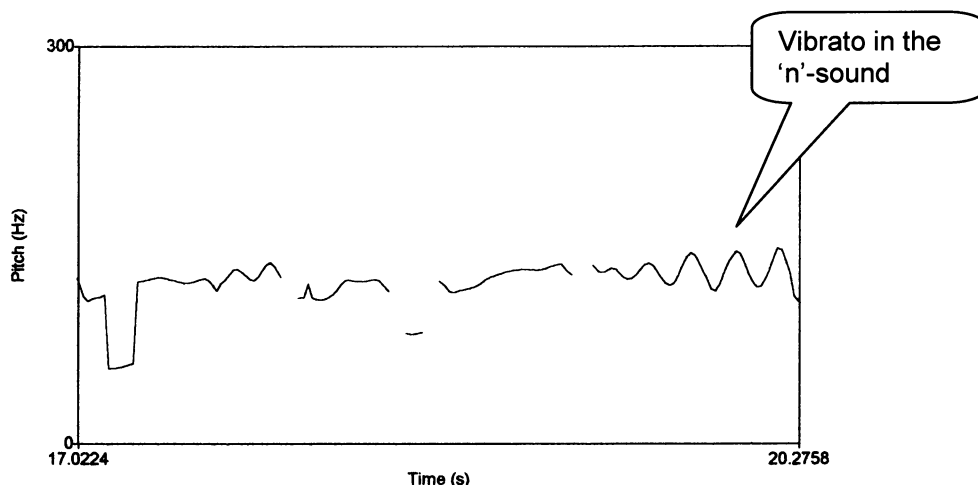


Figure 1.

use of the microphone. The pronunciation of lyrics is exemplary in its clarity, and the voicing is not aided with vocals that differ from those of spoken language, which often is not the case in classical singing. (*ibid.*, p. 425)

The aforementioned vibrato in the vowels starts to build up very quickly, and has time to light up even in very short voiced sounds. The vibrato is quite even in amplitude and frequency, which can also be registered visually in the graph made with the PRAAT-sound-analysis program¹¹ (see Figure 1).

In 'Luonasi jos oisin/Isn't It Romantic' the phrases tend to begin with a little glissando, which are also abundant elsewhere. Due to the glissandos and the wide vibrato, the pitch may become unstable at times without ever sounding out of tune – except for one flat note, a mistake that would have meant a new take during a normal session, but which was left here possibly to embellish the already quite spontaneous-sounding performance. The voice seems to have a soft attack, a feature which might explain the impression of 'softness' which is present in portrayals of Virta's voice. The 'l'-consonants are voiced, long and articulated with the tongue brought distinctly to the front to further increase the smoothness of the overall voice quality. Articulation is explicit and gives the listener a fair opportunity to listen to the semantic meaning and narrative progress of the lyrics. In the chorus, which is sung in English, the pronunciation is executed with a rather heavy accent, and includes clear mistakes. Extra noise and impure phonation are scarce; merely one explosive glottal closure in the first measure of the chorus in Finnish, and a slight crack in the voice in measure 25 of the chorus in English.

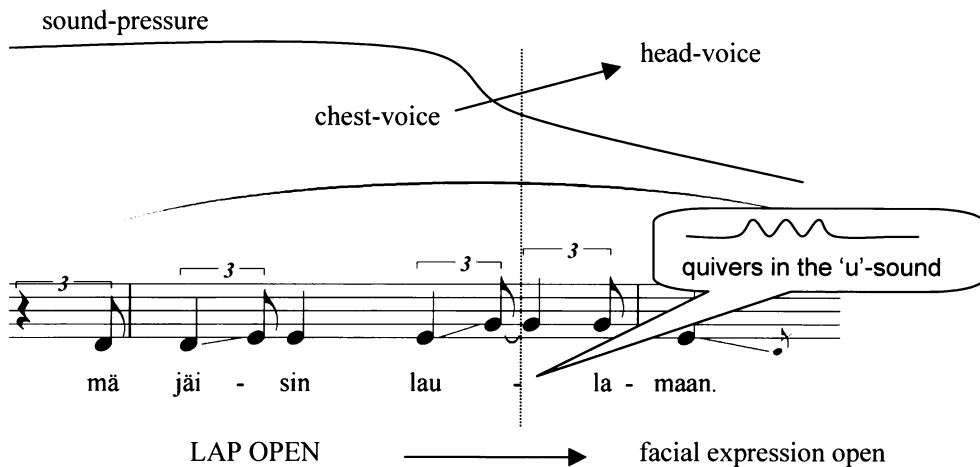
The phrasing breathes, both gaining and lagging in turn. In this sense, Virta's singing is 'swinging', thus creating a relaxed general feeling to the performance. Because of the constant rhythmic fluctuation, the tempo seems to change throughout, and this in turn gives the performance an occasional rubato character. All in all, Virta's rhythmic handling of the song is very relaxed and free. At the basis of the phrasing there is swing and although the rhythm might occasionally resemble even eights when notated, the underlying idea is 'swing-eights' triplets without the middle note. For example, in the first phrase of the song, the syllables 'luo-na-si jos' approximate a triplet rhythm, but both '-na' and 'jos' gain so that the result is somewhere between

You gave it to my heart, although I did not
know then.
I did not believe when you swore, I would be-
lieve now.
If I were with you,
I would not hope for anything else in the world.
If I were with you,
You protecting me, I would not be scared of the
night.
I would also allow you a kiss.
Despite I would be very small and fragile,
My dreams would be fulfilled.)

<i>Isn't it romantic?</i>	a	pick-up measure – 1. measure
<i>Music in the night, a dream that can be heard.</i>		2–4
<i>Isn't it romantic?</i>		4–5
<i>Moving shadows write the oldest magic word.</i>		6–8
<i>I hear the breezes playing in the trees above.</i>	b	9–12
<i>While all the world is saying you were meant for love.</i>		13–16
<i>Isn't it romantic?</i>	a	16–17
<i>Merely to be young on such a night as this?</i>		18–20
<i>Isn't it romantic?</i>		20–21
<i>Every note that's sung is like a lovers kiss.</i>		22–24
<i>Sweet symbols in the moonlight,</i>	b'	25–26
<i>do you mean that I will fall in love perchance?</i>		27–29
<i>It's romantic (actually Isn't it romance?)</i>		30–31

The basic unit of analysis is a phrase the length of one exhalation. Initially attention was given to details such as agogic variation ('phrasing'), the semantic meaning of the words, significant errors in pronunciation, noise, imperfect phonation (glottal closures, clicks of the tongue, creaky sounds, leaks), micro-intonation (vibrato, glissandos), dynamics, breathing and placement of the resonance. However, it soon became apparent that instead of going through the whole song mechanically – parameter by parameter – a better approach would be to go through each phrase individually and seek out the parameters that intuitively felt significant. In the examples below, the details that surfaced in the analyses are highlighted by notation examples, graphics, and verbal descriptions. With the aid of the PRAAT-sound-analysis programme, governing the multitude of small sound-excerpts was easy and illustrative graphics could be produced.

With graphics, however, it is possible to point to only some measurable features *correlating* with the real point of interest – the gestured content of the singing voice. They are not able to specifically illustrate which features of the sound as a measurable physical object make up a gesture. For example, while an open vowel might be analogous with 'arms opening wide', it does not reveal this gesture in particular as an open vowel might occur with many other gestures too. We do not have the technical means to directly extract gestures from sounds. What we can technically achieve, however, is to point to exact places in the temporal stream of the measurable sound. At a general level, we can also illustrate what takes place in the physical sound in connection with a found gesture. These features in the measured sound might be



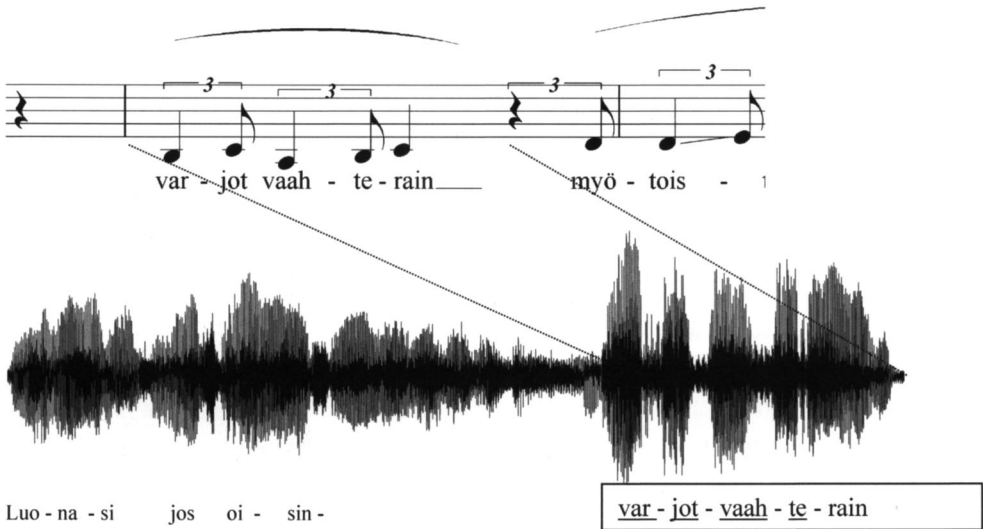
Example 2. 'Mä jäisin laulamaan' ('I would stay singing'); measures 3–4.

conceived, at best, as *traces* of the gestures. Notation and graphics are thus used to indirectly locate any gestures found in the stream of music; the graphic aids can then be used to guide the listening of the sound files which are presented on this article's website.

The first gesture-notable moment arrives with the third phrase in the first sentence of the lyrics (see Example 2): 'mä jäisin laulamaan' ('I would stay singing'). The phrase begins with a chest-swell, then a sudden decrescendo which soon changes into a thinner head-voice. Simultaneously, as the point of resonance ascends from the chest to the mask¹³ and the pressure drops from *forte* to *piano*, a sudden change in the level of the gesture occurs. One of the test people commented: 'relaxes and eases up, the singer is maybe leaning backwards and opening his arms'; 'a slight movement of the singer's right arm'. The gesture could be best characterised as a kind of *opening of the lap*. Until the dotted line in Example 3, one may sense the arms opening wide. The same gesture recurs in measures 23–24. After the dotted line, what becomes open is the *expression on the face*, as the voice becomes brighter and the eyebrows rise momentarily. In the middle of the decrescendo, with the U-sound in the word *laulamaan* ('to sing'), a vibrato lights up, giving Virta's voice its typical 'quivering' quality.

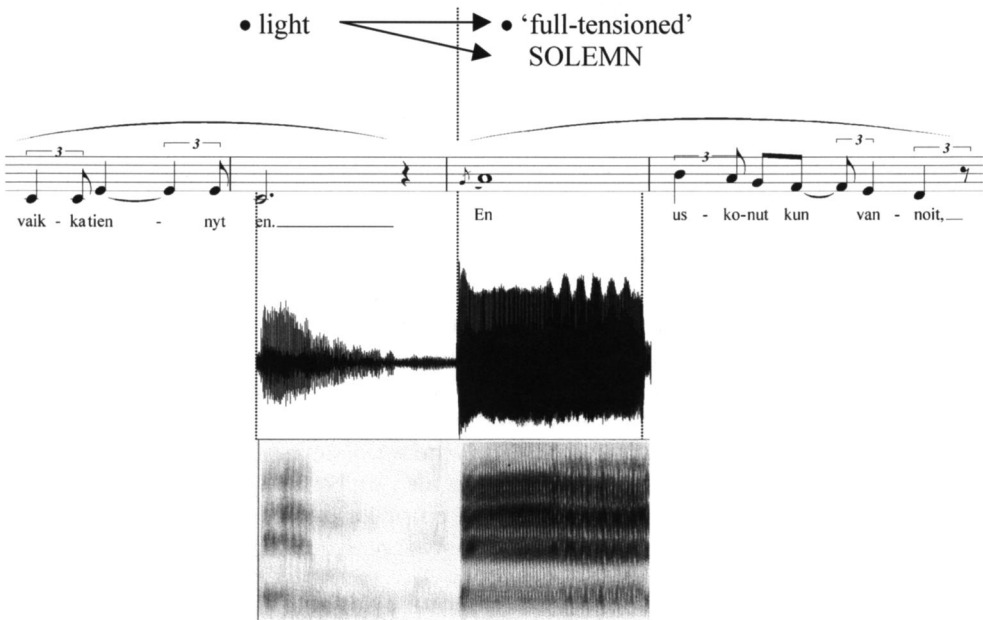
An emphasis on the words *varjot vaahterain* ('shadows of the maples') shows itself in the graphic description of the amplitude changes (see Example 3); an even stress on each syllable can be seen as well as heard. Compared to the words in the immediate vicinity, the stressed syllables are also delivered with a slightly deeper positioned voice and thus become even more emphasised. What is the effect created here? 'An air of self-confidence' at the very least is present, as one test subject testified. The tone of voice in question might be described as *convincing*. With this sincere tone, a masculine voice in a radio advert might try to convince the listener to buy insurance or make a safe and sober decision: 'ACME insurance' is a *reliable choice*. Note-worthy bodily or facial movements are not included; the face assumes a positive, basic expression as it would if the person is trying to convince someone.

A frequent expressional device of Virta is a sudden change of force between two consecutive phrases (see Example 4). This change of force is accompanied by an analogous change of facial expression; here, with *En uskonut kun vanhoit* ('I did not believe when you swore'), the facial expression changes from a faint smile to solemn



Positioning of the voice:

Example 3. 'Varjot vaahterain' ('*shadows of the maples*'); measure 6.



Example 4. 'En uskonut kun vannot' ('*I did not believe when you swore*'); measures 13–14.

('a rigid gesture, certain seriousness'; 'bulging of the chest, face turning into a slightly agonised expression'). 'The singer has a fierce look on his face . . .' as the delivery becomes 'full-tensioned', as the common epithet for Virta's singing goes. The abruptness in the change in force is clearly illustrated in the graphic illustrations of both

The image shows a musical score for a vocal line. The melody is written on a single staff with a treble clef. It begins with a half note 'Myös' followed by a series of eighth notes. There are three triplet markings (indicated by a '3' in a bracket) over groups of eighth notes. The lyrics 'suu - kon sul - le soi - sin,' are written below the staff. An arrow points from the word 'suu' to a black and white photograph of a person's mouth, which is pursed together in a 'kiss' shape.

Example 5. 'Suukon sulle soisin' ('I would allow you a kiss'); measure 26.

The image shows a musical score for a vocal line. The melody is written on a single staff with a treble clef. It begins with a half note 'täyt' followed by a series of eighth notes. There are three triplet markings (indicated by a '3' in a bracket) over groups of eighth notes. The lyrics 'tyis - u - nel - main.' are written below the staff. Below the score are two black and white photographs of a person's mouth. The first photograph shows the mouth open with the corners turned downwards. An arrow points from this photograph to the second photograph, which shows the lips pursed together.

Example 6. 'Täyttyis unelmain' ('my dreams would be fulfilled'); measures 30–31.

amplitude and spectrum in the section where the word *en* ('I do not') is repeated in consecutive phrases.

The 'U' sound becomes more forward and the lips purse to a 'kiss' as Virta is mimicking with his lips the semantic meaning of the word *suukon* ('a kiss') (see Example 5). Ekman (1977, p. 49) calls this kind of elucidating gesture a *kinetograph* – movements that show some function of the body. Something rather similar is taking place at the end of the long diphthong in the word *unelmain* ('my dream') in measures 30–31 (see Example 6) – the corners of the mouth turn downwards and for a fleeting moment, it feels as if Virta is about to burst into tears. In each case, some kind of 'getting seized by an emotion' is what is effectively communicated here.

'Isn't it romantic?' is the first line of the chorus in English. The voice here is somewhat thin and colourless. The effect created is hard to attach to any specific gesture, but pronounced pursing movements of the lips are nevertheless communicated ('eyes are lashing, chin is protruded, lips are pursed'; 'lips are extremely pursed . . .'; 'A small grin at the end of the word "romantic"'). It is difficult not to let a cultural stereotype, a certain character, creep into the mental image, so strong is the image of an erect little finger, subtly pursed lips in a mischievous smile – a habitus of a continental *debauché*. The impression culminates in the accent on the syllable '-tic' (see Example 8).

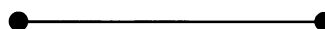
Micro-level: \rightarrow $>$

Isn't it ro - man - ti - c

Is - n't it ro - man - tic? _____

Example 7. 'Isn't it romantic?' (pick-up measure – 1. measure).

— a dream — that ca - n't be heard. —



'headshake'

Example 8. 'A dream that can't be heard' (measures 3–4).

'Eyebrows elevated, chin elevated' and 'tilts his head backwards, smiles' were two comments for the phrase 'a dream that can't be heard' (see Example 8). The orientation of the face seems to rise and one may sense a slight shake of the head. It is possible to detect a gesture here – a sensitive, friendly shake of the head, accompanied by a smile, as if someone is giving a negative answer to a question to which the inquirer well knows the answer. The gesture falls on a negative utterance, as Virta is articulating the original 'can' word of the lyrics in its negative form. As the word 'hear' arrives (see Example 9), and as the upward glissando is united with a slight rise of the point of resonance, one can easily visualise Virta's eyes closing briefly. Here, a 'dreamy' expression ('[on the word "hear"] the chin first rises a little, then falls back') seems to capture in some sense the dreamy atmosphere of the lyrics of the phrase and the whole sequence.

The exact gesture in measures 13–14 of the English chorus ('while all the world is saying', see Example 10) is a whisper. However, this phrase is also an example of skilful microphone techniques which has its effects on the listener: 'intimate atmosphere as the voice gets whispery'; 'sensitivity and humbleness'; 'eyes closed, a fragile expression, like in love or something'. Virta drops the sound pressure and phonation here so that the vocal cords are merely vibrated as he moves towards the microphone and onto the 'skin' of the listener. This is a conscious technique, a feast of crooning. The lyrics are meant to be intimate: 'While all the world is saying: you were meant for love'.

Isn't it romantic?
Music in the night, a dream that can be heard.
Isn't it romantic?
Moving shadows write the oldest magig word.
hear the breezes playing in the trees above.
While all the world is saying:
You were meant for love.

“Just as if the eyes were closing briefly”.

Glissando + resonance

Example 9. ‘I hear the breezes playing’ (measures 9–10).

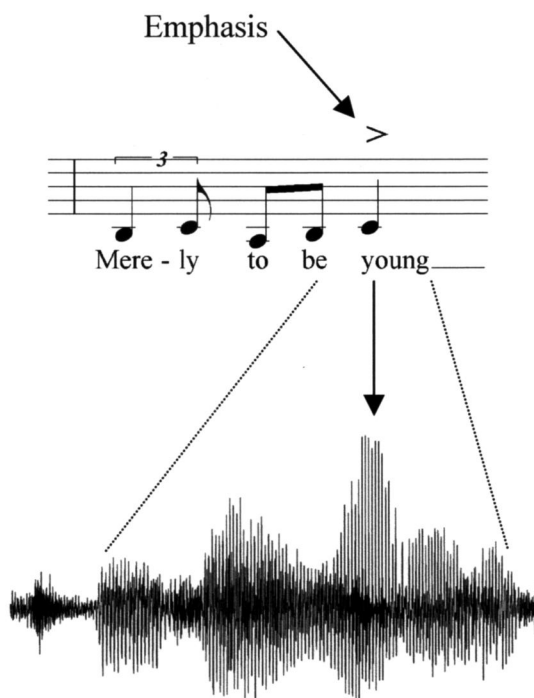
‘whisper’

While all the world is say - ing

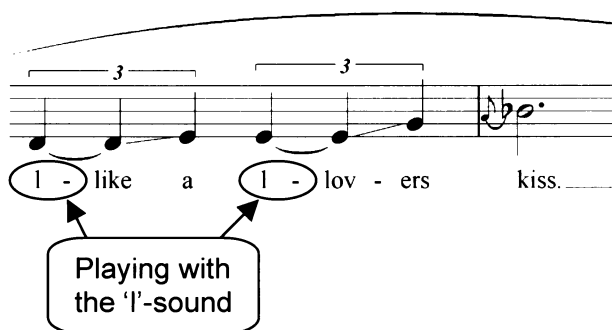
Example 10. ‘While all the world is saying’ (measures 13–14).

‘Young’ is the pivotal word of the phrase ‘merely to be young’ (see Example 11). The word is emphasised by hand-movement: ‘gesturing with his right hand, pointing at something, you get the feeling he is underlining the meaning of his words’ was one test-subject’s characterisation of this phrase. This is a brisk gesture, possibly made with a closed fist, as if to convince oneself – a gesture which might take place when reaching a conclusion after long consideration and with a sense of empowerment, as if one is suddenly absolutely sure: ‘Yes, I *can* do it’. This movement, made mostly with ‘musculus pectoralis’ and hands slightly bent, results in air being exhaled in a ‘puff’. This is synchronous with the fact that Virta’s voice cracks slightly on the accented word ‘young’. The gesture is, in some sense, analogous with the lyrics, if we consider the concept ‘young’ to implicate potency and ability.

‘The singer’s face looks anxious during the whole phrase’ was one description for the phrase ‘like a lovers kiss’ in measures 23 and 24 of the chorus in English (see Example 12). One of Virta’s mannerisms, present in most of his songs, is a particular ‘commanding tinge’. This is, for the most part, a by-product of the ‘full-force’ mode of his singing – as one aims for volume, the face turns naturally into an expression which resembles an ‘arrogant’ look (see also Example 4). While this effect might be



Example 11. 'Merely to be young' (measure 18).

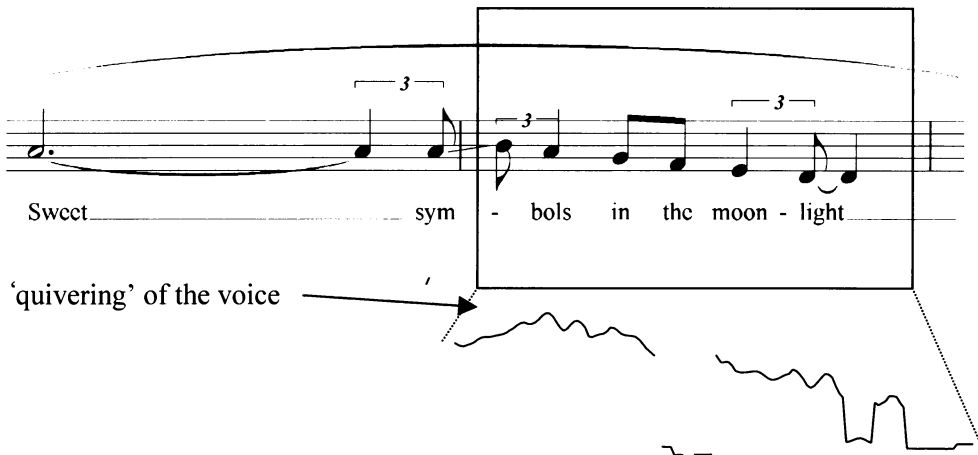


Example 12. 'Like a lover's kiss' (measures 23–24).

off-putting if overused, it may be important in creating a contrast for the softer gestures which, in this song, are heavily emphasised. The commanding tinge certainly has erotic connotations, despite the fact that this is not elaborated in the words of the phrase. The 'l'-sound seems a feasible platform for the 'commanding' gesture with Virta exploiting it to its full potential: 'L-like a l-lover's . . .'.

Again in measure 26 of the chorus in English, a sudden diminuendo follows a cracking forte (*sweet sym-*), this time with quivers that accompany the descent of the melody (See Example 13). The quivers are executed with a shrug of the head; the hands might easily be visualised being brought close to each other so as to describe a delicate and fine artefact: 'It was such a fine Tiffany-work'.

The last two phrases are again a feast of crooning. The piano backing begins a ritardando which evolves into a rubato in the last phrase ('Isn't it romantic?'), giving



Example 13. 'Symbols in the moonlight' (measure 26).

room to the looming playfulness of the singer's voice. One may sense the joy Virta feels at his own masterfulness as he displays his technique. The same air of self-confidence oozes from the performance as a whole – an indefinable sum of all the parameters of the vocal delivery.

Conclusions – Olavi Virta as an interpreter of a song

Although instrumental music occasionally reaches the charts, almost all popular music includes vocals. The human voice is, to most of us, more attractive than even the most pleasurable sounds from musical instruments. It is practically impossible to open up the full meaning the singing voice possesses for humans by treating it merely as a physiological, physical or acoustic phenomenon. The fact that computer-aided measuring devices were used in this study calls for a short comment. Computer-aided sound analysis all too easily leads the analyst to believe that the visible parameters are relevant – just because they are visible. Although computer-made measurements and the resulting visual reports are objective and precise, this precision does not transfer back to an audible sensation or musical experience. Computer-made graphics can provide us with information unobtainable by aural means: for example, we are unable to count how many pulsations per second there are in a pleasing vibrato or hear that there is a peak at 7,000 Hz in the upper partials of a voice that sounds somehow 'brilliant'. How the knowledge based on measurements increases our *understanding* of music remains unclear as the measurements are not included in the sphere of musical experience. How, for instance, would the above knowledge benefit a student of singing? It is for these reasons that I would like to emphasise that none of the transcriptions presented in this study were the object of analysis. Their function was totally different; they were used to illustrate and point out briefly passing moments in audible music.

The actual method of research employed here was emphatic listening and corporeal identification. The analysis was initially difficult because explicating even a couple of gestured moments was a laborious task. Why is this? Cubitt (2000, p. 154) has pointed out that in order to listen, the listening subject must be there to listen as an

ego, which is pleased by the face-value message of the singing voice. Because the ego is conscious, it tends to dominate the analysis even though the creative process – the active component in the act of listening – takes place at a much deeper level of consciousness. To even consider analysing by emphatic listening, one must be willing to momentarily shut out any self-criticism. For this reason, a test situation comprising several test persons would be problematic as one cannot force a certain mode of listening on anybody.

Despite this, the comments of six test subjects were included in this study. They were, however, not so much objects of analysis or the analysis of this study themselves, but a means for me to reflect on my own analyses and to become more *informed* about the subject of experiencing audible gestures. Creating an in-depth test seemed futile since one would not be able to avoid great variation, beginning with the intensity of listening and understanding of the object of the experiment. The most practical way was to let the test subjects listen to the test-material – Virta's performance as a whole and in separate phrases – in their own homes, even though this meant losing much of the control over the listening act. It was also difficult to express verbally exactly what I wanted. Eventually, five out of the six had the correct idea of the purpose of the experiment and their reports did show shared tendencies.

One realisation brought about by the experiment was that at the micro-level of the performance there remains plenty to consider even after the macro-level of the song ceases to present anything new. In listening intensively to the vocal delivery, the multitude of nuances seemed almost endless. The music passes quickly, and the short analysed sequences were heard for only a fleeting moment. We have often heard the claim that in order for a song to become a hit, there has to be something new and a lot that is familiar. When a song is heard a few times, it becomes familiar at the macro-level – one can 'remember how it goes'. Why do we not grow tired of the song at this moment? Could it be that it is partly because we have still not registered all the micro-level nuances even though we know the song by heart?

The point of departure for this study was the notion of the imperativeness of Virta's ability to express a song; this we learned from the testimonies of the fans. The resulting hypothesis was that this expressiveness would manifest itself at the micro-level of the vocal performance as non-verbal, albeit meaningful, gesturing. After close examination, a rich arsenal of gestures was detected. These included both recurring mannerisms and singular vocal effects. The mannerisms might include the 'commanding tinge' in the voice, and rarer gestures like the 'tearful expression on the face' at the end of the Finnish chorus. Part of all human behaviour is repetition. Therefore, one must abstain from assigning any value-hierarchy for mannerisms and more infrequent gestures.

The other hypothesis was that Virta 'acts out' the semantic meaning of the lyrics. Virta's gestures do not, however, possess a straightforward analogous relation to the semantic meaning of the lyrics. One can even find a slightly amusing example of this (see Figure 2) – Virta begins a swelling crescendo while he is singing *pienen pieni, hento* ('very very small, delicate'). Here, the meaning is born out of conflict; gesturing can be contrapuntal, polarising and estranging, not merely mimetic.

As one listens to the whole song, one immediately notices that the chorus in Finnish has some special quality – something the beginning of the chorus in English does not. What is this difference? One test-subject expresses it in this manner:

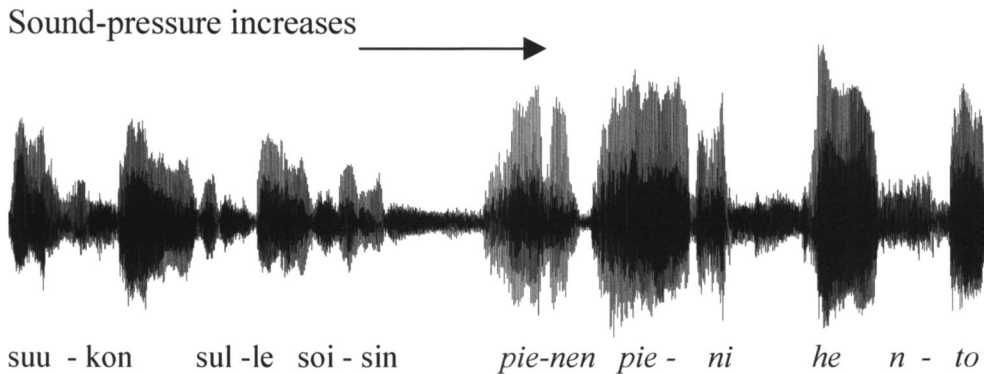


Figure 2.

The voice gets feeble as he sings in English. You get the impression that the expressions on his face are more insecure and mousy. The diction is tighter, the mouth more closed; for example, the I-vowels are tense. Finnish sounds are much broader – as is the face.

The difference is clear enough to be concretely detected. Only two of the test subjects comment on the second phrase of the chorus in English ('Isn't it romantic? Moving shadows write the oldest magic word') where Virta is clearly uncomfortable with the English language: 'the face is squirming as he is searching for words . . .' and 'the gaze is wandering about, the hands are in an unsettled if not restless movement'. The phrases in question are void of those gestured nuances which are so typical to Virta – the masculine gestured mannerisms which the fans are no doubt particularly fond of. To conclude this examination, one is obliged to bring forth the *erotic* nature of the relationship of the listener and the sonic representation of Olavi Virta. Potter's (1998, pp. 172–3) paper, which is otherwise critical of the classic genotext/phenotext division of Roland Barthes and Julia Kristeva, acknowledges that Barthes is correct in pointing out the erotic nature of singing. According to Potter, many singers (even if they cannot put it into words) are aware of the fact that at a fundamental level, there is a link between singing and sexuality (see also Cubitt 2000, p. 149). In the recording of 'Luonasi jos oisin/Isn't it Romantic?', the body of Olavi Virta can be heard in its full measure. It either promises the listener physical intimacy or the listener is actually experiencing the 'intimate atmosphere as the voice gets whispery'. There is an erotic charge between the listener and Virta due to the fact that his voice, produced mainly with parts of the body yet hidden from the eye, making the *inside* of Virta's body present. The technical solutions of the recording of this particular performance make the end product explicitly intimate. The awakened desire may or may not find its object in Virta's body. The voice may provide the listener with an object for identification, *through which* the listener may feel a desire analogous to the one felt by the singer. With Virta's voice, we may get the chance to whisper beautiful words into the ear of our loved one, or be whispered to.

A website with sound examples and other additional material has been constructed to supplement this article at: <http://www.uta.fi/~marko.aho/virta>

Acknowledgements

For this study, essential technical assistance and collegial support was provided by Anne Tarvainen.

Endnotes

1. Singing is essentially linked with breathing, which is a basic prerequisite of human organic processes.
2. The most read scholarly book on singing is probably still Johan Sundberg's *Science of the Singing Voice* (1988), in which singing is understood purely as a physiological phenomenon. An opposite example is provided, for example, by David Schwarz (1997, p. 146) in his graphic depiction of the voice of Diamanda Galas.
3. The author has not had access to the original postcards. Instead, the host Maarit Niiniluoto has provided the author with a photocopied selection, which supposedly comprises the most eloquent of the written responses.
4. All translations of straight quotations from the informants' comments in Finnish are by the author.
5. For example, probably the most important text with regards to the posthumous 'life' of Virta is the biography written by von Bagh, Koski and Aarnio (1995).
6. Kotlyar and Morozov (1976, according to Sundberg 1988, p. 152) executed an experiment in which professional singers were asked to convey in their singing different emotions (joy, sorrow, fear, anger, neutral). In the test situation the singers accomplished their goal very well. Fónagy (e.g. 1967) has had positive results in studying whether listeners can identify a facial expression on account of mere vocal delivery.
7. For example, a specific frequency-code seems to work on similar lines both with humans and animals: an animal expressing submissiveness and willingness to co-operate lifts the pitch and makes the voice brighter by lifting the corners of the mouth, whereas threat and cogency is sought from low pitch and dark sound colouration, which is achieved by expanding the vocal tract (Laukkanen and Leino 1999, p. 97).
8. All translations from works originally published in Finnish are by the author.
9. The test subjects were recruited from among the acquaintances and colleagues of the author. Two of the people were male, four female, and all of them were young adults (25–34 years old). In the selection of these persons, the criterion was an unassuming attitude towards both Olavi Virta and the idea of experiment itself. The recruiting included a conversation where the premises of the test were explained to the candidate, and where the criteria were proven filled. After this, the test-subjects received a CD, on which the performance of Olavi Virta could be listened to both in its whole and indexed to phrase-pairs. They were directed to take note of what kind of movement or posture of the body, facial expression or other tacit gesture they were able to detect with each sung phrase. They were then advised to go ahead with their task in a relaxed and positive mood, and preferably in an unlit room, and to use headphones if available. The instructions were reiterated in a piece of paper attached to the test CD.
10. Actually an Estonian who recorded extensively in Finnish.
11. <http://www.fon.hum.uva.nl/praat/>
12. However, it is impossible to notate swing itself. Swing as a heard phenomenon is tensions against the pulse in time. In a frozen visual notation these tensions cannot be experienced.
13. The area at the centre of the face.

References

- Aho, M. 2001. 'Iskelmädiskurssi. Tekstit vai lukijat?', *Musiikki*, 1/2001, pp. 5–21
- Von Bagh, P., Koski, M., and Aarnio, P. 1995. *Olavi Virta – legenda jo eläessään* (Helsinki, WSOY)
- Brackett, D. 2000 [1995]. *Interpreting Popular Music* (Berkeley, University of California Press)
- Coker, W. 1972. *Music & Meaning. A Theoretical Introduction to Musical Aesthetics* (London, Collier-MacMillan Ltd)
- Cubitt, S. 2003 [2000]. '“Maybellene”: meaning and the listening subject', in *Reading Pop. Approaches to Textual Analysis in Popular Music*, ed. R. Middleton (Oxford, Oxford University Press), pp. 141–59
- Dunn, L.C., and Jones, N.A. (eds.) 1994. *Embodied Voices: Representing Female Vocality in Western Culture* (Cambridge, Cambridge University Press)
- Ekman, P. 1977. 'Biological and cultural contributions to body and facial movement', in *The Anthropology of the Body*, ed. J. Blacking (London, Academic Press), pp. 39–81
- Fónagy, I. 1967. 'Hörbare mimik', *Phonetica*, 16, pp. 25–35
- Frith, S. 1996. *Performing Rites. On the Value of Popular Music* (Oxford and New York, NY, Oxford University Press)
- Jalkanen, P., and Kurkela, V. 2003. *Suomen musiikin historia. Populaarimusiikki* (Helsinki, WSOY)
- Kostenbaum, W. 1993. *The Queens Throat. Opera, Homosexuality and the Mystery of Desire* (London, GMP Publishers)
- Kotlyar, G.M., and Morozov, V.P. 1976. 'Acoustical correlates of the emotional content of vocalized speech', *Sov. Phys. Acoust.*, 22, pp. 208–11
- Laukkanen, A.-M., and Leino, T. 1999. *Ilmeellinen ihmisääni* (Tampere, Gaudeamus)
- Poizat, M. 1992. *The Angels Cry: Beyond the Pleasure Principle in Opera* (Ithaca, NY, Cornell University Press) [Orig. *L'opéra, ou, Le cri de l'ange* 1986]

- Potter, J. 1998. *Vocal Authority. Singing Style and Ideology* (Cambridge, Cambridge University Press)
- Sundberg, J. 1988. *Science of the Singing Voice* (Dekalb, IL, Northern Illinois University Press)
- Schwarz, D. 1997. *Listening Subjects. Music, Psychoanalysis, Culture* (Durham, Duke University Press)
- Tarvainen, A. 2004. 'Laulaminen liikkeenä', *Musiikin Suunta*, 3, pp. 5–15
- 2006a. 'Björk ja lihan arvoitus', in *Ääniä äänien takaa. Tulkintoja rock-lyriikasta*, ed. T. Lahtinen and M. Lehtimäki (Tampere, Tampere University Press), pp. 253–72
- 2006b. 'Käheys: Laulajan äänenlaadun tarkastelua fysiologisesta, kokemuksellisesta ja kulttuurisesta näkökulmasta', *Etnomusikologian vuosikirja 2006*, pp. 77–108

Discography

Olavi Virta, 'Luonasi jos oisin/Isn't It Romantic?' ESEK CD 1. Original matrice-number FIFMF5500017. 1992[1955]

Archival material

The audience's feedback with reference to the special 1998 broadcast of the program *Olavi Virran tie Suomen tangokuninkaaksi* ('Olavi Virta's journey to tango-kingdom of Finland') is the property of the Finnish Broadcast Corporation. The original postcards are in the custody of Maarit Niiniluoto.

The reports of the test subjects on the hearing of 'Luonasi jos oisin/Isn't It Romantic?' are in the custody of the author.