

# How Thomas A. Edison shaped today's singing ideal: Tracking his ambiguous concept of tremolo by analysing archival documents and sound recordings.

TILO HÄHNEL

*Musicology Seminar Detmold/Paderborn, Paderborn University, Detmold, Germany.*

KARIN MARTENSEN

*Audio Communication Group, Technische Universität Berlin, Berlin, Germany.*

**ABSTRACT:** We linked an analysis of vocal vibratos in early recordings with Edison's dismissive attitude towards singers' tremolos in the same recordings. We conclude that there are at least two different factors contributing to Edison's concept of tremolo, which include a) technical limitations and artefacts related to the recording process, and b) aesthetic judgements based on his taste for a good singing voice, but which are not necessarily related to the vibrato as such. In some cases, his comments on the tremolo even contradict the vibratos found in the recordings. Yet, Edison's concept of tremolo was as ambiguous as it was influential, for it was the tremolo on which Edison frequently based his decision to employ singers for commercial recordings. Since many of today's vocal artists still view early recordings as the gold standard of interpretation, we should consider the possibility that early recordings are by no means a representative sample of the artists of that time, but rather a selection due to the constraints of the technology at that time.

Submitted 2019 February 27; accepted 2019 June 12.

**KEYWORDS:** *Thomas A. Edison, Phonograph, Singing, Vibrato, Tremolo, Gramophone*

## 1 INTRODUCTION

OUR knowledge of historical performance practice is mainly based on written sources and artefacts. This restriction is a problem, at least as far as the time prior to the emergence of music recordings is concerned. This matter seems especially pertinent when it comes to singing. Since singing needs no musical instrument, there are no historical instruments to be analysed for deriving additional information related to singing practice. Thus, it seems a relief that music performances were recorded since the late 19th century. Recordings of singers seemingly conserve the style and sound of the singing practice since the early 20th century; this may help to evaluate certain questions, like the application of ornaments, coloraturas or expressive timing. Nevertheless, the usability of music recordings is limited, since they have several drawbacks, three of which should be mentioned briefly. Firstly, we cannot assume that the performance in a recording studio equaled a performance on stage. Secondly, after the recording and reproduction, the sonic image we are presented with is never identical to the sound of the source recorded, let alone recreating the space and situation in which the performance took place; alongside these general limitations, early acoustic recordings show such a narrow bandwidth and distinct resonance characteristics, that the sonic image is largely distorted. Thirdly, we must be aware of the possibility that the singers who are preserved in early recordings might not represent a cross section of the singing practice of that time. It seems unlikely that the general intention of record companies was to document singing practice (at least at that time). For instance, as a record producer, Thomas Alva Edison (1847–1931) had very concrete ideas of a singing voice worth recording; hence, it would be unwise if we were to pursue a recreation of, say, 1910th singing practice, based on recordings from this time without being aware of the limitations. If singers were to imitate Caruso, they may be imitating the sound of a Caruso record, but they would certainly fail in imitating the singing practice of Caruso's time in general.



This problem seems particularly true for the use of vibrato, which is regarded as one of the most important means of musical expression. Therefore, in the present paper we attempt to investigate how Edison exerted influence on the use of vibrato. We will query the assumptions that the singers recorded were a representative cross section, and that the vibrato we hear in the recording was sung in the studio. To do so, we will present some archive material, and relate it to historical recordings using metrological means. Before that, we need to explain the terms “tremolo” and “vibrato” as they are used in this paper. The introduction will also touch on the implications of Edison’s record policy on the selection of singers, which is a crucial point for evaluating sound recordings as historical sources.

### 1.1 Vibrato and Tremolo

We use the term “vibrato” to denominate a periodic undulation of a tone (here a singing tone) in pitch, volume, and timbre. For a better differentiation, we distinguish between a pitch vibrato, a volume vibrato, and a timbre vibrato respectively, although the last one will not be discussed further. The shape of the vibrato resembles a sine, by which it can and will be described mathematically. There are at least two important parameters shaping the vibrato, which are its rate (or speed, which is usually between five and nine hertz, meaning between five and nine vibrations per second) and its range (which will be given in pitch cent or sound pressure level, dB SPL, respectively). In the following, we refer to the term “range” to describe the distance between the upper and lower reversal point of the vibration, whereas the “amplitude” describes the deflection from the zero or fixed point in each direction. Therefore, it is the same if we say that a pitch vibrato has a range of two semitones, or an amplitude of plus minus one semitone. To avoid confusion, we will henceforth use the term “range.”

Pitch and volume changes of vibrato do not behave strictly proportionally to one another. This means that two vibratos can have the same pitch range but very different volume ranges. Thus, commonly used adjectives to describe the vibrato, such as “great” or “small,” are problematic, for it remains unclear whether these adjectives relate to the rate or the range, or if the adjective refers to both dimensions, volume and pitch.

The term “tremolo” is used as a historical term only, and we assume that both vibrato and tremolo were used to describe the same phenomenon around 1900. One might object that a tremolo is a volume vibrato without a pitch vibrato, because this is precisely what today’s effect units for tremolo do. There are also certain early singing techniques that substantiate this viewpoint; the *trillo* for instance, which is a rapid sequence of notes of – or a trill on – the same pitch, is particularly important in so called historically informed performances. It was described by Caccini in the 17th century and has been mentioned by Hugo Goldschmidt (1890, 87f). Therefore, it was known at the time in question, but as we will explain later, most sources of the late 19th and early 20th century do in fact describe what we call vibrato, when they use the term “tremolo.” Like many others, Edison only used “tremolo.” In the following, we use the term tremolo if the source we refer to used it, otherwise we use the term “vibrato.”

### 1.2 Edison and the Implications of His Record Policy

Edison did not only invent the phonograph; his record company was also one of the most important providers of records until its decline in the late 1920s, and its insolvency in 1929. Edison’s records are still valued as exemplary documents of early 20th century singing, including artists such as Frieda Hempel, Leo Slezak, Lucrezia Bori, and Jacques Urlus. For Edison, a recording of high quality was the most important criterion (see Martensen, 2019). He explicitly made that clear in a letter to Thomas Graf, head of his recording studio in Berlin, in which he wrote:

I propose to depend upon the quality of the records and not on the reputation of the singers. There are, of course, many people who will buy a distorted, ill-recorded and scratchy record if the singer has a great reputation, but there are infinitely more who will buy for the beauty of the record, with fine voices, well instrumented and no scratch. (Harvith & Edwards Harvith, 1987)

This attitude towards singers and their reputation was quite unique. To ensure that his staff knew what he wanted, he wrote a general policy in 1912, in which he demanded:

All that we desire is that the voice shall be as perfect as possible, free of conspicuous tremolo, clear and without ragged sustained notes, free of subsidiary and false waves on these notes. (Edison, 1912)

It is unclear whether the adjective “conspicuous” was meant as an attribute or a modifier. It could be read as “free of tremolo, which is always conspicuous” or as “free of those particular tremolos that are conspicuous”; as we will demonstrate further, the former interpretation is more likely.

Edison’s policy is only one instance showing that his selection of artists was based on his idea of a perfect voice, and he ensured personally that his criteria were met. In doing so, he pressed for technical improvements and evaluated singers and their voices. Thus, the voices we hear on old Edison records are not a cross section of the artists of that time, nor do they provide a representative sample of the singing technique of the early 20th century. This makes the use of Edison records problematic, if we use them as historical sources and take them as exemplary instances of a cultural practice. What we are listening to is a distorted image of the past. As this paper will show, that distortion refers to the fact that the artists selected became representatives for their generation, as well as for the sound of the voices that we hear in those recordings.

Edison inspected each sound recording made in his studios over the years, listening to them and writing memoranda for the evaluation of his singers. All artists who wanted to record in Edison’s studios had to undergo voice trials. Only those artists who had passed the trials successfully could make recordings for the market. Edison frequently demanded that singers or the accompaniment were replaced, and he even dreamt of freeing himself from the laborious selection of singers, when he wrote that “[w]ith a regular corps of singers we will be enabled to rehearse, change the voices and style of accompaniments and make several duplicates of the tune and thus adopt the most perfect one” (Edison, 1912). Edison and his staff went even further and made analyses of records from competitors, such as Victor Records, Columbia Records, and Fonotipia. If Edison cylinders or discs shall be used as historical sources, it is crucial to know more about Edison’s concept of vocal quality. For him, one of the most important aspects of voice quality was the tremolo, the meaning of which will be central to this paper.

### 1.3 Outline

In the following section, we look at different sources from around 1900 to figure out what was meant by “tremolo,” and especially how Edison used that term. The investigation of writings alone, however, cannot clarify what Edison meant, but it will raise questions which serve as the basis for our main approach. The main aim of this paper is to combine archival work and performance analyses to investigate both Edison’s comments, and the recordings Edison referred to. In Section 3, we will describe the archival material we gathered, and how we edited the comments in Edison’s voice trials for further use. With an analysis of the pitch vibrato described in Section 4, we will take a first step towards a better understanding of the relationship between the two vibrato parameters, rate and range, and Edison’s positive or negative reactions he recorded in his voice trials. Given that the audio files we used were of different origin, we distinguish two corpora of the voice trials, namely Records Trials and Probe Trials, which are later merged into one larger corpus called Test Trials. The preliminary analysis of the pitch vibrato allows for constructing hypotheses regarding the volume vibrato, which we analyze in Section 5. Finally, the main results are summarized and discussed in Section 6.

## 2 THE CONCEPT OF TREMOLO AND VIBRATO, 1861–1936

### 2.1 Edison’s Perspective

As Ann-Christine Mecke (2016, p. 672) argues, the nomenclature for “vibrato” or “tremolo” was not standardized until around 1900. And even after the turn of the century, there is hardly any agreement amongst the sources dealing with the vibrato. Therefore, it is not surprising that Edison could not rely on a consistent definition, as in his time there was little information on the vibrato.

For Edison, the tremolo was one of the most important features of the singing voice, as it was one of the singers’ most serious flaws. This reading is based on statements in several sources, including his personal pocket notebooks. For instance, in one of these notebooks, Edison referred to the Grove’s “A Dictionary of Music and Musicians” (probably the first or the second edition and presently known as “The New Grove Dictionary of Music and Musicians”):

Walter Miller[:] Groves Dictionary of Music[,] page 327 article Cavalieri middle of column on right says[:] “Certain vibratory production of the voice, a fault common among bad singers and known as Tremolo, shake.” (Edison, 1910–1912)

Here, Edison obviously addressed this entry to his staff member Walter H. Miller, head of sound recording in the Music Room in New Jersey. One possible explanation would be that he wanted to help define the term tremolo. It is also possible, however, that he wanted to provide some arguments supporting his personal view on the matter. The latter interpretation rests upon the observation that Edison altered certain aspects of the original article by omitting information. Originally, the article deals with Emilio de’ Cavalieri (1550–1602), and the author of the article asks, whether Cavalieri’s “trillo” must be regarded as a “shake” or as a “vibratory production of the voice”:

It may be questioned, however, whether the last-mentioned [trillo] was the true “shake”; that is to say, a rapid oscillation between two tones or semitones; or whether it was only a certain vibratory production of the voice, probably considered an elegance in early times, but now more fitly as a fault common among bad singers, and known as the “tremolo.” (Grove, 1879, 1900, both p. 327)

Here, Edison did not refer to the difference between “tremolo” and “shake.” He simply treats both equally as “a fault common among bad singers.” Further remarks suggest that Edison thought of the tremolo as a defect:

The tremolo is the greatest defect of the human voice and is never put in by the singer ... Singers are unaware that they have a tremolo ... As a rule if the strength of the voice is called 100% the defects outside of the tremolo seldom exceed 10%. Hence if singer is in a very large Auditorium the defects in many cases are not heard if the Orchestra is large and the hum of the audience Considerable [sic]. (Harvith & Edwards Harvith, 1987)

In an interview, Edison tried to explain the tremolo as a faulty alignment of the vocal cords, which underlines his view on the tremolo as a defect rather than a means of expression:

[B]oth cords must be in the same tension. If one cord is a little looser, it will have fewer vibrations than the other. This conflicting of the vibrations of two different tones produced at the same time is what causes the tremolo with its “out of tune” effect. (Anonymous, 1916b)

In another interview, he differentiated between a “trill” and a “tremolo,” again suggesting that the tremolo is a fault in comparison to the trill.

A singer’s trill is quite a different thing from a tremolo. A trill can be and is controlled by the brain, but a tremolo is not within brain control and, so far, there has been no means found for correcting it. Most singers cannot sustain a note, without breaking it up into a series of chatterings, or tremolos. The number of waves varies from two per second to as high as twelve. When at the latter rate, the chatter can just be heard and is not particularly objectionable. When at a lower rate, it is very objectionable. (Anonymous, 1916a)

The range of the vibrato rate between two and twelve hertz, to which Edison refers, is very large. His remark that a slower tremolo is worse than a faster tremolo is, however, noteworthy, and it is a rare case in which he makes a more differentiated statement about the tremolo and the perception of its quality.

The “tremolo” is an object of examination in most of his voice trial memoranda. Too strong a tremolo seems to be the main reason for Edison to expel a singer from further trials, and from making a master disc for the market. A few examples may provide an insight into Edison’s brief but clear diction:

MISS FANNY ANITIRA MEZZO-SOPRANO. Scala. This lady has a very pure scale – Has some tremolo but not very great – [...] but her voice is pure and if she didn’t have the tremolo she would be a GOOD SINGER. (Edison, 1912–1917)

FRANCO TURNENELLO. TENOR. [...] – A fine sharp extra set of vibrations all through his voice. His tremolo is too much for us. [...]. (Edison, 1912–1917)

PHILLIPS, JOSEPH A. [...] You can use this man. [...] His tremolo is so rapid that it will not be objectionable under many conditions. [...]. (Edison, 1912–1914b)

FERN, HARRY [...] This man we can use. He has no tremolo at all. [...]. (Edison, 1912–1914b)

88335 VICTOR [RECORDS. ENRICO] CARUSO is getting a bad tremolo. Record is horribly distorted. This record is a parody on any voice. Tune N[ot]. G[ood]. Victor loud record N[ot]. G[ood]. (Edison, 1912–1914b)

Murphy, Lambert This voice is fine – but spoiled by his tremolo. Voice is so rare you had better get him to come in and ask him if he cannot improve his tremolo. If he can we can use him extensively. If he cannot we don't want him. His voice is exceedingly sweet and rich. (Edison, 1912–1914b)

Adolph Fowler [...] Tune N[ot].G[ood]. Voice awful sharp, perhaps it's the recorder. Seems to have no tremolo, might make a trial record of good tune having sustained notes. (Edison, 1912–1914b)

The last example shows that Edison was aware of the possibility that the recording procedure, particularly the recording devices, might influence the sound and timbre of the voice. He also assumed that the recording apparatus could exaggerate or even cause audible vibrations. In a letter to his staff member Thomas Graf, Edison wrote that the competing record company Victor Records:

make[s] another mistake and think[s] that because a singer is good on the stage he will be good on the phonograph. This is not the case, because the phonograph exaggerates the tremolo which is present in nearly every voice. (Harvith & Edwards Harvith, 1987)

There are nevertheless a few examples indicating that Edison accepted the necessity of a vibrato for artistic reasons. In an interview, he said: “Such a use of the tremolo [for dramatic purposes] should be tolerated. It is like a trill. It intensifies. It is the constant use in lyric passages where a beautiful even tone is desired that I deplore” (Anonymous, 1916b).

In conclusion, we can say that Edison's comments raise more questions as opposed to clarifying what the tremolo might have meant to him. Although the tremolo served as an important criterion for the assessment of singers, his remarks lack stringency, or are contradictory. Obviously, for Edison the tremolo was a fault, and it was mainly related to the singer's voice and singing style. However, he showed a certain degree of awareness that the recording process (including the recording equipment) might also play a role. This is further substantiated by his insistence on listening to the voice trial recordings and not to the singers on stage instead. Nevertheless, he remains somewhat unclear about his ideas on how these two factors, voice and recording process, might interact. From the way he judges his singers, it seems reasonable to believe that he attributed his perception of tremolo mainly to the singer's voice, despite his awareness that the recording process distorts the voice.

For a better understanding of the tremolo as used by Edison, it is important to understand what was known about the tremolo and the vibrato in his time in general, and to assess what could have been known at that time. Therefore, other contemporaneous sources might help, or at least narrow down the scope of possible meanings.

## 2.2 Treatises on Singing

The literature reviewed here reflects the positions on tremolo as they circulated around and before 1900. Certainly, these views on vibrato and tremolo originated in the 19th century. We will therefore reflect on some 19th-century literature but refrain from older sources, because we are interested in the view on the

matter, at the time in question. Edison's writing suggests an affirmation of more physiologically based schools, which are most prominently represented by Manuel García. The intense use of German authors may seem unbalanced, but these works were highly influential, and many were translated and published in the U.S. immediately.

As mentioned above, vibrato and tremolo were not defined consistently before 1900, and even in the years after 1900 there is hardly any clear explanation for either term. In his analysis of treatises on singing written by 100 authors, Brent Monahan (1978, p. 222) investigated the use of the terms vibrato and tremolo. The meanings of both differed, and often no distinction was made at all. Tremolo was frequently described as an exaggerated vibrato; the latter showing barely noticeable pitch oscillations but loudness oscillations, the former showing apparent pitch undulations. Monahan (1978, 96f) also refers to the Harvard Dictionary of Music of that time, in which the frequency of the vibrato is specified at eight hertz. The physiologist Henry H. Curtis wrote in his "Voice building and tone placing, showing a new method of relieving injured vocal cords by tone exercises" (1914), that the vibrato is produced involuntarily through the process of breathing, whereas the tremolo results from a weakness of the muscles related to the vocal cords (Monahan, 1978, p. 97). John Andrew van Broekhaven's "Some unfamiliar facts concerning the functions of the larynx in singing" (1908) and George Dodds' "Practical hints for singers" (1927) are directly opposed to one another: Van Broekhaven (1908, p. 45) distinguishes between an avoidable tremolo, which results from a weakness, and an artistic vibrato, which is the result of voice control. Dodds (1927), on the other hand, describes both in an opposite fashion (Monahan, 1978, p. 97). However, von Broekhaven's view seems to be in accord with most other sources, according to which the tremolo is the worse vibrato.

In treatises on singing, statements about the use of vibrato are rare; if there are any comments, they are mostly about avoiding it. For instance, the soprano singer and singing teacher Lilli Lehmann (1902a) thought of the vibrato as a dangerous symptom and first step towards the worse tremolo. In her view, both were the result of a faulty vocal technique:

Big voices [...] are easily disposed to suffer from the tremolo, because the outflow of the breath against the vocal chords occurs too *immediately*. (Lehmann, 1902a, p. 170)

Even the vibrato, to which full voices are prone, should be nipped in the bud, for gradually the tremolo, and later even worse, is developed from it. (Lehmann, 1902a, p. 174)

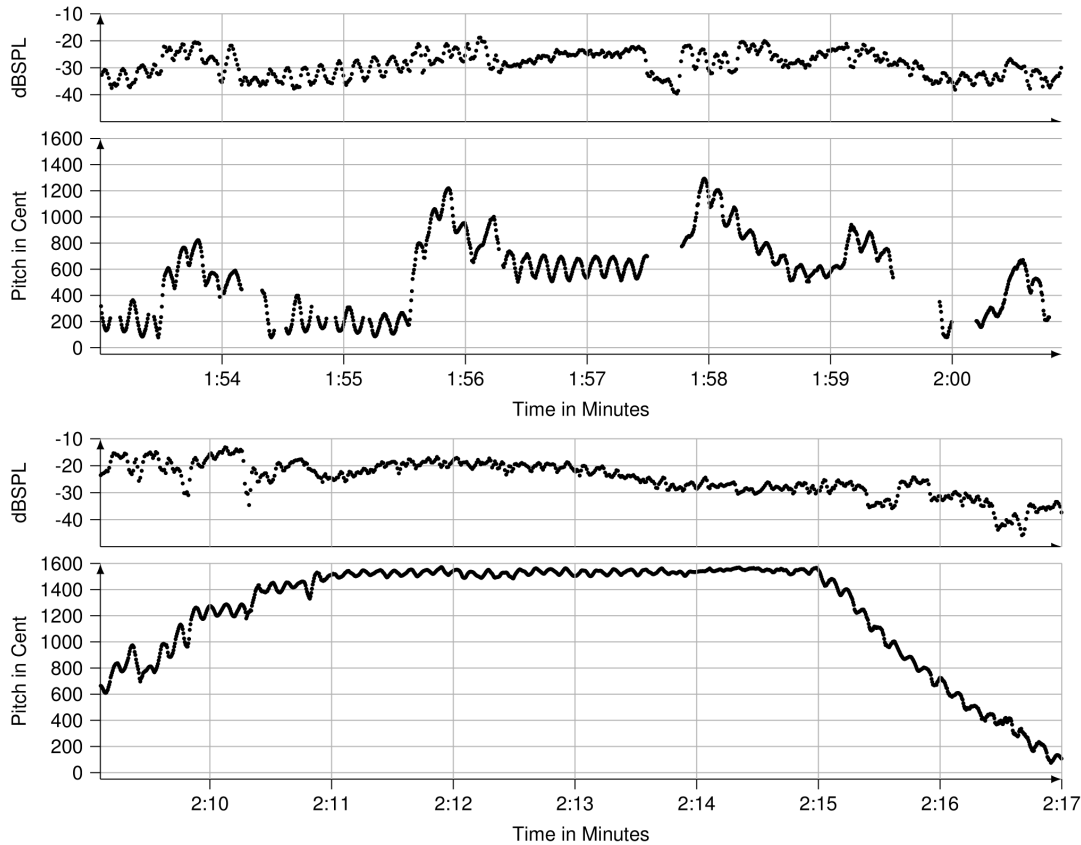
Vibrato is the first stage, tremolo the second; a third and last, and much more hopeless, shows itself in flat singing on the upper middle tones of the register. (Lehmann, 1902a, p. 174)

If the first and second stages of tremolo are difficult to remedy [...], the repair of the damage is nothing less than a fight, in which only an unspeakable patience can win the victory. (Lehmann, 1902a, p. 175)

But then, she opens the following section "The Cure" with the sentences:

There are no magic cures for the singer. Only slowly, vibration upon vibration, can the true pitch be won back. (Lehmann, 1902a, p. 175)

When listening to Lehmann's recordings, one can easily find that she made use of the vibrato and did not modulate it by chance; Figure 1 shows the pitch and dBSPL of two segments of her rendition of the aria "Casta Diva."



**Figure 1.** Pitch and dB SPL curves of two segments from Lilli Lehmann’s “Casta Diva” (*Norma*), recorded in 1907 (original Odeon Cat. 52698, Matrix xB 2978, taken from “Lilli Lehmann [1848–1929]. Prima Voce,” Nimbus Records, NI 7921, 2005). For the pitch curve, the fundamental frequency was computed by the Smoothed Pitch Track algorithm and then converted into cent with 0 ct = 460 Hz. The first segment clearly shows pitch vibratos with ranges of about two semitones, which are in parts accompanied by a distinct level vibrato (around 1:55). The first segment is representative for the whole recording, although Lehmann deliberately warned against the vibrato and the tremolo in her writings. The second segment shows an exception, which is a single high-pitched tone starting with a fast but very narrow vibrato that fades continuously until no vibrato is present for about one second (around 2:14).

This example demonstrates that she indeed used the vibrato as a means of expression, despite her condemnation of it in her writings. It is unlikely, however, that Lehmann omitted dealing with the positive side of the vibrato deliberately. Probably she was not aware of her own vibrato’s intensity, or she – as others, too – took the vibration we call vibrato for granted, and mainly spoke against an intense tremolo (whatever this would mean). A hint might be found in her passage about regaining the voice when damaged. Although the quotation above is taken from the official translation of Lehmann’s book, some subtle nuances were lost in translation. The original German version of the last sentence is:

Wunderkuren giebt es nicht für den Sänger. Ganz langsam nur kann Schwebung für Schwebung wieder gewonnen werden. (Lehmann, 1902b, p. 26)

Usually, the German term “Schwebung” is not used for the vibrato but for the beat that occurs by the superimposition of two very close pitches. Even though the phrase “vibration upon vibration” is somewhat unclear, for it can mean the vibration of the vocal folds, the term “Schwebung” clearly indicates something, the impression of which we would relate to the vibrato. Unfortunately, neither Lehmann nor other writers used the term “Schwebung” to designate a vibrato; however, the phrase indicates that Lehmann might have had an idea of a “good vibrato,” which she could not explicate.

Other authors were equally clear as Lehmann in their dismissal of the vibrato. Manuel García (1894) wrote that the tremolo is “[t]he trembling of the voice. This intolerable fault ruins every style of singing,” and referring to the slur, he added that “[n]ext to the tremolo, it is the most universal and most distressing fault.[...]” (García, 1894, p. 18). August Iffert (1906) wrote that “Das Tremolieren ist ein schlimmes Anzeichen davon, daß das Organ übermäßig angestrengt wird; [...]” (Iffert, 1906, p. 73, translates into: “The trembling is a severe sign of the fact that the organ is exerted excessively”). Oscar Guttmann (1861) wrote:

Bevor die Muskeln der Stimmwerkzeuge die gehörige Kraft erreicht haben, wird die Stimme immer mehr oder weniger nach oben oder unten abweichen (*detonieren*), sowie *tremolieren* (*zittern*) (Guttmann, 1861, 28f)

Before the muscles of the vocal organs have attained the necessary strength, the voice will always more or less vary from the correct pitch, as well as tremble (Guttmann, 1884, p. 77)

One rare exception is Ferdinand Sieber (1862), who speaks positively about the vibrato. In his view,

[...] every radical note should be accompanied with a swelling of the tone [...] The Italians call this swelling and sinking, this ebb and flow, the *vibration* of the voice, and they never find a song beautiful in which this is wanting. (Sieber, 1872, 146f)

Not only did he differentiate between tremolo and vibrato, he also used the vibrato positively as a necessary ingredient of the voice and as a means of expression:

Es herrscht in Betreff der Ausdrücke für die verschiedenen Gesangsthätigkeiten noch immer eine große Begriffsverwirrung. So lesen wir täglich, es sei zu rügen, daß der oder jener Sänger mit seiner Stimme *vibrare* u. dergl. m., während mit dem getadelten Fehler das *Tremolieren* gemeint ist. Die *Vibration* der Stimme ist ihr innerster Lebenspuls, ihre Triebfeder, ihr Alles - ohne sie giebt es nur Monotonie, Töne, wie wir sie aus den Orgelpfeifen gewinnen. Wird aber aus dem *vibrando* ein hörbares Beben und Erzittern des Tones, ein *tremolando*, gemacht, so verfällt der Sänger in einen (als Gewohnheit) unerträglichen Gesangsfehler, der nur in den seltenen Fällen eine Berechtigung erlangen kann, wo das Tremolieren eben als Ausdrucksmittel für die höchste Erregung gebraucht werden soll. (Sieber, 1862, p. 130)

It causes great confusion in regard to the expression among singers of different degrees of ability.

We read daily that it is reprehensible in this or that singers to indulge in this *vibration*, while it is really the *tremolando* which is blamed. The vibration of the voice is its inmost life-throb, its pulse, its spring; without it there is only monotony, simple notes, as we obtain them from the pipes of an organ.

But if the vibration is changed into a *tremolando*, the singer falls into an intolerable fault, which is warranted only in very rare cases, when it serves as a means to express the very highest degree of excitement. (Sieber, 1872, p. 147)

The interesting point here is that this passage is found in the section on dynamics. This leads to the suggestion that the vibrato was perceived as a periodic change of volume rather than of pitch.

### 2.3 The Vibrato in Music Psychology

It was not until the rise of empirical musicology, particularly the intense vibrato studies of Carl E. Seashore (1936) and his colleagues (Easley, 1932; Hattwick, 1932; Hollinshead, 1932; Lindner, 1932; Reger, 1932), that tremolo and vibrato became more precisely defined. In their empirical investigations, they used a “*strobograph camera* and a *vacuum tube voltmeter*, with which it was possible to gain reliable data about the pitch and amplitude of a sound” (Tiffin 1932, as quoted in Hähnel 2013, p. 11). Simply put, Seashore and his colleagues defined the vibrato as a sine-like fluctuation of pitch, sound pressure and timbre quality, the last of which they called “sonance”. For their investigations, they used what they called a *vibrato apparatus*, which allowed for independent control of the two parameters of loudness and pitch variance.



Most importantly, the experiments of Milton Metfessel (1926) showed that listeners were unaware of the vibrato parameters. When not told what they were listening to, they rather reported a certain timbre. When told, they could not say whether pitch or loudness was manipulated. Since Metfessel (1926) wanted the term “timbre” to be reserved for the description of the multiple frequencies and their proportions within a static sound, he tried to introduce the term “sonance”, which should describe the dynamic changes of the parameters constituting the vibrato, including rapid changes in fundamental frequency, amplitude *and* timbre. Unfortunately, the term “sonance” was never established, so today we are still facing similar terminological problems. When listeners were told that they were hearing pitch vibrations and that they should estimate the range between the upper and lower boundaries of these vibrations, they underestimated the pitch range of the synthetic vibratos heard. They perceived about half the range, which corresponds to the deviation from the mean (i. e. the amplitude) and not the overall range; the same was true for loudness. Moreover, the effect of the underestimation was more pronounced when the pitch or volume change became larger, when the vibrato rate became faster, and when the tone had been richer (Seashore, 1932).

#### 2.4 Towards a Better Understanding of Edison’s Tremolo

It seems likely that around 1910 – and therefore for Edison – there was no way to understand the characteristics of the vibrato. Thus, there was no solid base for a differentiation between vibrato and tremolo, at least not until the late 1920s, when Seashore and his colleagues started publishing their findings. As stated earlier, Edison’s impression of the “tremolo” served nevertheless as a selection criterion for singers. In the following sections, we will therefore investigate the vibrato in recordings Edison mentioned in his voice trials. We assume that if we find some systematic characteristics of certain vibrato parameters, which can be related to Edison’s comments, we will be able to better understand Edison’s perception.

### 3 ARCHIVAL MATERIAL AND SOUND RECORDINGS

To gain access to Edison’s memoranda, Karin Martensen visited the Thomas Edison National Historical Park in 2016. The collection includes several boxes of the Edison voice trials, from which she inspected Box 25–27, including the voice trials scrapbooks 1912–1914, 1912–1917, and 1915–1925. She took a photo of each page of his voice trials scrapbooks. She additionally took photos of other documents that she found in these boxes, such as advertisements and pages from his personal pocket notebooks. In total, Martensen took around 10,000 photos for later analysis. Generally, we distinguish between two different voice trials. The trials that referred to non-commercial recordings made to evaluate singers or candidates are henceforth called *Probe Trials*. Trials that were based on records that had already been on the market, and which included material from other competing companies as well, are called *Records Trials*. Probe Trials and Records Trials were initially treated as independent corpora or memoranda.

Firstly, we had our assistant check the Records Trials for any comment and select passages that provided some substantial information about the recording. This included sufficient information to track down the recording (Matrix Numbers, Titles, Artists, Record Companies, Dates), as well as meaningful information. Passages which only stated a simple “N. G.” (not good) or similar were left out. The remaining comments were transcribed into a table with columns for the name of the singer, the matrix number and the title of the recording, as well as Edison’s comment, which often consisted of a combination of typewriter text and handwritten additions.

Secondly, the assistant looked for any recordings already available online, and downloaded them if, and only if, the audio file came from an official source and was of acceptable quality. Although there are numerous private videos playing early recordings on the internet, we think that these sources are not as reliable and trustworthy in terms of correctness and quality. Most of the recordings selected were taken from the cylinder audio archive of the Library of the University of California, Santa Barbara (UCSB). The files were downloaded, and their respective source was transcribed and added to the table.

In a third step, we analyzed Edison’s comments regarding two conditions in two separate columns. Under the henceforth called *Overall condition*, we analyzed comments on the voice and vocal performance

**Table 1.** Overview of Edison’s comments in Records Trials (left side) and Probe Trials (right side), showing the amount of “Good” (G.), “Bad” (B.) and empty (N/A) labels for both conditions, Overall (rows) and Tremolo (columns). The respective sums show the groups of labels under the same condition, whereas the cells show the combination of the labels in both conditions. There are only a few instances of G. labels in the Overall condition that have B. labels in the Tremolo condition.

Records Trials						Probe Trials					
		Tremolo						Tremolo			
		B.	G.	N/A	Sum			B.	G.	N/A	Sum
Overall	B.	3	0	9	12	Overall	B.	33	0	6	39
	G.	1	4	15	20		G.	3	0	1	4
	N/A	0	0	1	1		N/A	1	1	0	2
Sum		4	4	25	33	Sum		37	1	7	45

in general. The second condition, henceforth called *Tremolo condition*, was an analysis that was particularly restricted to comments addressing the tremolo. Under both conditions, we assigned each comment one of three possible labels: A negative comment was labelled “Bad” (B.), a positive comment “Good” (G.), and in case the comment was neutral, unspecific, or contradictory, the respective cell was left blank (i. e. labelled *N/A*). In some cases, where the assignment was unclear, we discussed it in a group before making a final decision. This was particularly necessary in the Tremolo condition, because it seemed that there was no such thing as a good tremolo. However, in some rare cases Edison perceived very little tremolo. In these cases, we labelled the comment with “Good,” meaning a hardly noticeable, not disturbing, or at least acceptable tremolo. In two of these he referred to the tenor Ernst Krauss and the mezzo-soprano Margarete Matzenauer. About Krauss, Edison wrote “Some tremolo in voice. Pretty fair singer. Better than Slezak or J[ö]rn [...]” (Edison, 1912–1914a). Although “Some tremolo” can mean a considerable amount of tremolo, we decided that the alternative meaning “a little tremolo” would be more likely because of the positive context of the voice in general, and because at this point Edison preferred Krauss over Leo Slezak, about whom Edison wrote “Slezak has a horrible tremolo” (Edison, 1912–1914a). The fact that Slezak became an Edison artist after all is because Edison changed his mind about him in later years. About Matzenauer, Edison wrote “Good voice—has a little tremolo on some notes. [...]” (Edison, 1912–1914a). Here, it seems that although Edison made clear that there is a tremolo, he accepted it, because it was little and appeared only on some notes. Considering that he tolerated a tremolo for dramatic purposes (Anonymous 1916b, see Section 2.1), we assigned the label “Good” in this case too.

For the analysis of the Probe Trial corpus we used a commercial re-publication of 79 recordings of Edison’s voice trials (“The Edison Trials. Voice Audition Cylinders of 1912–1913,” Marston 52025-2, 2000). The booklet listed recording dates and respective artists; however, it did not include Edison’s comments on them. Therefore, we had our assistant analyze the photographed documents again, looking for entries that match the information given in the booklet. The final Probe Trial corpus comprised 45 recordings, of which we had Edison’s comments as well.

Table 1 provides an overview of the labels we assigned to the comments. The 33 samples of the Records Trials included four “Bad” labels in the Tremolo condition (henceforth “Bad” Tremolos), four “Good” Tremolos and 25 N/A Tremolos. A “Bad” Tremolo mostly went along with a “Bad” Overall label. Only once in the Records Trials and three times in the Probe Trials did we label the combination “Bad” Tremolo and “Good” Overall (the comments of these four instances are given fully in Table 2). The frequent association of “Bad” Tremolo and “Bad” Overall suggests that the tremolo was indeed one of the main reasons why a voice was rejected in general. We only assigned five statements about the tremolo with the label “Good,” one among the Probe Trials and four among the Records Trials. This shows how rarely Edison perceived a tremolo that did not disturb him. In both the Records Trials and Probe Trials, there was no case in which a “Bad” Overall met a “Good” Tremolo. Instead, there is a relatively large number of N/A Tremolos and “Good” Overalls in the Records Trials, which could be interpreted as a special case of the “Good” tremolo, under the assumption that only an unperceivable tremolo is an acceptable, hence good one.

**Table 2.** Selection of ambivalent comments suggesting a positive voice but a bad tremolo. The comments clearly show that the tremolo was a disqualifier even for good voices.

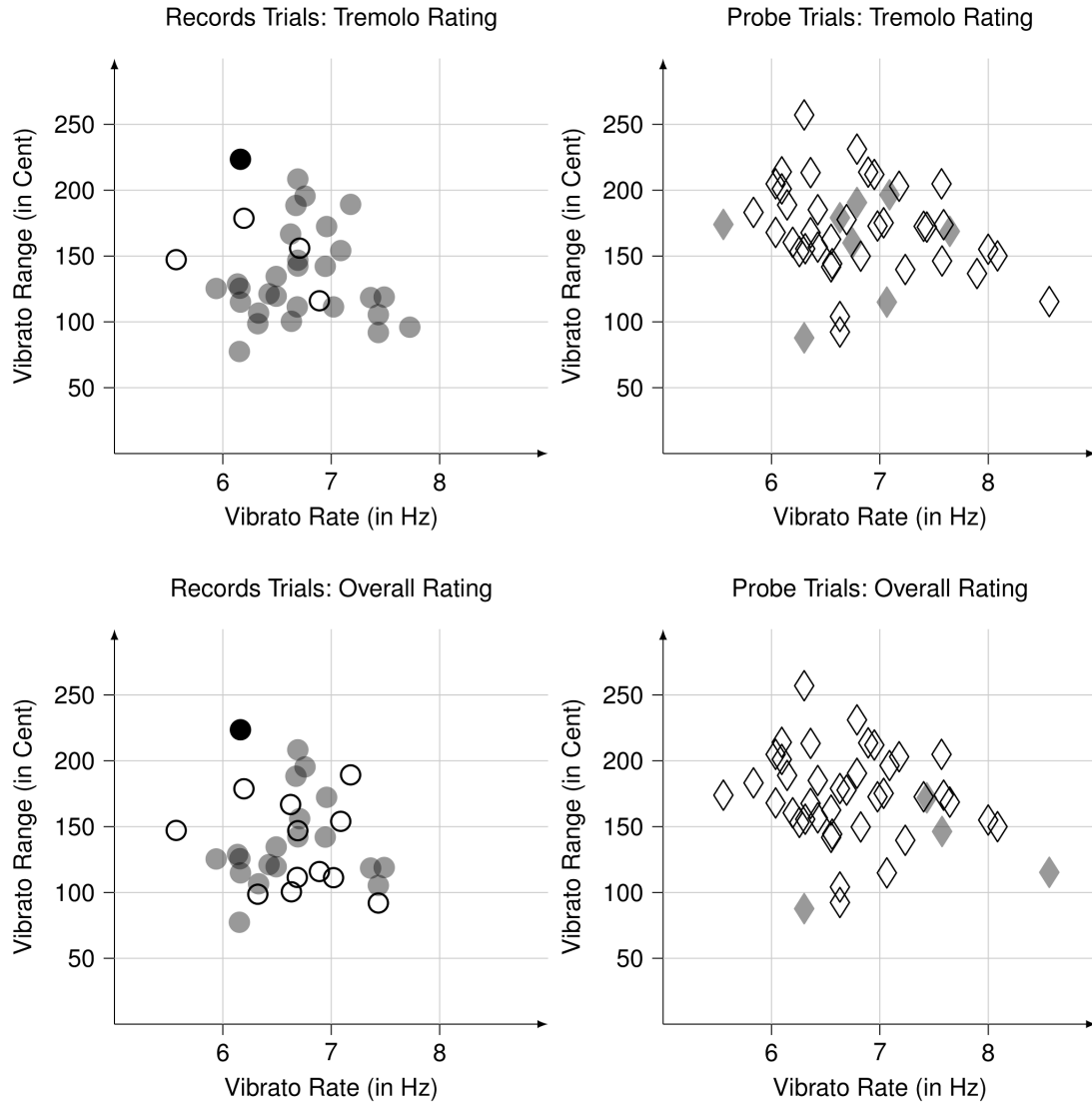
Name	Comment
Von Skilondz, Adelaid A.	Pretty fair coloratura - has bad chattering rapid tremolo high notes – catty not clear – mellow coloratura not near good as the Milan Coloraturas
Dupre, Pierre E.	Pretty fair basso – but has tremolo – Too much for us
Ponzano, Adela	Nice even scale but awful tremolo. Timbre good. Voice fair but awful tremolo
Caronna, Ernesto	Probably must get this. Voices here beat. But one voice must be changed so as to get rid of the horrible beats when they come together in the first half of record. Its blasty on cylinder. Must have this good. Lots bad tremolo in voices (Edison, 1912–1914b).

#### 4 ANALYSES OF VIBRATO RELATED PITCH CHANGES: A PILOT STUDY

Our preliminary analyses started with the Records Trials corpus because we had this data first. The Probe Trial corpus was used in a second step in the preliminary analyses. Following this, we merged both corpora for further analyses.

##### 4.1 Method of Pitch Vibrato Analysis

For analyzing the vibrato, we used a two-step approach, which consisted of (1) the extraction of a pitch curve related to the singer's voice, and (2) the automatic detection of sine-like undulations, which were interpreted as vibratos. Pitch curves were extracted using the Sonic Visualiser and the Smoothed Pitch Track algorithm of the 'pYin' plugin (Mauch & Dixon, 2014; Mauch et al., 2015), which estimates the fundamental frequency of the signal as an indication for pitch. Afterwards, the pitch curves were manually cleaned of any trills and areas unrelated to the vocal melody, because in these areas the data could be mistaken for a vibrato in the following automatic procedure (e.g., if a trill is played on a flute and recognized by the pitch extraction algorithm). The detection of vibratos within the pitch curve was done in a script written in Gnu Octave.<sup>1</sup> After converting the hertz values into pitch cent, the script narrows down the areas of possible vibratos by measuring the distance between succeeding local maxima. Since vibratos are assumed to have a rate between 4 and 10 Hz, which corresponds to wavelengths between 0.25 and 0.1 seconds, the script selects maxima that show these distances to their next maximum as well. Since the pitch curve is in fact a series of discrete points, the true maximum is likely to be positioned between the measured maximum point and one of its neighbors; therefore, the maximum and its two neighboring points are used to construct a quadratic function, the maximum of which is taken as the position of the corrected maximum. The distance between the corrected maxima is then used to calculate the frequency of a sine, which functions as the ideal vibrato, to which the pitch curve will be compared later. The comparison takes place within a short time window the size of 2.5 vibrato cycles, and the range of the pitch curve within this time frame is used as the range of the sine. Before the correlation of both curves, they are convolved with a Hanning window, which fades out the borders so that the resemblance of both curves is more important near the center of the time frame, which is the local maximum. The resemblance of the pitch curve and the sine wave is measured with the coefficient of determination  $D$ , expressed in percent, where  $D = 100 \cdot r^2$  and  $r$  is the Pearson correlation coefficient. To count as a real vibrato, the  $D$  must be at least 90 percent; when this criterion is matched, frequency, amplitude,  $D$ , and time position of the vibrato are all printed into a file. All pitch curves were analyzed in one step in a batch process, which provided a summary showing several location parameters and the total vibratos found in each recording.



**Figure 2.** Median rate and pitch range in vibratos. Left (○): Probe Trials. Points related to Non-Bad comments (Good + N/A) are filled grey. The black circle shows the recording of Emanuel Hollander (“Ne brany mena rodnaja,” 1911); right (◇): Records Trials: Diamonds related to Good comments are filled grey.

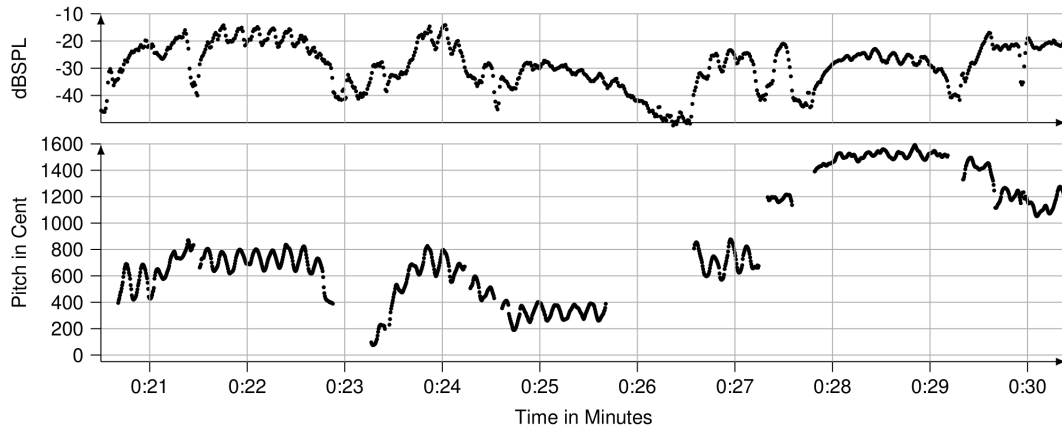
## 4.2 Exploration of the Pitch Vibratos of the Records Trials Corpus

### 4.2.1 OVERVIEW

For each of the 33 recordings we extracted the median vibrato rate and median vibrato pitch range. Figure 2 (left) shows the distribution of vibrato rate and range for the Records Trials. The point cloud shows no prominent features. This was a bit surprising, since we assumed that at least either the range had to be large for a “Bad” Tremolo or Overall, or the rate had to be extreme (either very fast or very slow).

### 4.2.2 THE HOLLANDER EXAMPLE: DOUBTING EDISON’S PERCEPTION AND A FIRST STEP TOWARDS AN ARTEFACT THEORY

In contrast to our expectations, the recording with the largest range (black circle in Figure 2), which even shows one of the slower rates, was surprisingly labelled with “Good” Tremolo. Any explanation for this is speculative; it is possible, however, that for Edison the tremolo is no pitch vibrato at all, but an irregular pitch deviation, which might indeed be a flaw. A closer appraisal of the recording, however, reveals that this



**Figure 3.** Pitch and dB SPL curve from a segment of Emanuel Hollander’s “Ne brány mená rodnája,” recorded 1911 (original Edison Amberol, Cat. No. 11202, taken from <http://cylinders.library.ucsb.edu>, visited 2017). Pitch is converted from hertz with  $0 \text{ ct} = 125 \text{ Hz}$ . The segment shows distinct pitch vibratos (0:21–0:26) with slow rates and pronounced ranges partly exceeding two semitones. The combination of gliding and vibrato (0:23–0:25) particularly intensifies the impression of the vibrato. Additionally, the voice shows an unstable, jittery pitch (0:28) and a clear volume vibrato (0:22). Yet Edison commented “Tune N[ot]. G[ood]. Scarcely any tremolo. Might use him.” Even though Edison disliked the tune, he decided to use Hollander for a series of “Russian Records” with ten songs, including “Ne brány mená rodnája” (Anonymous, 1911).

recording even showed a jittery voice, an unstable vibrato and distinct amounts of volume vibrato, as shown in Figure 3. The Hollander example contradicts nearly any hypothesis on the tremolo. A further explanation was provided by Tobias A. Weege, who did extensive measurements on gramophone parts to investigate how resonances of the recording and playing devices might influence the volume of the record when it is played (Weege et al., 2018). According to the henceforth called *Artefact Theory*, the devices react with different sensitivity upon different frequencies. If a voice of even volume is recorded on a gramophone disc, then, when the recording is played back, the output volume depends on the pitch that was sung. When singing a vibrato, its periodic pitch shift triggers an additional periodic volume shift, hence, a volume vibrato being an artefact produced by the devices. This effect depends on the interdependencies of all parts of the gramophone and the recording device that was used in the recording procedure. Thus, it is virtually impossible to reconstruct the precise effect of a specific recording without knowing all parts used in the recording procedure; yet, it is safe to say that those loudness modulations did happen, even though we cannot say exactly where. Given the *Artefact Theory*, a vibrato showing a large pitch range does not automatically result in an artificial volume vibrato, but it increases the chance that such a vibrato emerges.

In the Hollander example, the reverse case might have taken place: Suppose that despite the intense pitch range of the vibrato, Edison’s playing devices did not produce large artefacts; then Edison might not have been disturbed. Nevertheless, we do not know why Edison perceived “Scarcely any tremolo,” because the vibrato is clearly perceivable. We could consider that Edison might have missed the intensity of the vibrato because it fitted the song. In this case, the intense vibrato might be required for the Russian folk song; in other words, it is a necessary means for the interpretation. Then again, the comment “Tune N.G.” is problematic. Had Edison been impressed by the song, the assumption of a consistent interpretation would have been more conclusive. Obviously, Edison did not recognize the large vibrato, but it is unclear as to why; maybe because it did not disturb him in this case. It is also plausible that Edison may have become more aware of the tremolo when in fact something else disturbed him, like the timbre or the phrasing. The circumstance that he used a differentiated language in his memoranda does not rule out the possibility that his perception was less differentiated.

**Table 3.** Summary of the pitch vibrato analysis. Rate and Range are averaged values of the individual median values. E.g. the recordings with “Bad” Tremolos showed a median rate of 6.8 Hz and a median pitch range of 169 ct.

Condition	Group	Records Trials			Probe Trials			Test Trials (=Total)		
		n	Rate	Range	n	Rate	Range	n	Rate	Range
Tremolo	Non-Bad	29	6.7 Hz	135 ct	8	6.7 Hz	159 ct	37	6.7 Hz	140 ct
	Bad	4	6.3 Hz	148 ct	37	6.8 Hz	172 ct	41	6.8 Hz	169 ct
	Total	33	6.7 Hz	136 ct	45	6.8 Hz	169 ct	78	6.8 Hz	155 ct
Overall	Bad	12	6.7 Hz	133 ct	39	6.7 Hz	175 ct	51	6.7 Hz	165 ct
	Good	20	6.6 Hz	140 ct	4	7.5 Hz	130 ct	24	6.8 Hz	138 ct
	Total	32	6.6 Hz	138 ct	43	6.8 Hz	170 ct	75	6.7 Hz	156 ct

### 4.3 Pitch Vibrato Range in the Probe Trials

#### 4.3.1 CONSIDERATIONS FOR A FIRST HYPOTHESIS

So far, we can assume that there was little knowledge of pitch and volume vibrato. Considering the artefacts based on the production-reproduction chain of the gramophone (henceforth *transformation chain*), we have reason to believe that Edison’s use of the term tremolo might be a mixture of unclear perception and the artefacts resulting from the transformation chain. In fact, the small number of “Bad” Tremolos was a problem, which we hoped to solve by using the Probe Trials. For the following analyses, we decided to reduce the three possible labels (Good, Bad and N/A) to alternative data only. In the Overall condition, we left the N/A comments out, leaving 32 samples in the Records Trials and 43 samples in the Tremolo Trials, which were either labelled “Good” (G.) or “Bad” (B.). In the Tremolo condition instead, we merged the Good (G.) and N/A labels into one “Non-Bad” (N.-B.) group. This was based on the consideration that for Edison the tremolo was per se a negative quality and he was likely to comment on any tremolo that disturbed him. In the few “Good” comments he always emphasized the small degree of the tremolo. Therefore, we interpreted no comment with the absence of a bad tremolo.

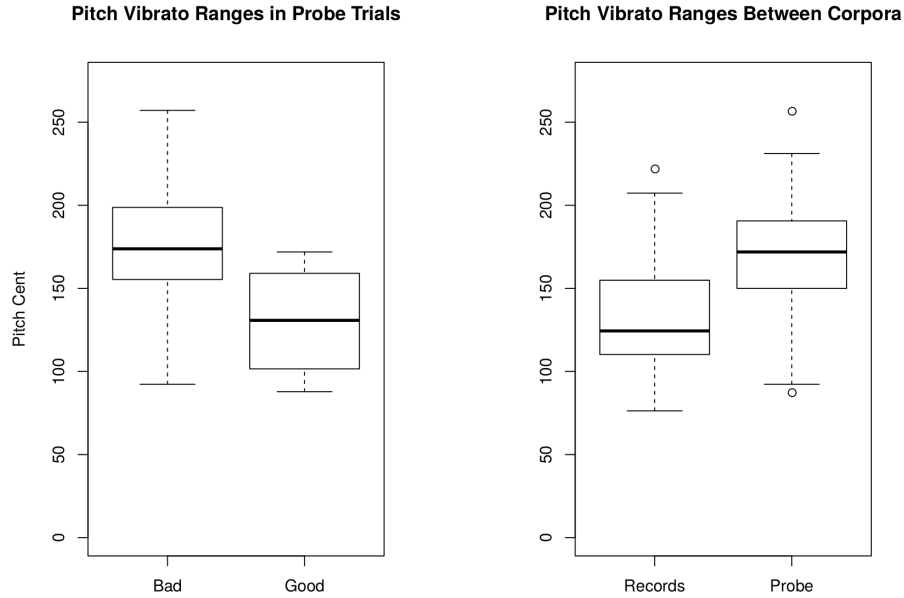
We derived a one-tailed hypothesis from our considerations: that in both conditions in the Probe Trials the vibrato range is larger in the respective “Bad” groups than in the “Non-Bad” (Tremolo) and “Good” (Overall) groups. We did not specify a hypothesis about the vibrato rate, because the Artefact Theory only relates to pitch changes and not to the vibrato rate.

#### 4.3.2 TESTING THE PROBE TRIALS

Table 3 shows the mean values of the vibrato rates and ranges for each corpus, condition, and group. All statistical analyses were carried out in R. Because of the small sample size in the Probe Trials, a Shapiro-Wilk normality test was used to decide if the data were normally distributed, with the criterion that *p* had to be larger than .2. The test revealed that the vibrato rate was not normally distributed ( $W = 0.973, p = .096$ ), but the vibrato range was ( $W = 0.986, p = .536$ ). This allowed the range to be analyzed with a parametric test (t-test), whereas we had to use a non-parametric alternative for the vibrato rate (Wilcoxon test).

We tested differences in vibrato range for “Bad” versus “Non-Bad” Tremolo, as well as “Bad” versus “Good” Overall, with a two-sample t-test. An F test was used beforehand to analyze if the variances of the respective samples differed significantly, which turned out not to be the case. The effect of the test was computed with Cohen’s *d*. As suggested by Platz et al. (2012, pp. 169–71), we additionally calculated the 95% confidence interval  $CI_{95}$  for the effect size *d* based on its standard error  $SE_d$ :

$$CI_{95} = d \pm SE_d \cdot 1.96 \quad \text{with} \quad d = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{(n_x - 1) \cdot s_x^2 + (n_y - 1) \cdot s_y^2}{(n_x - 1) + (n_y - 1)}}} \quad \text{and} \quad SE_d = \sqrt{\frac{n_x + n_y}{n_x \cdot n_y} + \frac{d^2}{2 \cdot (n_x + n_y)}}$$



**Figure 4.** Comparison of median vibrato ranges. Left: “Bad” versus “Good” groups in the Overall condition in the Probe Trial corpus, showing that Edison commented “Bad” more often when the pitch vibrato range was larger. Right: general differences in range between Records Trials and Probe Trials. In the Probe Trials, in which Edison rated more often “Bad” in both conditions, the ranges were generally larger than in the Records Trials.

Because of the Shapiro-Wilk tests, which indicated that the distribution of the vibrato rates was not normal, the respective differences in vibrato rate were tested with a Wilcoxon rank sum test with continuity correction as an equivalent to the Mann-Whitney test.

#### 4.3.3 RESULTS AND DISCUSSION

The results confirmed the hypothesis that “Bad” Overalls showed a larger vibrato range than “Good” Overalls ( $t_{41} = 2.51, p = .008, d = 1.32 [0.25; 2.38]$ , see Figure 4 left). There were, however, no significant differences in range under the Tremolo condition, neither were the differences in vibrato rate significant. Despite the significant difference in vibrato range under the Overall condition, the unbalanced size of the samples (39 “Bad” versus 4 “Good”) is problematic. It might rather be interpreted as an indicator encouraging further testing.

A comparison of the two corpora, the Probe and Records Trials, showed that the vibratos in the Probe Trials had a much larger range (169 ct) than in the Records Trials (136 ct, see Figure 4 right). A two-tailed t-test revealed that this difference was highly significant, and that the corpus had a large effect upon the vibrato range ( $t_{73} = -4.07, p < .001, d = 0.93 [0.46; 1.41]$ ). Figure 2 and Table 1 also show that the Probe Trials were often rated “Bad,” whereas the Records Trials included more “Good,” or at least more “N/A” comments. This is not surprising, and one can easily imagine that in sum, Probe Trial recordings were of lesser quality than commercial recordings for both technical and artistic reasons. Firstly, the material in the Records Trials already underwent some technical selection process. We know from Edison’s Notebooks (Edison, 1911–1919), but also from other Record Companies, such as Victor Records (*Red Seal Recording Book*, 1906), that sound engineers frequently changed parts of the recording device – such as the horn, its connection tube and the sound box, which transfers the air pressure into the movement of the needle – before the recording was released for commercial pressings (Edison, 1911–1919). It is very likely that commercial pressings showed less disruptive artefacts than test records did. The effort of making a Probe Trial recording was probably less intense, since the Probe Trials were recordings for candidates and were not made for commercial purposes. Secondly, the vocal ability of artists who did commercial recordings had also passed some criteria, whereas mere candidates had not necessarily undergone a selection before recording for a Probe Trial. Although other record companies did not make Probe Trials the way Edison did, they surely

selected their candidates as well for commercial recordings, excluding candidates with rather limited capabilities.

#### 4.4 Merging Records and Probe Trials into a Combined Test Trials Sample

After merging the corpora of the Records and Probe Trials, we had a better-balanced number of samples in the henceforth called *Test Trials* corpus. The mean values for range and rate in the Test Trials are shown in Table 3, but differences between the groups in the Test Trials corpus could not be tested. This would have been a methodological flaw, since our hypothesis for range differences was partially based on the exploration of the Records Trials corpus. Since the Records Trials are part of the merged corpus, logic forbids that a hypothesis is tested on the same material used for its construction (Section 4.2). The merged corpus is thus only used for the following analysis of the volume vibrato.

### 5 VIBRATO RELATED CHANGES OF SOUND PRESSURE LEVEL

#### 5.1 Preliminary Considerations

Measuring volume of vibrato is virtually impossible for different reasons. First and foremost, the voice is mostly accompanied by instruments, which have a volume of their own and are intermixed (particularly in mono records). Secondly, in the recording process, the level of the singing voice is not transformed evenly, which means that the sound pressure level (dBSPL) produced in the playback of the record does not reflect the real volume that had been produced in front of the recording horn. The volume we hear depends on the frequency which is put in, or in other words, how a tone is sung. The pitch of the voice may be of lesser importance than its timbre and the vowels it produces, because a tone has many overtones, which are all influenced by the transformation, but only one fundamental frequency. The third major problem is the noise of the recording, which always accompanies the singer. What we can measure is the dBSPL the record provides. Since the singer's voice is usually louder than the accompaniment, it is possible to detect changes in dBSPL that correspond to a vibrato. These level undulations may nevertheless be more important for the impression of the recorded voice than the actual volume vibrato; for instance, one can imagine that an intense volume vibrato might be perceived to be less prominent when it is embedded in a loud orchestral accompaniment, so that the volume produced by the singer does not stick out of the general level of the recording. A similar voice might be perceived as having a strong vibrato, when the accompaniment is only soft or even absent, because the general level of the recording is highly influenced by the singer's voice alone.

#### 5.2 Measuring dBSPL Ranges and $2f$ -Vibratos

In contrast to the pitch vibrato, which could be analyzed automatically, the loudness level of the vibrato was analyzed manually to a greater extent, because it requires a listener to decide if the loudness level corresponds to a vibrato, or to something else. This personal estimation was based on a combination of close listening and observing loudness and pitch curves. We had our assistant use the Sonic Visualiser's spectrogram view, the pitch curve and the RMS curve, which was created by the RMS Amplitude Plugin (Version 4) of the libXtract Plugin by Chris Cannam and Jamie Bullock. The assistant marked each local maximum of the RMS curve if it corresponded clearly to a loudness-vibrato of the RMS amplitude curve.

##### 5.2.1 $2F$ -VIBRATOS

After some experience with the procedure, our assistant reported that the level-vibrato frequently showed a doubled rate compared to the pitch curve. The occurrence of a volume vibrato twice as fast as the pitch vibrato (henceforth  $2f$ -vibrato) was surprising at first because we assumed that, by definition, the frequency and phase of the volume and pitch vibratos should be the same. We also observed a slight phase-shift of the (normal) loudness vibrato, the local maximum of which was roughly at the following zero-crossing point of the pitch curve, but not always. Therefore, we decided not only to check the ranges of dBSPL vibratos, but also to analyze their phase shift compared to the pitch vibrato, and the number of  $2f$ -vibratos.



### 5.2.2 DBSPL RANGE ANALYSIS

With a script programmed in R, we converted the RMS data into sound pressure level with  $dB_{SPL} = 20 \cdot \log_{10} RMS$ . The time positions for loudness maxima had been set manually. In case the manually set markers were not placed exactly, they were corrected to the maximum of the curve within a distance of 1/3 of the space between two maxima. In special cases, where the RMS curve showed two nearby peaks due to distortion effects, the markers were set carefully and labelled to indicate that they should not be corrected automatically to the nearby maximum. Within a window spanning three successive local maxima, corresponding to two complete cycles of the vibrato, the frequency of the volume-vibrato was computed by taking the mean distance between the three successive markers. The vibrato range was computed as the range of dB SPL values within the window. It is important to recall that the range of the dB SPL vibrato measured expresses the difference between the upper returning point of the level vibrato and the background level of the accompaniment or recording; it does not necessarily express the complete range of the vibrato.

### 5.2.3 PHASE SHIFT AND $2f$ -VIBRATO ANALYSIS

To estimate the phase shift between the dB SPL vibrato and the pitch vibrato, a sine representation of the dB SPL vibrato was created as an ideal, based on the dB SPL analysis so far, and then compared to the pitch vibrato. The sine was phase-shifted until a maximum correlation between the model vibrato and the pitch curve within the window was found. The stepwise phase shift was reduced to  $0.1 \pi$ , which means an accuracy of the procedure of less than  $\pm 0.1 \pi$ . To detect a  $2f$ -vibrato, the frequency of the model vibrato was halved before it was correlated with the pitch curve; if this halved model exceeded the correlation of the normal model, it was decided that the frequency of the model was approximately twice as large as the pitch vibrato. Either way, a minimum correlation of .8 was necessary to account for a vibrato; otherwise, the result for the window was not considered any further. A manual control showed that the correlation method was reliable. Whether the decision was made for or against the  $2f$ -vibrato, the respective alternative always showed a correlation distinctly lower than .8, and in most cases no correlation at all.

## 5.3 $1f$ -Vibrato Analysis

For clarity, we use the term  $1f$ -vibrato for the normal dB SPL vibrato and  $2f$ -vibrato for the special case, in which the dB SPL vibrato has a frequency that is two times the frequency of the pitch vibrato. The following subsections deal only with the  $1f$ -vibrato.

### 5.3.1 ANALYSING DIFFERENT LOCATION PARAMETERS

For the analysis of the  $1f$ -vibrato range we considered several location parameters of each recording, which were:

- **Q1:** The first quartile, indicating that 3/4 of all  $1f$ -vibratos had a dB SPL range larger than or equal to Q1. Q1 represents the threshold for most  $1f$ -vibratos, which might be more important than the average – particularly when we consider Edison's statement that a tremolo here and there might be acceptable if it is used as a means for interpretation.
- **MD:** The median as a measure of the central tendency should nevertheless be ignored, because it represents the point that divides the larger and smaller half of all ranges found in one recording.
- **Q3:** In contrast to the Q1, the third quartile (Q3) represents the threshold of the largest quarter of the  $1f$ -vibratos in terms of range. Q3 therefore indicates the vibratos with large range compared to all the ranges found in the recording. Whereas the assumption for Q1 is that the mass of all volume vibratos is responsible for the general impression of the tremolo, Q3 is tested, because it is also possible that large instances – rather than the mass – shape the impression of the tremolo in general.
- **Q90:** The ninth decile is a more extreme alternative to the Q3, for the Q90 represents the threshold of the largest tenth of the vibrato ranges in a recording. Or in other words, the largest ten percent of all  $1f$ -vibratos in the respective recording have a range larger than or equal to Q90. The maximum range is not used, because a single outlier might not be perceived at all and may be the result of a measurement error (e.g. a crack in the record or of unknown origin).

**Table 4.** ANOVA of location parameter Q3 with dB SPL range as dependent factor, and both Tremolo condition and corpus as independent factors. Edison’s judgement for “Bad” Tremolo is related to an increased dB SPL range of the 1*f*-vibrato. Here, the corpus (Records versus Probe Trial) shows barely any effect, but what is more important, is that no interaction effect between Tremolo groups and corpus was observed.

ANOVA: Q3 * Tremolo * Corpus:	Df	Sum Sq	Mean Sq	F	p	
Tremolo: ‘Bad’ vs. ‘Non-Bad’ groups	1	10.50	10.501	4.076	.0472	*
Corpus: Records vs. Probe Trials	1	8.56	8.560	3.332	.0724	.
Interaction: Tremolo:Corpus	1	0.33	0.323	0.127	.7224	
Residuals	73	188.08	2.576			

Signif. codes: 0 ‘\*\*\*’ .001 ‘\*\*’ .01 ‘\*’ .05 ‘.’ .1 ‘ ’ 1  
1 observation deleted due to missingness

**Table 5.** ANOVA of location parameter Q3 with dB SPL range as dependent factor, and both Overall condition and corpus as independent factors. In contrast to the Tremolo condition, there are no significant differences between the groups in the Overall condition. Here, the corpus has a significant effect, but again no interaction effect between groups and corpus was observed.

ANOVA: Q3 * Overall * Corpus:	Df	Sum Sq	Mean Sq	F	p	
Overall: ‘Bad’ vs. ‘Good’ groups	1	3.63	3.629	1.356	.2483	
Corpus: Records vs. Probe Trials	1	15.91	15.915	5.946	.0173	*
Interaction: Overall:Corpus	1	0.18	0.180	0.067	.7961	
Residuals	70	187.38	2.677			

Signif. codes: 0 ‘\*\*\*’ .001 ‘\*\*’ .01 ‘\*’ .05 ‘.’ .1 ‘ ’ 1  
1 observation deleted due to missingness

### 5.3.2 PREPARING THE DB SPL ANALYSIS

The dB SPL analysis was carried out with the Test Trials corpus; therefore, the number of samples was large enough to use a t-test under the assumption that the distribution of samples was unimodal and that the variances between the test groups did not vary significantly. Although these criteria were fulfilled for all location parameters, all of them were tested additionally for normality with a Shapiro-Wilk normality test. The Q1 and MD showed a significant difference to a normal distribution ( $p < .05$ ), Q90 showed no significant difference, but the difference was still too large to assume the contrary ( $W = 0.974, p = .113$ ); the same was true for Q3 ( $W = 0.977, p = .167$ ) but with a closer tendency to a normal distribution. The decision was to test all location parameters for differences between groups with a t-test and, additionally, to run two analyses of variance with the Q3 to test interaction effects of the corpus (Probe versus Records) and the two groups for each condition (“Bad” versus “Good” Overall or “Bad” versus “Non-Bad” Tremolo). The Q3 was selected because it showed the least non-normal distribution in the Shapiro-Wilk test. The one-tailed hypothesis for the t-tests was that 1*f*-vibratos had a larger dB SPL range in “Bad” groups than in the respective alternative groups in both the Tremolo and Overall conditions.

### 5.3.3 RESULTS AND DISCUSSION

One instance of the corpus had to be excluded from the analysis because it showed large distortions due to an obviously large crack in the recording, which produced intense volume jumps. This distortion had no influence on the pitch analysis, but it was of no use for the dB SPL analysis.

A summary of the analyses of variance is shown in Tables 4 and 5. Differences between the groups were significant only under the Tremolo but not the Overall condition. Although there were differences between the corpora (Records and Probe), no interaction effects were found between the respective condition and the corpora.

**Table 6.** Results of the t-test of location parameter differences between the two groups “Bad” ( $n = 41$ ) and “Non-Bad” ( $n = 36$ ) in the Tremolo condition. Differences in dB SPL ranges between Bad and Non-Bad Tremolos were significant for all location parameters. Despite some large effects, expressed as Cohen’s  $d$ , its confidence interval  $CI_{95}$  is unfortunately large, which relativizes the effect. Nevertheless, all results support the hypothesis constructed earlier.

Location Parameter	Mean		S.D.		t	df	p	d	CI <sub>95</sub>
	B.	N.-B.	B.	N.-B.					
Q1	4.63	4.04	1.02	0.97	2.559	75	.006	0.58	0.13 – 1.04
MD	5.74	5.21	1.23	1.41	1.747		.042	0.40	–0.05 – 0.85
Q3	7.11	6.37	1.59	1.65	2.000		.025	0.46	0.00 – 0.91
Q90	8.36	7.46	1.91	1.88	2.090		.020	0.48	0.02 – 0.93

All location parameters showed differences between the groups under the Tremolo condition, which were significant at the  $\alpha = .05$  level (see Table 6 and Figure 4). This means that 1*f*-vibratos in “Bad” Tremolos showed a larger dB SPL range than “Non-Bad” Tremolos. Notably, among the location parameters, the Q1 showed the largest effect ( $d = 0.58$ ), supporting our hypothesis that the judgement of tremolo might be based on many of the 1*f*-vibratos rather than on their central tendency. Q90 showed the second largest effect with  $d = 0.48$ . Due to the small sample size, the confidence interval relativizes the interpretation of the effect size; the differences are nevertheless significant. In the Overall condition, none of the location parameters showed a significant difference between the “Bad” and “Good” group.

The results support our hypothesis that the volume vibrato was indeed the “problem” for Edison: The effects are evident in the Tremolo condition but not in the Overall condition. Whether produced artificially or not, when perceived they are problematic. Considering the limited loudness capabilities of the records between noise and overdrive, loudness changes were generally difficult. The still pronounced effect of the Q90 is also interesting. The large mass of dB SPL vibratos might be most responsible for the impression of tremolo, but this impression is also influenced by the largest ten percent of the vibratos, which can turn a judgement from bad to worse.

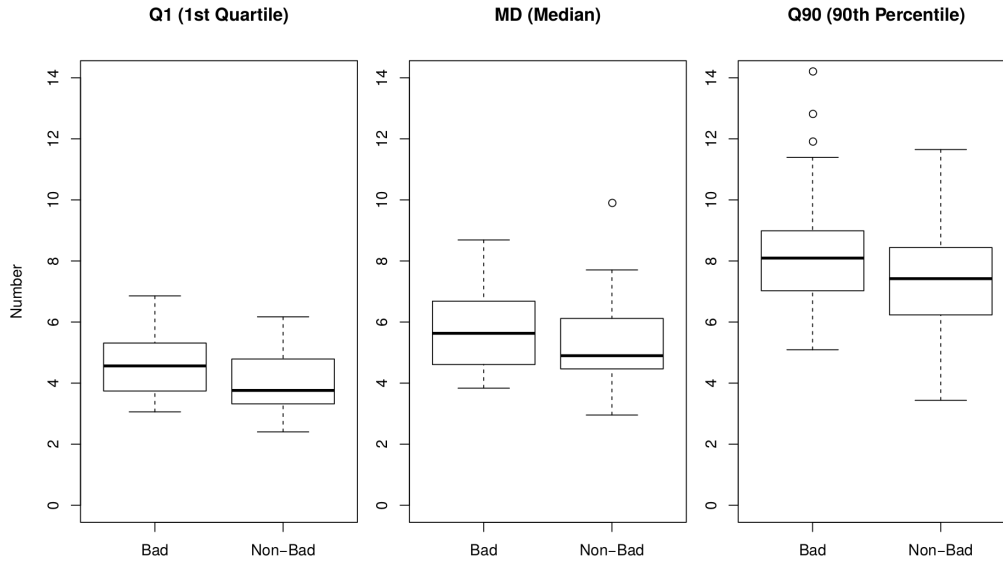
## 5.4 2*f*-vibratos

Up until now, we have analyzed the differences in dB SPL ranges as an indication for volume vibrato. Now we also have a look at the phenomenon of the 2*f*-vibrato. Before the analysis, we will elaborate the Artefact Theory further, to which the 2*f*-vibrato seems to be related.

### 5.4.1 THE ARTEFACT THEORY

Boris-A. Bolles (Kob et al., 2018) investigated the effect of the recording and reproduction chain on the output volume of a vocal signal based on the measurements of Weege et al. (2018). He could demonstrate that it is theoretically possible to create an artificial volume vibrato from a pitch vibrato with constant volume by simulating its transformation in the recording and reproduction process. As a stimulus, he used a modelled saw-tooth signal roughly imitating the vowel “a” by adding two formant resonances. When sweeping the pitch of the signal through frequencies from 0.2 to 1 kHz, he observed large differences in the output level and thus obtained a response curve for this transformation chain and the vocal signal. At one point (approximately 0.6 kHz), this response curve showed a steep increase of about 10 dB SPL within a half tone interval. Using a pitch vibrato of constant volume as input signal, he could show that there are indeed pitches at which point a distinct volume vibrato was added to the pitch vibrato. Moreover, when the input vibrato’s mean pitch was located at a local minimum of the response curve, any change of pitch, higher or lower, would result in a level increase. In this case, within one cycle of the input pitch vibrato, the loudness response passes through two local maxima, one at the upper and another at the lower turning point of the pitch vibrato, hence, creating a 2*f*-vibrato.

The demonstration that it is possible to artificially add a loudness vibrato to a pitch vibrato, even with a doubled frequency, does not rule out the possibility that such a thing can be done vocally. It could be possible that a quasi-2*f*-vibrato happens during a trill, when both reversal points are distinct tones between which the trill changes back and forth. The mental targets, to which the singer aims, are the upper and lower



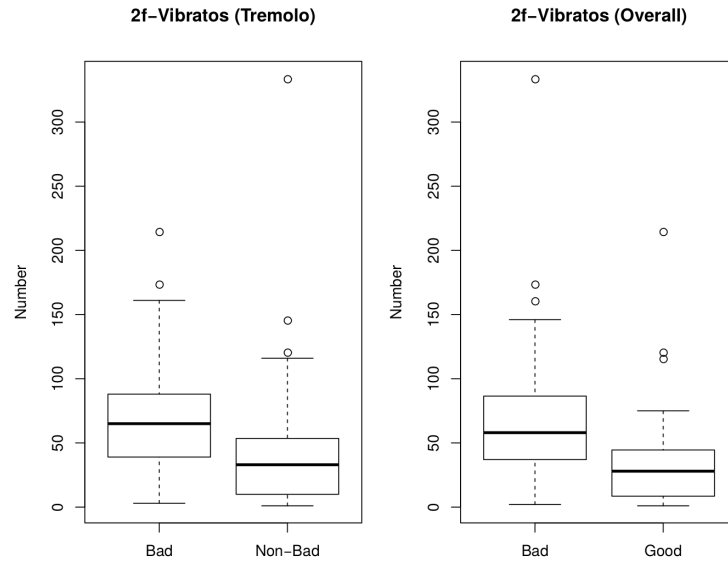
**Figure 5.** Comparison of location parameters of dB SPL ranges between “Bad” and “Non-Bad” Tremolo groups. The differences are significant, but most pronounced for the Q1, which supports the hypothesis that Edison based his comment on the tremolo on the ranges of most vibratos, rather than on a central tendency. The Q90 shows a larger effect than the MD, which indicates that a few very large ranges might have contributed to his decision. As the location parameters are highly correlated with each other, a distinction between the effects of the location parameters, however, remains speculation.

turning points of the trill, and it is possible that the singer somehow gives them a slight impulse. Since we analyzed vibratos and not trills, this possibility, however, does not apply here. Another point is that the loudness of the pitch vibrato is modelled by the resonance of the vocal tract. After all, the formant-based resonance is but another resonance modelling the output loudness, and this resonance also depends on the frequencies created by the vocal folds. A vibrato sung with the right proportion of formants (or wrong proportion, depending on the perspective), may also result in a modulation of the loudness of a pitch vibrato. Finally, we do not have the means to detect the reason for the loudness vibrato in old recordings, since we do not know the exact equipment that had been used – and the singers, who did the recording, cannot be asked to sing again. Yet, the Artefact Theory is appealing and will therefore be considered further in the following subsections.

#### 5.4.2 ANALYSING THE $2f$ -VIBRATO

In the analysis of the  $2f$ -vibratos, we compared the total number of  $2f$ -vibratos between the groups in both conditions. A look at the distribution showed that we could neither assume a normal distribution for the  $2f$ -vibrato amounts, nor an equal variance between groups. Therefore, we used the Wilcoxon rank sum test with continuity correction.

Additionally, we considered that the total number of  $2f$ -vibratos in a recording does not only depend on the singer’s vibrato and the transformation chain. After all, songs with many long tones and a sparse accompaniment have more vibrato and therefore are more likely to show  $2f$ -vibratos; and even simpler: A very short song with larger instrumental sections has a smaller number of vocal vibratos than a long recording in which the singer sings all the time. Thus, it might be insufficient to compare only the sum of  $2f$ -vibratos without considering the number of  $1f$ -vibratos as well. Therefore, we constructed three (one-tailed) hypotheses. The first hypothesis was that Edison generally disregarded loudness vibratos as tremolos; thus, the total number of  $1f$ -vibratos alone should be larger in the “Bad” groups than in the alternative group of the respective condition. The second hypothesis was that  $2f$ -vibratos particularly indicate a distortion, which can be explained best with the Artefact Theory. Therefore, the number of  $2f$ -vibratos is even more critical and should be greater in the “Bad” groups. The third hypothesis regarded the consideration that the absolute total of vibratos might be affected by the character of the songs. Therefore, we computed the ratio between the number of  $2f$ -vibratos and  $1f$ -vibratos and hypothesized that this ratio should be larger in the “Bad” groups.



**Figure 6.** Box plot of the number of  $2f$ -vibratos in both groups for both conditions.

#### 5.4.3 RESULTS AND DISCUSSION

Table 7 and Figure 6 show a summary of the main results.

The total number of  $1f$ -vibratos was not larger in the “Bad” groups; it was even smaller. This clearly and surprisingly contradicts the hypothesis that Edison disliked loudness vibratos in general; the high  $p$  value is due to the one-tailed hypothesis. Although it is forbidden to change the hypothesis *a posteriori*, we tested the alternative just out of curiosity, with the result that the difference would have failed significance when tested two-tailed, or even with the opposite one-tailed assumption.

The second hypothesis is supported, because the larger number of  $2f$ -vibratos in the “Bad” groups was highly significant. The occurrence of  $2f$ -vibratos seems to have a large effect on Edison’s prejudice against the tremolo. In the light of the rejected hypothesis concerning the  $1f$ -vibrato before, this result is the more striking. It is possible that indeed the absolute total of  $2f$ -vibratos leads to a negative comment.

The third hypothesis could also be confirmed. Negative comments are related to recordings with relatively more  $2f$ -vibratos than positive or neutral comments. The fact that in the “Bad” Tremolo group the ratio even exceeds 1.0 is notable. These recordings showed even more  $2f$ -vibratos than  $1f$ -vibratos; however, the numbers are somewhat misleading, because there are, by definition, twice as many  $2f$ -vibratos at the same time as  $1f$ -vibratos. Nevertheless, the sheer number of  $2f$ -vibratos is still remarkable, but with the previous two results in mind, it is not surprising that the third hypothesis was confirmed.

### 5.5 Phase-Shift-Exploration

The analysis of the phase shift between the pitch vibrato and the volume vibrato ( $1f$ -vibrato) was exploratory in nature, since we had no basis on which we could create a hypothesis. Figure 7 shows the distribution for all median phases from the recordings of the Test Trial corpus. Except for the “Non-Bad” Tremolo group, in which the median phase shift was  $0.525\pi$ , the median was always  $0.5\pi$ . The results are in accordance with our subjective impression. However, the median hides the fact that the phase is not always  $0.5$  and, for now, we do not have a plausible theory explaining why the phase differs in so many cases. Small deviations from the median phase shift may be explained by the accuracy of the measurement or the alignment procedure, but not to that extent, because there are substantial amounts of recordings with a median phase shift

**Table 7.** Results of the analysis of  $2f$ - and  $1f$ -vibratos. Contrary to the hypothesis, the number of  $1f$ -vibratos was even less, but not significantly less in the “Bad” (B.) groups, than in the “Non-Bad” (N.-B.) groups. The number of  $2f$ -vibratos is significantly increased in the B. groups. The ratio of  $2f$ -vibratos/ $1f$ -vibratos is significantly increased in B. groups, indicating that the proportion of  $2f$ -vibratos influenced Edison’s dislike of the tremolo. Together the results support a hypothesis which is based on the Artefact Theory, according to which the volume changes that are caused by the recording and playback devices are probably more responsible for the “Bad” comments than the vibrato as it was sung.

	Tremolo				Overall			
	B.	N.-B.	W	p	B.	G.	W	p
$1f$ vibratos (amount)	57	66	677	.735	59	65	516.5	.795
$2f$ -vibratos (amount)	65	33	1018	.002	58	28	840.5	.002
$2f$ -vibratos / $1f$ -vibratos	1.09	0.57	988.5	.005	0.83	0.29	813.0	.004

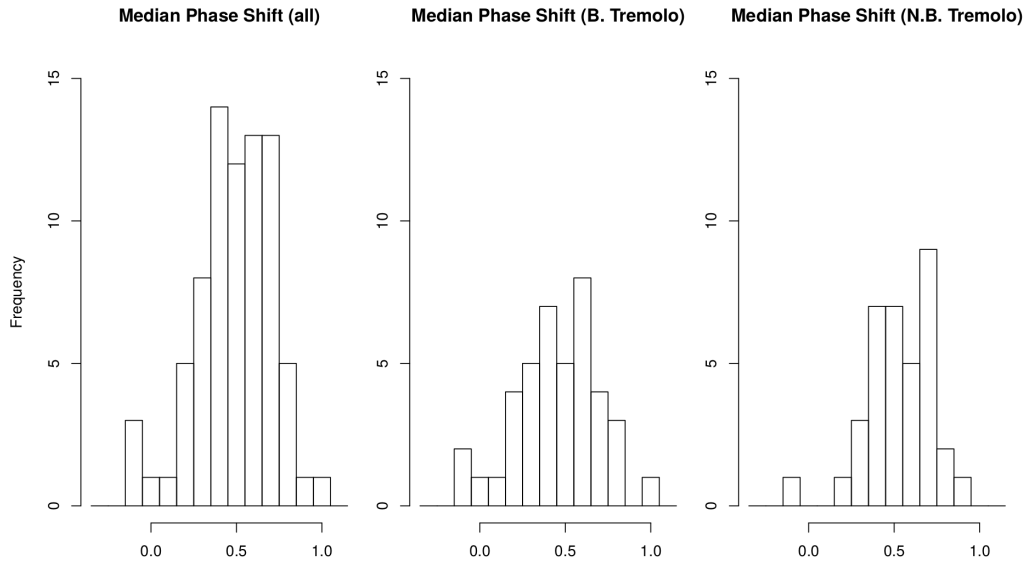
between  $0.25\pi$  and  $0.75\pi$ . If we suppose that the phase shift is determined biophysically by the production of the vibrato in the vocal tract or other parts of the body, then deviations from this ‘natural’ phase shift may be the result of additional resonances caused by the recording and reproduction chain, or other influences. If produced artificially, we could further assume that such phase shift deviations might sound implausible or strange. However, since there is no difference in median phase shift between the groups, this assumption is not supported by the data. We present these observations regardless, as they might encourage further research on this matter.

## 6 DISCUSSION

In his voice trials, Edison made numerous comments on singing voice recordings, particularly referring to what he called tremolo. We linked his comments to the rate and range of the vibratos in these recordings, both with respect to pitch and sound pressure level. In doing so, we are a step further towards understanding Edison’s dislike of the tremolo. It seems as if he was mostly troubled by mechanical artefacts that distorted the loudness of the vibrato that was originally produced by the singers. This conclusion is based on several observations.

Firstly, the Artefact Theory predicts artificial dB SPL vibratos, which are caused by the specific response of the recording and playback devices of that time. The Artefact Theory is derived from the observation that the output loudness of a record differs with the frequency of the input signal. Due to large differences in the response of the devices on different frequencies, a frequency change is accompanied by an artificial loudness change when the recording is played back. A pitch vibrato is likely to trigger artificial and audible volume changes. The intensity of the artefact depends on the transformation chain and on the frequencies, hence, on the pitch and timbre of the vocal vibrato. The probability of increased artefacts should therefore be higher when the pitch range of the vibrato is larger, because a larger vibrato covers a larger frequency range. In Section 5.3.3 we showed that the recordings Edison rated “Bad” had vibratos with higher dB SPL ranges. According to the Artefact Theory, the emergence of additional dB SPL vibratos depends on the timbre with which the tone is produced and is directly caused by the pitch undulation of a vibrato. A larger pitch range was partly found in recordings Edison rated “Bad” (Section 4.3.2). Although the statistics were rather weak because of very unbalanced samples, the results did point in the same direction and at least did not contradict the theory.

Secondly, the Artefact Theory also explains the phenomenon of  $2f$ -vibratos, which is currently hard to explain otherwise. If the output volume changes with the frequency put in, then it is possible that the pitch varies around a point of a local response minimum. Around such a local minimum, the sound level increases each time the pitch deflects upwards and downwards from the vibrato’s central pitch, causing a full sound level vibrato cycle during half a pitch vibrato cycle, hence a level of periodicity which is twice as fast as the pitch vibrato. At a local maximum, the volume decreases each time the pitch deflects from the central pitch and increases again towards the central point, also causing a doubling of the vibrato rate. The significant increase of  $2f$ -vibratos in “Bad” groups (Section 5.4.3) thus supports the Artefact Theory further. Interestingly, in the “Bad” group the numbers of  $1f$ -vibratos did not differ significantly and were even less



**Figure 7.** The median phase shift between the  $1f$ -vibrato and the pitch vibrato is roughly  $0.5\pi$  for all groups. However, there are always substantial amounts of vibratos with different phase shifts, which still lack an explanation.

than in the other groups. That may be consistent with the Artefact Theory, but that is beside the point: It indicates that Edison might have even had no problem with vibratos at all, even when they showed distinct amounts of dB SPL ranges (if they were  $1f$ -vibratos). The conclusion that Edison was rather disturbed by technical artefacts is consistent with his idea of a perfect recording and is also in accordance with his statements discussed in Section 2.1. Still, the question remains, to what degree Edison had been aware of that? The Hollander example discussed in Section 4.2.2 is an impressive instance of a recording with distinct and diverse vibratos, in which Edison perceived “Scarcely any tremolo”, probably because it showed only small amounts of artefacts to him. Since the timbre, and hence the enunciation of vowels, is an important factor in the Artefact Theory, slight accents or differences in language can have large (and unpredictable) effects on the occurrence of these artefacts; maybe Hollander was lucky. Many of Edison’s comments let us assume that the reverse case also happened frequently, which would be that Edison tended to blame artefacts on the singers, who were not responsible.

From the perspective of record production, changes in volume were a problem in the recording procedure, and even more so in a mechanical recording, where it was impossible for the producer to control the level changes during a performance, except to move the singer closer or farther away from the recording horn. The dynamic range between noise and overdrive was very limited. Changes in dB SPL emerge artificially through the transformation of the input signal, but the characteristics of the response curves, and therefore of the artefacts, depend on all components and their interactions. It is unknown what components Edison used for recording and playback and considering that in the Records Trials he listened to records from several competitors, it is out of the question that we will ever know all the components used from the recording up to the playback.

Edison’s judgement can neither be fully explained by differences in dB SPL ranges of  $1f$ -vibratos, nor by the number of  $2f$ -vibratos alone. Figures 5 and 6 make clear that although the respective features that have been compared show a significant difference in their distribution, they also share substantial amounts of overlapping areas. For instance, in Figure 5 the median range of the location parameter Q1 is only 0.59 dB SPL larger in the “Bad” group than in the “Non-Bad” group. But what is more important, is that many vibratos in both groups show range sizes between about 3.2 and 6 dB SPL. Similarly, in Figure 6, the “Bad” Tremolo group has a median amount of 65  $2f$ -vibratos, which is nearly twice as much as the 33  $2f$ -vibratos in the “Non-Bad” Tremolo group. However, there are instances between zero to 100  $2f$ -vibratos in both groups. It seems reasonable to assume that Edison’s impression of the tremolo was at least influenced by a mixture of vocal quality, actual vibrato, and artefacts. How much influence each of these three factors had is currently unknown. Edison’s writings suggest that he had a theoretical knowledge which did not fall behind

the knowledge of his time; he even showed some instinct for technical relationships that have not been fully understood. What he might have lacked was bringing together knowledge, instinct, and his auditory perception. It is quite possible that he mixed (a) aesthetic judgements of vocal vibratos, (b) perceptions of artefacts based on the transformation chain, including artificial volume vibratos as well as other distortions (e.g. overdrive phenomena occurring at the peaks of loudness vibratos), and (c) aesthetic judgements of vocal features that were independent from the vibrato. In the end, the tremolo was blamed for this mixed perception. In other words, Edison's prejudice against the tremolo, which he shows clearly in his writings (as did most of his contemporaries) very likely influenced his judgement. The probable result was that he made the tremolo responsible for vocal or even interpretative characteristics that he disliked.

Although Edison's memoranda suggest that he was not very clear about what he perceived, his lack of clarity should not automatically be viewed as ignorance, an attribute Peters (2003) related to Edison when comparing him to Hermann von Helmholtz. We should bear in mind that a precise analysis of the complexity of the vibrato did not start until Seashore and his colleagues undertook enormous efforts in the late 1920s. With today's tools, the analysis of volume level and pitch poses no challenge, and it is easy to forget how difficult, if not impossible, it was then even to get an idea of what a vibrato is, how it can be described mathematically, and how it interacts with resonances and formants. Once we have seen some vibratos in a spectrogram, its characteristics are clear. Still, Joseph Tiffin (1932) complained about the difficulty of getting a loudness curve to interpret. He wished

to obtain an oscillogram of the sound wave, divide it into hundredth of a second periods, compute the relative energy [...] by making a mean-square analysis of the curve [...], but at his time [...] the labor involved in such a scheme can easily be seen to be prohibitive. (Tiffin, 1932, p. 162)

Tiffin asked for nothing but a computer, only he was not aware that such a thing could possibly exist. However, he knew exactly which operations would have been necessary. This example shows that knowing the solution mathematically or in principle does not automatically solve the problem.

Finally, we want to stress that we did not aim to reveal Edison's gaps in knowledge. Instead, our goal was to call for a certain consciousness in dealing with historical sources, particularly records in this case, which are by no means a representative selection of (the singing of) the time, but rather a selection that follows specific conditions. In our case, the conditions that steered the selection are not just based on the taste of a few, or their power, or luck.

The present paper was a first step towards a better understanding of how these selection processes took place and which influences might have played an important part. Therefore, it was necessary to understand Edison's comments, particularly those concerning the tremolo. Although the documents suggest the assumption that Edison's focus on the tremolo was personally motivated, and his definition of the tremolo seems as elusive given the inconsistent records in his memoranda, our analyses show another important factor, which is related to technology. The role which technological constraints play in the steering of the selections has been underestimated, at least in historical science and historical musicology. With our approach, we could show that behind a mere terminological problem ("What did Edison mean by 'tremolo'?") probably hides a technological problem ("What caused the vibrato to sound distorted so that Edison thought of it as a 'Bad' tremolo?"). Now we can assume that not only was tremolo a selection criterion, but also the timbre. Even the slightest change in enunciation could have made the difference between triggering disturbing artificially loudness vibrations in the recording and the impression of a smooth and barely noticeable tremolo. The timbre that was wrong for the recorder might not have been wrong on the stage at all, because it was not an aesthetic issue; it was a question instead of how the apparatus responded. As Eugen Fisher put it in the preface of his "Neue Gesangsschule," which was a treatise on singing that included accompanying gramophone records of then-distinguished artists singing the exercises:

Nicht jeder große Künstler ist auch ein guter Grammophonsänger oder -sprecher und – um nur ein Beispiel anzuführen – Sonnenthal konnte trotz aller seiner Kunst vor dem Apparat nicht verwendet werden. (Fischer, 1910, translates into "Not every great artist is a good gramophone singer or speaker, and – to give but one example – despite all of his art, Sonnenthal could not be used in front of the apparatus.")



Further investigations will be the future task for us and others, particularly for acousticians and physicists, who should be – and already are – included in this interdisciplinary undertaking.

### ACKNOWLEDGEMENT

We would like to thank Tobias Andreas Weege and Boris-Alexander Bolles for their contribution to the investigation of the Artefact Theory and their contribution to numerous discussions about it. We also thank David Astinet for researching the audio-files and transcriptions of Edison's notes and for his immense help in providing us with the necessary pitch-curves, and Gregor Schmitz for his careful and observant work on the RMS-curves including the time-consuming marking of every local maximum of a volume vibrato. This article was copyedited by Scott Bannister and layout edited by Diana Kayser.

### FUNDING

This work was supported by the Deutsche Forschungsgemeinschaft under Grant No. 289601849: "Technologies of singing. Investigations of the dispositive singing – body – media at the dawn of sound recording."

### NOTES

[1] This script is available at [www.tilohaehnel.de](http://www.tilohaehnel.de).

### REFERENCES

*Red Seal Recording book* (1906). Sony Archive, New York.

Anonymous (1911). Edison Russian Records by Emanuel Hollander. *Edison Phonograph Monthly*, April, 19.

Anonymous (1916a). An Interview with Thomas A. Edison regarding the imperfections of the human voice. *Edison Companies, Box 28: Advertisements [Primary Printed]*. Folder Edison Re-Creations 2, Thomas Edison National Historical Park.

Anonymous (1916b). Making Music more musical. Thomas Edison National Historical Park.

Easley, E. (1932). A Comparison of the Vibrato in Concert and Opera Singing. In C. E. Seashore (Ed.), *The Vibrato, vol. I of Studies in the Psychology of Music* (pp. 269–275). Iowa: University Press.

Edison, T. A. (1910–1912). *Pocket Notebooks 10-04-03 – 12-08-16*. Folder 11-12-01. Thomas Edison National Historical Park.

Edison, T. A. (1911–1919). *Notebook series 1911 – 1919*. Microfilm part V, Reel 237, NBK\_61. Thomas Edison National Historical Park.

Edison, T. A. (1912). Edison General File. Microfilm Part V, Reel 253, excerpt. Thomas Edison National Historical Park.

Edison, T. A. (1912–1914a). *Voice trials, Scrapbooks 1912–1914*. Box 25. Folder (without signature). Thomas Edison National Historical Park.

Edison, T. A. (1912–1914b). *Voice trials, Scrapbooks 1912–1914*. Box 25. Folder Singers E 6276 I. Thomas Edison National Historical Park.

Edison, T. A. (1912–1917). *Voice trials, Scrapbooks 1912–1917*. Box 26. Folder E 6276–2. Thomas Edison National Historical Park.

Fischer, E. (1910). *Neue Gesangsschule mit praktischen Beispielen auf dem Grammophon*. Leipzig: Breitkopf & Härtel.

Garcia, M. (1894). *Hints on Singing* (2<sup>nd</sup> ed.). New York: Ascherberg & Co. Translated from the French by B. Garcia.

Goldschmidt, H. (1890). *Die italienische Gesangsmethode des XVII. Jahrhunderts und ihre Bedeutung für die Gegenwart*. Breslau: Schlesische Buchdruckerei, Kunst- und Verlags-Anstalt.

Grove, G. (Ed.) (1879). *A Dictionary of Music and Musicians*, (1<sup>st</sup> ed.). vol. 1. London.

Grove, G. (Ed.) (1900). *A Dictionary of Music and Musicians*, (2<sup>nd</sup> ed.), vol. 1. London.

Guttman, O. (1861). *Gymnastik der Stimme, gestützt auf physiologische Gesetze. Eine Anweisung zum Selbstunterricht in der Uebung und dem richtigen Gebrauche der Sprach- und Gesangsorgane*. Leipzig: J. J. Weber.

Guttman, O. (1884). *Gymnastics of the Voice. A System of Correct Breathing in Singing and Speaking, based on Physiological Laws. A Practical Guide in the Training and Use of the Singing and Speaking Voice. Designed for Schools and for Self-Instruction*. New York: Edgar S. Werner. Retrieved from <http://www.archive.org/details/cu31924022418374>.

Harvith, J. & Edwards Harvith, S. (1987). *Edison, Musicians and the Phonograph: A Century in Retrospect*. New York.

Hattwick, M. (1932). The Vibrato in Wind Instruments. In C. E. Seashore (Ed.), *The Vibrato, vol. I of Studies in the Psychology of Music* (pp. 276–280). Iowa: University Press.

Hollinshead, M. T. (1932). A Study of the Vibrato in Artistic Violin Playing. In C. E. Seashore (Ed.), *The Vibrato, vol. I of Studies in the Psychology of Music* (pp. 281–288). Iowa: University Press.

Hähnel, T. (2013). *Baroque performance—a research study on characteristic parameters of 18th Century music*. Studies in Cognitive Musicology 2. Osnabrück: Electronic Publishing Osnabrück.

Iffert, A. (1906). *Allgemeine Gesangsschule*.

Kob, M., Amengual Garí, S. V., Bolles, B. A., Ritter, D. M. & Pirch, P. (2018). Influence of Early Recording and Playing Devices on Voice Sounds: Modification of singing voice formants. *Fortschritte der Akustik – DAGA 2018* (pp. 1703–1706). Berlin: Deutsche Gesellschaft für Akustik e.V. (DEGA).

Lehmann, L. (1902a). *How to Sing [Meine Gesangskunst]*. New York: Macmillan. Translated from the German by Richard Aldrich.

Lehmann, L. (1902b). *Meine Gesangskunst*. Berlin: Verlag der Zukunft.

Lindner, F. E. (1932). Measurement of the Pitch Extent of the Vibrato on Attack, Release, and Transition Tones. In C. E. Seashore (Ed.), *The Vibrato, vol. I of Studies in the Psychology of Music* (pp. 245–249). Iowa: University Press.

Martensen, K. (2019). “The phonograph is not an opera house.” *Quellen und Analysen zu Ästhetik und Geschichte der frühen Tonaufnahme am Beispiel von Edison und Victor*. Technologien des Singens 1, Munich: Allitera.

Mauch, M., Cannam, C., Bittner, R., Fazekas, G., Salamon, J., Dai, J., Bello, J., & Dixon, S. (2015). Computer-aided Melody Note Transcription Using the Tony Software: Accuracy and Efficiency. In *Proceedings of the First International Conference on Technologies for Music Notation and Representation*.

- Mauch, M., & Dixon, S. (2014). pYIN: A Fundamental Frequency Estimator Using Probabilistic Threshold Distributions. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2014)*. <https://doi.org/10.1109/ICASSP.2014.6853678>
- Mecke, A.-C. (2016). Vibrato. In Mecke, A.-C., Pfeleiderer, M., Richter, B. & Seedorf, T. (Eds.), *Lexikon der Gesangsstimme. Geschichte, wissenschaftliche Grundlagen, Gesangstechniken, Interpreten*. Laaber: Laaber.
- Metfessel, M. (1926). Sonance as a form of tonal fusion. *Psychological Review*, 33(6), 459–466. <https://doi.org/10.1037/h0070587>
- Monahan, B. J. (1978). *The Art of Singing. A Compendium of Thoughts on Singing Published Between 1777 and 1927*. London: Scarecrow Press.
- Peters, J. D. (2003). Helmholtz und Edison. Zur Endlichkeit der Stimme. In F. Kittler, T. Macho & S. Weigel (Eds.), *Zwischen Rauschen und Offenbarung. Zur Kultur- und Mediengeschichte der Stimme* (pp. 291–312). Berlin: Akademie Verlag.
- Platz, F., Kopiecz, R & Lehmann, M. (2012). Statistische Poweranalyse als Weg zu einer ‘kraftvolleren’ Musikpsychologie. In Auhagen W., Bullerjahn C., Höge H. (Eds.), *Musikpsychologie – Populäre Musik*, vol. 22 of Jahrbuch der Deutschen Gesellschaft für Musikpsychologie, (pp. 165–179). Göttingen, Hogrefe.
- Reger, S. N. (1932). The String Instrument Vibrato. In C. E. Seashore (Ed.), *The Vibrato, vol. I of Studies in the Psychology of Music* (pp. 305–343). Iowa: University Press.
- Seashore, C. E. (1936). *Psychology of the Vibrato in Voice and Instrument, vol. III of Studies in the Psychology of Music* (pp. 94–106). Iowa: University Press.
- Seashore, H. (1932). The Hearing of the Pitch and Intensity in Vibrato. In C. E. Seashore (Ed.), *The Vibrato, vol. I of Studies in the Psychology of Music* (pp. 213–235). Iowa: University Press.
- Sieber, F. (1862). *Katechismus der Gesangskunst*. Leipzig: J. J. Weber.
- Sieber, F. (1872). *The Art of Singing*. New York: William Pond & Co.
- Tiffin, J. (1932). Phonograph Apparatus. In C. E. Seashore (Ed.). *The Vibrato, vol. I of Studies in the Psychology of Music* (pp. 118–133). Iowa: University Press.
- Weege, T., Habasinska, D. & Kob, M. (2018). Influence of Early Recording and Playing Devices on Musical Sound: FRF Measurements of Horn, Soundbox and Tonearm. In *Fortschritte der Akustik - DAGA 2018* (pp. 1707–1710). Berlin. Deutsche Gesellschaft für Akustik e.V. (DEGA).

## APPENDIX

### Abbreviations and Terminology

- **Corpus:** The corpus of samples used in the analysis, consisting of Edison’s comments on recordings and the (digitalised) recordings themselves.
  - **Records Trials:** The corpus based on Edison’s comments on commercial recordings of competing companies and his own company.
  - **Probe Trials:** The corpus of unpublished recordings, which were made to evaluate singers, often as a part of an application process.
  - **Test Trials:** The merged corpus, including the Records Trials and the Probe Trials.
- **Condition:** The condition, under which Edison’s comments were analysed.
  - **Tremolo condition:** Edison’s statement regarding the tremolo.
  - **Overall condition:** Edison’s statement regarding the singing voice in general.
- **Group:** valence of Edison’s statements.
  - **G.:** “Good,” which means an affirmative statement by Edison.
  - **B.:** “Bad,” which means a disapproving statement by Edison.
  - **N/A:** No explicit statement.
  - **N.-B.** “Non-Bad,” which includes all G. and N/A statements. Used in the Probe Trial corpus under the Tremolo condition. Thus, analysing Edison’s statements about the tremolo in Probe trials we only differentiated between Bad and the remaining statements.
- **2f-vibrato:** A vibrato with distinct pitch and volume changes, where the volume changes show a vibrato rate twice as fast as the pitch vibrato range.
- **1f-vibrato:** A vibrato with distinct pitch and volume changes, where the volume changes show the same vibrato rate as the pitch changes.