NTNU – Trondheim
Norwegian University of
Science and Technology

An acoustical study of trombone performance, with special attention to auditory feedback deprivation

Eirik Kristensen

Physics
Submission date: May 2014
Supervisor: Jon Andreas Støvneng, IFY
Co-supervisor: Jan Tro, Institutt for elektronikk og telekommunikasjon

Norwegian University of Science and Technology
Department of Physics

Preface

The following paper is a Master's thesis from NTNU, as part of the study program Master of Science in physics. It was carried out between August 2013 and May 2014. My experience from jazz and big band music has been the background for this work, in combination with my interests in physics. Acoustics is an interdisciplinary field of science, something that is an extra inspiration for the work, since it offers an opportunity to view the problem from many different standpoints. My hope is that this work will be of interest to both acousticians and musicians, and perhaps trigger an interest among musicians to further explore the science of musical performance.

Trondheim 2014-05-15

Eirik Kristensen

Acknowledgement

I would like to thank my supervisor Jan Tro, for his guidance and enthusiasm for this project, and Jon Andreas Støvneng, for his brilliant proofreading. I would also like to thank Tim Cato Netland for superb technical assistance. For the many trombonists who have endured the pain of playing in the laboratory, I would like to direct an apology and a big appreciation. Your musical contribution, perspective and curiosity have been a huge contribution to the research, and serve as the foundation for this work. To my friends and others who have helped me with proofreading, lending me their headphones and trombones, thank you! Lastly, I would like to thank my girlfriend Sandra for her patience, healthy criticism and smart advise.

E.K

Abstract

When discussing how instruments work and how performers play them, the matter can be viewed from the standpoint of the musician or the scientist. In an attempt to bridge the gap between what science and musical experience know, acoustical methods are used to analyze trombone performance. The results provide a visual description of musical factors in trombone performance. Glissandi, sustained notes, transients, crescendos and decrescendos are studied through measurements of envelopes, spectra and spectral centroids for 9 players of different experience level. The analysis is done with the use of a Matlab toolbox for Music Information Retrieval (*MIRtoolbox)*. The results shed light on previous studies on trombone performance, and give insight into the matter from a performance perspective. The discussion is focused on combining acoustic theory with musical expertise in order to better understand the full model of trombone performance.

Further on, the effect of auditory feedback deprivation upon trombone performance is investigated. Eight trombonists, out of whom four were experienced professionals, and four amateurs, played musical pieces from notation and from memory. In addition, they performed a pitch-bending exercise (lipping). They did this under two playing conditions: A: With a masking noise of 95 dB on headphones, to completely remove auditory feedback from the instrument, and B: Without masking. Measurements of timing and RMS energy yielded statistically significant effects of auditory feedback deprivation on some amateur players, but not on professionals. The results indicate a correlation between musical experience and dependency on auditory feedback. A reason for this is suggested to be a more developed "inner ear." It is also found that the effects of auditory feedback deprivation were greater for music performed from notation than music performed from memory. This is in agreement with Finney and Palmer (2003), who found that auditory feedback deprivation affects learning conditions, but not retrieval.

Previous studies have concluded with no significant effects of auditory feedback deprivation upon keyboard performance (Gates and Bradshaw, 1974; Banton, 1995; Finney, 1997). In this study, observed tendencies towards disruption of performance in pitch bending exercises are

found. Many players also failed to hit the correct resonant mode in the performance from notation, and some played ‘out of tune.’ This suggests that auditory feedback deprivation has a more disturbing effect upon trombone performance than in the case of keyboard performance. A likely explanation for this result is that brass instruments provide the player with the possibility of producing several notes for one position or fingering, thus also the possibility of producing a wrong note. In lip-reed instruments, tuning and timbre are controlled by the player, which also introduces more room for disruption of performance. Slight differences on expressiveness in performance, tempo and dynamics were also observed, extending the findings of Repp (1999), who found small, but statistically significant results of auditory feedback deprivation upon expressive keyboard performance.

Sammendrag

Når man skal forklare hvordan musikkinstrumenter fungerer eller hvordan man spiller på dem, kan man benytte seg av enten en vitenskaplig eller en utøvende innfallsvinkel. Begge leirer sitter på mye kunnskap om emnet. I dette studiet forsøkes det å forene kunnskap fra disse ved å bruke akustiske metoder for å visualisere musikalske faktorer i trombonespill. Ved å foreta malinger av envelope, spektrum og spektral sentroide, undersøkes musikalske elementer som glissando, ansatser og crescendo. Ni trombonister med forskjellig erfaring har spilt inn materialet i ekkofritt rom. Analysen er gjort ved hjelp av *MIRtoolbox,* som er en programpakke designet til Matlab, for å hente ut musikalsk informasjon av lyd (Music Information Retrieval). I diskusjonen trekkes det linjer mot tidligere forskning, og resultatene diskuteres fra et utøvende perspektiv. Dette blir gjort med mål om å øke forståelsen for lydproduksjon i tromboner.

Senere i rapporten blir effekten av *fjernet akustisk tilbakekobling* undersøkt. Fire profesjonelle trombonister og fire amatører har spilt musikkstykker fra noter og fra hukommelsen. I tillegg har de utført en en *pitch-bending* øvelse. Dette ble tatt opp mens musikerne spilte med og uten øretelefoner påsatt 95 dB rosa støy, for å fjerne den akustiske tilbakekoblingen. Deretter ble RMS-energien av stykkene, samt varigheten av dem, målt. Resultatene for RMS-energi og varighet av musikkstykkene gav statistisk signifikante utslag av fjernet akustisk tilbakekobling for noen av amatørene. Ingen signifikante effekter ble påvist for de profesjonelle. Dette indikerer at det er en korrelasjon mellom musikalske ferdigheter og avhengigheten av akustisk tilbakekobling. En forklaring på dette kan være at profesjonelle utøvere har et mer utviklet "indre øre." Fjernet akustisk tilbakekobling resulterte også i større utslag for musikken som ble fremført fra noter, enn den som ble fremført fra hukommelsen. Dette samsvarer med Finney and Palmer (2003) som oppdaget at effekten av manglende akustisk tilbakekobling påvirket læringssituasjoner i større grad enn situasjoner hvor informasjon hentes fra hukommelsen.

Tidligere forskning på pianister har ikke funnet noen merkbare konsekvenser av at akustisk tilbakekobling blir fjernet (Gates and Bradshaw, 1974; Banton, 1995; Finney, 1997). I dette forsøket har det blitt funnet tendenser som tyder på at intonasjon blir påvirket, og da særlig i

pitch-bending øvelsen. Under musikkfremførelse etter noter hadde mange utøvere problemer med å spille rent og traff ofte feil moder på instrumentet. Dette tyder på at konsekvensene av fjernet akustisk tilbakekobling er større for trombonister enn for pianister. En sannsynlig forklaring kan ligge i selve instrumentet. I motsetning til et piano, hvor hver note har sin tangent, har brassinstrumenter flere mulige noter for hvert ventilgrep eller slideposisjon. Dette medfører også flere kilder til feil. Nevneverdige forskjeller i musikalsk uttrykk og dynamikk ble også registresrt. Dette stemmer overens med Repp (1999) sin forskning som påviste små, men signifikante konsekvenser av fjernet akustisk tilbakekobling hos pianister som spilte ekspressivt.

Contents

1 Introduction

1.1 Background

The acoustics of brass instruments have been studied since the beginning of the 19th century. But musical instruments have been made for ages before we started to study them, and the products that are being used today are not made this way because years of science have concluded that this is the best way to go. The result stems from years of practice of making these instruments, and a lot of trials and errors. The design of the trombone has shifted from a narrow bore with a small bell (sackbut), to a more open and wide variant. The science behind instrument making can be explained now, but the people who made these instruments a few hundred years ago did not have a PhD in acoustics. Science is catching up with the genius of handcrafting, but not accelerating miles ahead of it. The author has a great motivation for making scientific work available for more people. Therefore, this work is intended for both acousticians and musicians, and perhaps also sound engineers.

Part I

In the world of musical acoustics, much time has been spent on assessing different parameters in the production of sound in wind instruments. The main bulk of these studies focus only on specific parameters, such as the effect of surface coatings upon brass instrument tone quality (Pyle, 1981) and the influence of tuning resonances in the vocal tract (Chen *et al*., 2012). These studies generally include in their conclusion a hint about whether the property is significant or audible. There is even little research on the influence of different parameters compared to each other. This yields the still unanswered problem; which parameters and processes in the production of sound in the trombone are the *most significant* when it comes to tone quality, transients and steady state. One of the aims of this project is to give insight into the research that has been carried out so far in the field of trombone acoustics, and to describe the instrument-player system. The goal is that the full model of sound production in the trombone should be further understood.

The same can be said for performance. There have been written a number of books and training booklets on brass instrument performance. Much of what musicians learn in school is inherited knowledge that is shaped by the people reproducing it. To develop a fuller understanding of sound production in musical instruments, attention must be paid to both scientific knowledge and the vast amount of experience that performers possess.

Part II

In the second part, a more specific subject is brought to attention, namely the effect of auditory feedback deprivation. This is an important area of study because it imitates the difficult playing conditions of on-stage performance with acoustically harsh environments. Musicians are often faced with the challenge of playing with bad monitoring, in large halls and other difficult situations. In these circumstances, the question can be raised whether performers really are dependent on hearing their own sound.

Auditory feedback is the simple process of hearing ones own performance. In musical performance, it is arguably a prerequisite to hear what you are playing, especially in training. But when a player gets older and more experienced, the whole system of sound production starts to take a new form. The player uses the instrument more like an extension of his or her body. Trombone performance involves a number of sources of perceptual information, including auditory feedback. The player also gets information from tactile interaction with the vibrating instrument and air column, kinesthetic motor feedback from the hand moving the slide, and visual information from the music notation. It is plausible that auditory feedback is essential to music performance. Intuitively, this makes sense. And if you ask a musician, he will probably confirm this. Studies have shown, however, that auditory feedback (absence or presence of) may not be essential to music performance (Lashley, 1951; Gates and Bradshaw, 1974; Banton, 1995; Finney, 1997). Repp (1999) found small, although statistically significant, effects of feedback absence concerning expressiveness in piano performance. Previous studies have only investigated auditory feedback deprivation on keyboard performance, and it is of interest to further investigate instruments that include lip-reed interaction. Brass instruments are of particular interest, because the player has a very dominant influence on the produced sound. The player controls timbre and tuning by adjusting parameters of the lips and vocal tract. The instrument also provides multiple resonant modes for each instrument length (slide position or finger setting).

1.2 Aims

The broad aim of this work is to give insight into the matter of trombone performance, and discuss this from multiple viewpoints. The problem descriptions for the two parts of this report are as follows:

PART I

The aim is to give an overview of parameters that contribute to sound production in a trombone; which are there, what do they affect and how important are they? Parameters concerning both the instrument, the player and the whole system are discussed. The report should also give a visual description of trombone performance with a discussion that considers multiple viewpoints.

PART II

Investigate the effect of auditory feedback deprivation in trombone performance: How dependent are performers of auditory feedback, and which areas of musical performance does auditory feedback deprivation affect. Are the effects of auditory feedback deprivation dependent on musical experience and playing conditions?

1.3 Limitations

To perform laboratory exercises with musicians poses a great challenge, as the environmental parameters may affect the player's performance. It may also affect the individual players differently. As this happens, something is lost; individualism and expressiveness are key ingredients in musical performance. When telling a person to play one way or another, the balance is shifted towards a more rigid and preset situation, and perhaps differently for different performers. To achieve useful recordings of different musicians, the acoustic parameters of the environment should be as stable as possible and the dynamic level, microphone distance, and the instructions that are given should be the same. These requirements pose a challenge both to the experimenter and the player. As a result, the fight for physical perfection in the experiments may lead to big compromises in performance.

The acoustical analysis of music performance also faces great challenges. Music cannot be described only in numbers, and both the computing skills of the author and the computers themselves are limited. Musical features are difficult to quantify. The *MIRtoolbox* is one of the best tools available, but still there are a lot of musical factors that are much better analyzed through listening than through acoustical analysis.

It must be stressed that this is an interdisciplinary project, with a focus on making the results available to many different readers. It aims to give an overview of what physical and human processes contribute to sound production, with a limited amount of in-depth analysis. The discussion views the results from the standing points of both the musician and the scientist.

1.4 Approach

The acoustic analysis is performed using the *MIR Toolbox (Music Information Retrieval* (Lartilliot, 2008). The *MIRtoolbox* is a toolbox designed for Matlab that consists of more than 50 different audio and music feature extractors. It uses generalized and established acoustical theory and allows the user to adjust how the information is extracted. A fuller description of the *MIRtoolbox* can be found in the appendix.

For part I, all recordings were made in an anechoic chamber. For part II, the location was the Aura Lab at NTNU. This room is not as non-reverberant as the anechoic chamber, and more comfortable to play in.

1.5 Structure of the report

Part I

Part I consists of a theoretical chapter about trombone acoustics, and an experimental part, with results from the recordings of 9 trombonists. The theoretical chapter gives an introduction to the acoustics of the trombone, including parameters concerning the player and performance. It is a short review of what is state of the art, but not in any way a full description of the trombone-player system. In the experimental chapter, basic trombone performance phenomena, such as glissandi, attack transients, and crescendos are studied.

Part II

Part II consists of chapter 4, and focuses on the effect of auditory feedback deprivation, with an introduction to relevant literature, results, and discussion.

The overall conclusion is found in chapter 5. The appendices contain additional data and results from both part I and part II, and also relevant codes, information about *MIRtoolbox* and musical scores. The original data, including more than 1000 sound samples and other experimental data are available on request.

PART I

This part consists of a theoretical chapter concerning general trombone acoustics, including a literature review of some of the most significant research done in the field during the last 100 years, and a chapter with results from experiments conducted on trombone performance.

2 Trombone acoustics

2.1 The instrument

2.1.1 Speed of sound

The speed of sound waves in an ideal gas is given by,

$$v = \sqrt{\frac{\gamma R T}{M}} \tag{1}$$

where $R = 8{,}31\ \frac{\mathrm{J}}{\mathrm{mol \cdot K}}$, is the molar gas constant, T, is absolute temperature in Kelvin, $\gamma = 1.4$, is the adiabatic index for air at $0°\mathrm{C}$ and $M = 2.88 \cdot \frac{10^{-2}\,\mathrm{kg}}{\mathrm{mol}}$, is the molar mass. The relationship between speed of sound and temperature over the temperature range we normally encounter (around $20°\mathrm{C}$) can be linearly approximated by the following formula:

$$v = 331.3 + 0.6 \cdot T \text{ m/s} \tag{2}$$

This is to a first approximation negligible in the analysis of sound production in brass instruments for moderate dynamic levels. When the sound pressure inside the instrument rises (it can be as high as 170 dB), the temperature rises, so that the speed of sound is different for portions of air at pressure maxima and vice versa. A consequence of this is a distinct change in timbre, referred to as the *brassy* sound. Section 2.1.9 will further describe this phenomenon.

2.1.2 Trombone sound production parameters

Figure 1 shows a sketch of a trombone, with the main parameters that are significant in the production of sound listed. This is not in any way a complete model, as the instrument-player system also acts differently for different dynamic levels.

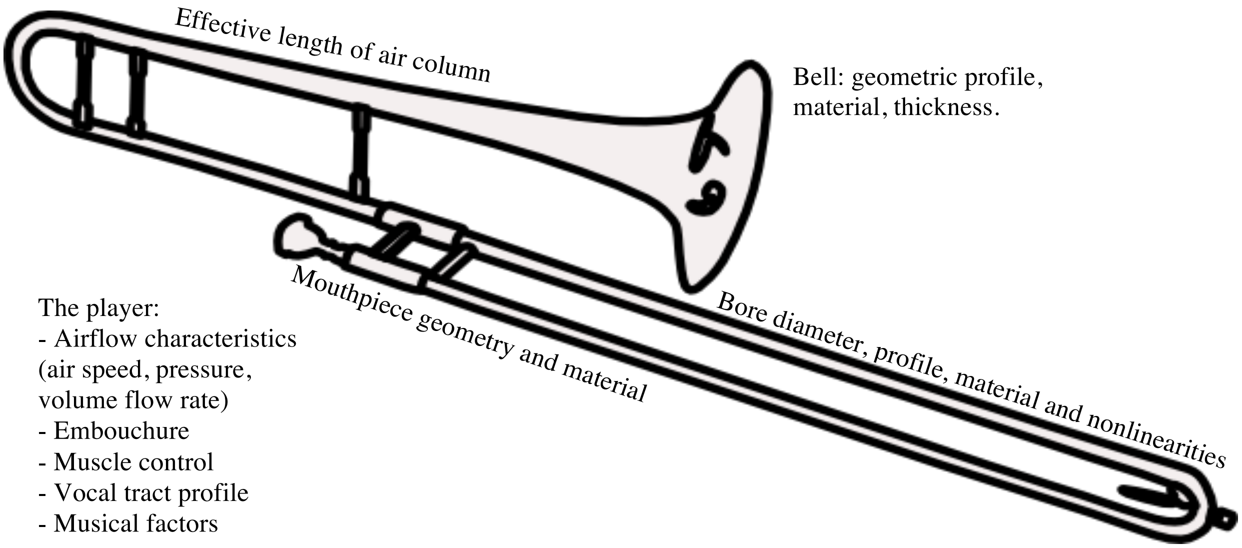


Figure 1: Parametric model of trombone sound production.

2.1.3 The resonator – a pipe

The fundamental idea of a brass instrument is that of a resonating air column with a closed end at the mouthpiece and an open end at the bell. The air column of a tenor trombone with the slide in 1st position has a length, L, of approximately 275 cm, which at temperature 20°C would give rise to resonances at

$$f_n = n\frac{v}{4L} \approx 31, 43, 156, \ldots \text{ Hz} \tag{3}$$

with v being the speed of sound. As seen, this does not constitute a complete harmonic series corresponding to the resonance modes of a trombone. The (almost) perfect harmonic series is achieved by attaching the mouthpiece and the bell. Also, the effective acoustical length is not fixed, but changes with frequency. In other words, the bell makes the acoustical length of the air column vary, and as a result the wave front can be reflected at different *turning points* (Fletcher and Rossing, 1998).

The most important parameter to describe the resonating system is the acoustic input impedance, which tells us how the instrument will respond to excitation at various frequencies. The acoustic input impedance, Z is defined as the ratio of pressure, p, and acoustic volume flow, U, through the surface, S. Both p and U are functions of angular frequency, ω.

$$Z(\omega) = \frac{p(\omega)}{U(\omega)} \tag{4}$$

In the case of a brass instrument, the volume flow will be the amount of air that flows through the cross-sectional area, S, of the pipe, $U(\omega) = u(\omega)S$, with u being the particle velocity. The pressure is measured at the entrance plane of the lead pipe, right after the mouthpiece. Figure 2 shows the acoustical input impedance of a bass trombone.

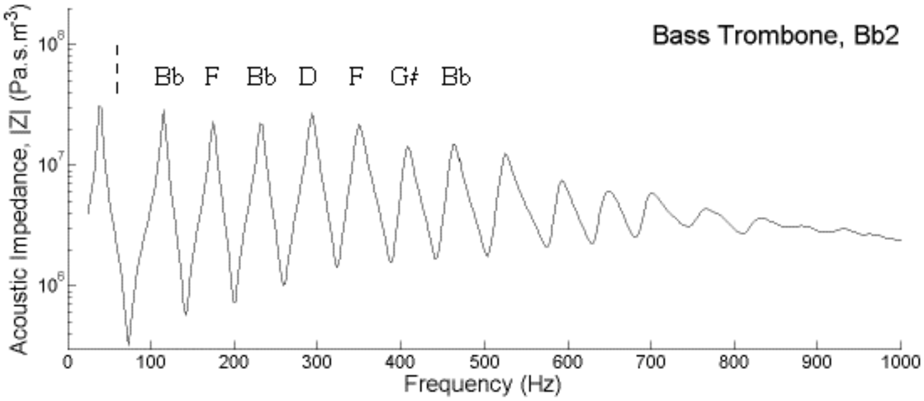


Figure 2: Typical input impedance for a bass trombone with the first 8 resonant modes shown. The dashed marking at 58.3 Hz displays the pedal note Bb0, which trained players are able to produce even though it does not coincide with a resonant mode of the instrument.[1]

---

[1] From http://www.phys.unsw.edu.au/jw/brassacoustics.html (Accessed Dec 15, 13)

2.1.4 Mouthpiece effect

The effect of mouthpieces have been been devoted much research. Individual preferences among musicians differ greatly and the choice of mouthpiece is somewhat personal. The mouthpiece consists of the cup and the shank. The join between the mouthpiece and the bore of the instrument should be as smooth as possible to avoid reflections inside the instrument. This is partly achieved by having the shank radius vary, as shown in figure 3. The mouthpiece acts as a Helmholtz resonator (Campbell, 1987) and has its own resonance determined by the volume of air that is enclosed, and the area opening in the shank. It strengthens the resonance of the instrument around the mouthpiece resonant frequency, usually around 700 Hz, and reduces the frequencies of the higher modes. The resulting effect is that a certain mouthpiece for example can make the instrument more easily played in a high register (with a small cup volume).

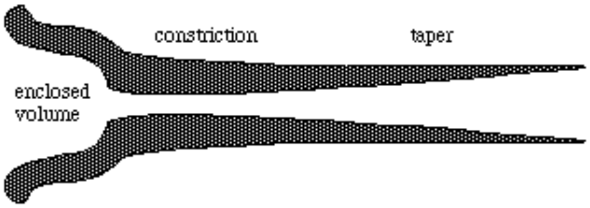


Figure 3: Mouthpiece geometry. The enclosed volume acts as a Helmholtz resonator. The enclosed volume and shank length determine the "popping" frequency. The tapered section has its form to smoothen the transition between the mouthpiece and the bore.[2]

Experience shows that both the cup volume and the diameter of the constricted passage have significant effect upon the performance of a given mouthpiece (Fletcher and Rossing, 1998). In addition to the acoustic characteristics of the mouthpiece, the ergonomic features of it are also important features.

---

[2] From http://www.phys.unsw.edu.au/jw/brassacoustics.html (Accessed 2013-12-03)

2.1.5 The Bell

The flaring bell acts as a high pass filter and improves the radiation of upper harmonics. This is a great contribution to the characteristic timbre of brass instruments. It makes the instrument highly directional, especially for high frequencies, which can be confirmed by any violinist who has sat in front of the brass section. The effect the bell has on lower frequencies was described by Campbell (1987). The effective length of the air column is shorter for lower frequencies where the wavelength is longest and cannot 'see' the bell, as the bell radius is smaller than the wavelength. This contributes to raising the lower frequencies. The bell is also one of the factors contributing to tuning differences. It affects a whole range of frequencies, both flattening the higher notes and sharpening the lower notes, or vice versa (Webster, 1949). The combination of the mouthpiece lowering the higher mode frequencies, and the bell raising the lower mode frequencies, gives the instrument resonances which are very close to a complete harmonic series.

Some of the latest research concludes that the most prominent musical effect of the bell diameter and shape is the playability in certain frequency ranges. A study done by Murray Campbell, Arnold Myers and John Chick confirmed the general experience of a wider bell sounding more full and less bright than a small and narrow bell. (Campbell *et al.*, 2013). They tested many different bell shapes by continuously lengthening the tapered section in the end, and measuring the transfer function between input and output pressure.

2.1.8 Wall Material

It is perhaps no surprise that the leading brass instrument material is brass (an alloy of copper and zinc). A fact that is generally agreed upon among trombone performers is that the red brass, yellow brass and gold brass differ slightly in tone color. Manufacturers often say that the red brass (90% copper, 10% zinc) produces a warmer, darker sound, whereas the gold/rose brass (85% copper, 15% zinc) gives a more defined articulation and a brighter sound. In the other end of the scale is the yellow brass (70% copper, 30% zinc) with its very bright and resonant sound. This is the most common material and is especially preferred by big band lead trombonists who like to "cut through" in a section. Jazz solo trombonists who like a covered, soft sound, more and more prefer the warmer alternatives, like red brass. Trombones have also been made out of steel, nickel and plastic (Pyle, 1998).

In the scientific forum, the effect of materials has been debated for many years, and the scientists seem to disagree slightly with the musicians on the matter. Experimental brass instruments made of wood and other materials have been reported by listeners to be indistinguishable from instruments of brass. It is clear that the material choice is not the most important factor compared to others. However, it should be noted that the radiated sound along the axis of the playing direction is less affected by material choice than the lateral radiation. The effects are therefore more noticeable for the player than the audience. The sound that comes out of the bell comes from the vibrating air column, and does not depend as much on the material the instrument is made of. Richard Smith found that in trumpets, it is the thickness of the wall material, more than the material itself that affects the vibration of the instrument (Smith, 1978).

2.1.9 The *brassy* sound

An area of trombone acoustics that has been given much attention is the *brassy* or *cuivré* sound. Composers often utilize this as musical expression, and it is a key ingredient for a composer who wants to make a dramatic effect in the music. Many researchers have investigated this phenomenon (Norman *et al.*, 2010; Myers *et al.*, 2012; Ebihara and Yoshikawa, 2013b; a; Gilbert *et al.*, 2013). Especially interesting is the Japanese experiment from 1996 where they managed to film shockwaves from a trombone (Hirschberg *et al.*, 1996). There is even a video of a trombonist who sneezes in the instrument during a performance that has gone viral and has 250 000 hits on YouTube. The pioneer of this research however, is James Beauchamp (1980) who discovered one of the reasons for the brassy sound in brass instruments. He found that an explanation lies in the formation of shockwaves and wave steepening in the bore. At moderate playing levels, the harmonic content of trombone sound can be explained by the non-sinusoidal wave generated by the lip-reed (Backus and Baskerville, 1977; Elliott and Bowsher, 1982; Fletcher, 1999; Backus and Hundley, 2005), but when the instrument is played at loud levels, the pressure inside the bore increases, and as a result the physical parameters inside the bore are changed. When the pressure increases enough, the propagation speed varies at different points on the waveform. The waves start to steepen, as can be seen in figure 4. When this effect gets more and more pronounced, a shock wave is formed, and a distinct change in timbre occurs.

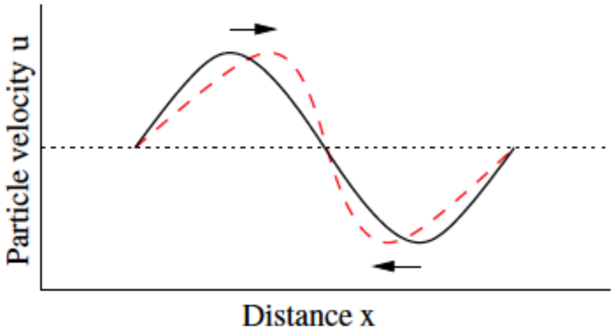


Figure 4: Wave propagation from left to right. The thick line is a particle velocity curve inside an instrument bore at a normal playing level. The dashed sinusoidal curve depicts the wave steepening that occurs because the speed of sound changes due to high-pressure formation at a loud playing level.

The area is of such interest that there has been created a parameter called the *brassiness potential*, which describes an instrument's potential for making brassy sounds as a number between 0 and 1 (Myers *et al.*, 2012). For a tenor trombone, this value is around 0.7 – 0.8, which is the highest number of any instrument. For a trumpet, the number is 0.65 and euphonium has brassiness potential of 0.40. The reason why these numbers differ is that the trombone has a much longer cylindrical section. The cylindrical section makes it easier for the pressure to build up towards a critical value. The euphonium is a conical instrument, which provides the air with more and more 'room' through the air column because the cross-sectional area increases along the length of the instrument. The energy is therefore spread across a larger area, thus limiting the brassiness potential of the instrument.

2.2 The player

Trombone sound production takes place in a coupled system, which relies heavily on human parameters. Skilled performers use muscle control to overcome limitations of the instrument, e.g., when playing pedal notes, which do not correspond to a resonating mode of the instrument (see figure 2). To explain this, we must also consider the influence of the lip-reed generator and the vocal tract.

2.2.1 The lips

The lips act as a pressure valve that lets through air when the pressure inside the mouth reaches a high value, thus opening the lips and letting a portion of high-pressure air through. The lip motion has been described in many ways, ranging from a simple outward striking pressure valve acting in one dimension to a more complex 3D model with several degrees of freedom. In the research at hand, it shall suffice to regard the lips as a simple outward striking reed.

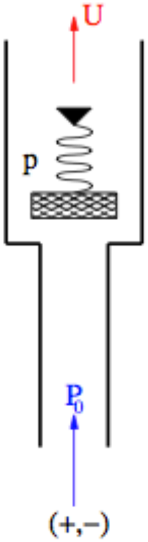


Figure 5: Outwards striking reed. The static overpressure, p_0, overcomes the pressure, p, of the lip tension, generating a pressure difference over the valve. U is the volume flow.[3]

---

[3] From http://www.acoustics.ed.ac.uk/wp-content/uploads/Theses/Logie_Shona__PhDThesis_UniversityOfEdinburgh_2012.pdf (accessed Feb 25, 14)

This does not give the complete picture of the mechanic properties of the lip-reed mechanism. There is also a feedback loop, as can be seen in Figure 6. The air moves through the length of the instruments and reflects back to the mouthpiece. To maintain a stable oscillation regime, the player has to add a new puff of air at the right time, that is, when the pressure maximum is at the mouthpiece. This pressure actually helps the mouth do the work. We can say that the instrument and the lips cooperate to maintain a tone. Skilled players can force oscillations in the lips via mouth pressure and tension in the lips so as to create oscillation regimes that fall in between the resonant modes of the instrument. This is often referred to as "lipping" and is sometimes called for in orchestral scores (e.g. in Gershwin's Rhapsody in Blue, the clarinet solo in the beginning of the piece). This phenomenon is investigated further in part II.

Since the standing wave cannot form without the waves first travelling to the end of the instrument and reflecting back, the lips do not "know" the resonance of the instrument during the first few cycles. Here, a considerable difference is found between those that are trained musicians and beginners. Beginners will take more time to find the resonance of the horn and adjust the lip tension accordingly.

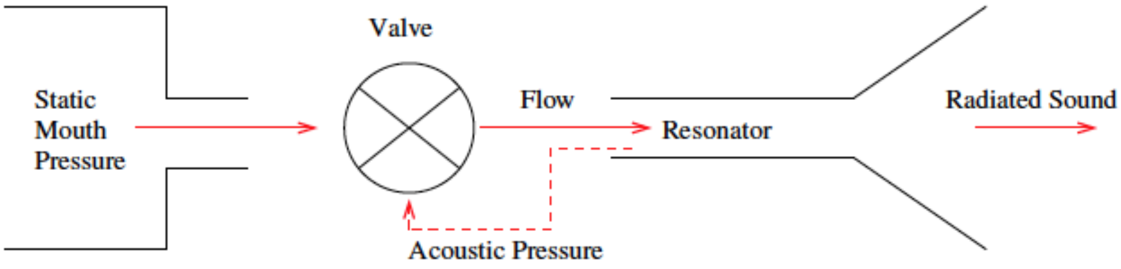


Figure 6: Schematic representation of the brass wind instrument. The system consists of three parts: The resonator, the valve and the flow through the bore.[4]

---

[4] From http://www.acoustics.ed.ac.uk/wp-content/uploads/Theses/Logie_Shona__PhDThesis_UniversityOfEdinburgh_2012.pdf (accessed Feb 25, 14)

2.2.2 Vocal Tract

Imagine what a brass instrument player would say if you took his or her instrument, and dented it several places, beyond repair. They would certainly claim that the instrument is no longer good enough. What many do not realize is that the vibrating air column in a player-instrument system is far longer than the instrument itself. Imagine how many "dents" and imperfections there are for the air along the way from the bottom of the lungs to the end of the mouth. The argument for this is rather hard to convey, but the point is that the vocal tract and throat play a huge role in controlling the resulting sound of the system. To put this into practice, imagine yourself playing a tone with the tongue and mouth in the settings for an extreme "*ee",* as in the word *see*. The output would be a thin and edgy tone. The exact opposite would happen if you do the same whilst holding the tongue and mouth in a situation simulating saying "*aw*." The tone is darker and more full. The most preferred output would be if you find something in the middle, because only then will the tone have a balanced content of frequencies (Fox, 1982). This is closely related to breath and musculature, and the professional musicians are experts in this manner. When finding the best constellation of mouth and tongue, musicians often talk about finding the core of the tone. On a trombone, the core of the tone will be present when simulating something between an "*aw*" and an "*ee*." For a trumpet, the same goal is more towards the "*ee*." The differences come from the different spectra that the instruments are able to produce.

The player's vocal tract constitutes an acoustical resonator connected to the instrument via the reed generator. At frequencies near the resonance of the instrument bore, the input impedance of the vocal tract is usually smaller than the impedance of the bore (Kaburagi *et al.*, 2011). It can therefore in a first approximation be ignored. But this does not mean that the acoustic contribution of the vocal tract is inaudible. In 2003, a study was done with the aim of finding out what the vocal tract does with the sound of the instrument (Wolfe *et al.*, 2003). They compared the effects of vocal tract configuration on trombone and didgeridoo. The research confirms what we have learned from Fred Fox. The main findings were that the vocal tract contributes greatly to intonation (as is later proven in the pitch bending in part II). It also has an effect on timbre (ref. Fox). Lastly, tongue position can increase the heights of the impedance peaks of the vocal tract. This makes the coupling between the vocal tract and instrument resonances more effective, and gives players a method for fine pitch adjustment. It also explains why there often is a problem with intonation when using double-tonguing to play rapid successive notes.

2.2.3 Transient Analysis

Transience means passing with time, and in acoustics, we talk about the transient being the high amplitude, short duration sound at the beginning of a waveform. When sounding a note on a brass instrument, the player must force the lips to oscillate with a frequency at or near a resonating frequency of the instrument. When the wave has travelled the length of the instrument and partly reflected back to the lips, a constructive reinforcement is established, if the frequency and phase are correct (Benade, 2012). It can contain a high degree of non-periodic components and inharmonic features. In a brass instrument, the transient can tell us much about the attack of the tone, the envelope of the harmonics and in the end, transient analysis can address the playing quality of the instrument when testing under playing conditions, as well as something about the player (Hoekje, 2013). The importance of transient phenomena cannot be neglected as they play a huge part in the perception of timbre, musicality and overall sound quality. It is hard to recognize the instrument only by listening to the steady-state tone, but when the transient is heard, the task is much easier. It should also be noted that there is a difference in transient behavior for partials above and below the cutoff frequency of the instrument (Luce and Clark, 1967), (the frequency where the radiation efficiency stops to rise, usually around 1000 Hz for the trombone (Beauchamp, 2012)). Partials above the cutoff frequency build up more slowly and the time it takes for them to reach a steady state increases with frequency, whereas the partials below the cutoff frequency build up and reach steady state simultaneously. The approximate time in which this occurs is about 50 ms (Logie, 2012).

Skilled players are able to slur between resonant modes of the instrument. This is referred to as lip-slurs. There is a strong coupling between the lips and the resonating air column when a note is established. The player has to break this coupling and establish a new one at a different resonant mode of the instrument. Some players are capable of performing lip-slurs between non-adjacent resonant modes of the instrument, for example between the 4th and 8th resonant mode (octave slur). When doing a lip-slur, the player must first decouple from a resonating regime and then start a new oscillating regime as smooth as possible.

Transients are also important factors upon which musicians judge the quality of an instrument. That is, how quickly the instrument “jumps” into the nearest resonant mode, which can be a desirable quality in an instrument.

2.3 Trombone performance

Can musical performance really be measured in a lab? The question remains to be answered, but in part, yes. At least some musical parameters and aspects of performance can be separated, measured and treated scientifically. An example is tonality. Frequencies can be measured and science can tell us how much “in tune” a performance is. There is even an app for it. There are other examples, but many musicians will argue that their performance is something that cannot be described scientifically. Scientists will claim another thing and the discussion continues. As musical training is not a very scientific subject compared to acoustics, some of the information in this chapter derives from the author’s own experience and from discussions with the many trombonists involved as participants in this project.

2.3.1 Basic training

Every musician receives some form of musical training concerning technique. The different teachers often agree on the basic philosophy, but still the variation is large. The first thing a trombonist learns is how to make a sound. *Make your lips form an “m” and blow air out between them as to produce a humming sound*. In trombone performance, often the techniques themselves are different, but they are all focused on a musical goal. The goal can be different for different performers and different musical tasks. The training often consists of playing etudes and exercises, which are designed to improve the performance of certain musical aspects. E.g. arpeggio-exercises to strengthen the lip muscles and exercises played on the mouthpiece alone, to improve breathing. Overall, breathing technique is one of the most prominent factors contributing to good sound.

2.3.2 Breathing technique

According to a majority of expert performers and scientists, the breath is one of the most important elements of brass instrument performance (Kleinhammer, 1963). The breathing directly relates to many factors in the performance, including timbre and transients. When talking about a good breathing technique, one often talks about naturalness. Natural processes should guide the airflow. When the last of the air leaves the lungs, the natural thing is to fill them up again. A good example of the opposite of a good breathing technique is a balloon that is filled up with air, and in need of a tight pinch against the opening to stop the balloon from inflating. The same happens when a person tries to hold his breath and seals off the passageway in the throat. If a trombonist does this before playing a note, the result is an unsteady airflow and a bad tone attack. All “unnatural” processes, like squeezing the air out of the lungs, are regarded as wrong. A bellow used to breathe air to keep a fire going, will also be an example of what is incorrect. The correct thing is to be relaxed and breathe naturally. We know that smoking, exercise, temperament and anxiety can influence how we breathe. This will also influence how we play. It is expected that some of the variations in performance in this project will be due to different breathing techniques.

3 Acoustical analysis of trombone performance

In this section, several different aspects of trombone performance are examined. The aim is to give an overview and visual description of trombone sound, with a qualitative discussion. The main focus is on the variation and spread in performance.

3.1 Method

The testing consists of a series of basic technique exercises, performed by 9 players with different background and musical training. The exercises were designed to investigate specific musical factors present in trombone performance, such as glissando and crescendo. The score is available in the appendix. The participants were asked to perform as if on stage, and with as consistent a performance as possible, without adjusting their technique and musical choices. However, the instructions made it clear that the main objective was to capture their individual performance, not to make them play a certain way. As a result, there is a big spread in performance, especially in dynamic levels.

The tests were carried out in an anechoic chamber at NTNU. They were all performed using a large-bore tenor trombone (*Conn 88 H0*) with a *Greenhoe* F-valve and a standard *Bach 5G* mouthpiece. The sound was recorded with a *BSWA MP216* prepolarized free field condenser microphone with a flat frequency response. The distance to the microphone was approximately 2 m and the sound was amplified with a *Yamaha HA8* amplifier and sent via a *LynxTWO-C* sound card into *Audacity*. The sampling frequency was 96 kHz and the resolution 16 bit. The analysis was done in *Matlab* with the aid of *MIRtoolbox*. Figure 7 shows the signal pathway.

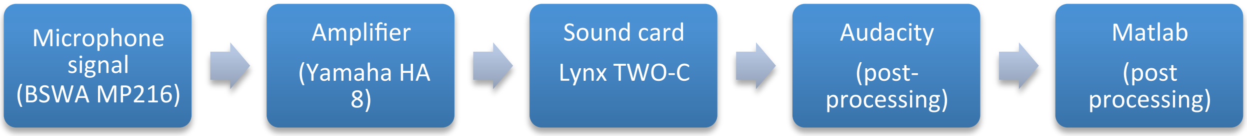


Figure 7: Signal pathway. The microphone signal is amplified before it is sent to the sound card. The amplification gain was adjusted to different levels for different players.

3.1.1 Acoustical approach

The acoustical analysis consists of measurements of normalized energy envelopes and spectral contents. The spectral content is shown both in the time domain and on average for whole notes. All the calculations for all temporal measurements are done with *MIRtoolbox*. For all temporal measurements, the *MIRtoolbox* standard setting of 50 ms window and a hop rate of 0.5 are used. The frame length of 50 ms is selected for its correlation with the *precedence effect* (Wallach *et al*., 1949). Details about *MIRtoolbox* can be found in the Appendix.

3.2 Results

3.2.1 Simple tone – temporal and spectral contents for one player

Figures 8 - 12 show different characteristics of a sustained fundamental trombone note on first position, with the slide pulled all the way in. The note is a Bb with the fundamental frequency of 116.5 Hz, which is the 2nd resonating mode of a trombone.

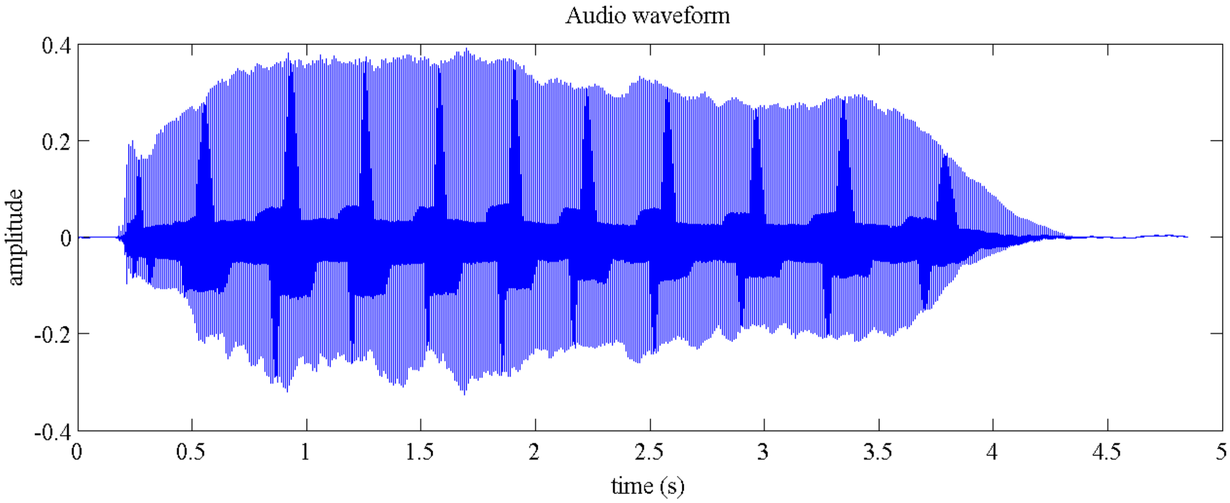


Figure 8: Audio waveform of a Bb1 with fundamental frequency 116.5 Hz, played on a trombone. (*miraudio*). Player 9.

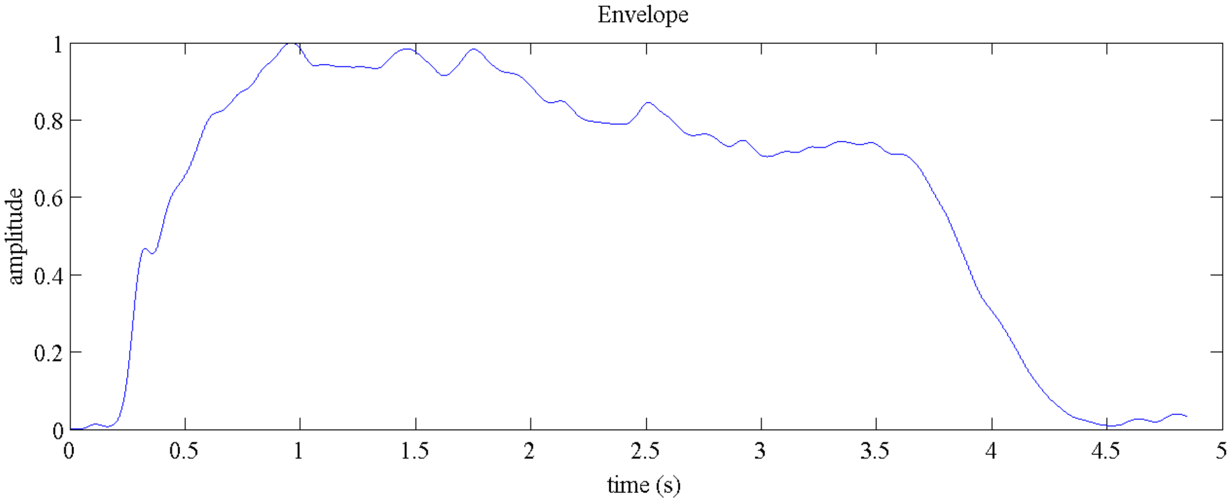


Figure 9: Normalized envelope of a Bb1 with fundamental frequency 116.5 Hz, played on a trombone (*mirenvelope*). Player 9.

The harmonic content of this tone is shown in the spectrum in figure 10, and in the time domain in figures 11 and 12, the latter through measurement of the spectral centroid.

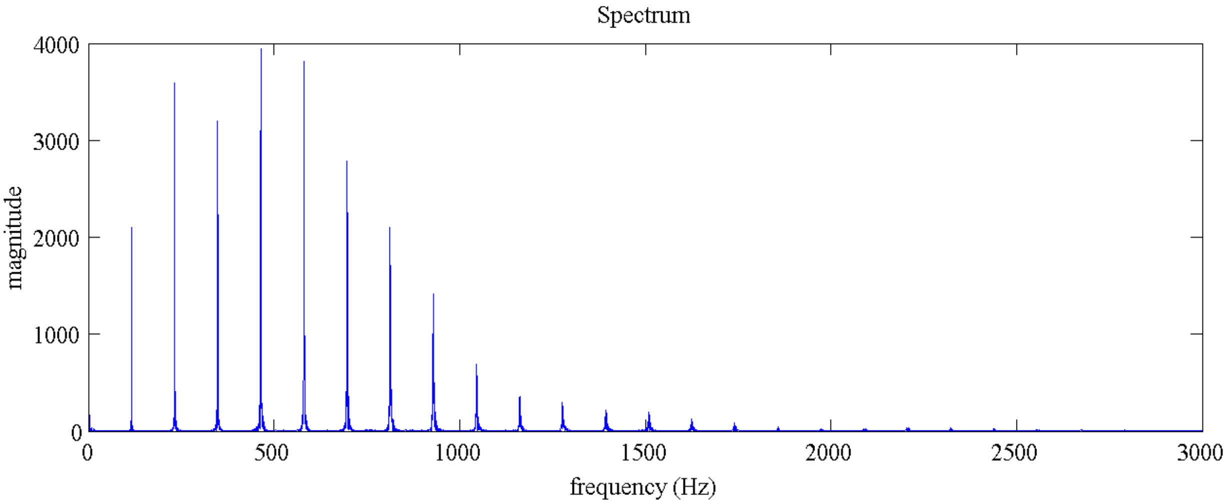


Figure 10: Spectrum of a Bb1 with fundamental frequency 116.5 Hz on a trombone. The y-axis shows magnitude, as the repartition of the amplitude of the frequencies (*mirspectrum)*, player 9.

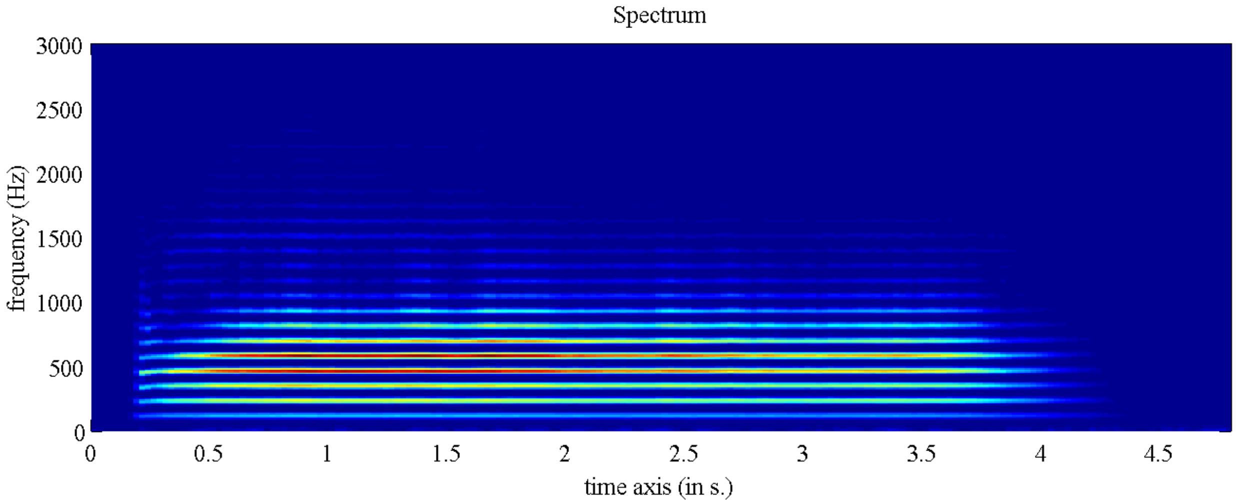


Figure 11: Spectrum in the time domain of a Bb1 with fundamental frequency 116.5 Hz (*mirspectrum*). Player 9.

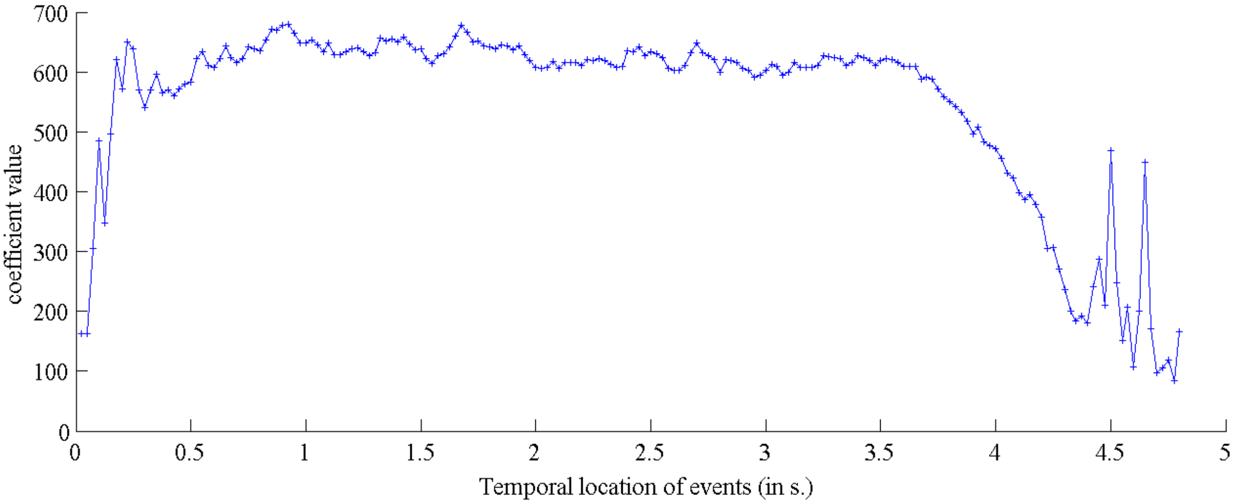


Figure 12: Spectral centroid in the time domain of a sustained Bb tone with fundamental frequency 116.5 Hz. The y-axis displays frequency in Hz. This shows the spectral center of gravity of the frequency content (*mircentroid)* for player 9. The peaks at 4.45 and 4.65 seconds occur because the centroid function picks up noise from the instrument in the end of the tone.

3.2.2 Simple tone – temporal and spectral contents for all players

Players were asked to play as normal as possible, in the dynamic level of *mezzo forte*. Figure 13 shows the envelope of the Bb1 note with fundamental frequency 116.5 Hz, for all players.

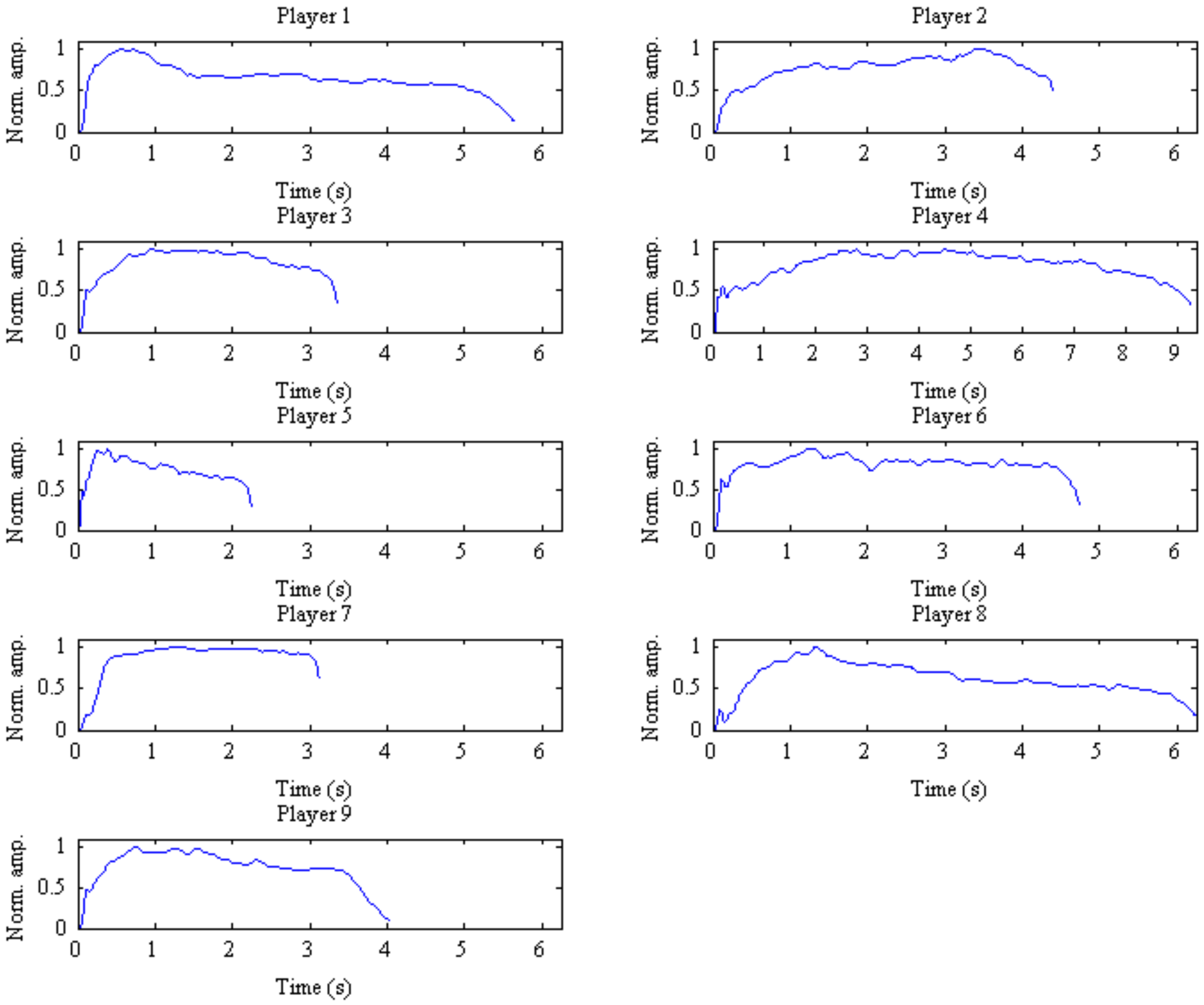


Figure 13: Normalized envelopes of a sustained Bb1 note with fundamental frequency 116.5 Hz for 9 players (*mirenvelope*). Player 7 is a professional trombonist.

The spectral centroid was computed by time-averaging the centroid in a 2 second window in a stable part of the tone. The are shown in figure 14. The dynamic levels were not equal for all players, each of them played in their own interpretation of a *mezzo forte* level. The chart indicates a big variance of spectral contents. Player 7 is a professional performer. The first note, at 61.7 Hz fundamental frequency, is the *pedal note*.

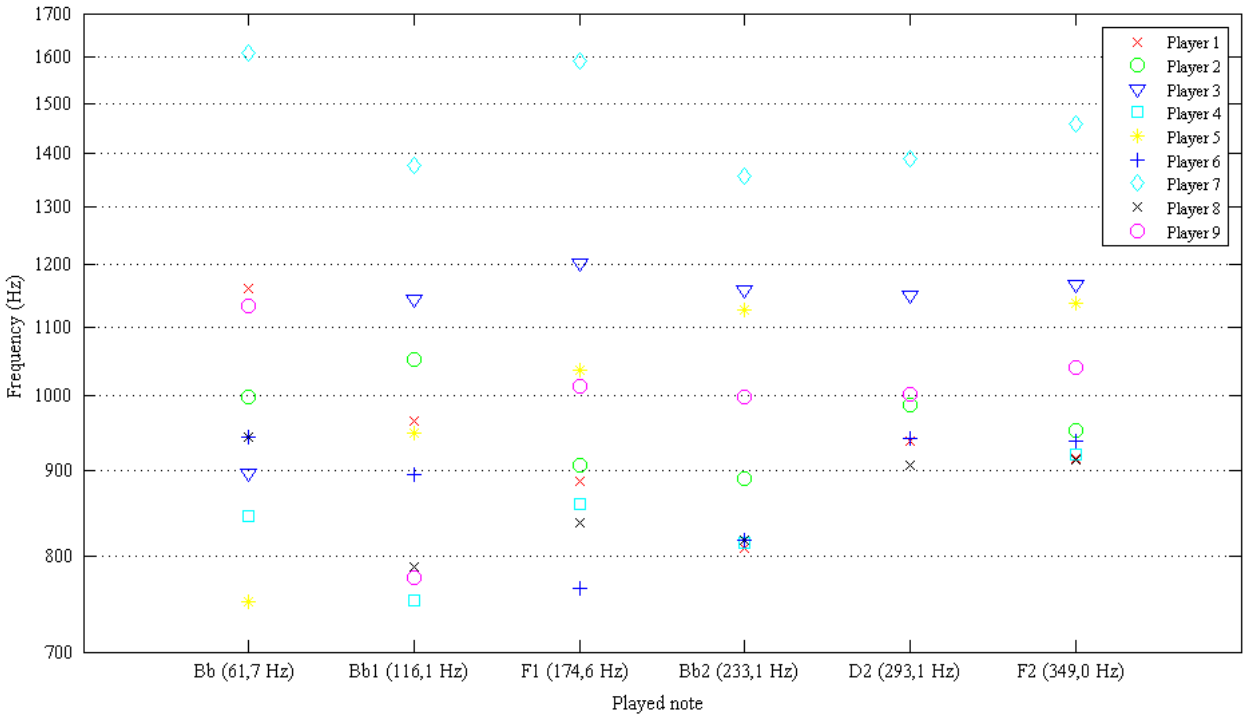


Figure 14: Time-averaged spectral centroid for the 6 first resonant modes of the trombone in 1st position (*mircentroid, 2s window)*. Fundamental frequencies are marked on the x-axis. The y-axis shows spectral center of gravity.

3.2.3 Slurred transients

Figure 15 shows a lip-slur, or slurred transient from Bb2 to D3 (resonant modes 4 and 5 with fundamental frequencies of 233.1 Hz and 293.1 Hz) on 1st position performed by three different players. As can be seen in the figure, the transient time and the amplitude drop vary for each player. These effects are audible. In terms of musicality, the hard attack transient and short transient time for player 3 (top), is not an ideal transient. The result is that the transition sounds uncontrolled, and not soft. According to trombone teachers, the production of a soft slurred transient requires a stable flow of air. Measurements of envelope for slurred transients on 4th and 7th position were also made. These results are available in the appendix for all players.

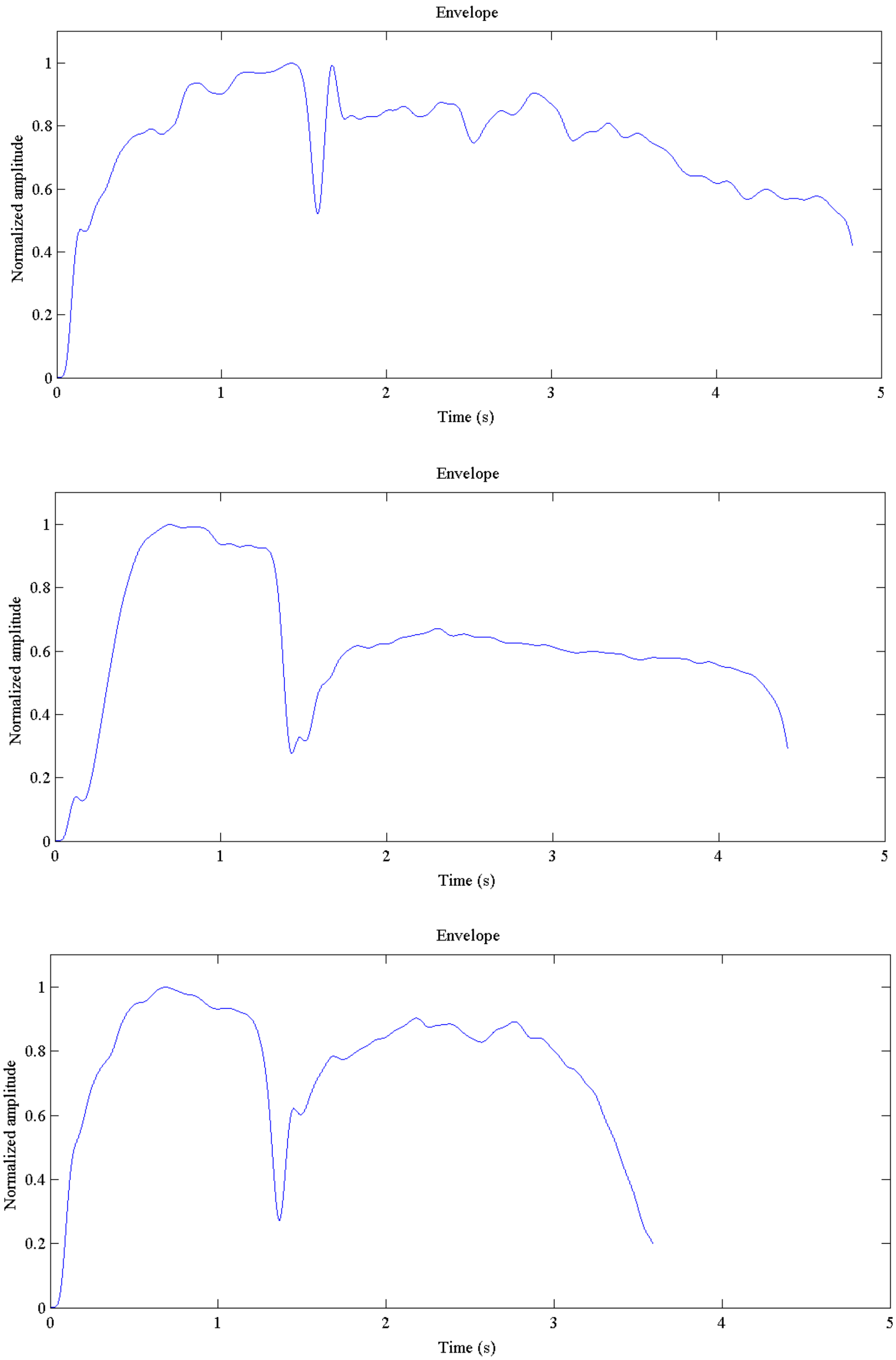


Figure 15: Normalized envelope of slurred transient between Bb2 and D3 with fundamental frequencies 233.1 Hz and 293.1 Hz on the 1st position. Top: Non-professional player (3), Middle: Professional player (7), Bottom: Experienced non-professional player (9).

Figure 16 shows normalized envelopes of the same slurred transient for all players. As can be seen, the individual differences are large. Some performers tend to try to maintain the same amplitude on both tones, and some vary the amplitude. There is also a spread in transience times for both attacks and transience between notes. Here, the players from figure 15 are players 3,7 and 9. Notice players 4, 7 and 9, who are the only classically trained players, all produce a rather slow, or soft, attack on the second note.

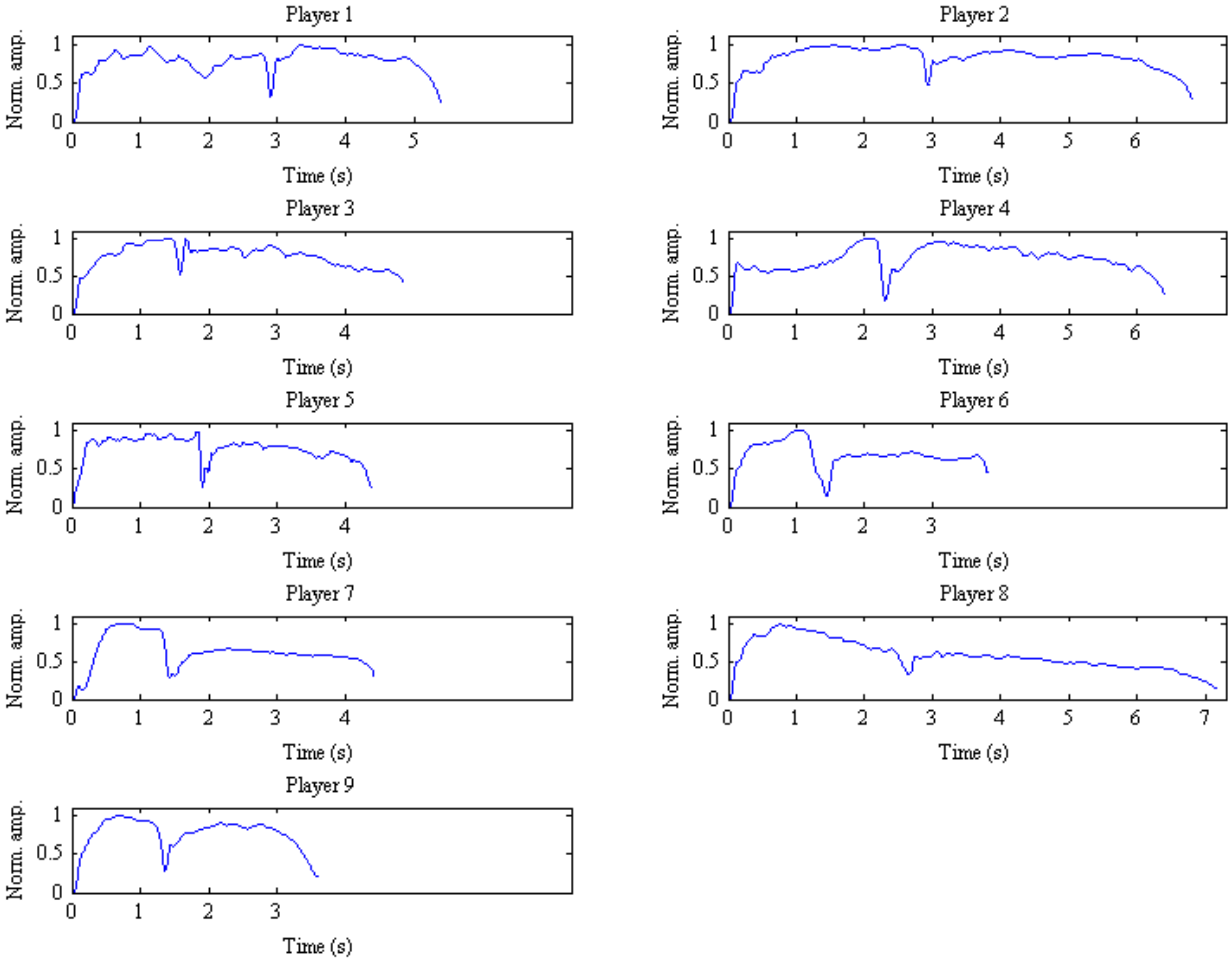


Figure 16: Normalized envelope of slur between Bb2 with fundamental frequency 233.1 Hz and D3 with fundamental frequency 293.1 Hz in 1st position for 9 players (*mirenvelope*).

Figure 17 shows the pitch of the slurred transients from Bb2 to D3. The dotted reference lines are the pitch of nearby (target) semitones. There is a wider spread in pitch on the D3, but this is due to the logarithmic scaling of the y-axis.

The pitch is calculated by a software for pitch tracking, called "*Yaapt" (Yet Another Algorithm for Pitch Tracking),* originally designed for speech communication. The program used a frame length of 15 frames. The code for the pitch tracking is listed in the appendix (Zahorian and Hu, 2008).

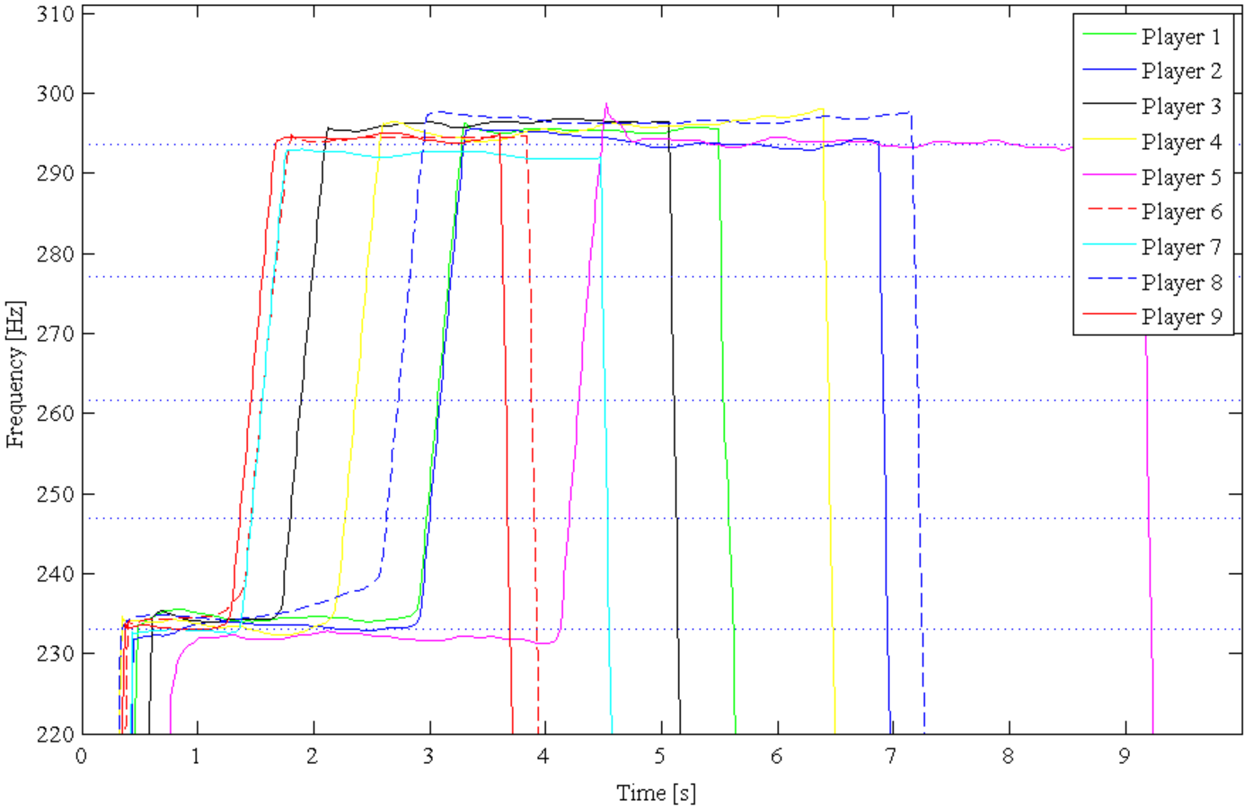


Figure 17: Pitch of slurred transients between Bb2 with fundamental frequency 233.1 Hz and D3 with fundamental frequency 293.1 Hz on 1st position for all players. Calculated with *Yaapt*.

The harmonic components of the slurred transients for a professional player (player 7) are displayed in the spectrograms in figure 18 and 19. The first figure is a slur from Bb2 to D3 on 1st position and the second is from E1 with fundamental frequency 164.8 Hz to Ab1 with fundamental frequency 207.6 Hz, on 7th position. This is considered to be a more difficult lip-slur, because the length of the air column has increased, requiring more energy to sustain tones.

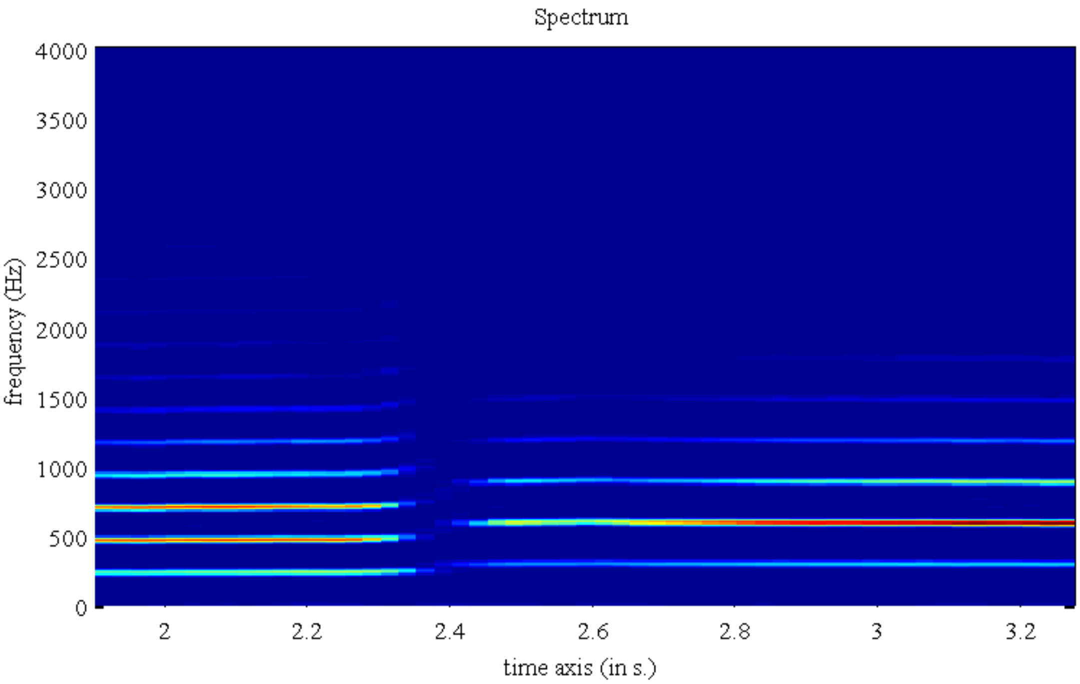


Figure 18: Spectrogram of slurred transient from Bb2 with fundamental frequency 233.1 Hz to D3 with fundamental frequency 293.1 Hz on 1st position. *(mirspectrum)*. Player 7.

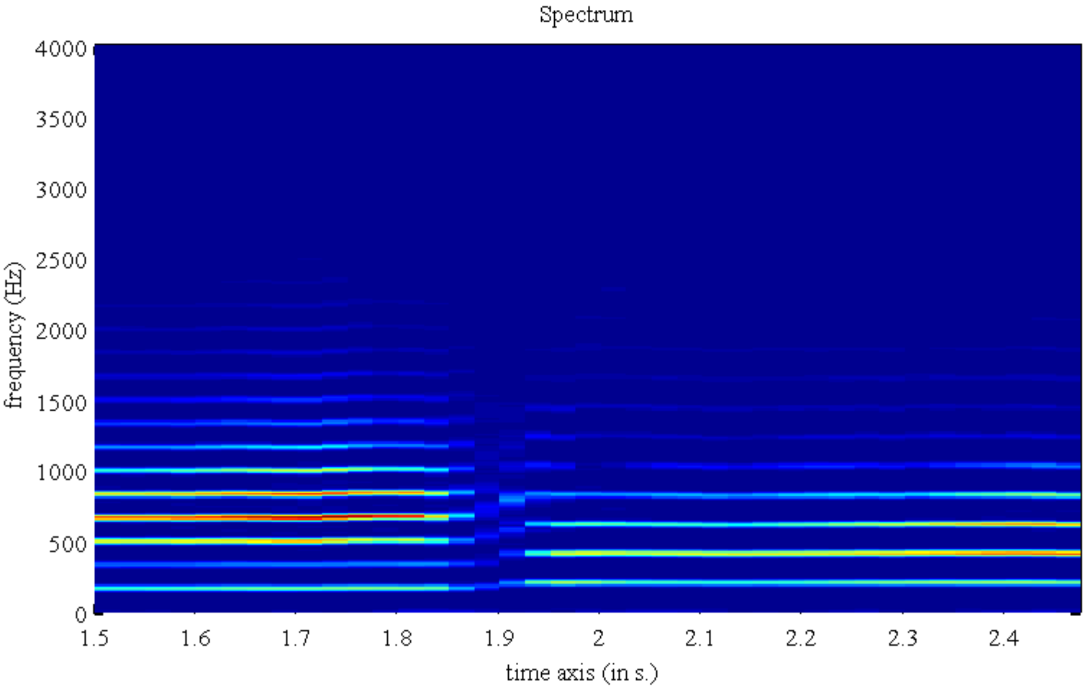


Figure 19: Spectrogram of slurred transient from E1 with fundamental frequency 164.8 Hz to Ab1 with fundamental frequency 207.6 Hz, on 7th position. *(mirspectrum)*. Player 7.

3.2.4 Glissando

Figure 20 shows a glissando from E1 with fundamental frequency 164.8 Hz on 7th position to Bb2 with fundamental frequency 233.1 Hz on 1st position, performed by a professional trombonist.

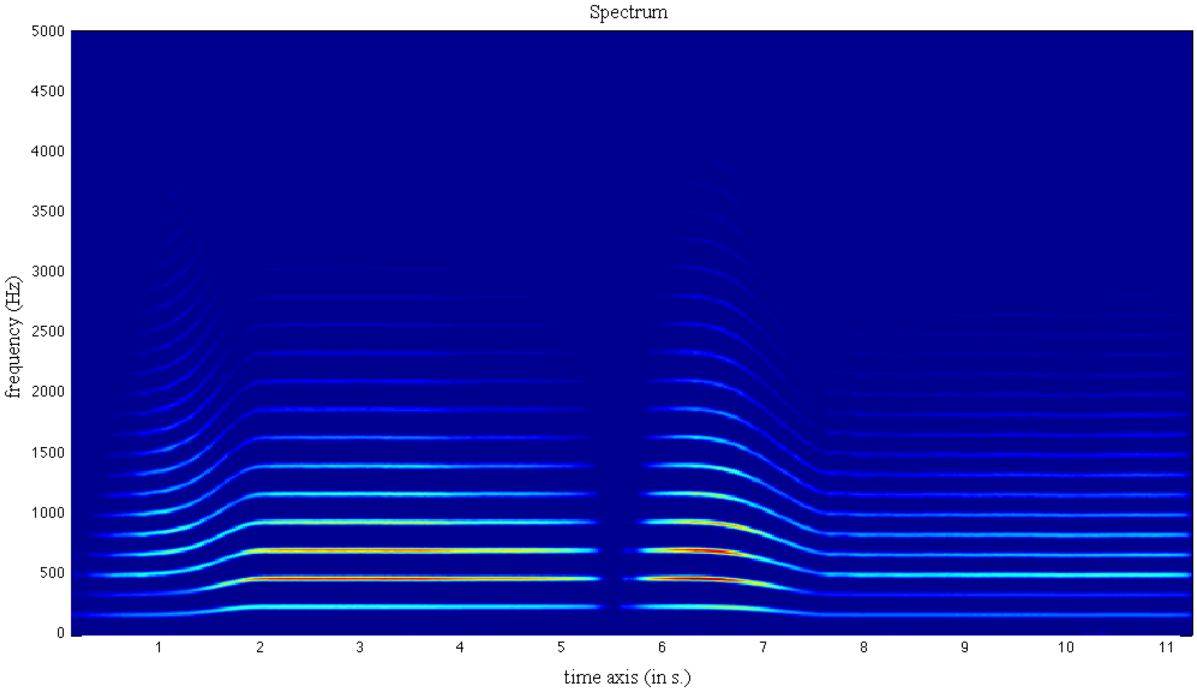


Figure 20: Glissando up and down from E1 with fundamental frequency 164.8 Hz to Bb2 with fundamental frequency 233.1 Hz. *(mirspectrum)*. Player 7.

3.2.5 Attack transients

There are a number of ways to start a note on a brass instrument. The most common is to deliver the tone as if pronouncing the letter t. This is what we call a *–tu* attack. Another version of this is the -*ku* attack, which is mostly used in combination with the *–tu* attack, pronouncing '*tukutukutu*' to play fast, successive notes. The third method investigated in this experiment is the excitation of the instrument without the use of the tongue. That is, just blowing air through the lips and not using the tongue to mark a clear starting point. Figure 21 shows 8 different players doing a *tu-attack* on a Bb1 note with fundamental frequency 116.5 Hz on 1st position. Figure 22 shows a comparison of the three investigated attack techniques for three players. The results for all attack envelopes are found in the appendix.

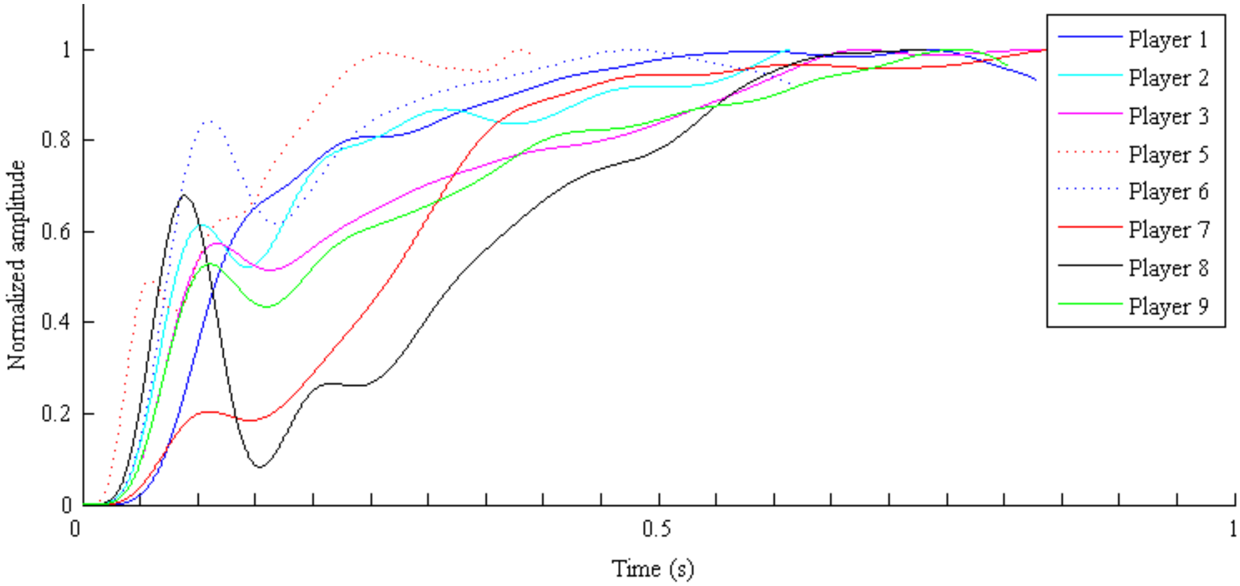


Figure 21: Normalized envelopes of a *tu-attack* on Bb1 with fundamental frequency 116.5 Hz for 8 players. (*mirenvelope*). The figure shows the first second of the notes.

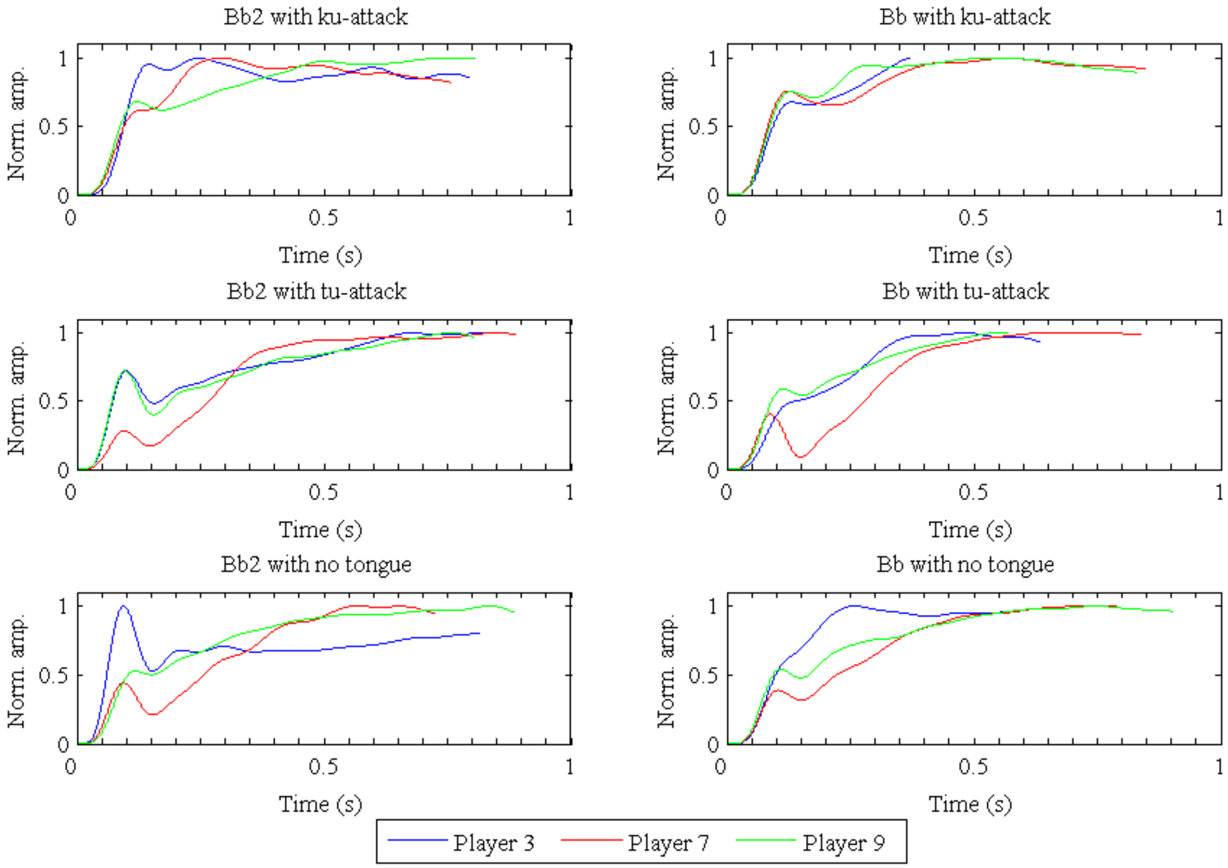


Figure 22: Normalized envelopes of attack transients for three different players, using three different attack techniques. Left column: Bb2 (fundamental frequency 233.1 Hz). Right column: Bb (fundamental frequency116.5 Hz). Red: Professional musician (7), Blue: Non-professional (3), Green: Experienced, non-professional musician (9).

3.2.6 Crescendo and decrescendo

An airstream that continuously increases in volume is the foundation of a crescendo. This leads to a wider and wider mouth opening and a higher pressure. There are no right or wrong ways to play a crescendo. A linear or exponential curve could be desirable, but the slope of the curve is a musical choice. However, it is generally agreed that the curve should be as continuous as possible, with few dents and imperfections. The players were instructed to play a crescendo from *pianissimo* to *fortissimo* in one breath. Figures 23 and 24 show the envelope and spectrum of a crescendo on Bb1 with fundamental frequency 116.5 Hz for player 9. The especially strong harmonics around 2000 Hz that are produced after about 2.5 seconds, at the peak of the crescendo, are what we refer to as the *brassy* sound of brass instruments.

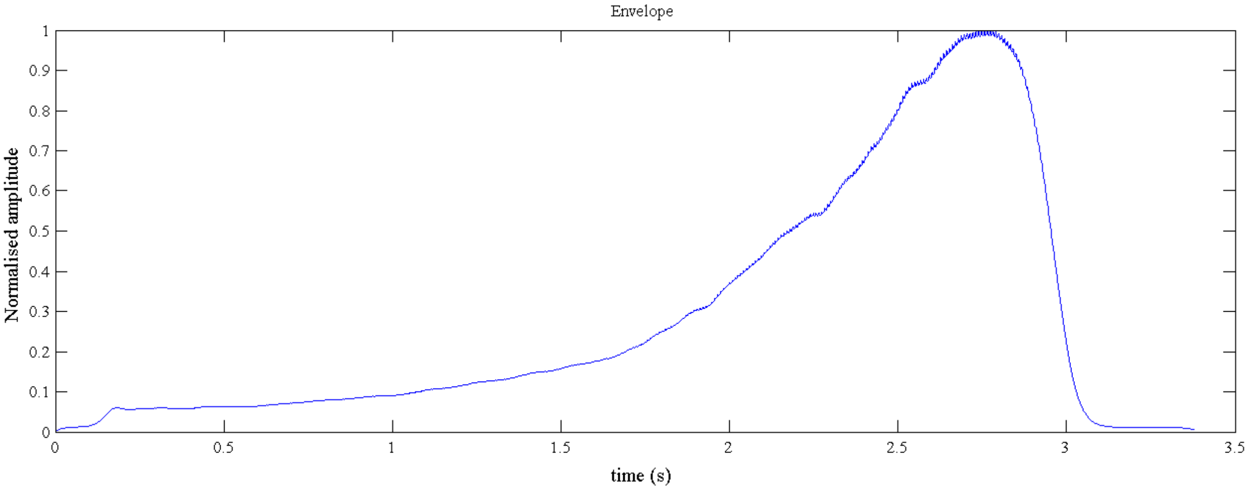


Figure 23: Normalized envelope of a crescendo on a Bb1 with fundamental frequency 116.5 Hz for player 9. *(mirenvelope)*.

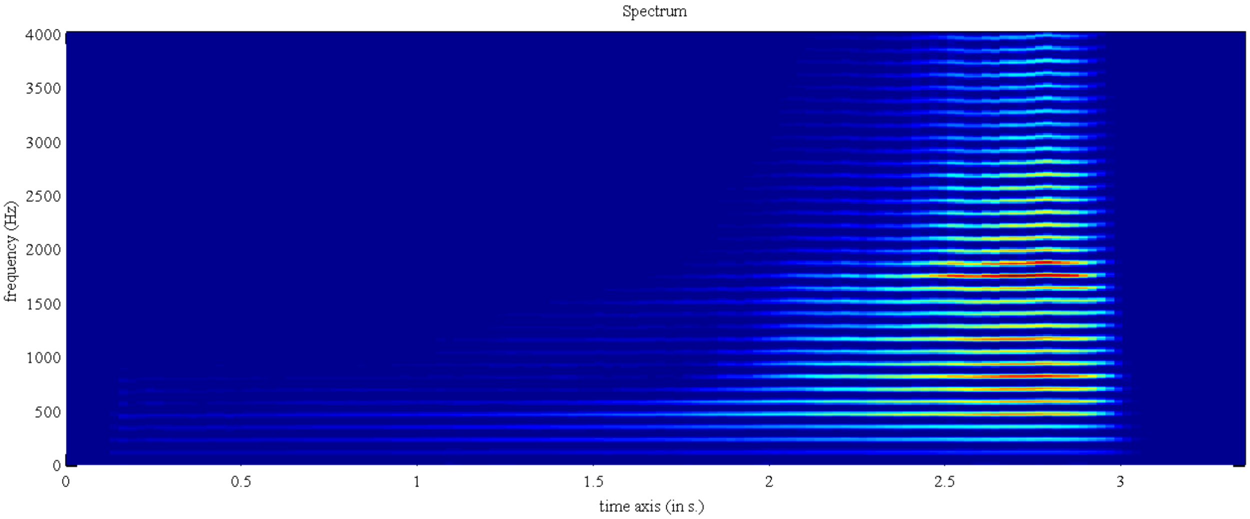


Figure 24: The spectrum of a crescendo on Bb1 with fundamental frequency 116.5 Hz for player 9, in the time domain. *(mirspectrum)*

Figures 25 and 26 show the opposite of a crescendo, a *decrescendo*.

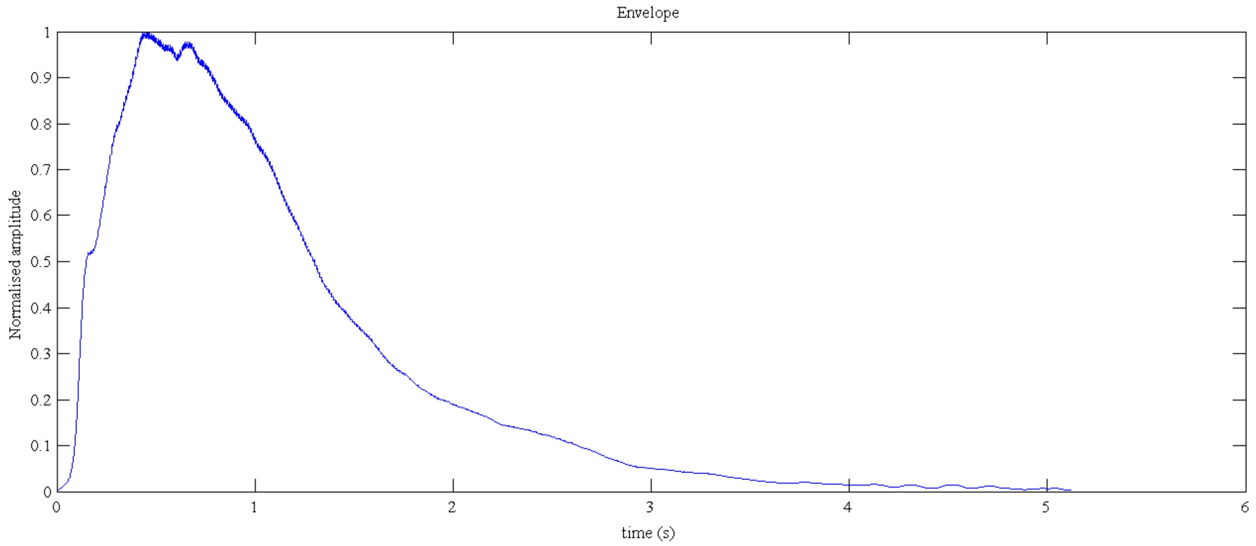


Figure 25: Normalized envelope of a decrescendo on Bb1 with fundamental frequency 116.5 Hz. Player 9. *(mirenvelope)*.

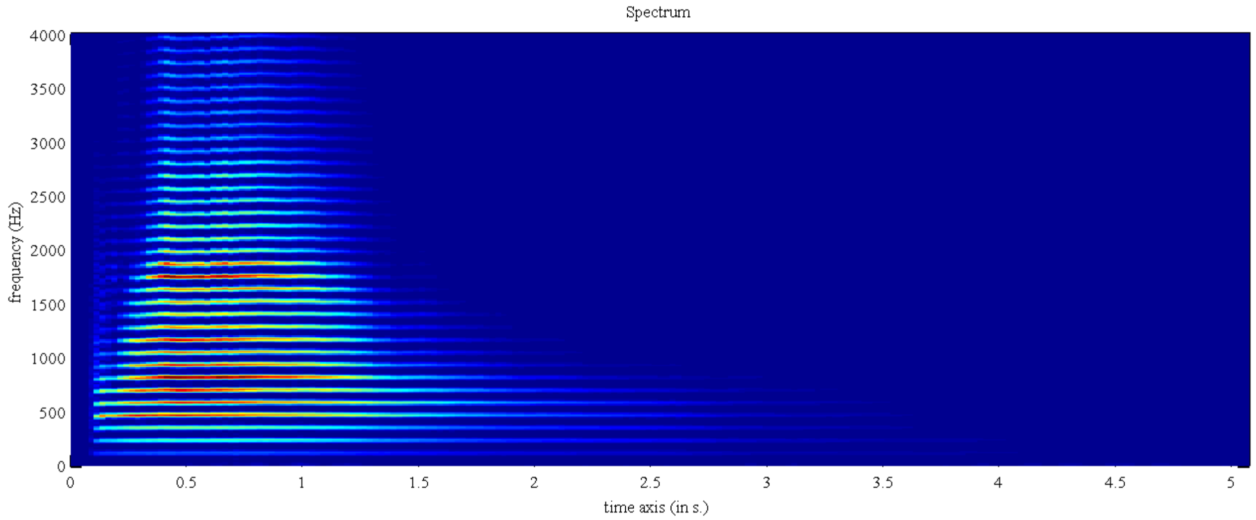


Figure 26: Spectrum of a decrescendo on Bb1 with fundamental frequency 116.5 Hz for player 9. *(mirspectrum)*.

In figure 27, we observe the envelopes of the crescendos for all players.

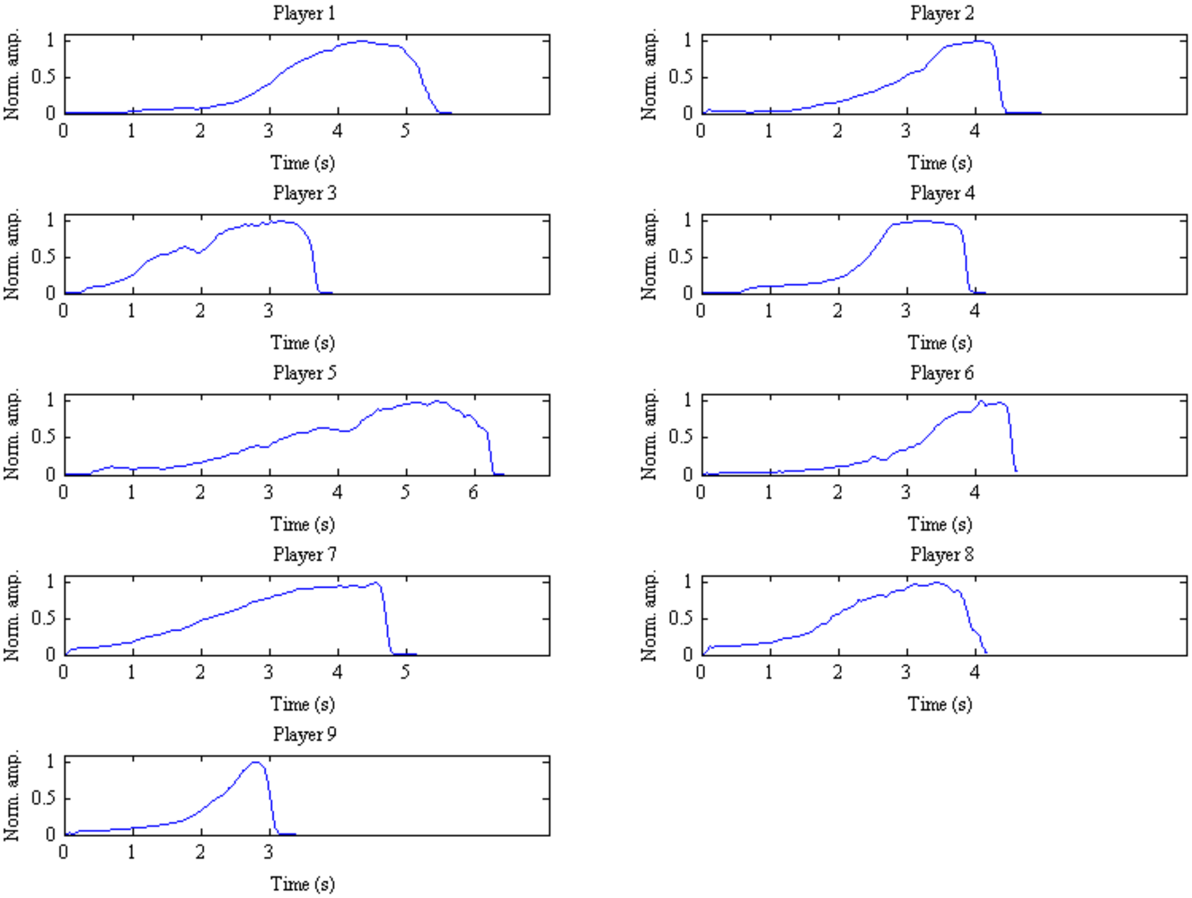


Figure 27: Normalized envelopes of crescendos on Bb with fundamental frequency 116.54 Hz for 9 players. *(mirenvelope)*.

When playing a crescendo, not only the dynamic level is increased. The rise in dynamic level does not mean that the already existing partials get stronger; it results in the formation of new partials and much stronger higher partials. This means that the spectral gravity (centroid) also increases. When a crescendo is performed e.g. with the voice (try yourself), the partials are mostly the same, they only get stronger. But in a trombone and other brass instruments, the spectral content is dramatically changed. Figure 28 shows the temporal evolution of the spectral centroid in crescendos for all players. As can be seen, the curves of the centroids take almost the same form as the energy envelopes. This is why brass instruments are particularly interesting in orchestral scores to produce dramatic and powerful musical elements. The crescendos are both *amplitude- and timbre crescendos*.

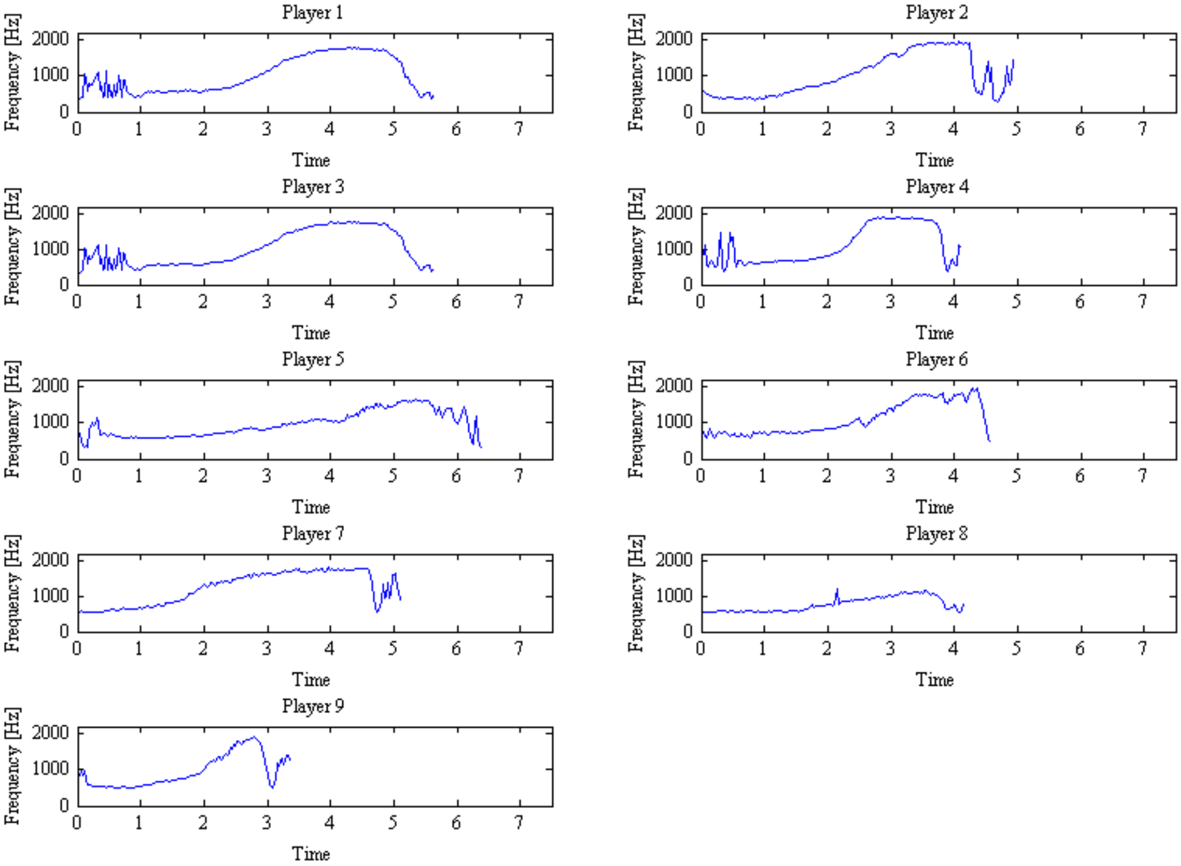


Figure 28: Spectral centroid of crescendos for 9 players. The y-axis shows the center of gravity of the frequency contents. The note is a Bb1 with fundamental frequency 116.5 Hz. *(mircentroid)*. All the peaks that occur after the crescendo has stopped are due to noise from the instrument.

3.3 Discussion

The results give an insight into the playing characteristics of the trombone and the players. Some results are merely depicting basic concepts and their characteristics, while some are more informative. Historical results from brass instrument research have been visualized and are hopefully more easily understood.

3.3.1 Slurred transients

In figure 15, player 7 (professional trombonist) produces a more stable amplitude in the sustained tones, and perhaps surprisingly, a longer transient time than the non-professionals. This could suggest that he has a more conscious performance and tries to smoothen the transition, so as to produce a softer attack on the second note. In figure 16, we can observe that the players make different musical choices. For most players, the second note is approximately equal with the first in amplitude. This is true for players 1,2,8 and 9. However, two players (4 and 7) insist on a softer second note. This observation is consistent through all the slurred transients (see appendix), and could suggest that these players make a conscious choice in the performance. Both these players are experienced professionals. The reason for this choice is that it sounds more subtle and delicate when the second note is softer than the first.

3.3.2 Attack transients

The ideal attack transient is a curve as smooth as possible. As figure 21 and figure 22 show, the players have a tendency to “overshoot” at the start, meaning that the amplitude in the beginning of the note is higher than the rest, or there is a pulse-like amplitude in the beginning of the waveform. An explanation for this effect can be found in a study from 2013, which introduced the matter of pulse-like lip vibration in starting transients of brass instrument tones (Kemp and Smith, 2013). The occurrence is due to an insufficient lip tension or an overlap of the lips, and found especially present for beginners. When the lips overlap, a greater mouth pressure is needed to force them to open, resulting in a burst of high-pressure waves in the start of the note. This only happens in the start of the attack.

The reason for this behavior could be that during the first cycles of oscillation, the wave has not reached, and reflected from, the end of the instrument and begun to help sustain the tone (Benade, 2012). This is also true for the other attack transients, where the performer has even less control of the air stream. You can try yourself to pronounce a ‘k’ and a ‘t’ followed by a stream of air. You will notice that the *–tu* attack gives you more control of the airstream. Generally, the “*tu-attacks”* have the smoothest transients, indicating that this is both the easiest and most commonly used way of starting a note.

3.3.3 Glissando

In the spectrum of the glissandi in figure 20, the amount of harmonics increases in the upward glissando. This is because the instrument’s length decreases and therefore needs less power to be excited. The extra energy contributes to a richer harmonic spectrum. The opposite can be seen on the way down. We can also see that it takes some time for the higher partials to establish themselves in the beginning of the tone.

So what does this mean for the performance? In trombone performance, moving the slide back and forth at a rapid pace produces vibratos. Since this contributes to a lengthening and shortening of the air column, and we already know that this length change causes energy differences, we can prove that a vibrato also causes an amplitude difference (tremolo).

3.3.4 Crescendo and decrescendo

As is evident from figures 23-26, this is an experienced player, who produces very smooth curves. There is a dent in the envelope for the decrescendo in figure 25. A possible explanation for this imperfection is that it is harder to start a note with high dynamic level than with soft dynamic level, because when the instrument is presented with very high pressures, non-linear effects occur (wave-steepening and shock wave formation), that interfere with the resonances of the instrument. Also, it is a “long way to go” for the lips between the non-oscillating state (before the attack) and the *fortissimo level,* and the lips have to go all this way in just half a second, instead of a soft attack.

As the figures of spectral centroids for all players and one player show (figures 12 and 28), there are peaks in the end of each recording, even though the tone has faded out. To examine this, the "silent" parts in the end of each recorded note were amplified 40 dB, and the reason was found to be noise generated by water drops inside the instrument, and the noise floor in the room. The former of these was recognized as the sound brass instrument players often hear when they need to open the water key, or spit valve, which is a valve situated somewhere on the instrument to clear out residual water from the instrument. The sound has a beating characteristic, and often occurs on crescendo performances because the large pressure variations inside the instrument makes the moist air condenses more efficiently. This also leads to viscothermal wall losses as described by Chick *et al*. (2012).

3.3.5 Individual differences

We observe many different individual choices of performance. The duration of each note, amplitude envelopes and transients all differ from each other. It is therefore hard to find strong correlations between for instance classical training and jazz. It is also quite difficult to prove that professional players and amateurs execute the tasks differently. The human parameters seem to have a very strong influence of the produced sound.

In a comparison of two trombones, it would be almost impossible to separate them when analyzing the performance of a musician on said instruments, because the player parameters are more dominant on the sound. However, the player would perhaps be able to tell the difference. Sterile measurements of acoustic input impedance would also be sufficient to characterize the instruments.

The individual musical choices and playing styles of each musician were far more important for the resulting sound than the difference between professionals and amateurs. But one key aspect where the professional musician and the semi-professionals differed from the rest was the level of tone control and consistency of choices throughout the session. The reasons for this can be many, and chapter 4 will further elaborate upon the aspect of control.

PART II

4 Auditory feedback deprivation

Auditory feedback aids the musician in playing. You are able to hear what you are playing and therefore adjust your performance along the way. But most musicians are skilled enough on their instrument to play it without hearing what comes out. A study done in 2003 concluded that auditory feedback is a contextual factor that affects learning but is relatively independent of retrieval, that is, playing a piece from memory (Finney and Palmer, 2003). They also found that the removal of auditory feedback had little influence on the performance on well-learned music. They further concluded that auditory feedback may not be as universally important or dominant a phenomenon as is usually assumed.

4.1 Auditory feedback, bone conduction and auditory imagery

4.1.1 Feedback loops

When playing an instrument, it is not a linear process where the player produces an air jet and sends it through the instrument on its journey to the audience. The process consists of several feedback loops. An example is the aural feedback that consists of a tone being produced by the player and the player's perception of that note. This feedback is greatly dependent on acoustic parameters of the environment, such as reverberation and radiation characteristics of the instrument and of the room. Other feedback loops can be that of the audience's response to the music and tactile feedback between the instrument and the mouth. Figure 29 shows an overview of the different feedback loops in brass instrument performance. The leftmost feedback is only prominent when an audience, listener or other musicians are present. The figure shows 4 different feedbacks. Note that the figure suggests that the loops work individually, but the truth may be more complex. For instance, perception bone-conducted sound from the tactile interaction blends with the perception of air-conducted sound.

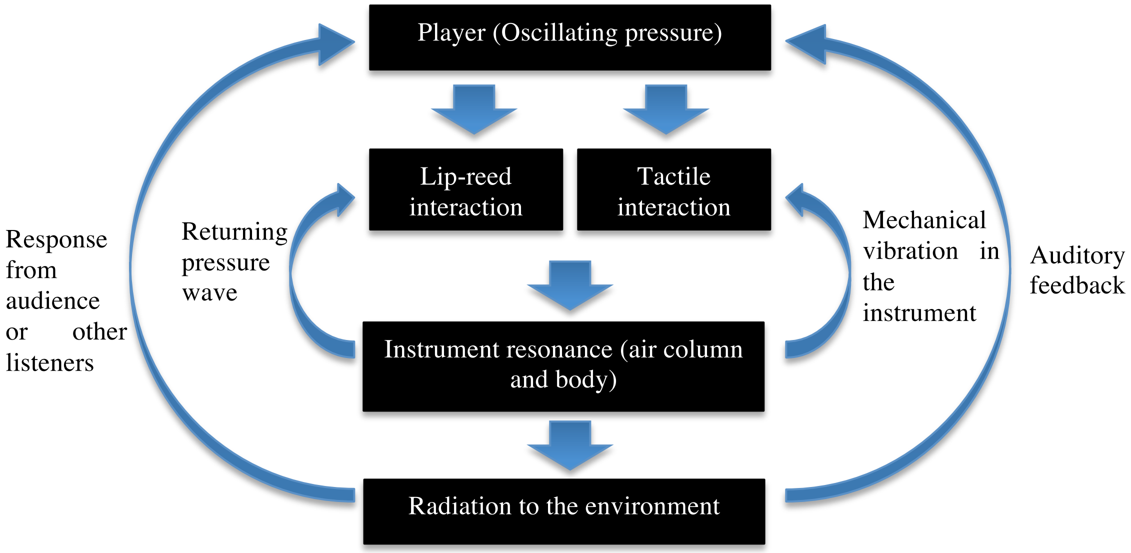


Figure 29: Trombone feedback model. The player provides the instrument with an oscillating pressure wave. When the oscillating regime is established, several feedbacks are present. The model does not take into account any visual feedback.

4.1.2 Feedback influences

A relevant question is to which degree the individual musician utilizes each of these feedback loops. This paper attempts to provide an answer by removing auditory feedback and comparing the results of performance *without* auditory feedback with performances *with* auditory feedback. Musicians use these feedbacks to enhance their performance in some ways. Intonation, tone stability and tone quality to mention some. But not all feedbacks are relevant for all the performance attributes. Figure 30 shows a suggested relationship between them.

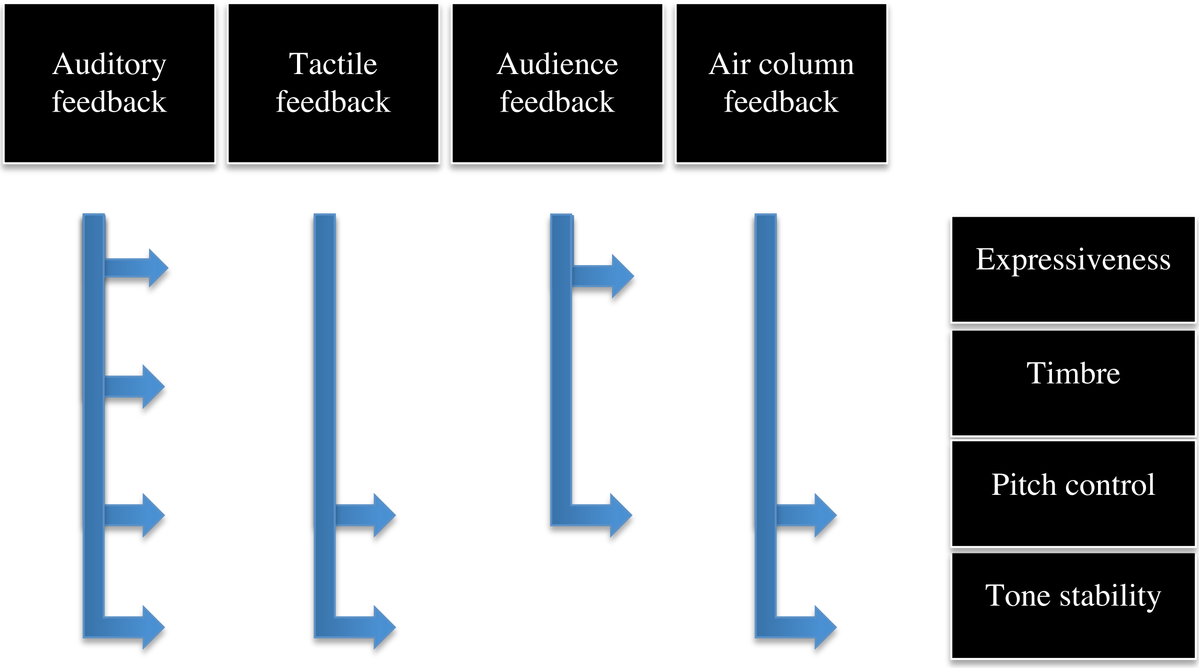


Figure 30: Suggested relationship between feedback loops and sound production in wind instruments.

4.1.3 Bone conduction

We may not know it, but we hear bone conducted sound all the time. If you put your fingers in your ears and start talking, the experienced voice sound is the sound that travels as bone conducted sound from your mouth through your bone and flesh to the cochlea. This sound behaves differently than normal air conducted sound. A number of papers have been written on the matter of bone conduction (v. Békésy, 1949; Corso, 1963; Stenfelt *et al*., 2002; Stenfelt *et al*., 2003; Stenfelt, 2011; Muramatsu *et al*., 2013), the latter focusing on sound conducted via teeth and cheek. The study proved that bone conduction is extra effective through the teeth, since they are not covered with skin and flesh. This can also be verified by any brass instrument player, or by putting a tuning fork to your teeth. The experienced sound is very clear. Bone conduction is utilized in hearing aids when the damaged part of the ear is either the outer or the middle ear. The bone conducted sound bypasses these and travels directly to the cochlea. Special waterproof earplugs are developed for swimmers to listen to music underwater, as the sound comes directly from tactile contact with the skull (Agus *et al*., 2012).

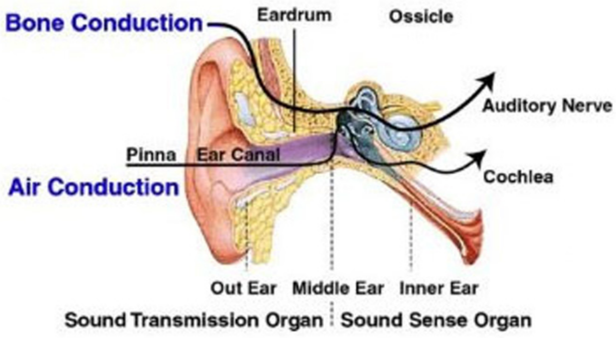


Figure 31: Bone conduction. The air-conducted sound travels through the ear canal and the middle ear, whereas the bone-conducted sound can bypass these and travel directly to the cochlea. As a result, the experienced sound is filtered through resonance characteristics of the skull and flesh surrounding the ear.[5]

The bone-conducted sound behaves differently than the air-conducted sound. If you think of a sound consisting of several frequencies being played in a room, your ears will boost this sound in certain frequency bands due to resonance in the ear canal and directivity. The sound that is picked up by your skin and bones will be affected in a different way. We can approximate this by saying that two different filters filter the sound. The bone-conducted sound will have a bass boost due to resonance in the bones and flesh in your head. The result of this is that the perceived bone-conducted sound will differ from the air-conducted sound.

4.1.4 Attenuation of bone-conducted sound from a trombone

The bone-conducted sound is filtered through the acoustic properties of the skull, bones and skin. The filtering properties of these are very complex, as they rely on head size, angle of attack, and many other individual properties. However, the effects of these parameters are small and we are only interested in finding out how musicians perceive this sound. It is therefore sufficient to say that the sound is changed so that it does not sound 'as good' as air-conducted sound. This is why many wind instrument players prefer to play without earplugs. The sound can be described as

---

[5] From http://siliconstation.com/aftershokz-listen-bones/ (Accessed March 14, 14)

"tinny," referring to the lack of resonance, or reverberation, in the sound. The expression is derived from the metal tin, which makes a non-resonant and damp sound when excited.

In this paper, we investigate the effect of auditory feedback deprivation. In order to design an experiment, which makes all sound inaudible, one has to attenuate or mask the bone-conducted sound as well as the air-conducted sound. Figure 32 shows attenuation level for sound when the ears are masked, open and plugged with an earplug. As the figure shows, bone-conducted sound under 1500 Hz is easier heard when the ears are plugged. This is due to resonances in the small, enclosed volume between the earplug and the tympanic membrane. When attempting to attenuate or mask bone-conducted trombone sound, plugging the ears would only make it worse. As can be seen from 100 – 1000 Hz, the playing frequency of a trombone, the attenuation level increases with frequency, meaning that the high-pitched sound is more difficult to mask.

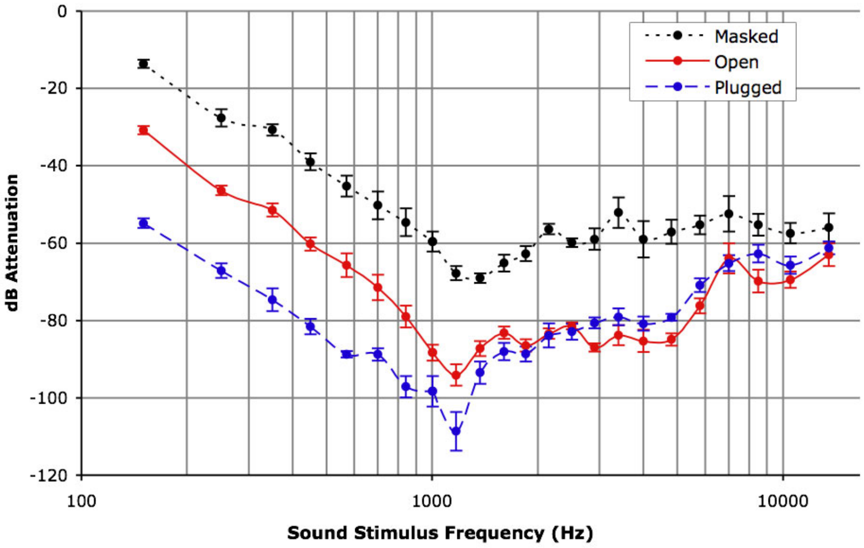


Figure 32: Bone-conducted sound, attenuation level for the sound when the ears are masked, open and plugged with earplugs. The chart indicates how much the bone-conducted sound will have to be attenuated in order for it to be inaudible.[6]

There is no practical way to attenuate bone-conducted trombone sound for this experiment; hence the answer will be a masking noise played through headphones (not earplugs). The masking noise will have to be quite loud in order for it to be an efficient masker. In the end, there is no

---

[6] From http://sonify.psych.gatech.edu/publications/pdfs/2005ICAD-WalkerStanley.pdf (accessed May 6, 14)

fail-proof way of determining if the sound is completely masked, since we cannot measure the bone-conducted sound from the instrument in dB SPL. To determine if the masking is good enough, the experiment has to rely on the musicians being honest in their reports.

4.1.5 Auditory imagery

In Pink Floyd’s recording of “Wish you were here” from 1975, the lead guitarist, David Gilmour, hums his guitar solo while playing it. This is an example of what we call auditory imagery, showing itself through the voice of David Gilmour.

Auditory imagery is defined as the ability to perceive sound in the absence of the physical presence of sound. We could call it the inner ear (Allen, 2007). We can also imagine ourselves moving a hand or walking into the kitchen. This is called mental representation of motor processes, and happens unconsciously. In this study, the ability to create and use of mental representations are assumed to be skills that everyone possesses. However, people both train and use these skills differently. It is a high achievement in jazz music to be able to play what you are thinking. It is expected, that musicians with more experience, will have a much more developed skill in auditory imagery. This will lead them to play what they think, and when the auditory feedback is taken away, the performance will be unaltered, because they still hear it in their heads. This was proven correctly for some piano players in a test performed by Bruno Repp (1999) He asked pianists to play a piece of music a repeated number of times for different playing conditions; Expressive vs. metronomic, and with vs. without auditory feedback. The expressive vs. metronomic performance means that the musicians were asked to play with and without all the annotations in the musical piece. The musical piece was the first 5 bars of Chopin’s Etude in E major, op. 10, no. 3.

The results revealed significant effects of auditory feedback deprivation in both metronomic and expressive performance, suggesting that auditory imagery is “not enough” to perform adequately. However, the results were of small magnitude and could be explained by other reasons. Finney and Palmer (2003) showed that absence or presence of auditory feedback affects learning processes, but not the recollection of that knowledge when performing a learned piece from memory.

4.1.6 JND in pitch of complex tones

The just-noticeable difference, JND, in pitch of complex tones is about 9 cents. One cent is defined as 1/100 of a semitone. This means, that people can hear a difference in pitch when two complex tones are presented with the pitch difference between them being as small as about 1/10 of a semitone. In the range of the lower trombone register (fundamental of 116.54 Hz) this corresponds to a frequency difference of 0.6 Hz (Carral, 2005).

4.2 Method

4.2.1 Participants

Eight trombonist's, four out of which professional with different experience, and four amateurs with different levels of experience, participated in this experiment. They will be referred to as 'P (professional) 1-4' and 'A (Amateur) 1-4,' respectively. This label is introduced to ease the reading of the results, and is not an accurate description of their musical performance level. It is based on their experience, which is judged by comparing the number of years they have been playing, and the estimated number of hours a week of practice during the last year. For the professionals, the average number of hours a week of practice was 23 (range 15-40), and they had played trombone for a mean of 31 years (range 17 – 50). For the amateurs, the mean number of hours a week of practice was 5 (range 3-6) and the mean number of years played was 16 (range 11-20). They all participated voluntarily and were given a gift card as a token of appreciation.

4.2.2 Experimental setup

The recordings were carried out in the Aura Lab at NTNU Department of Electronics. The experiment consists of performing different musical pieces, including some simple notes and exercises, whilst not hearing. This was achieved by playing a masking noise of 95 dB through noise-cancelling headphones. The masking noise was pink noise, filtered through a band-pass filter with cutoff frequencies at 100 Hz and 1000 Hz, as to best mask the spectrum produced by a trombone. The maximum exposure time for the masking noise was 3 minutes, which lies well within the limit for noise exposure. This limit is approximately 45 minutes for 95 dB noise according to ISO 1999 from 2013. The headphones were *Bose Quiet Comfort 15*. To make sure the participants did not hear their own sound, a number of testing exercises were done. The participants were also asked to tap their foot to the floor if they heard themselves at any time during the performance. Figure 33 shows the experimental setup.

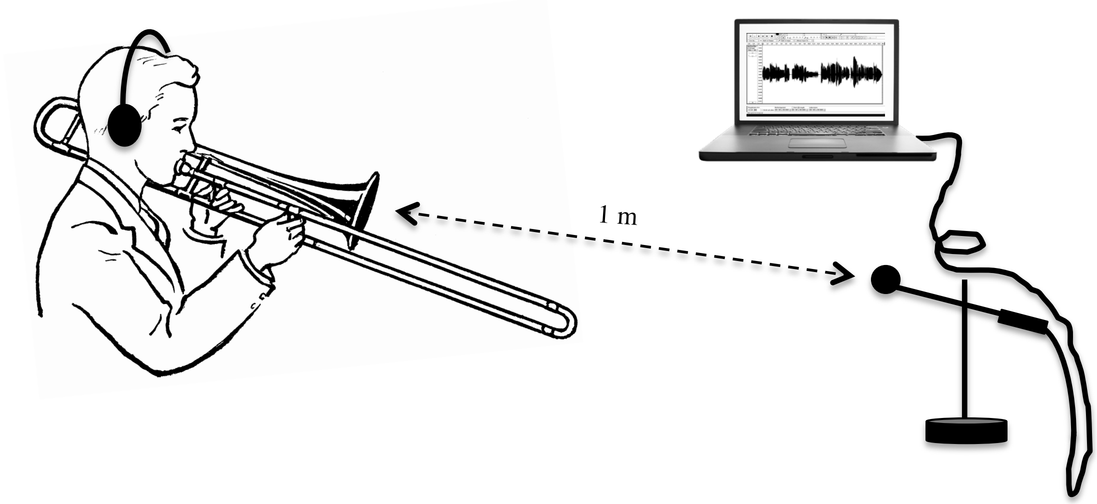


Figure 33: Experimental setup.

The performances were recorded with a ½” *BSWA MP-216* condenser microphone through a *Focusrite Forte* sound card. The sound was recorded with a sampling frequency of 44.1 kHz and 16 bit resolution. It was then fed into *Audacity*. The calculations were done in *Matlab* with *MIRtoolbox*. The players used their own instruments.

4.2.3 Exercises

The participants were asked to play two pieces of music, with durations of less than 30 seconds. The first piece was a simple melody written by the author of this thesis. It was written so that all participants would be able to play it *prima vista* (PV), and contained several musical aspects that could be affected by the removal of auditory feedback, such as vibrato, difficult intonation and more. This will be referred to as PV. The second piece was a piece the musicians chose themselves. They were told to pick a piece they knew by heart, and would not have trouble performing from memory. This latter piece is referred to as own piece (OP). They played each piece 5 times wearing headphones. Subsequently, they performed the piece without headphones. In addition to the two pieces, they were asked to play some exercises. The whole session is described in table 1. After the session, the participants were asked to answer a questionnaire about their performance, as well as their own experience and training.

Table 1: Feedback experiment, overview of exercises.

Chronological order of ex.	Description	Number of times performed with/without masking
1	Prima Vista piece	5/1
2	Own piece (here, they played the unmasked version first)	5/1
3	Semitone pitch bending without slide (lipping down)	3/3
4	Holding a steady note for 5 seconds	3/3
5	Major triads starting from Bb (116 Hz) and Gb (92 Hz)	1/1
6	High pitched F (fundamental frequency 349,2 Hz) on three different positions (1,4,6)	1/1
7	Hitting a randomly chosen, presented tone	1/1

4.2.4 Statistical significance

The statistical significance at was calculated by performing a two-tailed student t-test with $p < 0.05$. Since only one recording was made with headphones removed, it is assumed that the variance of the unmasked data is equal to the masked dataset. This assumption is made based on sample recordings. Some players performed the unmasked takes a repeated number of times, and showed excellent stability in performance, with a variance equal or less than the variance for masked performances. It should be noted that the t-test is not designed for sample sizes as small as 1. Studies have shown however (De Winter, 2013), that the t-test is reliable for small sample sizes as long as the effect size, D, is large. The effect sizes for this experiment can be found in table 3 in the appendix. The smallest of which is 3, which is considered to be a large number ($D > 3$ yields that the probability of Type 1 errors is less than 2 %).

4.3 Results

4.3.1 Prima vista and own piece performance

Figure 34 shows the piece written by the author for this experiment. To examine the effect auditory feedback deprivation has on the perception of time and dynamics, duration and RMS energy were measured. Note that some of the recordings in take 5 are missing. This is because they contained musical errors that influenced the timing or RMS energy in a disruptive way.

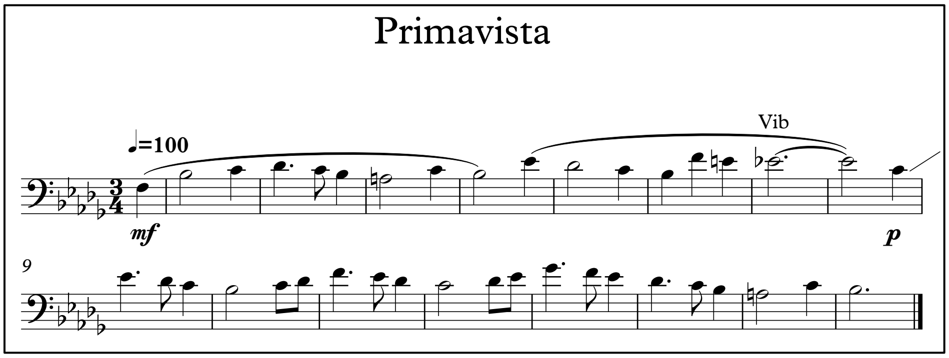


Figure 34: Prima vista piece to be performed with and without auditory feedback.

Duration

The duration was calculated after the recordings were trimmed in the start and end. The recordings were also inspected for musical errors that might affect total duration, e.g., failure to hit a note could result in half a second delay in the total duration. This is accounted for in the results. The note durations in the upstroke and endnote were also normalized, because participants had a tendency to play these too long. The calculation was done with the function *mirlength*. As can be seen in figure 35, the only players who significantly changed duration in the PV performance were two of the amateurs (A2 and A3), who played more slowly in the masked attempts. In the OP performance in 36, the same effect occurs for players A2 and A4. Player A1 plays *faster* in the masked attempt.

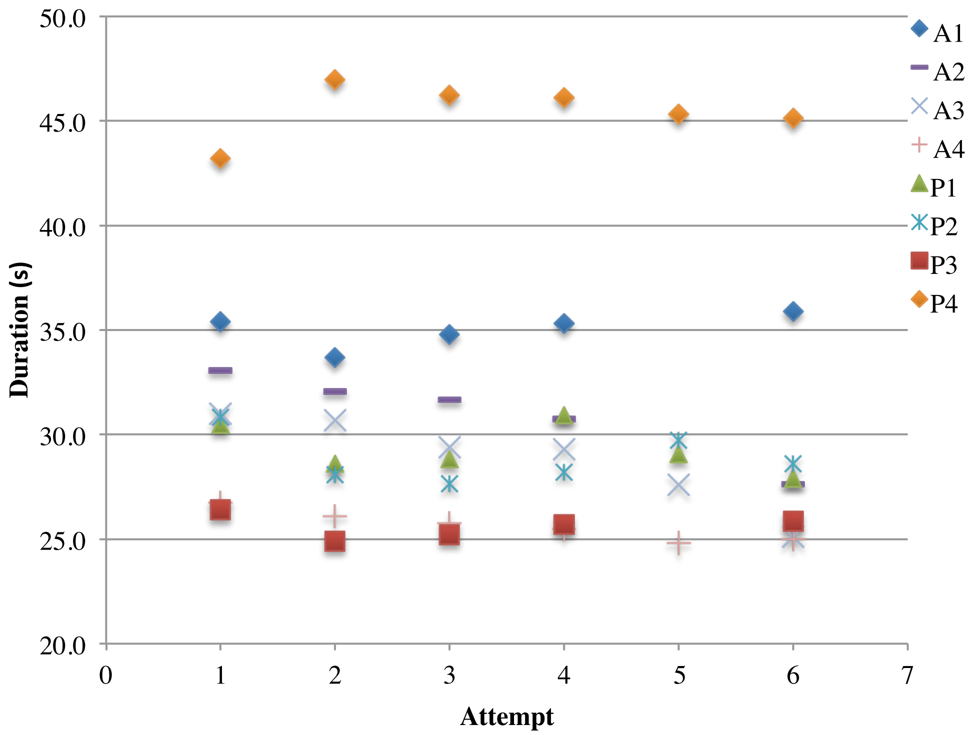


Figure 35: Duration of PV. Attempt 6 is unmasked. The results are calculated with *mirlength*, and the files are trimmed in the start and end to eliminate timing errors.

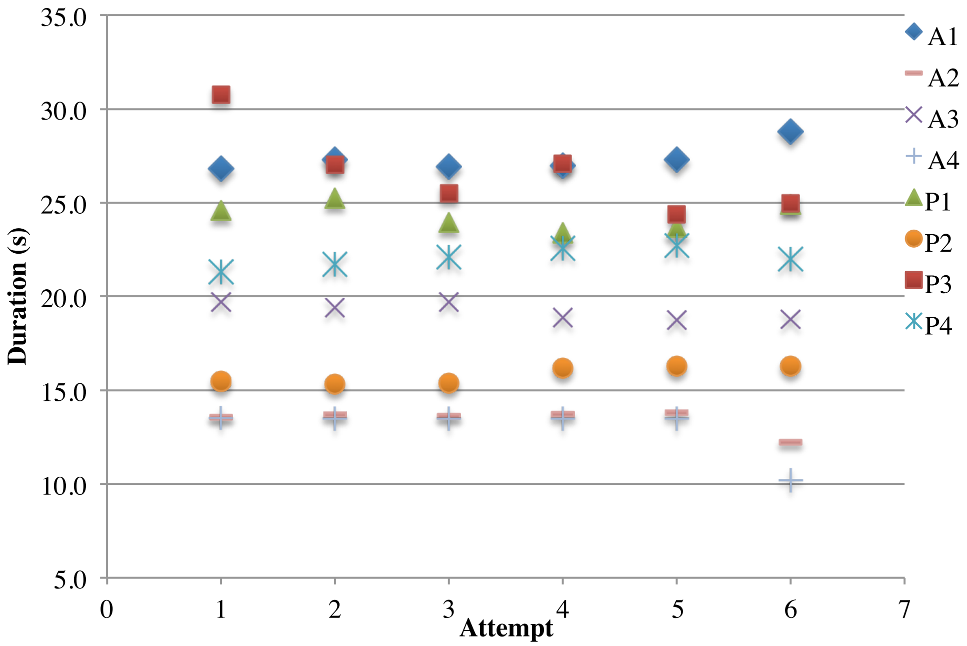


Figure 36: Duration of OP. Attempt 6 is unmasked. The recordings are trimmed in the start and end to eliminate timing errors. Calculated with *mirlength*.

RMS energy

To evaluate the dynamics in their performance, the RMS energy was calculated across a whole piece. The results are shown in figures 37 and 38. In the PV performance, only player A2 played significantly louder with the masking noise present. However, players P2 and P3 played significantly softer in the masked attempts.

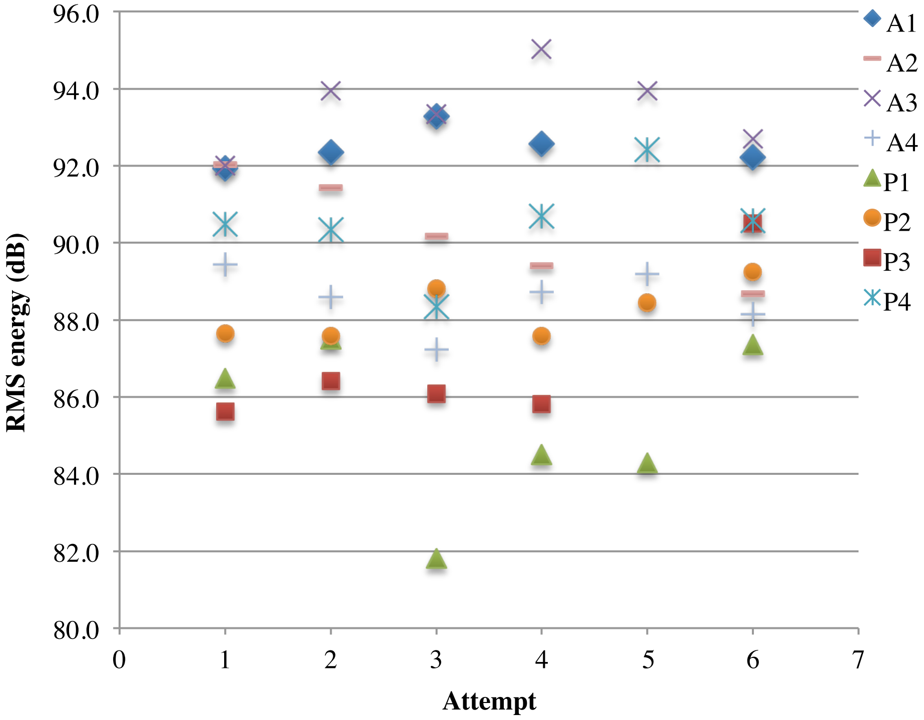


Figure 37: RMS energy of prima vista recordings. Attempt 6 is unmasked. Calculated with *mirrms*.

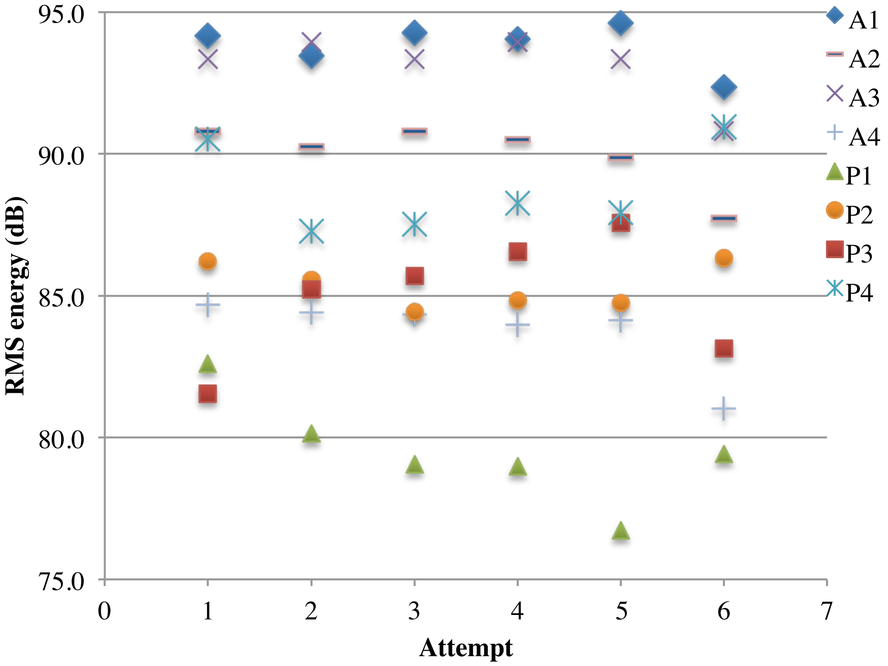


Figure 38: RMS energy of OP. Attempt 6 is unmasked. The results are calculated with *mirrms*.

Statistical significance of the results

The statistically significant results are as follows:

- Four amateur players played significantly louder in OP performance (mean difference 2.6 dB) and two professionals played significantly softer (mean difference 1.4 dB)
- Two amateur players played significantly slower in OP performance (mean difference 17 %), and one amateur player played significantly faster (6%).
- One amateur player played significantly louder in the PV performance (2.1 dB), and two professionals played significantly softer (mean difference 1.7 dB).
- Two amateur players played significantly slower in the PV performance (mean difference 14%).

Manual inspection

The performances of PV and OP have been inspected through listening. The inspection was carried out by the author, and the results are shown in table 2. The inspection consists of an evaluation of specific musical attributes. Four questions were asked:

- How expressive is the performance?
- How dynamic is the performance?
- How rhythmic is the performance?
- How "in tune" is the performance?

The performances are compared (unmasked vs. masked) according to these questions. The outcome is a description of the *total difference in performance*. This is described using four levels of magnitude:

- No noticeable difference
- Slight difference
- Some difference
- Large difference

Faults and errors in the performance, such as players missing a sign or failing to hit a note, are also mentioned in the table. The descriptions in the right column consider the performance in the masked recordings compared to the unmasked.

Table 2: Manual inspection of masked vs. unmasked performance. PV = prima vista, OP = own piece. Only the masked recordings are discussed in the right column.

Player	Manual inspection of performance for PV and OP
A1	PV: Some differences: Missed a b-sign in bar 13 in two takes of PV. More dynamic performance in unmasked take. OP: No noticeable differences.
A2	PV: Large difference: "Rough" sound. Hit wrong note in bar 6 on all attempts, but still hit the Gb in bar 13 every time. Less dynamic and expressive. Slightly out of tune. OP: Slight difference: Harder attack transients and generally less dynamic performance.
A3	PV: Slight difference: In bar 13, participant hit a semitone off, producing a very "wrong" note. OP: Slight difference: Played more unexpressive.
A4	PV: Large difference: Had trouble performing the prima-vista piece without auditory feedback. Performance was slightly out of tune and unexpressive. Several musical errors. OP: Slight difference: More subtle and expressive performance in unmasked attempt. Masked attempt was in tune and rhythmically good, but lacked dynamic variation.
P1	PV: No noticeable differences. OP: No noticeable differences.
P2	PV: No noticeable differences. OP: No noticeable differences.
P3	PV: Slight difference: little more dynamic, expressive playing in the unmasked attempt. OP: Slight difference: Performed worse on take 3-5 than on takes 1-2. Less dynamic in general and slight trembling on some notes.
P4	PV: No noticeable difference in performance. OP: No noticeable difference in performance.

4.3.2 Sustained notes

The players were instructed to hold a note as steady as possible for five seconds, in order to determine if tone stability is affected by removal of auditory feedback. Figure 39 shows the normalized energy envelopes of sustained notes on a trombone, for the two playing conditions. The note was a G1 with fundamental frequency of 196,2 Hz on 4th position. The envelopes are calculated with *mirenvelope*.

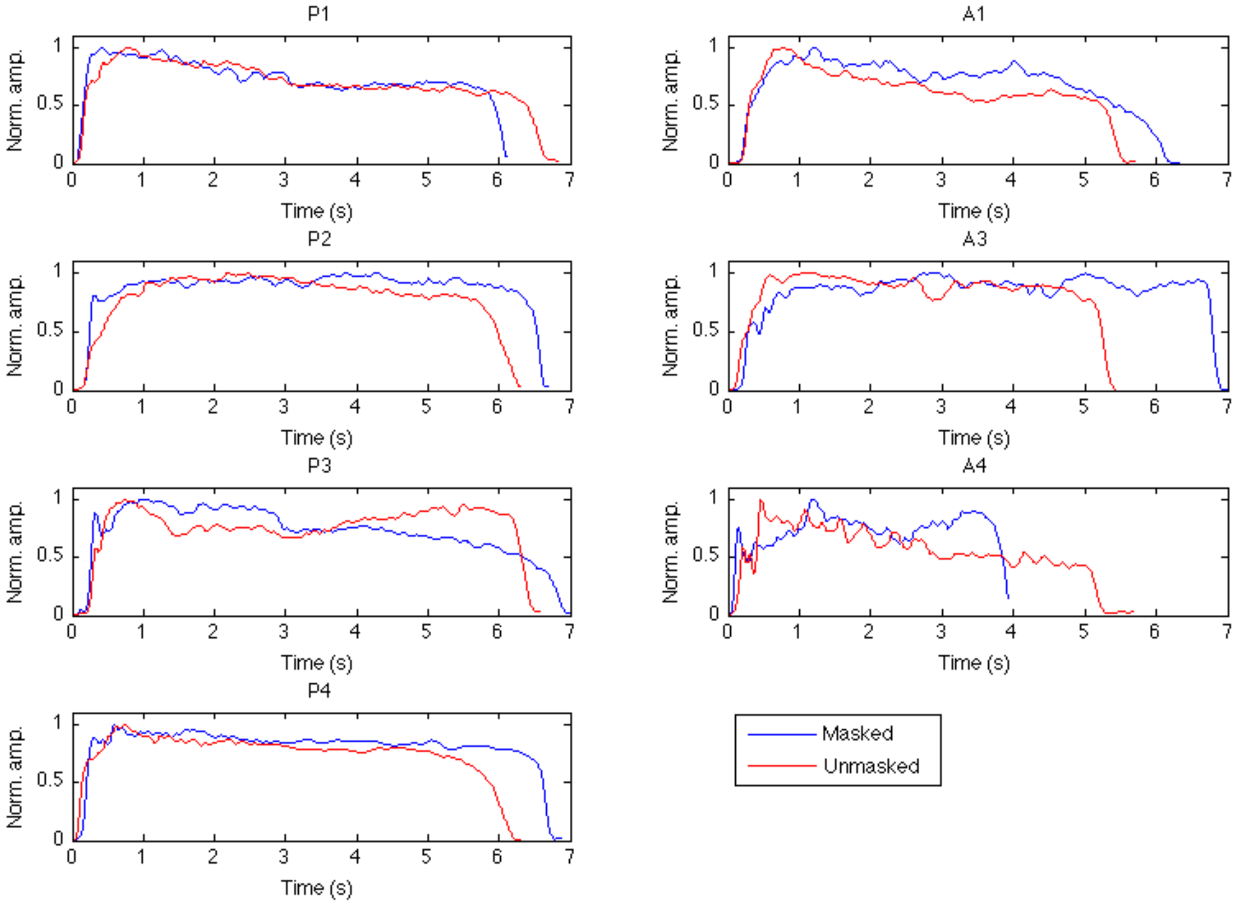


Figure 39: Sustained G1 with fundamental frequency 196.2 Hz on a trombone. Blue: Masked. Red: Unmasked. *(mirenvelope)*.

4.3.3 Pitch bending

Players were instructed to bend the pitch down a semitone from three different starting notes, using only their lips. Figure 40 shows pitch bending for a professional trombonist (P3), attempting to bend the pitch down a semitone from three different start notes (Bb1, A0 and Ab0). The rest of the pitch-bending results are shown in the appendix. The dotted reference lines are the pitch of the nearby semitones (Bb1 – 116.5 Hz, A0 – 110.0 Hz, Ab0 – 103.8 Hz etc.).

The pitch is calculated by software for pitch tracking, called "*Yaapt" (yet another algorithm for pitch tracking),* originally designed for speech communication. The program used a frame length of 15 frames. The code for the pitch tracking is listed in the appendix.

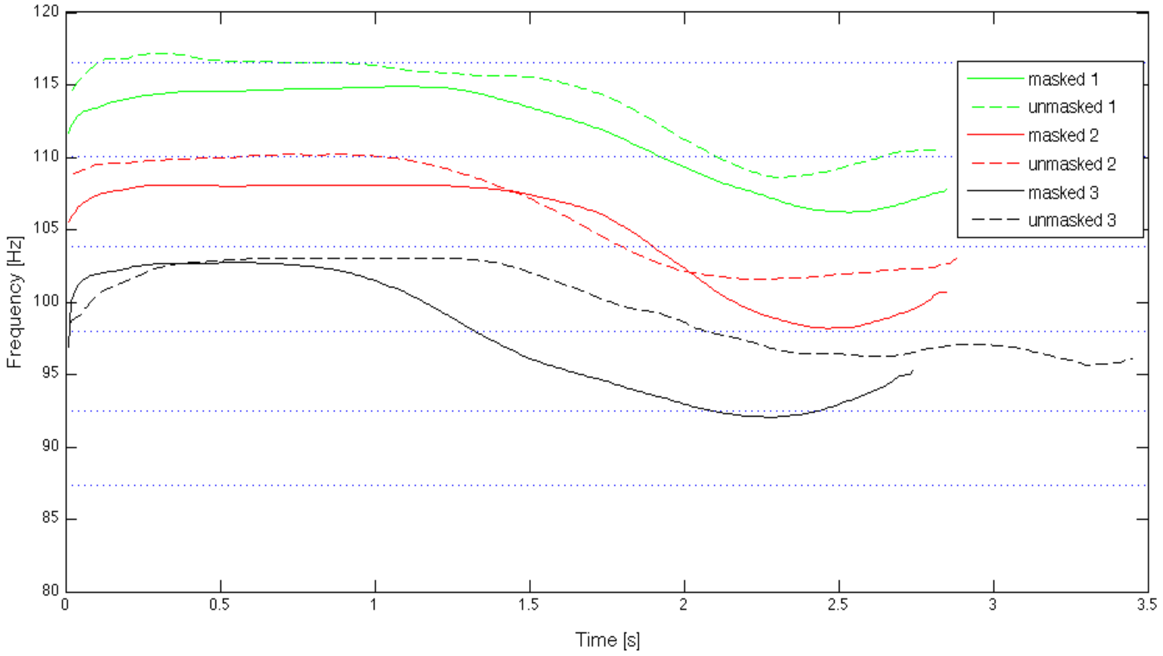


Figure 40: Semitone downward pitch bending from starting notes: Bb1 – 116.5 Hz, A0 – 110.0 Hz, Ab0 – 103.8 Hz on the instrument's second resonant mode. Player P3. Calculated with a pitch-detection software (Zahorian and Hu, 2008).

4.4 Discussion

The overall results suggest that auditory feedback deprivation has an effect on trombone performance. The removal of auditory feedback affects an array of musical aspects in various ways. However, it is unclear whether all these results stem from the removal of auditory feedback itself, or just from the fact that the participants are psychologically affected by the situation, or perform musical choices to deal with the unfamiliar situation. The large individual differences in the results could suggest that the two latter explanations are the case. A masking noise at 95 dB is not comfortable, and may have been a large influence on performance. In addition, musical performance is not a repeatable exercise; it is a process that includes a number of sensory mechanisms, cognitive processes, and conscious choices. This leads to multiple sources of errors. However, the results indicate that there are grounds for further research.

4.4.1 Duration of prima vista and own piece performance

The results for duration of performance show that two amateur players played slower in the OP performance and two amateur players played slower in the PV performance. No professionals changed their performance significantly. Player A1 played slower in the unmasked attempt, but claimed to do so because she wanted to sound as good as possible. This lead to a much more subtle and dynamic performance, and the timing paid the price for the increase in expressiveness in the performance. This result, and some others, is omitted in the conclusion because they were a consequence of conscious musical choices.

The results suggest that auditory feedback is a contributing factor to timing for some players. The tempo differences for these players are of magnitude 15 %, which is an audible difference. An explanation may be that some players use auditory feedback as confirmative feedback. When sounding a note on a trombone, it is important to know that the sounded note was the right one. This is normally achieved by listening to the note, but when feedback is removed, the player has to rely on sensory information from the lips in order to determine if the note is correct. The use of this ‘secondary’ sensory input may have led to a delay in timing. The reason why other players were not affected could be that they are more used to rely on sensory information, e.g., have more experience with on-stage performance. There could also be psychological reasons as to why this has occurred.

In general, it seems like timing is an aspect that is not greatly affected by auditory feedback deprivation, which is consistent with the findings of Gates and Bradshaw (1974).

4.4.2 RMS energy of prima vista and own piece performance

The results for RMS energy show that all the amateur players played significantly louder in the masked OP performance (mean difference of 2.6 dB) and one of them played significantly louder in the PV performance (difference of 2.8 dB). This effect is audible because the JND is about 1 dB for sound intensity.

Three of the professionals reported to deliberately adjust their performance in order to avoid auditory feedback. The RMS energy results for these professional players (P2 and P4 in OP, P2 and P3 in PV) are therefore not a consequence of auditory feedback. The reason why only these players made this adjustment is that they claimed to hear their own sound. However, in the interview that preceded the recordings, they all confirmed that they did in fact *not* hear themselves, but misinterpreted the tactile sensory information from the lips as sound.

The results from the amateur players indicate a strong correlation between dynamic level and auditory feedback. A natural explanation is that when the auditory feedback is removed, the natural thing to do is to play louder in order to hear. The fact that the differences for OP were larger than for PV could suggest that auditory feedback has a different affect on learning than on retrieval conditions, and that the effect is strongest for the retrieval situation. This would be in contradiction to the findings of Finney and Palmer (2003). However, a more likely explanation for this occurrence is that the variances were smaller in the OP performances, as is evident from figure 37, figure 38 and table 3, yielding a smaller standard deviation. Thus it cannot be concluded that the participants were more affected in the OP performance.

4.4.3 Manual inspection of prima vista and own piece

The results from the manual inspection lack statistical robustness, and should be extended by performing listening tests or a detailed acoustical analysis of the performances. However, some tendencies are present, and the results suggest that there are grounds for further research on the area. Supporting the findings of Finney and Palmer (2003), it seems that absence of auditory feedback affects learning (PV) more than it affects retrieval of information (OP). Extending the findings of Repp (1999), auditory feedback deprivation has a minor disruptive effect on expressiveness in performance. It seems like this effect may be larger for trombone performance than for piano performance. If a pianist doesn't hear himself, he still knows that he produces the correct note by pressing the correct key. In brass instruments, however, there are a number of available resonance modes for each finger setting or slide position, demanding a more preset embouchure configuration from the player. As a consequence, brass instrument players can hit a wrong note or play out of tune even though the fingering or slide position is correct. This introduces another dimension of possible errors and yields that brass instrument performance is more dependent on auditory feedback.

In the PV performance, the effects of auditory feedback deprivation were slightly bigger for some players. This is in agreement with Finney and Palmer (2003), who concluded with a greater dependency on auditory feedback upon learning. However, the performances of professionals and amateurs may be different, not because they were differently affected by the absence of auditory feedback, but because the *conditions* for them were different. Some musicians were observed to subtly hum the melody before they began to play, and they were all asked if they could be able to sing the melody from the score without rehearsal. The professional players and some of the amateurs (A1) answered yes to this question. Therefore, it may seem plausible that these players *did not* learn the piece as they were playing, but already knew how it sounded. The playing condition for these cases would therefore be the same as for the OP, that is, music performed from memory, with a mental audio representation present (the inner ear). In other words, some players were learning the piece as they played, and some "already knew it." This would lead to the same conclusion as Finney and Palmer (2003); that auditory feedback deprivation has en effect on learning, but not retrieval.

4.4.4 Individual differences

Recent work has explored the role of three theoretically motivated factors that contribute to individual differences (Benitez, 2005). Music perception and memory ability, intelligence and the players deliberate actions (from the "Big 5" personality inventory (Costa and McCrae, 1992)). The role of individual differences in performance might be larger than previously suggested, and the results must therefore take into account the variable personal attributes of the player in a thorough discussion. Some performances may have been affected in a psychological way i.e. player P3 became more and more unsure of his performance in the last takes, suggesting that he was confused after not receiving any reassurance from auditory feedback. This is not a result of him not remembering the piece, because the two first takes were correctly played. Other results from various performers show the same tendency. An explanation of this could be a fatigue effect in performance (poorer performance in later attempts), as discussed by Tro (2000).

4.4.5 Sustained notes

As figure 39 shows, the players had no difficulties maintaining a stable note with masking noise present. This suggests that to obtain a stable oscillation regime, a performer is not dependent on auditory feedback. However, 5 out of 7 players held the note longer with the masking noise present. This is not in any way a conclusive result, but may suggest, in addition to the results from PV and OP, that timing is slightly affected by auditory feedback. More research, with a larger amount of data, is needed to prove such claims. The feedback influence model in figure 30 suggests a relationship between auditory feedback and tone stability. This assumption seems to be incorrect, and this is further supported by the overall results. The players rarely had trouble producing stable tones.

4.4.6 Pitch bending

The participants seemed to have limited control over the produced fine pitch, suggesting that the feedback provided from the instrument (the returning pressure wave contributing to the oscillation of the lips) aids in telling the performer which resonating mode he or she is hitting, but not any information about fine pitch. The result is that almost all participants bended the note more than what was instructed, and the results are larger than the JND in pitch for complex tones, hence the effect is very much audible. The player in figure 40 is a professional jazz trombonist with a great deal of experience playing on stage with bad monitoring and acoustically difficult situations. The results from the other players (see appendix) support this finding. It should be mentioned that pitch bending is easiest on the lower resonant modes of the instrument. Even beginners are able to perform pitch bending in this frequency range. Pitch bending for higher modes would have resulted in smaller variations, and this is also the reason why exactly this resonant mode was chosen. This result is relevant for on-stage performance with difficult acoustical environments. It seems like pitch control is one of the areas that are strongly affected by absence of feedback - and musicians must rely on auditory information to play in tune, confirming the assumption in figure 30.

5 Conclusions and further work

Trombone performances have been studied through acoustical methods in order to visualize the mechanics of sound production. The results show high variations in temporal and spectral dynamics, suggesting that human parameters and individual differences are prominent factors in trombone performance. Attack transients and slurred transients are important characteristics in brass instrument sound production, and it is shown that players have different approaches and methods for producing transients. It is of interest to further study the characteristics of slurred transients and attack transients to understand the importance of these, especially in terms of expressiveness in musical performance. The study has attempted to bridge the gap between science and experience, and as a concluding remark, it is of importance to remember to view the matter from different perspectives, as the musicians themselves often have greater insight into the system than the researchers. In the future, a model of trombone sound production with special attention to player parameters should be possible to create.

Auditory feedback deprivation in trombone performance has been studied through music performed from notation and from memory. It is shown that auditory feedback deprivation has an effect on various aspects of trombone performance, but the magnitude of the effects depend on musical skill and individual human parameters. The results revealed significant effects on timing and dynamics for some amateur players, but not for any professional players. The results for the affected players were a mean decrease in tempo of 15 % and a mean increase in dynamic level of 2.3 dB, within a significance level of 0.05. Through manual inspection, it was found that players with less experience were more affected by removal of auditory feedback than professionals. The results suggest that musical skill and experience is a contributive factor to how dependent players are of auditory feedback in trombone performance. A reason for this could be trained musicians have a more developed skill in auditory imagery, i.e., “inner ear.”

Further, it was shown that the effects of auditory feedback deprivation on musical performance from notation tended to be greater than for performance from memory. This is in agreement with Finney and Palmer (2003).

The findings in this report suggest that absence of auditory feedback greatly affects pitch control in trombone performance. It is therefore relevant to further investigate the role this has on other lip-reed instruments where tuning is dependent on vocal tract configuration and embouchure. To further study the effects of auditory feedback deprivation upon brass instrument performance, listening tests should be performed to determine which effects are audible. It would also be of great interest to discuss the results from a psychological point of view, as music performance include cognitive processes.

Bibliography

Agus, T. R., Suied, C., Thorpe, S. J., and Pressnitzer, D. (**2012**). "Fast recognition of musical sounds based on timbre," *The Journal of the Acoustical Society of America* **131**(5), 4124-4133.

Allen, D. R. (**2007**). *Mental representations in clarinet performance: Connections between auditory imagery and motor behaviors*. MSc. University of North Carolina, Greensboro

Backus, J., and Baskerville, R. (**1977**). *The acoustical foundations of music.* New York: W.W. Norton and

Backus, J., and Hundley, T. (**2005**). "Harmonic generation in the trumpet," *The Journal of the Acoustical Society of America* **49**(2B), 509-519.

Banton, L. J. (**1995**). "The role of visual and auditory feedback during the sight-reading of music," *Psychology of music* **23**(1), 3-16.

Beauchamp, J. W. (**1980**) Analysis of simultaneous mouthpiece and output waveforms of wind instruments, *Audio Engineering Society Convention 66.*

Beauchamp, J. W. (**2012**). "Trombone Transfer Functions: Comparison Between Frequency-Swept Sine Wave and Human Performer Input," *Archives of Acoustics* **37**(4), 447-454.

Benade, A. H. (**2012**). *Fundamentals of Musical Acoustics.* New York: Courier Dover Publications.

Benitez, B. (**2005**). *Effects of altered auditory feedback in music performance : the role of impulsivity, musical aptitude, and general intelligence*. MSc. University of Texas, San Antonio

Campbell, M., Myers, A., and Chick, J. (**2013**) Influence of the bell profile of the trombone on sound reflection and radiation, *Proceedings of Meetings on Acoustics.* **19**, 035068. Montréal

Campbell, M. a. G., C. (**1987**). *The musician's guide to acoustics.* New York: Oxford University Press.

Carral, S. (**2005**). *Relationship between the physical parameters of musical wind instruments and the psychoacoustic attributes of the produced sound*. PhD. University of Edinburgh, Edinburgh

Chen, J.-M., Smith, J., and Wolfe, J. (**2012**). "Do trumpet players tune resonances of the vocal tract?," *The Journal of the Acoustical Society of America* **131**(1), 722-727.

Chick, J., Logie, S., Campbell, M., and Gilbert, J. (**2012**) Spectral enrichment and wall losses in trombones played at high dynamic levels, *Acoustics 2012.* Nantes, France

Corso, J. F. (**1963**). "Bone‐Conduction Thresholds for Sonic and Ultrasonic Frequencies," *The Journal of the Acoustical Society of America* **35**(11), 1738-1743.

Costa, P. T., and McCrae, R. R. (**1992**). "Normal personality assessment in clinical practice: the NEO Personality Inventory," *Psychological assessment* **4**(1), 5-13.

De Winter, J. (**2013**). "Using the Student's t-test with extremely small sample sizes," *Practical Assessment, Research & Evaluation* **18**(10).

Ebihara, T., and Yoshikawa, S. (**2013a**). "Nonlinear effects contributing to hand-stopping tones in a horn," *The Journal of the Acoustical Society of America* **133**(5), 3094-3106.

Ebihara, T., and Yoshikawa, S. (**2013b**) Nonlinear wall vibration and wave steepening contributing to tonal metallicness and brassiness in a horn, *Proceedings of Meetings on Acoustics.* **19**, 035071. Montréal

Elliott, S., and Bowsher, J. (**1982**). "Regeneration in brass wind instruments," *Journal of Sound and Vibration* **83**(2), 181-217.

Finney, S., and Palmer, C. (**2003**). "Auditory feedback and memory for music performance: Sound evidence for an encoding effect," *Memory & Cognition* **31**(1), 51-64.

Finney, S. A. (**1997**). "Auditory Feedback and Musical Keyboard Performance," *Music Perception: An Interdisciplinary Journal* **15**(2), 153-174.

Fletcher, N. H. (**1999**). "The nonlinear physics of musical instruments," *Reports on Progress in Physics* **62**(5), 723-764.

Fletcher, N. H., and Rossing, T. D. (**1998**). *The physics of musical instruments.* New York: Springer Verlag.

Fox, F. (**1982**). *Essentials of Brass Playing: An Explicit, Logical Approach to Important Basic Factors that Contribute to Superior Brass Instrument Performance.* Van Nuys: Alfred Music Publishing.

Gates, A., and Bradshaw, J. L. (**1974**). "Effects of auditory feedback on a musical performance task," *Perception & Psychophysics* **16**(1), 105-109.

Gilbert, J., Dalmont, J.-P., and Potier, R. (**2013**) Does the elephant trumpet like a trumpet?, *20th International Congress on Acoustics (ICA 2010).* Sydney, Australia

Hirschberg, A., Gilbert, J., Msallam, R., and Wijnands, A. (**1996**). "Shock waves in trombones," *The Journal of the Acoustical Society of America* **99**1754.

Hoekje, P. (**2013**) Comparing steady-state and transient phenomena in brass instruments, *Proceedings of Meetings on Acoustics.* **19**, 035070. Montréal

Kaburagi, T., Yamada, N., Fukui, T., and Minamiya, E. (**2011**). "A methodological and preliminary study on the acoustic effect of a trumpet player's vocal tract," *The Journal of the Acoustical Society of America* **130**(1), 536-545.

Kemp, J. A., and Smith, R. A. (**2013**) Modeling pulse-like lip vibrations in brass instruments, *Proceedings of Meetings on Acoustics.* **19**, 035055. Montréal

Kleinhammer, E. (**1963**). *The art of trombone playing.* Miami: Alfred Music Publishing.

Lartilliot, O., Toiviainen, Petri., Eerola, Tuomas (**2008**). "A Matlab Toolbox for Music Information Retrieval," in *Data Analysis, Machine Learning and Applications*. edited by C. Preisach, Burkhardt, L., Schmidt-Thime, L., Decker, R (Springer Verlag).

Lashley, K. (**1951**). "The problem of serial order in behavior," *Cerebral Mechanisms in Behaviour*(October), 112-135.

Logie, S. (**2012**). *An Acoustical Study of the Playing Characteristics of Brass Wind Instruments*. PhD. University of Edinburgh, Edinburgh

Luce, D., and Clark, M. (**1967**). "Physical Correlates of Brass - Instrument Tones," *The Journal of the Acoustical Society of America* **42**(6), 1232-1243.

Muramatsu, M., Kurosawa, J., Oikawa, Y., and Yamasaki, Y. (**2013**) Communication aid utilizing bone-conducted sound via teeth by means of mouthpiece form actuator, *Proceedings of Meetings on Acoustics.* **19**, 050090. Montréal

Myers, A., Robert W. Pyle, J., Gilbert, J., Campbell, D. M., Chick, J. P., and Logie, S. (**2012**). "Effects of nonlinear sound propagation on the characteristic timbres of brass instruments," *The Journal of the Acoustical Society of America* **131**(1), 678-688.

Norman, L., Chick, J., Campbell, D., Myers, A., and Gilbert, J. (**2010**). "Player Control of'Brassiness' at Intermediate Dynamic Levels in Brass Instruments," *Acta Acustica united with Acustica* **96**(4), 614-621.

Pyle, R. W. (**1981**). "The effect of surface coatings upon brass ‐ instrument tone quality," *The Journal of the Acoustical Society of America* **69**(S1), S37-S37.

Pyle, R. W. (**1998**). "The effect of wall materials on the timbre of brass instruments," *The Journal of the Acoustical Society of America* **103**(5), 2834-2834.

Repp, B. H. (**1999**). "Effects of auditory feedback deprivation on expressive piano performance," *Music Perception* **16**(4), 409-438.

Smith, R. A. (**1978**). "Recent development in trumpet design," *International Trumpet Guild* **42**(10).

Stenfelt, S. (**2011**). "Acoustic and physiologic aspects of bone conduction hearing," *Advances in Oto-Rhino-Laryngology* **71**10-21.

Stenfelt, S., Hato, N., and Goode, R. L. (**2002**). "Factors contributing to bone conduction: The middle ear," *The Journal of the Acoustical Society of America* **111**(2), 947-959.

Stenfelt, S., Wild, T., Hato, N., and Goode, R. L. (**2003**). "Factors contributing to bone conduction: The outer ear," *The Journal of the Acoustical Society of America* **113**(2), 902-913.

Tro, J. (**2000**) Aspects of control and perception, *Proceedings of the 3rd COST-G6 Workshop on Digital Audio Effects (DAFx00).* 171-176. Verona, Italy

v. Békésy, G. (**1949**). "The Structure of the Middle Ear and the Hearing of One's Own Voice by Bone Conduction," *The Journal of the Acoustical Society of America* **21**(3), 217-232.

Wallach, H., Newman, E. B., and Rosenzweig, M. R. (**1949**). "A Precedence Effect in Sound Localization," *The Journal of the Acoustical Society of America* **21**(4), 468-468.

Webster, J. C. (**1949**). "Internal Tuning Differences due to Players and the Taper of Trumpet Bells," *The Journal of the Acoustical Society of America* **21**(3), 208-214.

Wolfe, J., Tarnopolsky, A., Fletcher, N., Hollenberg, L., and Smith, J. (**2003**) Some effects of the player's vocal tract and tongue on wind instrument sound, *Proc. Stockholm Music Acoust. Conf., SMAC-03.* 307-310. Stockholm, Sweden

Zahorian, S. A., and Hu, H. (**2008**). "A spectral/temporal method for robust fundamental frequency tracking," *The Journal of the Acoustical Society of America* **123**(6), 4559-4571.

A Additional results

A.1 Results from Part I

A.1.1 Sustained notes

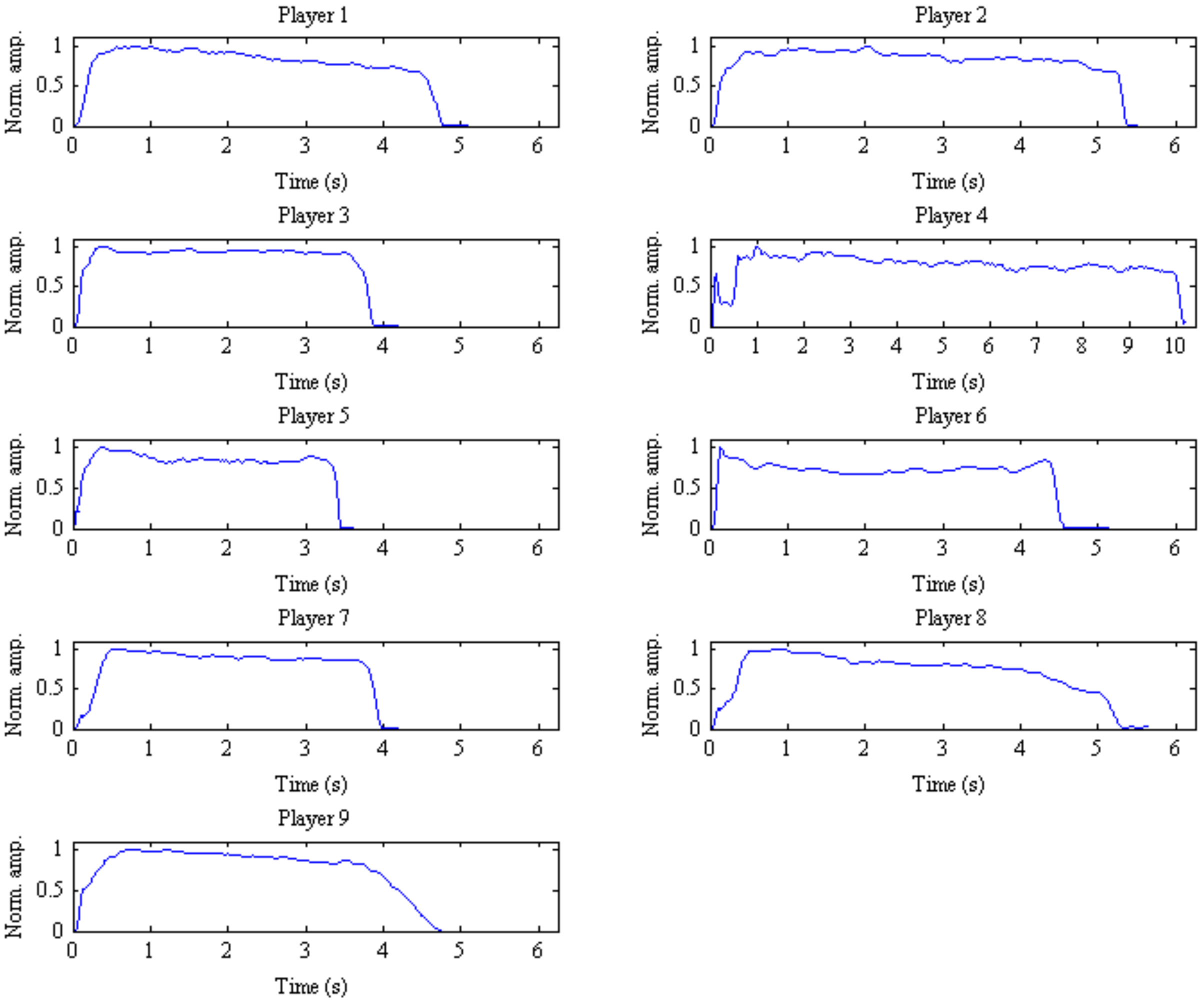
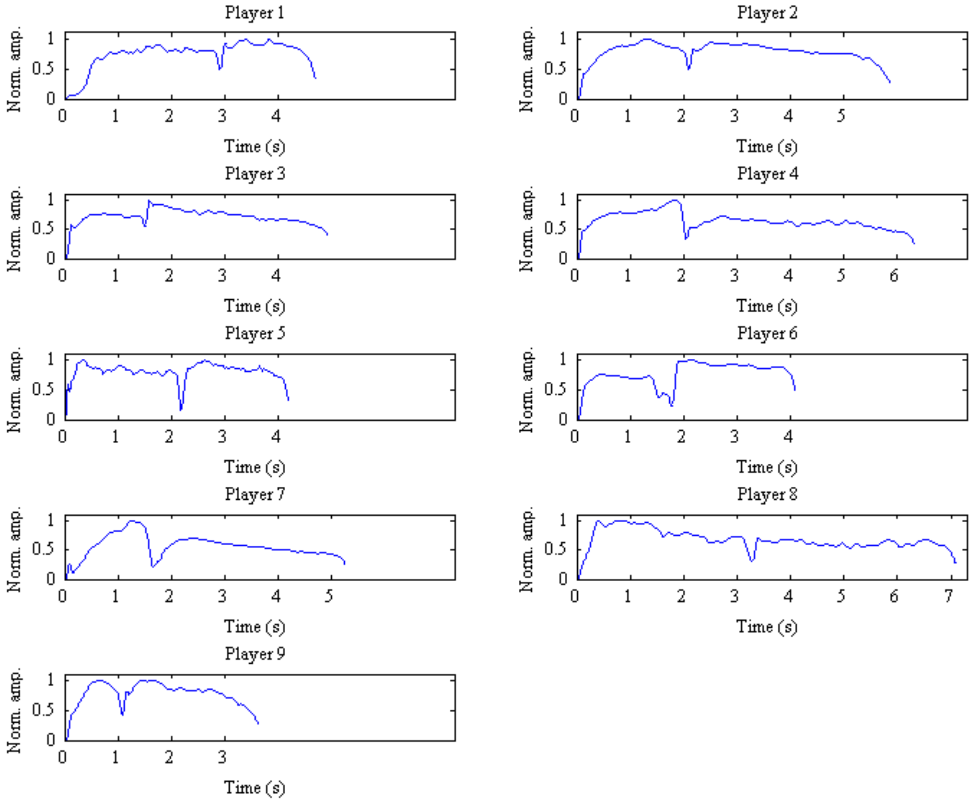
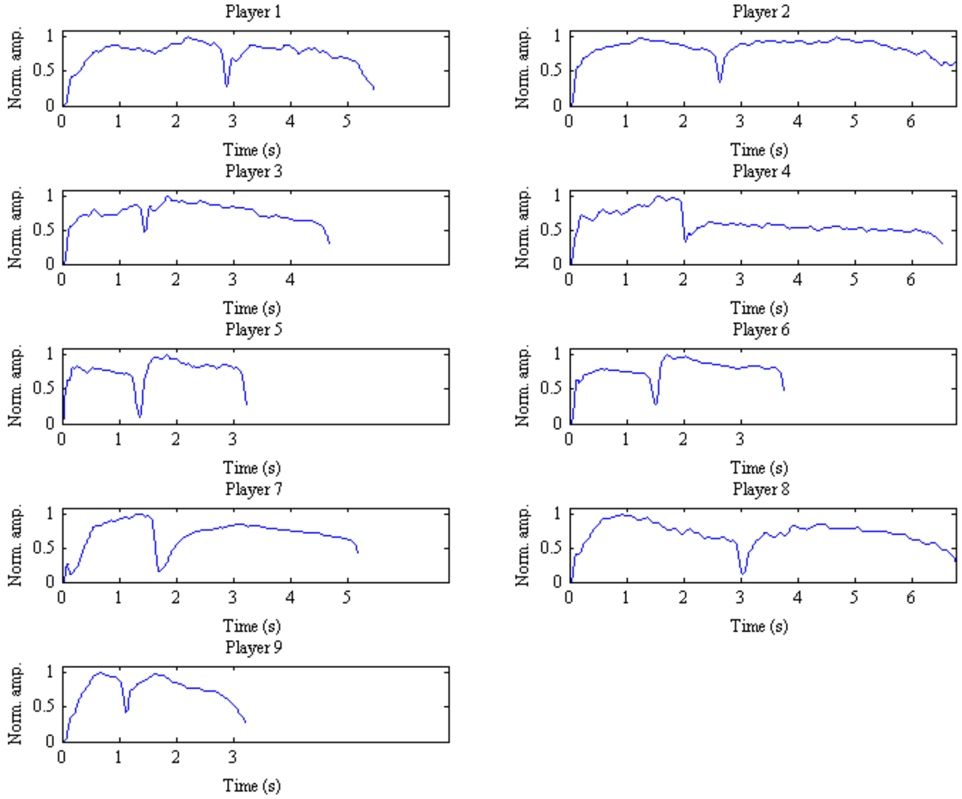


Figure 41: Envelope of sustained F2-note with fundamental frequency 349.2 Hz. *(mirenvelope)*.

A.1.2 Slurred transients





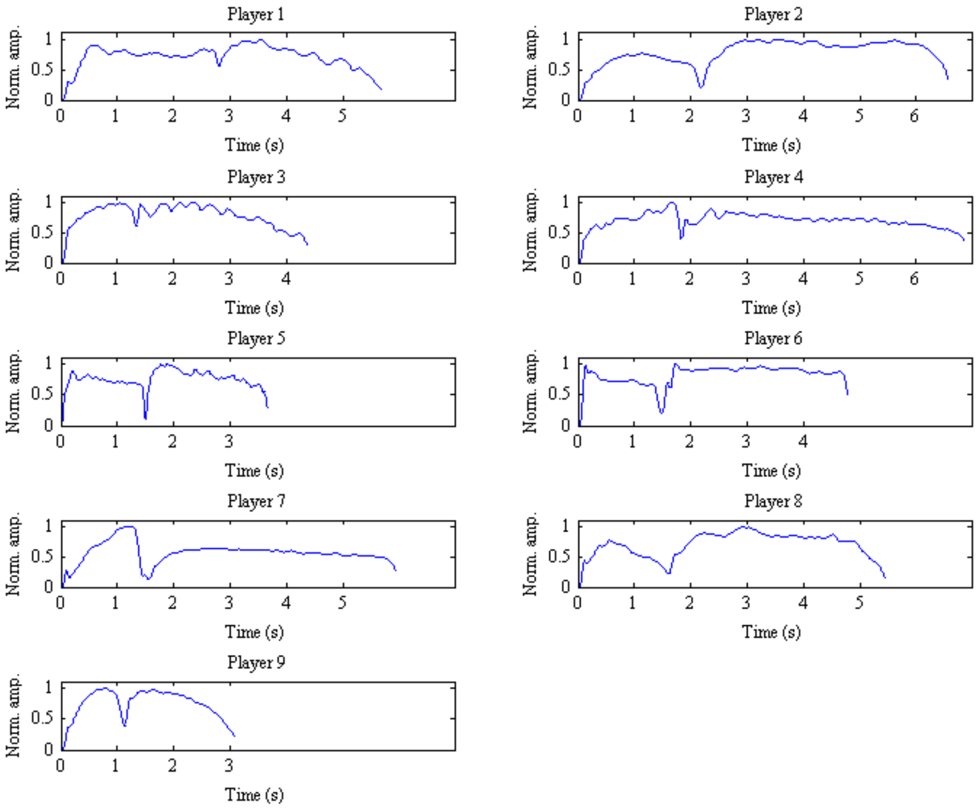


Figure 42: Envelope of slurred transients on three different positions. The slurring was upwards slur between adjacent resonance modes. *(mirenvelope)*.

Top: 1st position, Bb2 to D2 with fundamental frequencies 233.1 Hz and 293,6 Hz

Middle: 4th position, A1 to Db2 with fundamental frequencies 220,0 Hz and 277.2 Hz

Bottom: 7th position, E1 to Ab1 with fundamental frequencies 164.8 Hz and 207.6 Hz

A.1.3 Attack transients

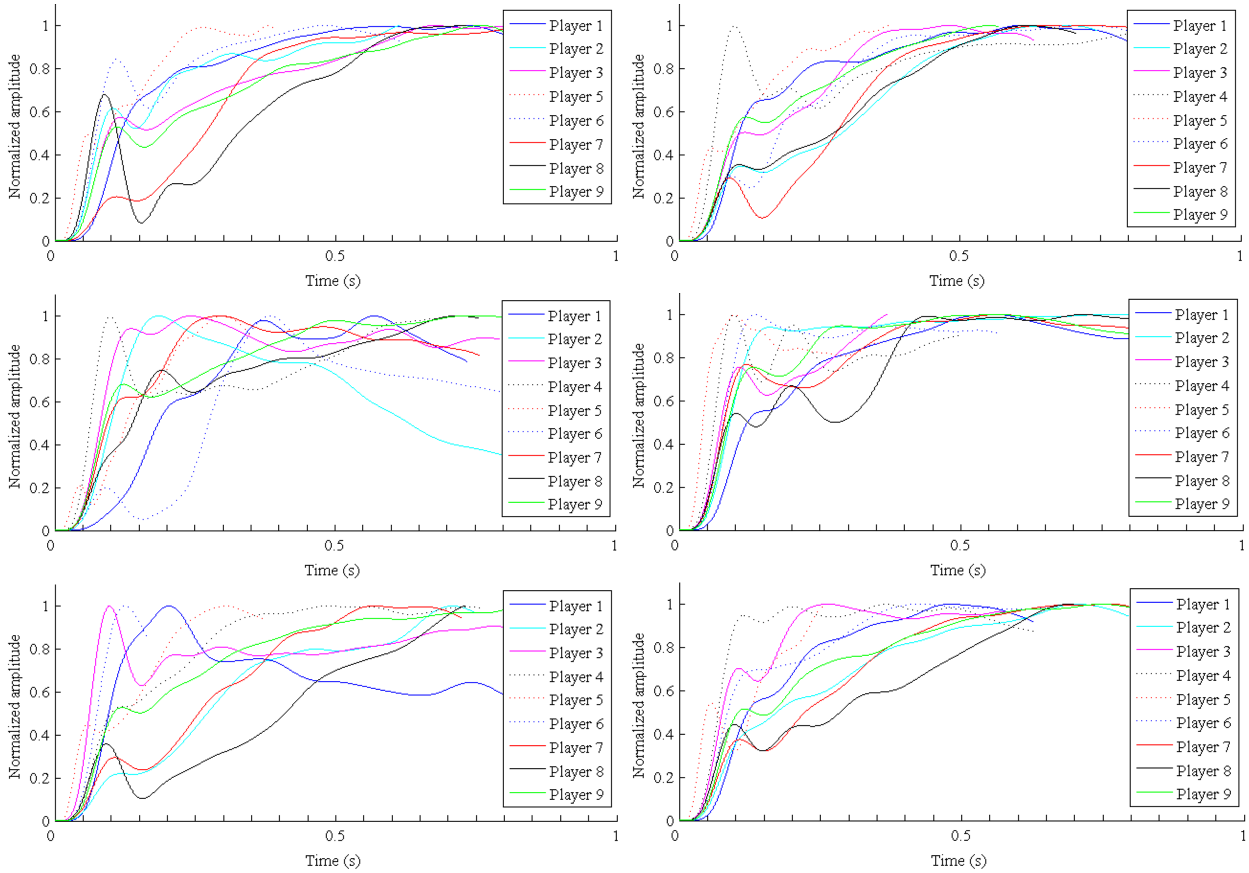
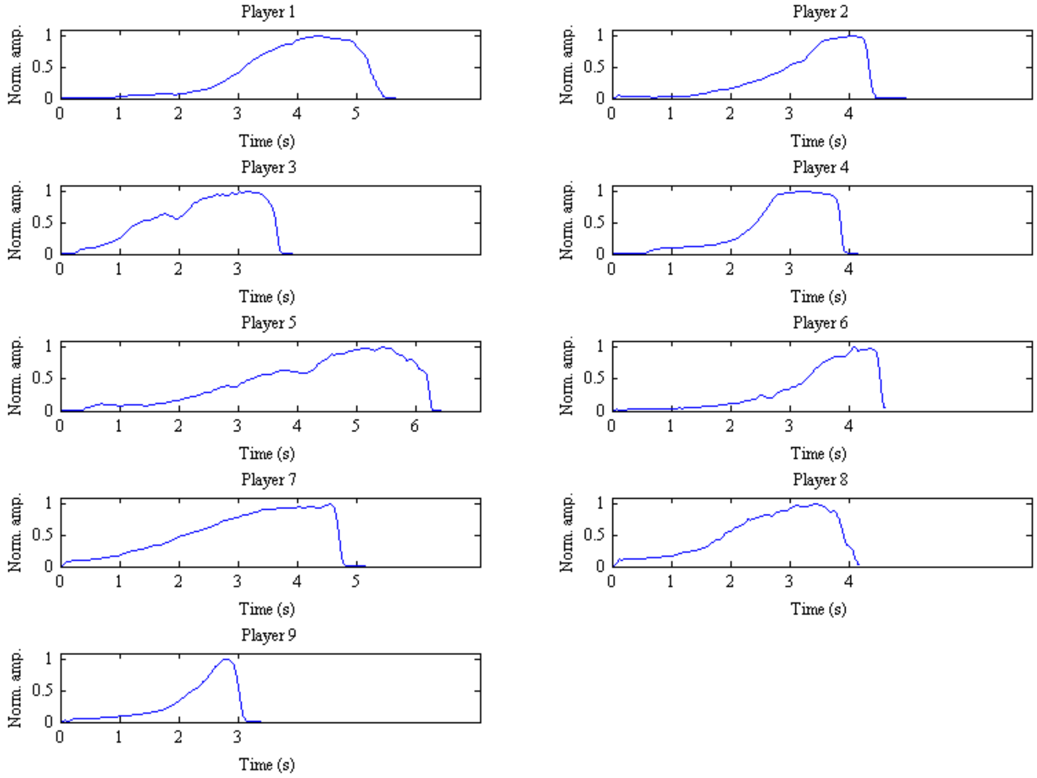
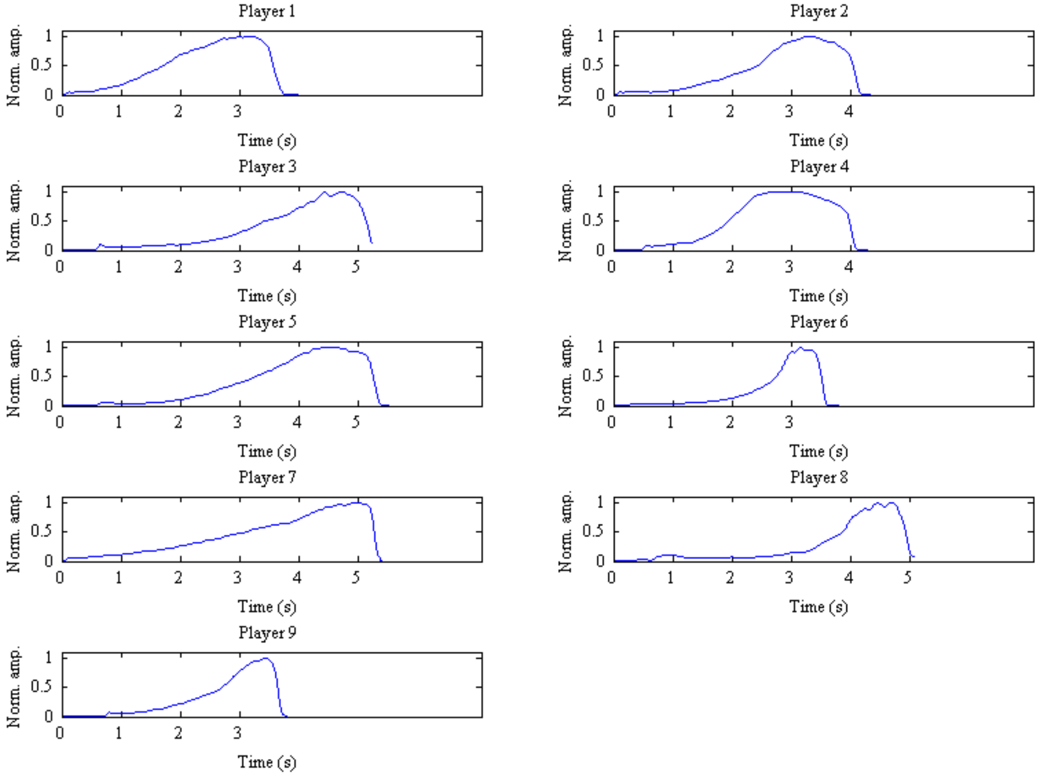


Figure 43: Attack envelopes: Top: *tu-attacks*, Middle: *ku-attacks*, Bottom: *No tongue-attacks*. Left column: Bb1 with fundamental frequency 116.5 Hz. Right column: Bb2 with fundamental frequency 233.1 Hz. (*mirenvelope)*

A.1.4 Crescendo and decrescendo

Figures 42 – 44 show envelopes of crescendos and decrescendos





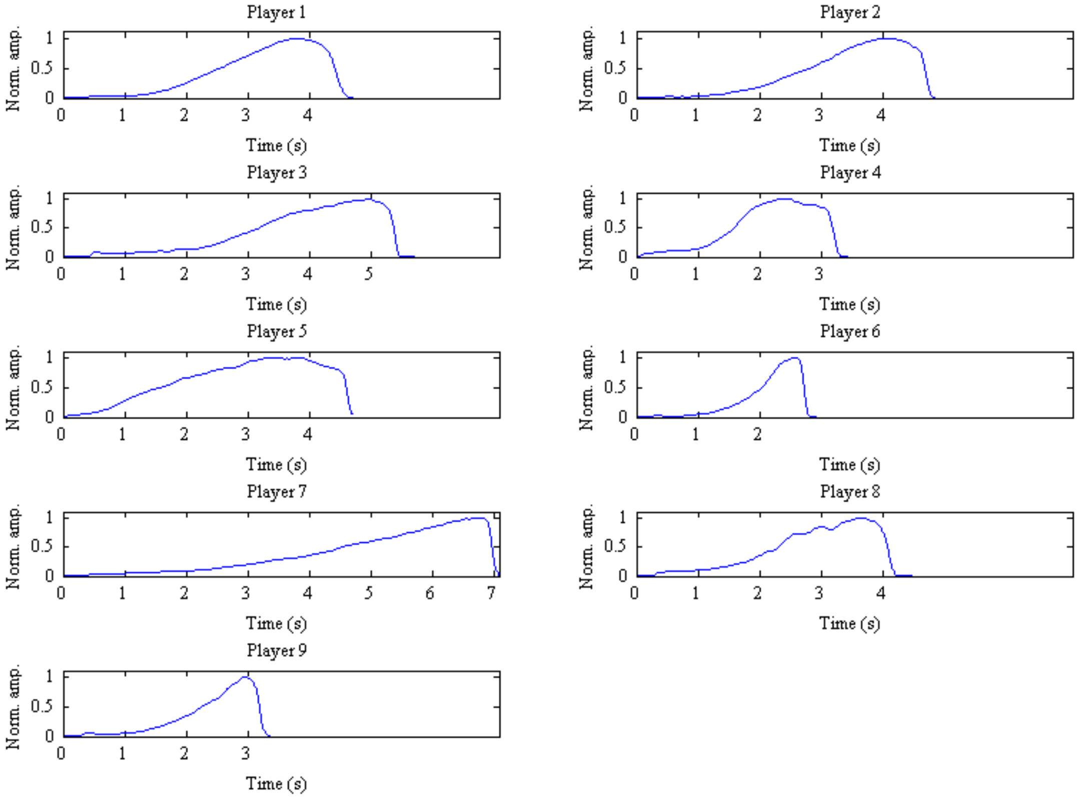
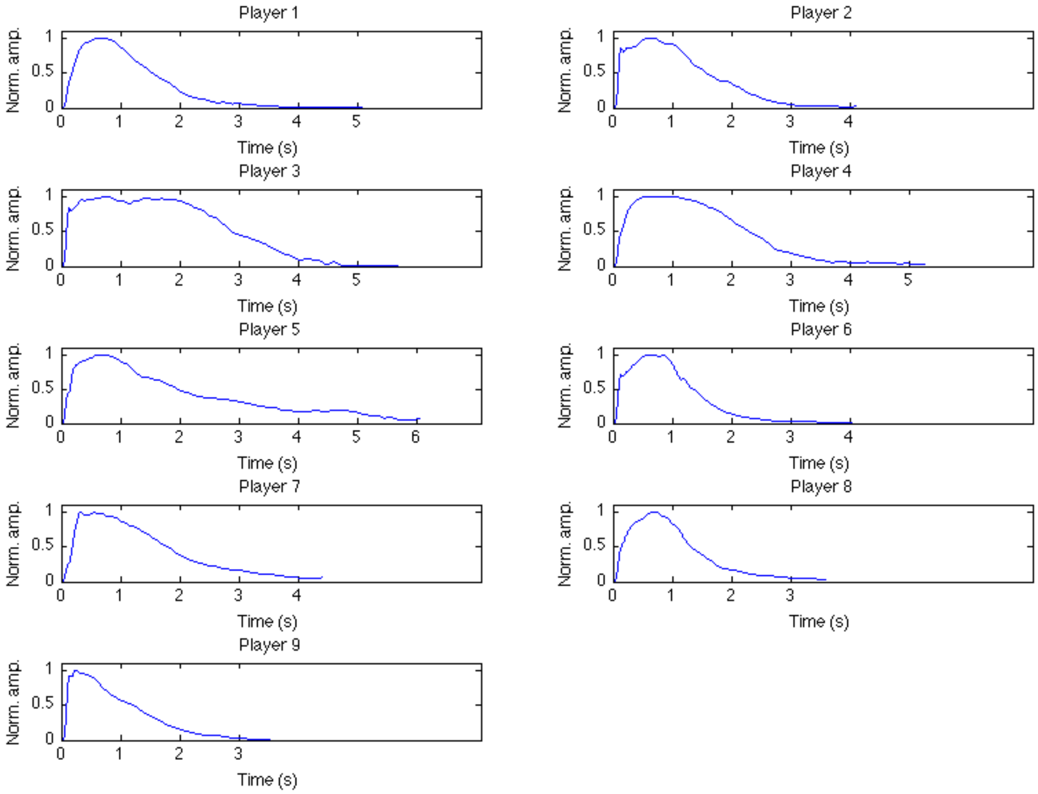
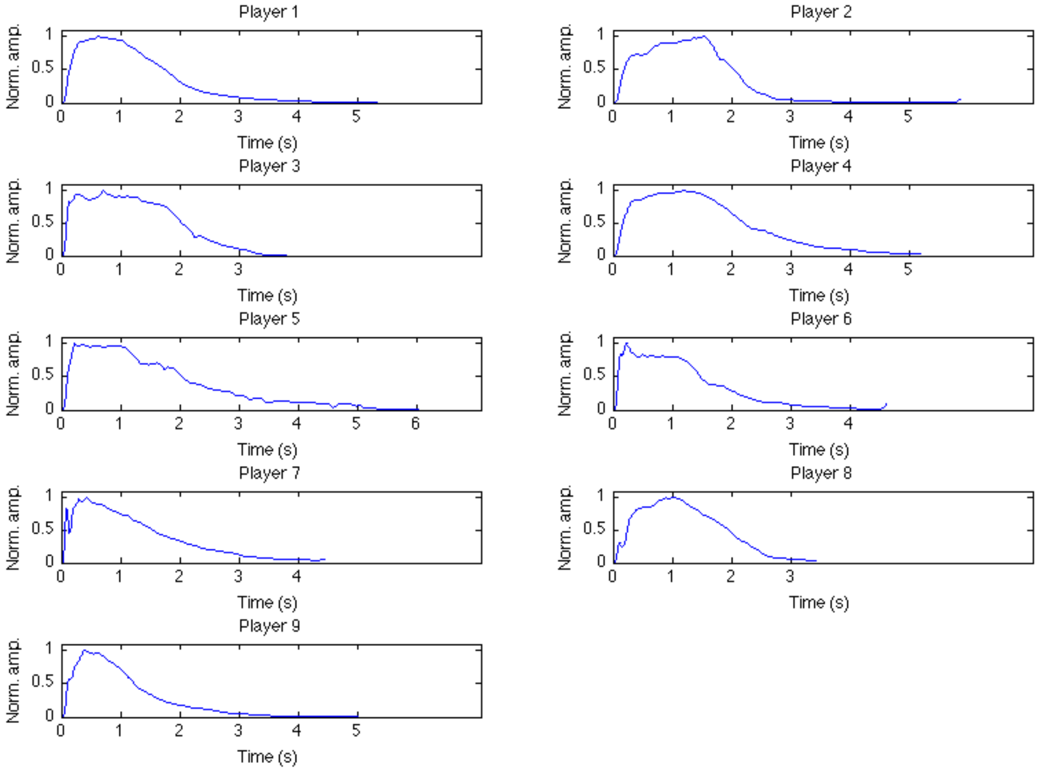


Figure 44: Crescendo envelopes. Top: Bb1, fundamental frequency 116.5 Hz, Middle: F3, fundamental frequency 174.6 Hz, Bottom: Bb2, fundamental frequency 233.1 Hz. (*mirenvelope*)



Player 1
Player 2
Player 3
Player 4
Player 5
Player 6
Player 7
Player 8
Player 9
Norm. amp.
Time (s)



Player 1
Player 2
Player 3
Player 4
Player 5
Player 6
Player 7
Player 8
Player 9
Norm. amp.
Time (s)

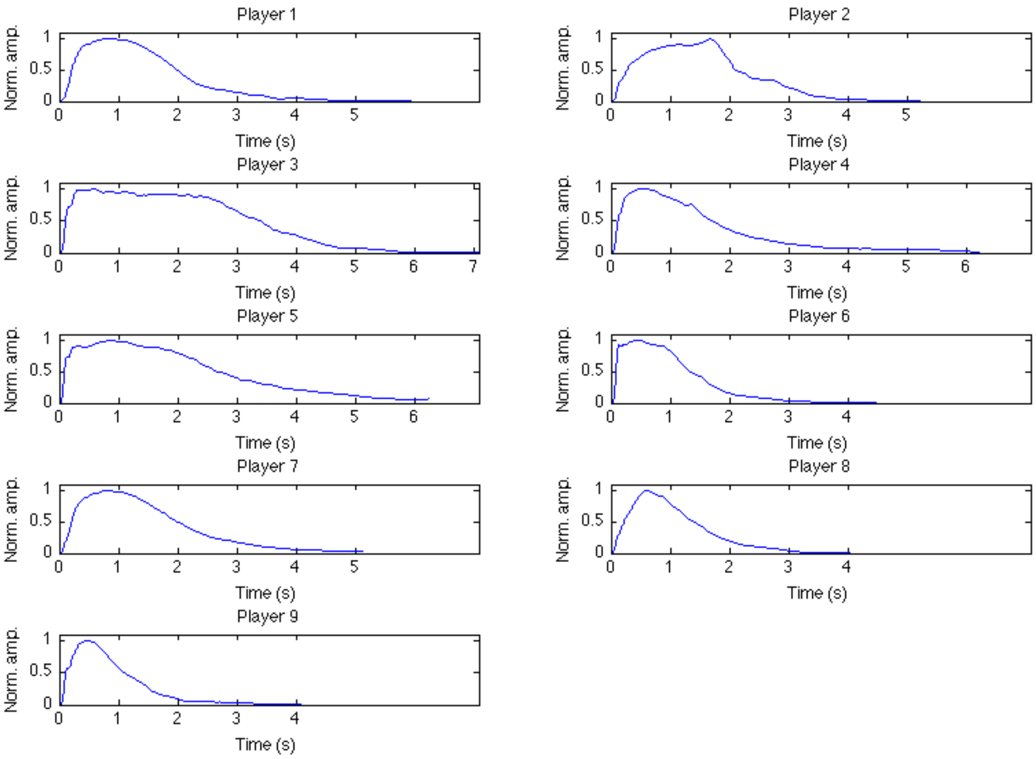


Figure 45: Decrescendo envelopes. Top: Bb, fundamental frequency 116.5 Hz, Middle: F fundamental frequency 174.6 Hz, Bottom: Bb2, fundamental frequency 233.1 Hz. (*mirenvelope*)

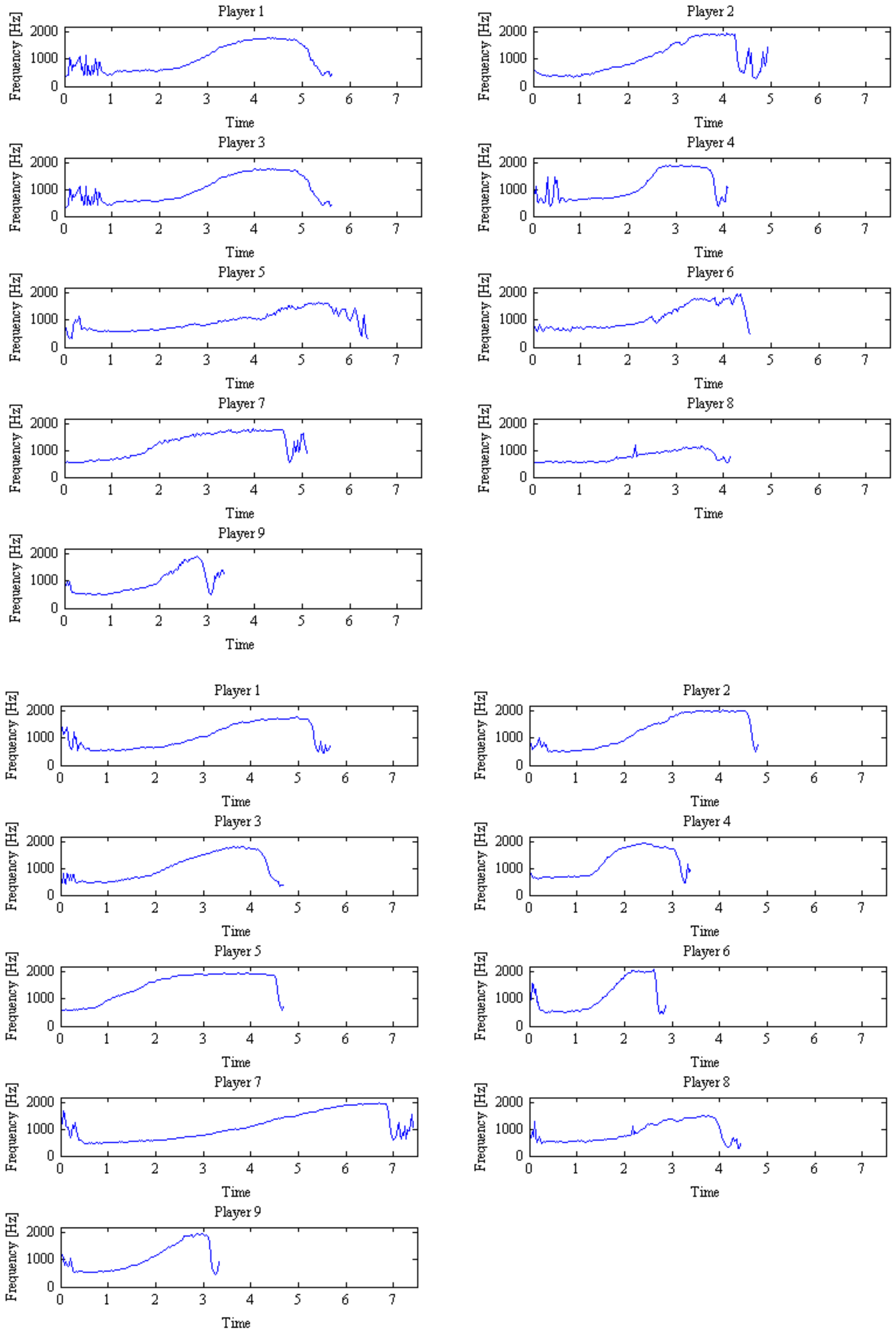


Figure 46: Spectral centroid of crescendos. Top: Bb, fundamental frequency 116.5 Hz, Bottom: Bb2, fundamental frequency 233.1 Hz. (*mircentroid)*

A.2 Results from Part II

A.2.1 Table of RMS energy and duration

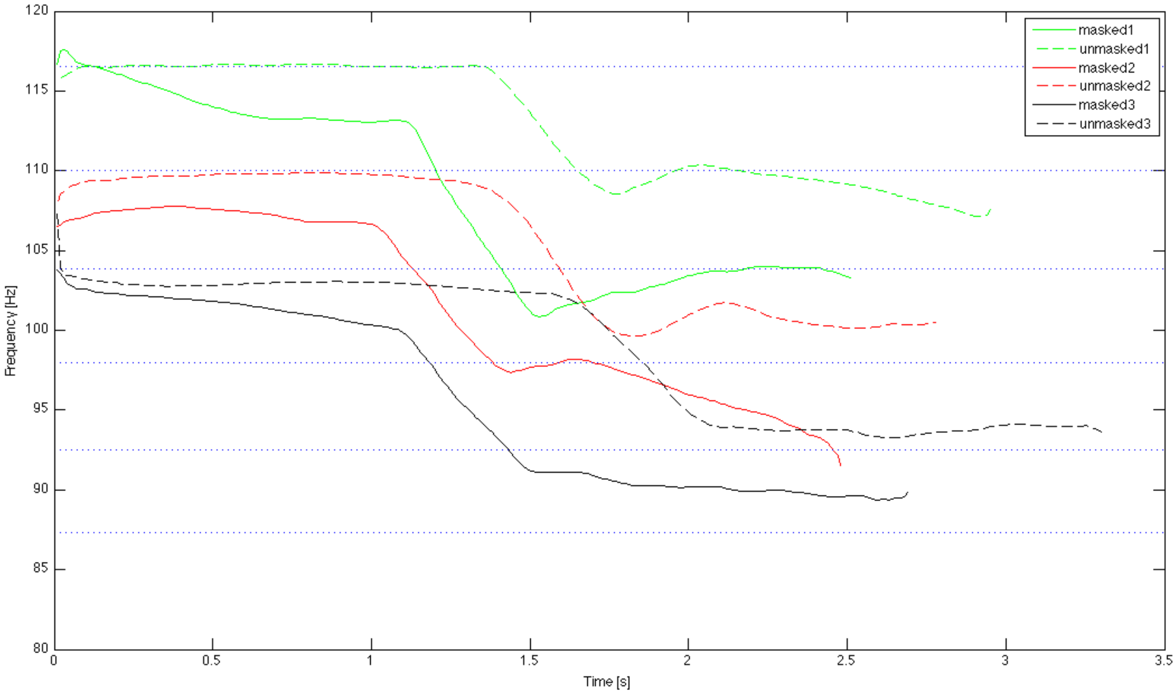
Table 3: Results from feedback experiment.

🟩	**Prima Vista**	MIRlength:	Length of signal calculated in *MIRtoolbox*
🟥	**Own Piece**	RMS energy (dB):	Root-mean-squared energy of the signal
		Effect size:	No. of std’s between masked and unmasked.

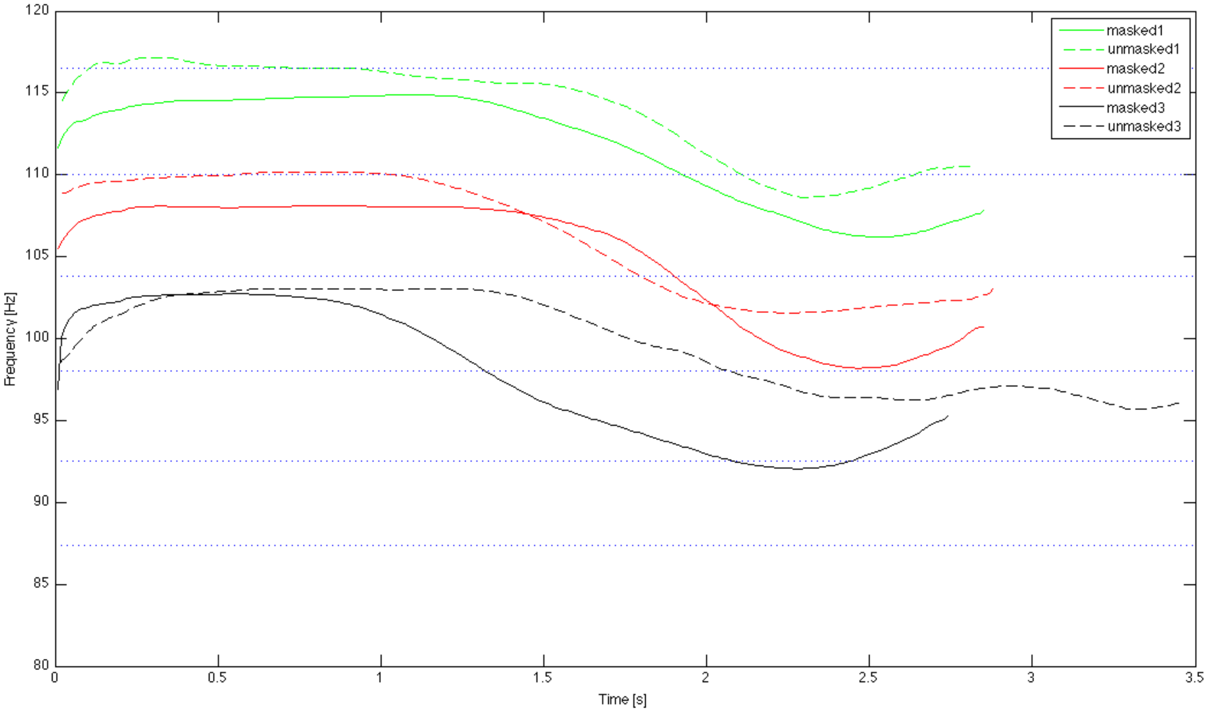
A1	**1**	**2**	**3**	**4**	**5**	**Avg**	**Std**	**Ref**	**Effect size**
RMS energy (dB)	94.2	93.5	94.3	94.1	94.6	94.1	0.37	92.4	-4.7
MIRlength (s)	26.8	27.3	26.9	27.0	27.3	27.1	0.21	28.8	8.5
RMS energy (dB)	91.9	92.4	93.3	92.6		92.5	0.49	92.2	-0.7
MIRlength (s)	35.4	33.7	34.8	35.3		34.8	0.67	35.9	1.6
P3	**1**	**2**	**3**	**4**	**5**	**Avg**	**Std**	**Ref**	**Effect size**
RMS energy (dB)	81.5	85.2	85.7	86.5	87.6	85.5	2.05	83.1	-1.2
MIRlength (s)	30.8	27.0	25.5	27.1	24.4	27.0	2.16	25.0	-0.9
RMS energy (dB)	91.9	92.4	93.3	92.6		92.5	0.49	92.2	-0.7
MIRlength (s)	26.4	24.9	25.2	25.7		25.6	0.57	25.9	0.5
P1	**1**	**2**	**3**	**4**	**5**	**Avg**	**Std**	**Ref**	**Effect size**
RMS energy (dB)	82.6	80.1	79.0	79.0	76.7	79.7	1.91	79.4	-0.2
MIRlength (s)	24.6	25.3	23.9	23.4	23.7	24.2	0.67	24.9	1.1
RMS energy (dB)	86.5	87.5	81.8	84.5	84.3	85.1	1.96	87.4	1.1
MIRlength (s)	30.5	28.6	28.8	30.9	29.1	29.6	0.95	27.9	-1.8
A3	**1**	**2**	**3**	**4**	**5**	**Avg**	**Std**	**Ref**	**Effect size**
RMS energy (dB)	93.3	93.9	93.3	93.9	93.3	93.6	0.29	90.8	-9.5
MIRlength (s)	19.7	19.4	19.7	18.9	18.8	19.3	0.40	18.8	-1.2
RMS energy (dB)	92.0	93.9	93.3	95.0	93.9	93.7	0.99	92.7	-1.0
MIRlength (s)	31.0	30.7	29.4	29.3	27.6	29.6	1.21	25.1	-3.7
P4	**1**	**2**	**3**	**4**	**5**	**Avg**	**Std**	**Ref**	**Effect size**
RMS energy (dB)	90.5	87.3	87.5	88.3	87.9	88.4	1.17	91.0	2.2
MIRlength (s)	21.3	21.7	22.1	22.6	22.7	22.1	0.52	22.0	-0.2
RMS energy (dB)	90.5	90.3	88.3	90.7	92.4	90.6	1.29	90.6	0.0
MIRlength (s)	43.2	47.0	46.2	46.1	45.3	45.6	1.29	45.1	-0.3
P2	**1**	**2**	**3**	**4**	**5**	**Avg**	**Std**	**Ref**	**Effect size**
RMS energy (dB)	86.2	85.6	84.4	84.8	84.7	85.2	0.64	86.3	1.8
MIRlength (s)	15.5	15.3	15.4	16.2	16.3	15.7	0.43	16.3	1.3
MIRlowenergy	0.54	0.53	0.50	0.52	0.52	0.53	0.01	0.52	-0.6
RMS energy (dB)	87.6	87.6	88.8	87.6	88.4	88.0	0.52	89.2	2.4

MIRrms	0.07	0.07	0.08	0.07	0.08	0.08	0.00	0.09	2.5
MIRlength (s)	30.8	28.1	27.6	28.2	29.7	28.9	1.19	28.6	-0.3
A4	**1**	**2**	**3**	**4**	**5**	**Avg**	**Std**	**Ref**	**Effect size**
RMS energy (dB)	84.7	84.4	84.3	84.0	84.1	84.3	0.24	81.0	-14.0
MIRlength (s)	13.5	13.5	13.5	13.5	13.5	13.5	0.01	10.2	-341.6
RMS energy (dB)	89.4	88.6	87.2	88.7	89.2	88.7	0.77	88.1	-0.7
MIRlength (s)	26.7	26.1	25.7	25.5	24.8	25.8	0.65	25.0	-1.2
A2	**1**	**2**	**3**	**4**	**5**	**Avg**	**Std**	**Ref**	**Effect size**
RMS energy (dB)	90.8	90.3	90.8	90.5	89.9	90.5	0.35	87.7	-7.8
MIRlength (s)	13.6	13.7	13.6	13.7	13.8	13.7	0.08	12.2	-17.9
RMS energy (dB)	92.0	91.4	90.2	89.4		90.8	1.03	88.7	-2.1
MIRlength (s)	33.1	32.1	31.7	30.8		31.9	0.83	27.6	-5.2

A.2.2 Pitch bending for all participants

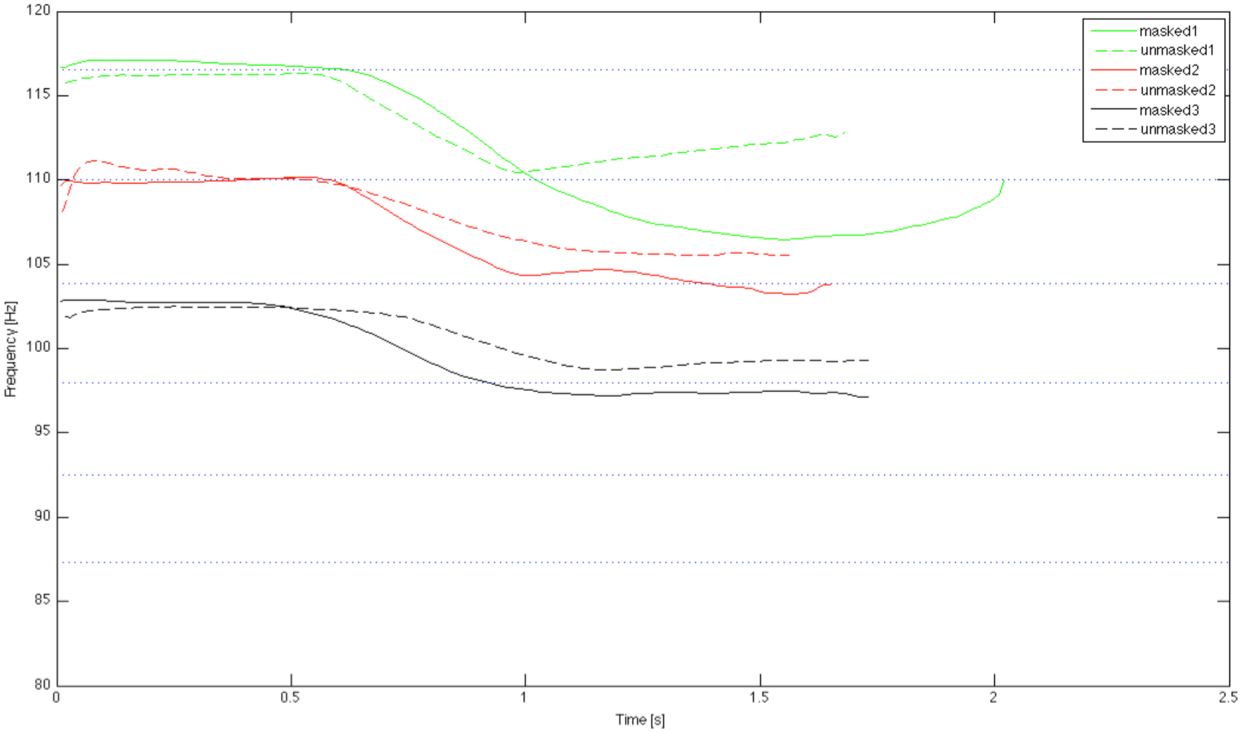


(a) P1



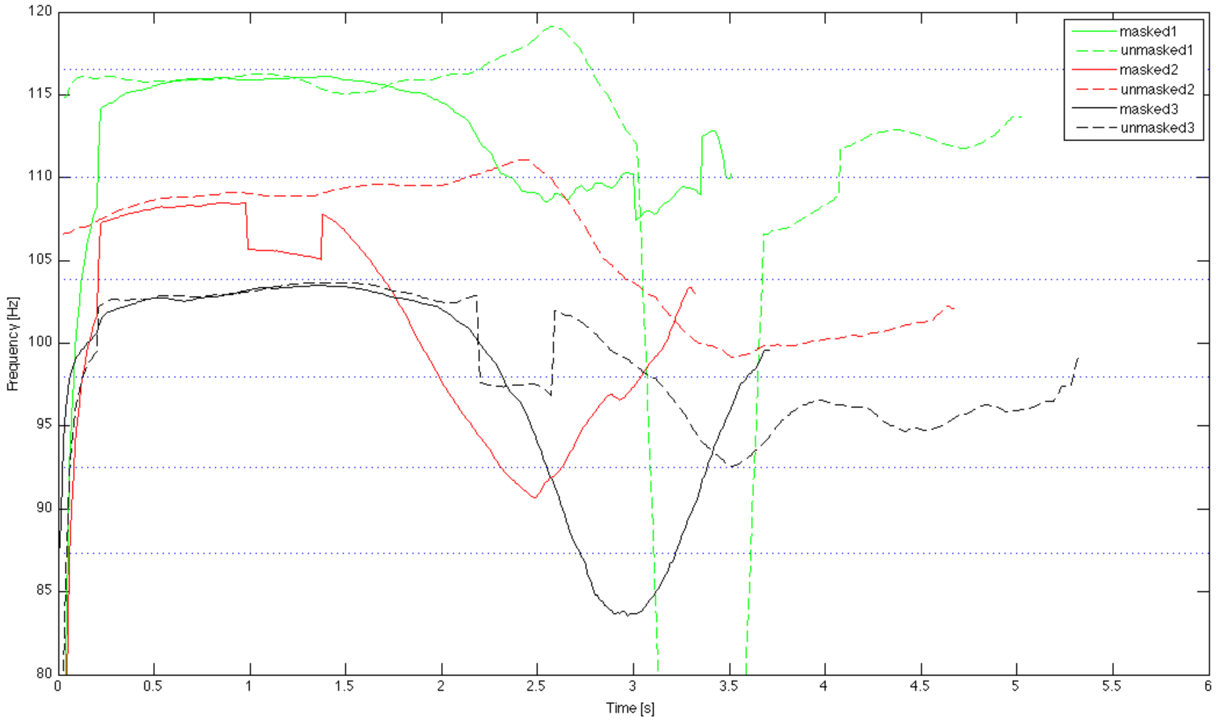
Frequency [Hz]
Time [s]
120
115
110
105
100
95
90
85
80
0
0.5
1
1.5
2
2.5
3
3.5
masked1
unmasked1
masked2
unmasked2
masked3
unmasked3

(b) P3

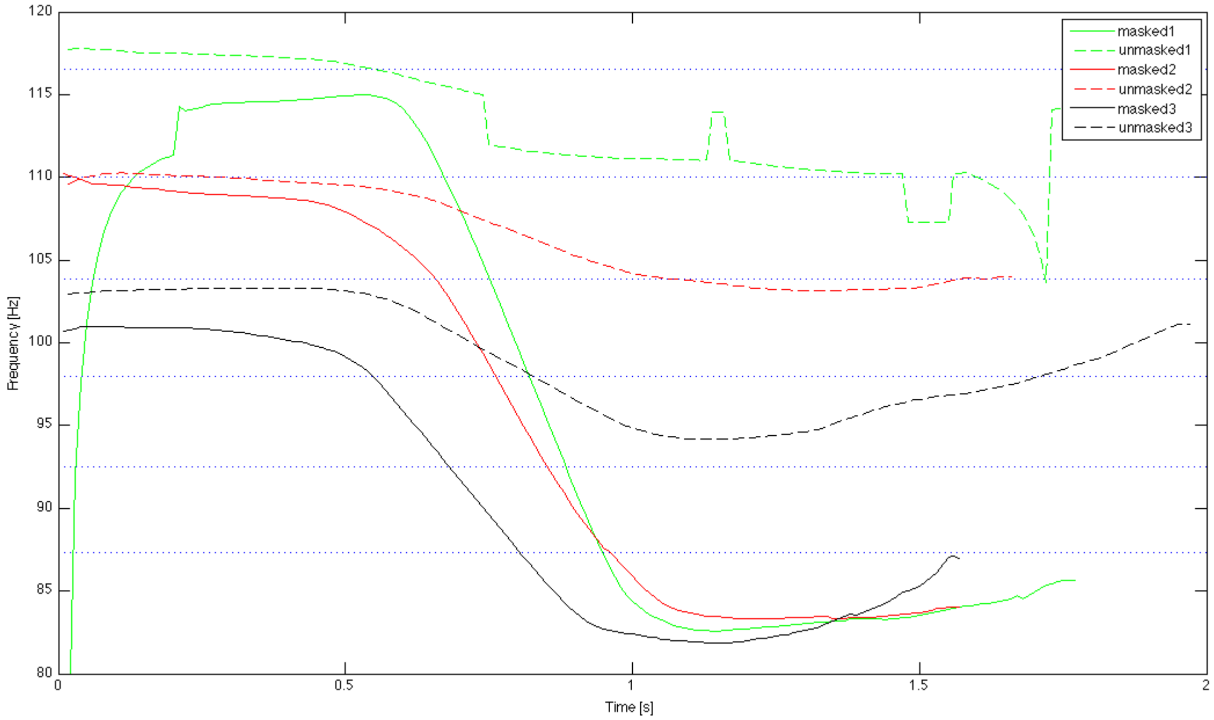


Frequency [Hz]
Time [s]
120
115
110
105
100
95
90
85
80
0
0.5
1
1.5
2
2.5
masked1
unmasked1
masked2
unmasked2
masked3
unmasked3

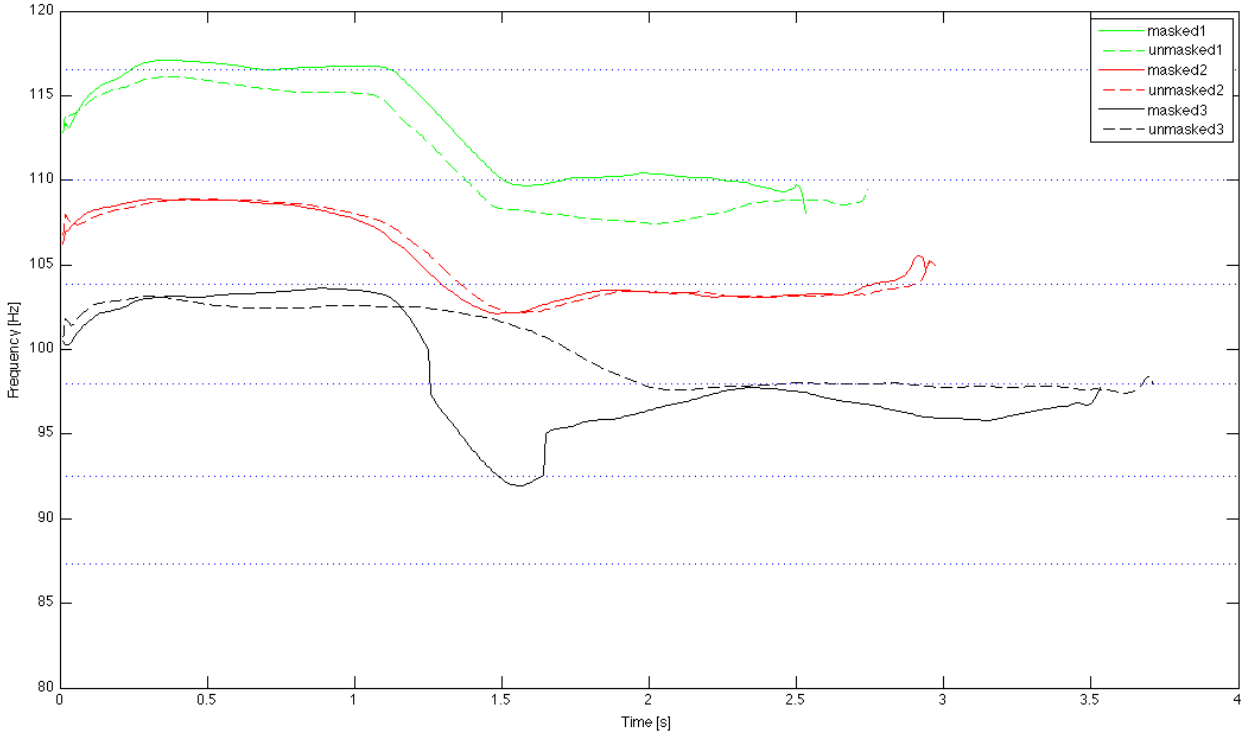
(c) A3


120
115
110
105
100
95
90
85
80
Frequency [Hz]
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
6
Time [s]
masked1
unmasked1
masked2
unmasked2
masked3
unmasked3

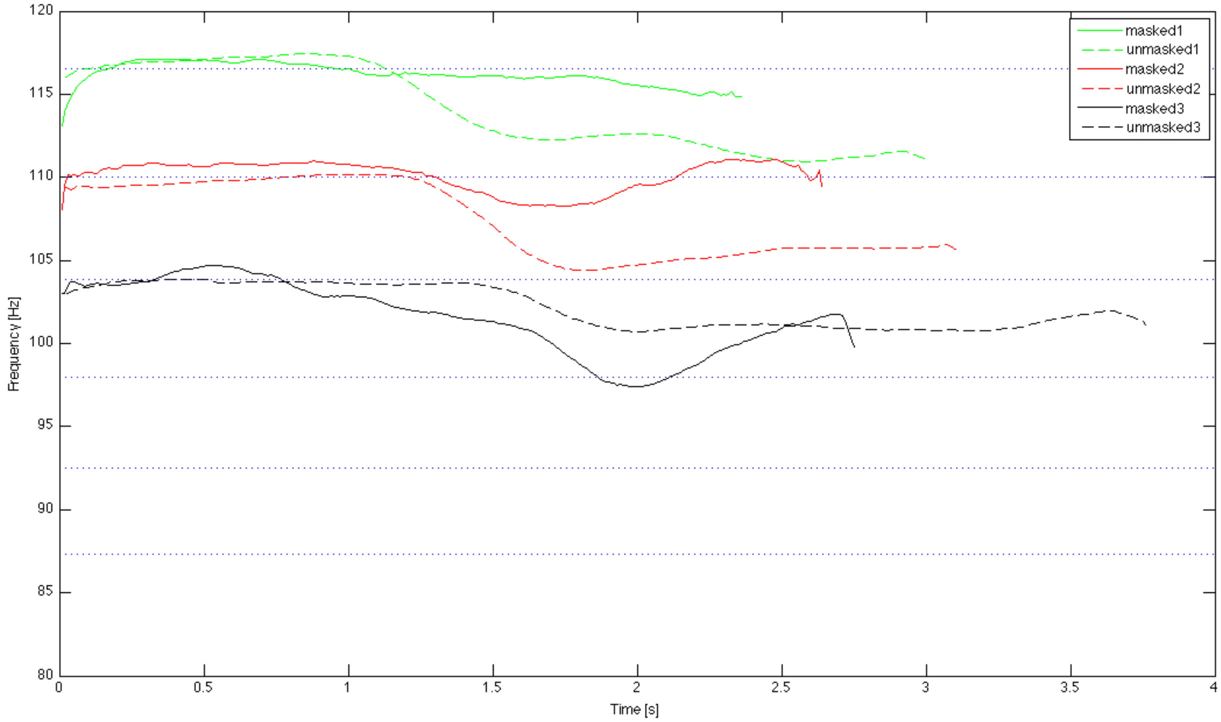
(d) P4


120
115
110
105
100
95
90
85
80
Frequency [Hz]
0
0.5
1
1.5
2
Time [s]
masked1
unmasked1
masked2
unmasked2
masked3
unmasked3

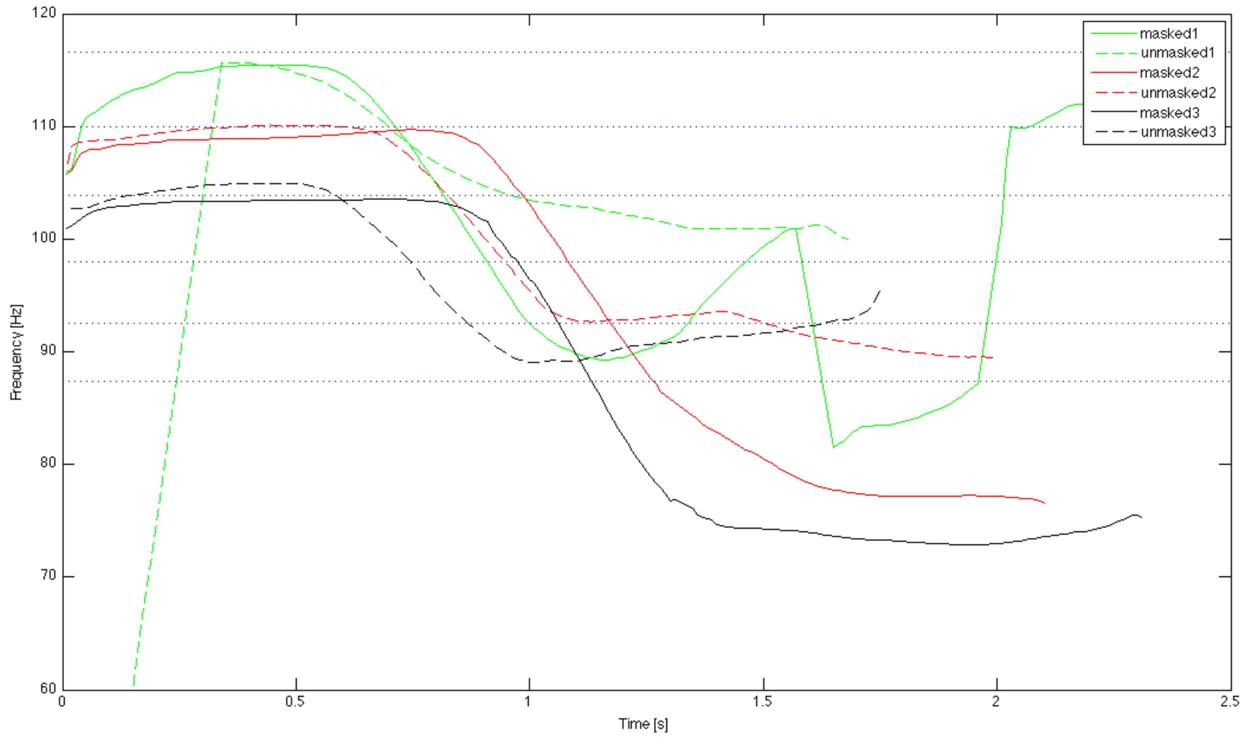
(e) A2


120
115
110
105
100
95
90
85
80
Frequency [Hz]
0
0.5
1
1.5
2
2.5
3
3.5
4
Time [s]
masked1
unmasked1
masked2
unmasked2
masked3
unmasked3

(f) P2


120
115
110
105
100
95
90
85
80
Frequency [Hz]
0
0.5
1
1.5
2
2.5
3
3.5
4
Time [s]
masked1
unmasked1
masked2
unmasked2
masked3
unmasked3

(g) A1



(h) A4

Figure 47: Pitch bending (lipping) a semitone from Bb1 to A1 (green line), from A1 to Ab1 (red line) and from Ab1 to G1 (black line). Pitch bending is performed without the use of the slide. Figures (a) through (h) shows the respective performers. Note figure (h), (e) and (d). Here, the software for calculating pitch had trouble determining pitch at some times because of the instability in the tones produced. The tones produced when the instrument is lipped down often have an unstable or beating characteristic. As we can see, almost every participant lipped the tone far too much in the masked recordings, suggesting a limited control of pitch without hearing oneself. The error was big for both non-professionals and professionals. The reason why some of the curves are starting to move up towards the start note in the end is that some of the participants bended the note down and then up again. Calculated with a software for pitch tracking (Zahorian and Hu, 2008).

B MIR Toolbox

The MIR Toolbox is a toolbox designed for Matlab by members of the *Finnish Centre of Excellence in Interdisciplinary Music Research*, University of Jyväskylä in Finland. The toolbox has been developed within the context of a European project called "*Tuning the brain for music*", funded by NEST (New and Emerging Science and Technology) (Lartilliot, 2008). The version used in this report was version 1.5. Figure 48 shows an overview of the functions available in MIRtoolbox.

B.1 Functions

The MIRtoolbox uses established theory to extract musical factors from music recordings. This project has used a limited number of functions provided by MIR. They are listed here. All functions use a standard frame length of 50 milliseconds and a hop rate of 0.5.

MIRaudio

Takes a wav-format audio file and converts it into a Matlab object.

MIRenvelope

Computes the global outer shape of an audio input. An available parameter adjustment in this function is normalization with respect to energy.

MIRspectrum

Computes the spectrum of an audio signal using a Discrete Fourier Transform

$$X_k = \sum_{n=0}^{N-1} x_n e^{\frac{-2\pi kn}{N}}, \qquad k = 0, \ldots, N-1$$

MIR calls the Matlab *fft*-function when computing the spectrum. This function also offers the choice of computing the spectrum in the time domain.

MIRcentroid

Shows the geometric center (centroid) of a variable distribution, e.g. a spectrum, and tells us something about the central tendency (mean) of a distribution.

$$\mu_1 = x f(x) dx.$$

MIRtempo

Finds the temporal length of an audio sequence.

MIRrms

Calculates the global energy of the signal by taking the root average of the square of the amplitude (Root-mean-square).

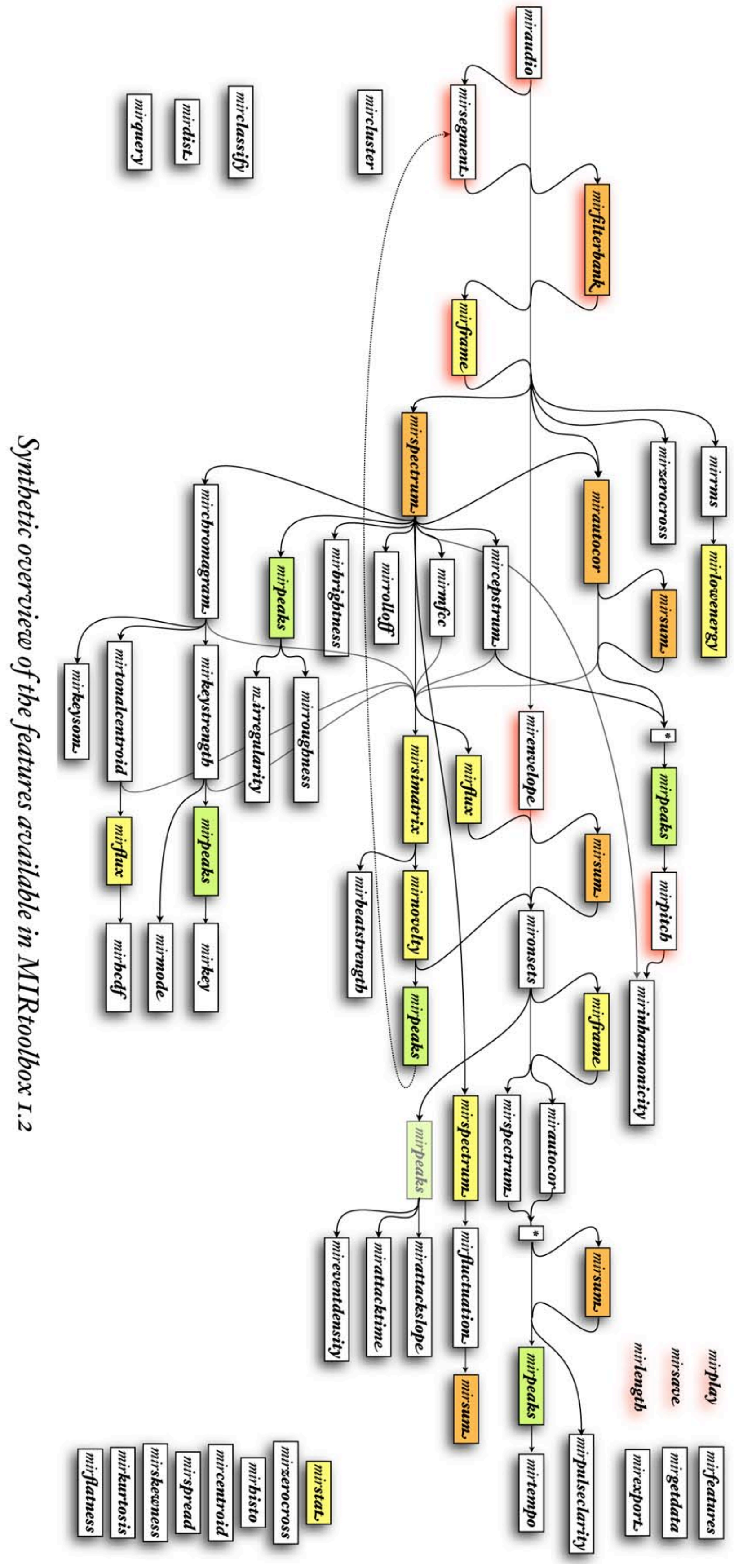


Figure 48: Overview of *MIRtoolbox* features. Accessed May 14 – 2014. From: https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox/FeatureOverview

C Experiment information

C.1 Score for part I

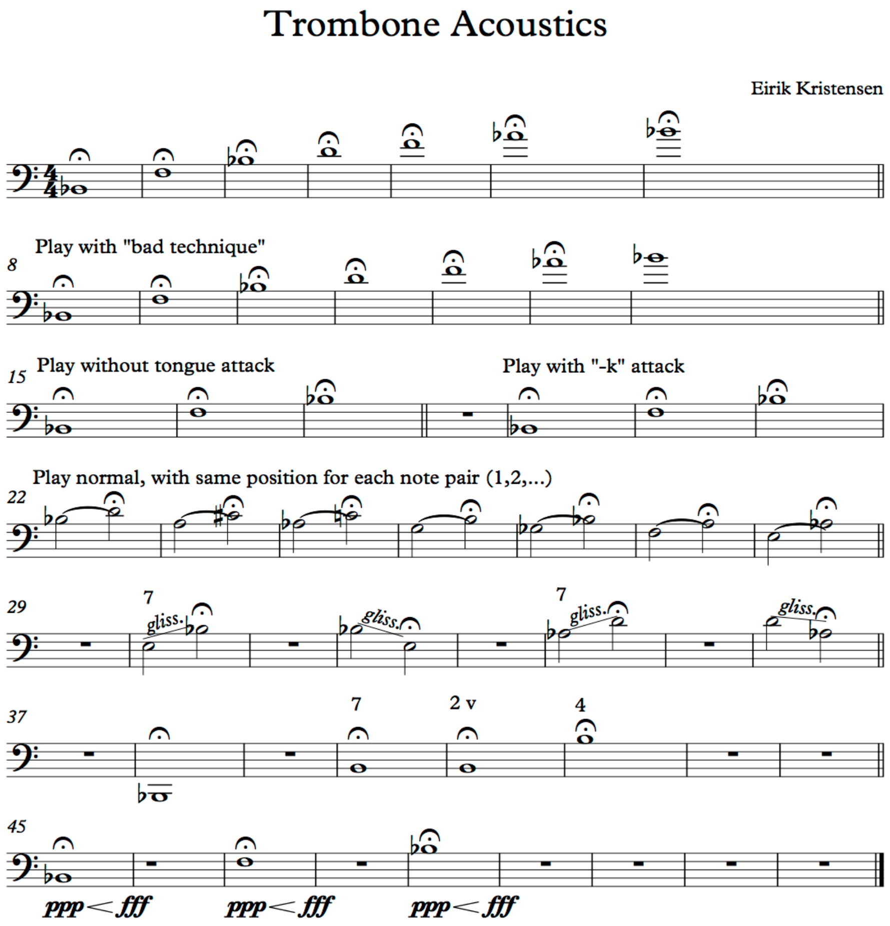


Figure 49: Score for part I.

C.2 Musical dynamics

pp	*pianissimo*	‘very soft’
p	*piano*	‘soft’
mp	*mezzo-piano*	‘moderately soft’
mf	*mezzo-forte*	‘moderately loud’
f	*forte*	‘loudly’
ff	*fortissimo*	‘very loudly’

C.3 Questionnaire for part II

The participants answered the first part of the questionnaire *before* the experiment was carried out, and the second part *after* the session. The experiment supervisor asked the questions in an interview-like style, and the answers were more like discussions than simple answers.

Personal information and experience

1. Name:
2. How many years have you played trombone:
3. Classical / Jazz performer or none/mixed:
4. Estimated hours a week of practice/performance:
5. Estimated hours a week of performance in genres jazz and classical:
6. Describe your experience with playing in acoustically difficult situations. E.g. on stage with bad monitoring, *fortissimo* parts with the orchestra etc.:

Experiment comments

1. Were you able to hear or in any way have a perception of your own sound during the experiment? If yes – when, how and how do you think it affected your performance?
2. In the overall performance, how would you describe the playing conditions, and how did it affect your musical performance?
3. In the OP performance, did you feel that the masking noise affected the expressiveness in the performance?
4. In the PV performance, did you feel that the masking noise affected the expressiveness in the performance?
5. Would you be able to sing or hum the Prima Vista melody directly from the score?
6. Other comments:

D Matlab Code

The Matlab code for the pitch tracking algorithm *Yaapt:*

```
%%Importing the audio
[a1,fs1] = wavread ('er_bend1');
[a2,fs2] = wavread ('er_bend1_unmask');
[a3,fs3] = wavread ('er_bend2');
[a4,fs4] = wavread ('er_bend2_unmask');
[a5,fs5] = wavread ('er_bend3');
[a6,fs6] = wavread ('er_bend3_unmask');


%PITCH DETECTION
ExtrPrm.f0_min = 60;
ExtrPrm.f0_max = 130;
ExtrPrm.f0_double = 100;
ExtrPrm.frame_length = 15;
ExtrPrm.fft_length = 8192;

[Pitch1, nf1] = yaapt(a1, fs1, 1,ExtrPrm);
[Pitch2, nf2] = yaapt(a2, fs2, 1,ExtrPrm);
[Pitch3, nf3] = yaapt(a3, fs3, 1,ExtrPrm);
[Pitch4, nf4] = yaapt(a4, fs4, 1,ExtrPrm);
[Pitch5, nf5] = yaapt(a5, fs5, 1,ExtrPrm);
[Pitch6, nf6] = yaapt(a6, fs6, 1,ExtrPrm);


%SMOOTHING OF THE CURVES
PS1 = smooth (Pitch1,40);
PS2 = smooth (Pitch2,40);
PS3 = smooth (Pitch3,40);
PS4 = smooth (Pitch4,40);
PS5 = smooth (Pitch5,40);
PS6 = smooth (Pitch6,40);


%PLOTTING
figure ('name','Pitch bending for a professional jazz trombonist');;
box on;
hold on;
ylim([80 120]);
xlabel ('Time [s]');
ylabel ('Frequency [Hz]');
set(gca,'XTick',[0 50 100 150 200 250 300 350 400] );
set(gca,'XTickLabel',[0 0.5 1 1.5 2 2.5 3 3.5 4] );
plot (PS1, 'g-','linewidth',1);
plot (PS2, 'g--','linewidth',1);
plot (PS3, 'r-','linewidth',1);
plot (PS4, 'r--','linewidth',1);
plot (PS5, 'k-','linewidth',1);
plot (PS6, 'k--','linewidth',1);
```