INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

‘Brightness’ Measures of Trombone Timbre

by

Jay C. Bulen

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Musical Arts

University of Washington

1995

Approved by ______________________
(Chairperson of Supervisory Committee)

Program Authorized
to Offer Degree School of Music

Date 18 May 1995

UMI Number: 9609595

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

In presenting this dissertation in partial fulfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to University Microfilms, 1490 Eisenhower Place, P.O. Box 975, Ann Arbor, MI 48106, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature Jay C

Date 5/18/95

University of Washington

Abstract

‘Brightness’ Measures of Trombone Timbre

by Jay C. Bulen

Chairperson of Supervisory Committee: *Professor Stuart Dempster*
School of Music

Trombone timbre is difficult to describe because of the multidimensional nature of timbre perception and because of inherent timbral variations with pitch and loudness. Following suggestions by Balzano (1986), a measure of trombone timbre is selected that is suited to the dynamic processes of trombone tone production. The measure is *Schärfe* (Aures, 1985), translated here as *brightness*, which is a weighted first moment of the spectral energy distribution. A descriptive study is undertaken to determine the usefulness of this measure, by testing its ability to discriminate between members of the trombone family (alto, tenor, and bass) and between performers using the same instrument.

Systematic differences are found between instruments and between performers. For notes at the same loudness and pitch, alto trombone is brighter than tenor, which is brighter than bass. Differences between performers using the same instrument are smaller than differences between instruments. Differences between performers tend to be concentrated in regions of loudness and pitch. The regions are specific to the performers being compared and independent of instrument. Brightness of trombone tones depends primarily on loudness, and to a lesser extent on pitch. The rate of change of brightness with loudness depends on pitch and sounding length. Measured log brightness vs. log loudness and frequency can be approximated with a simple surface. It is concluded that brightness vs. loudness and pitch is a useful measure of trombone timbre. It yields results consistent with musical practice and musical understanding. Compared with verbal descriptions of timbre, its results are compact and reliable.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

While it is impossible to adequately thank all those who have helped in the preparation of this dissertation, I would like to acknowledge three groups of people.

The first is composed of faculty members of the University of Washington, whose exceptional capacities and breadth of expertise have made this multidisciplinary study possible. I thank my advisory committee members Roy Cummings, Edward Curtis, David Kappy, Douglas Keefe and Stuart Dempster. I would especially like to thank Stuart Dempster, whose astonishing ability to make music (and help others make music) is accompanied by broad interests; and Douglas Keefe, whose vast knowledge and scientific expertise is matched by a passion for music. In short, I have been helped and encouraged by scholars whose vision of music and musicianship encompasses more than I can hope to realize.

The second group includes many friends and colleagues who have assisted me. Without their help and encouragement this project could never have been completed. Included in this group are Pamela Goad, Robert Ling and Chad Kirby.

The final group comprises my family, who have endured much in this process. I am grateful for the efforts of my wife Kathy, who has shouldered many additional responsibilities while never allowing me to lose sight of our ultimate goals. My son David and daughter Joy have been a great source of inspiration. This endeavor would have been entirely impossible, however, without the personal and material and personal support of my parents Herbert and Jan Bulen.

DEDICATION

This dissertation is dedicated to the memory of my father Herbert J. Bulen.

Chapter 1

INTRODUCTION

Jeffrey Reynolds of the Los Angeles Philharmonic once told me that he is a trombonist because there is nothing he would rather do than "make that sound." His sentiments are echoed by many trombonists. The characteristic sound we produce in conjunction with our instruments is of immense importance to brass players, and "tone quality" enters into every level of musical consideration. It distinguishes an individual performer, marking both a musical and a personal identity. It is one of the basic elements of interpretation, and its manipulation as a function of genre, piece and even individual passage manifests musicianship. Tone quality provides the impetus for many decisions concerning instrument and accessories, such as mouthpiece and mutes. Finally, it invariably forms one of the categories upon which a trombone or trombonist is evaluated, whether in informal comparisons or on audition score-sheets. In *The Art of Trombone Playing*, Kleinhammer says that "The quality of a musician's tone shows up in everything he plays" (1963, p. 36). Clearly, tone quality is important to a trombonist.

Given its importance, there should be reliable means for describing trombone tone quality. Descriptions should (1) signify the same thing for all users, (2) be independent of how (or by whom) the description is formed, and (3) be sufficiently accurate to reflect important differences. There is no such description for trombone tone quality.

A reliable description would impact many aspects of music making. It would facilitate discussion of appropriate tone qualities, much as appropriate tempi are discussed now. It would enable teachers to specify how students' tones differ from standards, as is done with intonation. Finally, it would inform selection of equipment.

1.1 Timbre: Subjective and Multidimensional

Tone quality (also called tone color or timbre) is a subjective quantity. Its most frequently cited definition, that of the American Standards Association, is based entirely upon the ability to make subjective discriminations: "timbre is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar" (American Standards Association, 1960, p. 45). Clearly, it is not possible to have an objective measure of a subjective experience. However, it may be possible to formulate an objective measure to correlate with aspects of experience.

An objective measure which attempts to mimic perceptions constitutes a model. Values produced by the model can be regarded as *predictions* of a perceived quantity. For example, measurements of a trombone tone might result in a predicted loudness of 40 sones, using Paulus and Zwicker's model for loudness (1972). By definition, this prediction suggests that a listener would judge the sound to be equally loud as a 1000 Hz sine tone, presented frontally at a sound pressure level of 40 dB. The prediction's accuracy can be assessed only by actually presenting listeners with the trombone tone and reference, and determining whether they are indeed perceived as equally loud.

Besides being strictly subjective, timbre is also multidimensional. In this sense it differs from pitch, which is unidimensional but multivariate. Pitch is perceived to vary on a single high–low continuum, but depends upon several variables (fundamental frequency, spectral envelope, intensity) to establish its position. Tone quality, although perceived as a single entity, has no unique subjective continuum along which it varies. Contributing factors can influence timbre in different ways, producing many varieties of tone colors. For a unidimensional attribute like pitch, contributing variables may offset each other while the perceived attribute remains the same. For example, a low-intensity, high-frequency sound can have the same subjective pitch as a high-intensity, low-frequency sound. With a multidimensional attribute, however, this does not occur: differences in one dimension are not eliminated by changes in another.

1.2 Selecting Dimensions for Modelling

Any model of trombone timbre conveys information about some number of subjective attributes or dimensions. A concise model is based on the most appropriate dimensions. These must first have relevance to trombone tones. Timbral dimensions that are important for gong tones, for example, may be uninformative when applied to trombone tones. Secondly, the dimension must be perceptually salient. Since timbre is robust and musically important, it is unlikely to be based on subtle attributes. Finally, the dimension must be essential to trombone timbre. Some sound characteristics are more essential than others. For example, key noise is a reliable cue for differentiating between 'cello and bassoon tones; but basing a description of bassoon timbre on key noise would be foolish. In short, an objective description of trombone timbre that attempts to reflect perception should be based on attributes which are relevant, salient, and essential. To the extent that a description mimics perception of a timbral attribute, it can be viewed as a model of that attribute's perception.

1.3 Synthetic and Analytic Listening for Timbre

In the previous discussion two different uses of the term *timbre* can be discerned. The first refers exclusively to the discrimination of sounds; this kind of listening we will call *analytic.* This is the kind of listening described in the previously-mentioned American Standards Association (1960) definition. The most significant work in analytic timbre perception is that of Green and his co-workers (1988), in their careful development and application of *profile analysis.* Their experimental paradigm consists of varying the level of a single component of a complex tone, and determining detectability of the alteration in terms of degree of amplification of the selected component, overall level, tonal complexity, temporal attributes, etc. The goal of this research is nothing less than the definition of the basic unit of timbral perception. However, in its current state profile analysis can say little about perception of trombone timbre.

The second use of *timbre* refers to the grouping together of sounds that come from a sound-producing system, and the ability to identify the system based upon what is heard. For example, a listener presented with a loud, low note and a soft, high note from an oboe identifies the timbre of the oboe in spite of differences between the tones. Risset and Wessel (1982) describe this as "the notion of timbral constancy", and it has great musical importance.

How can this ability be explained? It is sometimes explained with memory alone, by assuming that an auditor has stored up a comprehensive collection of oboe tones for comparison. As Balzano (1986) has pointed out, this explanation is unsatisfactory on at least two counts. First, it presupposes the existence of an unreasonably large amount of stored information. Secondly, and more importantly, it is unable to address the question of how a connection between sounds is established in the first place. A memory-based explanation works just as well for arbitrary collections of sounds as for those coming from an instrument; it thus "denies—or at best ignores—any necessary connection or principled variation among the different oboe sounds as an important source of perceptual information" (p. 309). Memory-based explanations thus fail to account for this important auditory skill in a parsimonious fashion.

Listening for unifying characteristics can be called *synthetic* in that an auditor subconsciously combines sounds into a coherent whole, connected not only by the sounds themselves but by a cognitive model of the system that produced them. One conceivable model is a mental representation of sound production. Listeners, for example, have no trouble discerning sounds emanating from bounced as opposed to breaking objects (Warren and Verbruggee, 1984). Similar representations may be involved in musical timbre. Balzano has stated that "perceiving timbre is more a matter of perceiving underlying dynamics of physical processes than perceiving the places of things in abstract ordered structures" (1986, p. 312). Synthetic listening for trombone timbre depends upon those elements that define "trombone-ness": inherent timbral characteristics, stemming from the physical processes of tone production, that enable a listener to identify a sound and predict surface changes based upon deeper invariances.

Clearly, a description intended to describe general characteristics of trombone tones should be based on elements that contribute to synthetic listening. These may be inferred from the physical properties of the sound-producing system and the characteristics of its sounds.

It is proposed in this study that the defining objective characteristic of trombone tones is high-frequency spectral content as a function of intensity and frequency. These will be shown to relate in a way determined by the underlying dynamics of tone production. The relationship is thus a likely cue for synthetic listening. Measurements performed in this study will show that variations in trombone tones, for example those produced by different trombones, will be reflected in it. The perceptual correlate of high-frequency spectral content is *brightness.*

1.4 Brightness and Trombone Timbre

As will be more thoroughly discussed in later chapters, subjectively judged brightness is useful for describing trombone tones (Pratt and Bowsher, 1978; Runyan, 1979). It appears to be related to perceived *size* (Mays, 1980), which is also salient (Edwards, 1978). Although individual listeners can make reliable subjective judgements of trombone brightness, ratings are not always consistent across listeners. For example, when comparing tones produced by two trombonists, some listeners will find one player's tones consistently brighter, while others consistently ascribe higher brightnesses to the other. Possible explanations for this include preference, comparison to different referents, and confounding by other aspects, which will be discussed in later sections. For now, it should be noted that an objective prediction of brightness would be constructive.

1.5 Statement of Purpose and Research Objectives

The purpose of this study is to describe trombone tones using objective means, in terms that are perceptually meaningful. Brightness appears to be an important dimension of trombone timbre. However, trombone timbre changes dramatically with loudness and pitch (Reitz, 1950). Therefore, a three-dimensional description of trombone tones is indicated, presenting predicted brightness as a function of predicted loudness and pitch.

Two research objectives will allow assessment of the usefulness of the description. First, can predicted brightness vs. loudness vs. pitch be used to discriminate tones from alto, tenor and bass trombones? Secondly, can it be used to discriminate between tones from the same trombone, but different players? Edwards (1978) found that listeners can accomplish these tasks. If consistent differences are manifested in measurements, it would indicate that the description's resolution is adequate for some musical applications.

1.6 Overview of the Study

In order to justify a particular representation of trombone timbre, both physical characteristics and perceptual attributes must be considered. Spectral content is known to be one of the primary determinants of tone quality (Risset and Matthews, 1969). It will be shown that the spectral content of a trombone tone can be quite accurately expressed as a function of frequency and intensity. The relationship between frequency, intensity and spectral content reflects the physics of trombone tone production. Intensity is the principle determinant of loudness, and frequency most strongly influences pitch. Thus, from both perceptual and physical perspectives, trombone timbre appears to vary with loudness and pitch.

It will also be shown that individual instruments vary systematically in the relationship between spectral content and intensity. As Benade states, "The spectrum itself changes with playing level in a fashion that is characteristic not only of brass instruments as a group, but also of the individual instrument" (1976, p. 424-425). Furthermore, performers' attributes and habits exert recognizable influences. Thus, a portrait of spectral content changing with amplitude and frequency might reflect differences between players and differences between instruments.

In this study, predicted brightness is measured over the entire practical ranges of pitch and loudness for alto, tenor and bass trombones. Each of the instruments is played by three highly skilled trombonists. Results are examined for patterns of predicted brightness vs. loudness vs. pitch that are typical of trombones in general. Additionally, the patterns are examined for differences between alto, tenor and bass trombones (same performer) and for differences between performers (same trombone). From this data the research objectives of using predicted brightness vs. loudness and pitch to discriminate tones from different trombones and from different trombonists can be addressed.

1.6.1 Terminology and Units

Predictions of loudness are expressed in units of *sones*, calculated according to Paulus and Zwicker's algorithm (1972). Pitch is described with musical note names (A, B, C♯ , etc.) with numbers denoting octave. C1 is the lowest C on the piano; C4 is middle C. Equal-temperament, referenced to A4 = 440 Hz, is assumed.

Brightness is expressed in units of *acums*, as calculated with Aures' (1985) formula for *Schärfe*. The translation of this term presents problems in the current context. A literal translation from the German is *sharpness*, and this is how it most often appears in English-language research literature (e.g., von Bismarck, 1974a, 1974b; Zwicker and Fastl, 1992; Goad, 1992, 1994; Lindevald, 1987). However, musically-oriented researchers acknowledge that *sharp* is not an entirely satisfactory translation. Ernst Terhard has written that

> Von Bismarck ... was fully aware of the problem of adequate translation. He translated *Schärfe* into "sharpness" ... Yet there indeed exists a drastic difference between the meaning of "sharf" with regards to musical pitch! While the notion tone "A" is sharp relative to tone "B" in English means that A is somewhat higher in pitch, the corresponding German phrase "Ton A ist scharfer als Ton B" does not at all refer to pitch but to timbre. (quoted in Kendall and Carterette, 1993a, p. 456)

It has been proposed that *Schärfe* is similar to *brightness* (Beauchamp, 1982) and *brilliance* (Pratt and Doak, 1976). Webster's (Gove, 1981) lists *sharp* as a synonymous cross-reference for *bright*. For present purposes, I feel that *brightness* is the most appropriate English translation, since it implies a relationship to tone quality and is familiar to trombonists. It will be used henceforth, with the understanding that it refers to *Shärfe* as defined by Aures (1985).

The term *sharpness* will be reserved for von Bismarck's earlier formulation. Beauchamp (1982) used *brightness* to describe a specific formulation of the centroid of spectral energy; this will be called *BR* in accordance with his own notation. The various formulations of spectral centroid, sometimes also referred to as *brightness* (Grey and Gordon, 1978), will be referred to by their method of calculation, e.g. "centroid of loudness function" or "centroid of linear spectral amplitudes."

1.6.2 Outline of Dissertation

The models used to generate predictions of loudness, pitch and brightness must be explained and assessed. Unfortunately, justifying a model requires a great deal of information not directly related to trombone performance. An important issue is the completeness of the description, which must be considered from the broadest possible perspective. I have tried to segregate topics into those that are essential as well as of interest to trombonists, and those that are of interest primarily to acousticians and psychoacousticians.

Chapters 1, 2, 3 and 4 are indispensable. Various models are presented, along with subjective and objective characteristics of trombone tones.

Chapters 5, 6, 7 and 8 are more arcane, and may be skimmed or skipped. These establish the centrality of brightness to trombone timbre, by investigating general timbre perception, timbre description, and sound production in the trombone. Finally, the influence of a room on timbre and spectral measurements is discussed.

Chapters 9 through 12 are again essential. These describe methods, results, analyses and conclusions.

Perceptions and Descriptions of Musical Tones

Chapter 2 provides a convenient reference for the calculation of predicted loudness in sones, predicted brightness in acums. and predicted pitch in Hertz (Hz). Although it may seem odd to describe models prior to explaining why they have been chosen, it is necessary since some reasons for the models' selection are based upon the appropriateness of their formulation.

Chapter 3 summarizes research into subjective descriptions of trombone timbre. Sources include adjectival descriptions, ratings of similarity, and ratings on verbal scales. Methods of extracting additional information from subjective evaluations are discussed, including multidimensional scaling and semantic differential scaling. Perceived size as well as brightness are shown to be important subjective aspects of trombone tones. It is also shown that, while consistent subjective evaluations are obtained from most listeners. responses tend to be idiosyncratic. An objective measure appears necessary.

In Chapter 4, objective descriptions of trombone tones are discussed. Time-domain and frequency-domain descriptions are presented. It is shown that steady-state trombone tones can be described fairly well using spectral envelopes. Bass trombone spectral envelopes differ from those of tenor trombones, as do envelopes from different trombone players using their own instruments. Formulations based upon formant structures are also presented. For trombone tones, the proportion of high- frequency to low-frequency energy is related to overall loudness. The problem of relating objective measures to subjective descriptions is presented.

Chapter 5 represents a broader view of timbre, made important by the findings of Chapters 1, 2 and 3. Research in timbre dimensionality is reviewed, showing that for musical and quasi-musical sustained tones in general (not just trombone tones), the relative proportion of high- to low-frequency energy is perceptually important. The attack transient is also found to be perceptually important. Techniques are discussed for relating perceptual dimensions to verbal and physical descriptions. It is concluded that timbre perception in general can be modelled in a low-dimensioned space. The precise number of dimensions and their relative importance depends upon the stimuli. For time-invariant tones the dominant dimension appears to be brightness.

Chapter 6 discusses physical processes of tone production in the trombone, thus providing a theoretical basis for timbral variations. This is essential because not only do physical processes determine general characteristics of trombone tones, but they also suggest ways in which individual instruments and performers influence timbre. It is shown that many of the important timbral attributes of a trombone can be related to the instrument's input impedance and pressure transfer function. Cutoff frequency is a simple and influential parameter. It is also shown that performers have a number of means for influencing trombone timbre, including intonation, embouchure impedance, and vocal tract configuration. These influence the radiated spectrum in various ways, all of which impact brightness.

Chapter 7 considers the question of whether an adequate description of trombone timbre can be made without including the attack transient. Transients are difficult to measure and all but impossible to describe concisely. From a review of literature concerning the relative importance of various parts of tones, it is concluded that an adequate description based on steady states may be possible, if a range of pitches and loudnesses are included.

Chapter 8 discusses the accuracy of loudness and brightness measurements, from a theoretical perspective. A foundation is laid for evaluating recording conditions for trombone tones and for estimating accuracy of predictions.

Measurements

Chapter 9 is concerned with methods. It describes materials, participants, and procedures involved in data collection.

Chapter 10 presents results in general terms, in order to characterize trombone tones in terms of brightness, loudness, and pitch.

Chapter 11 describes further analyses, including averaging and smoothing. Comparisons between trombones (same player) are then undertaken. Finally, comparisons between players (same trombone) are made. Various methods of visualization facilitate the comparisons.

Chapter 12 presents conclusions and suggestions for further research.

Chapter 2

PREDICTED BRIGHTNESS, LOUDNESS AND PITCH CALCULATIONS

In order to provide a convenient reference, models for predicted loudness, brightness, and pitch are described below. The reference is placed here since the models are implicated from numerous points of view. For example, spectral measurements of trombone tones described in Chapter 4 suggest the form of model that would respond to general characteristics of trombone tones. A different model could be appropriate for the effects of vocal tract resonances described in Chapter 7. Additionally, it is important to understand the model to compare it to others discussed in Chapter 5. Although this ordering is awkward it is necessary.

The models used to predict loudness, brightness, and pitch are discussed individually below.

2.1 Loudness Prediction

The algorithm used to predict loudness is based upon Paulus and Zwicker's model (1972). International Standards Organization (ISO) standard R532B for calculating loudness level (1975) is also based upon this model. Since it underlies both brightness and loudness prediction, it will be discussed somewhat in depth.

ISO R532B predicts loudness by modelling basic psychoacoustic principles. Since human ears generally function as frequency analyzers (Risset and Wessel, 1982), calculations are performed in the frequency domain; that is, complex sounds are analyzed into the sum of sine tones with individual amplitudes and relative phases. This is implemented via a digital Fourier transform (DFT). Figure 2.1 shows a portion of a trombone tone analyzed in this way. Only amplitudes are shown, since relative phases are less important for hearing (op. cit.). Each peak in the plot represents the amplitude of a particular component; this tone is composed primarily of multiples of 233 Hz.

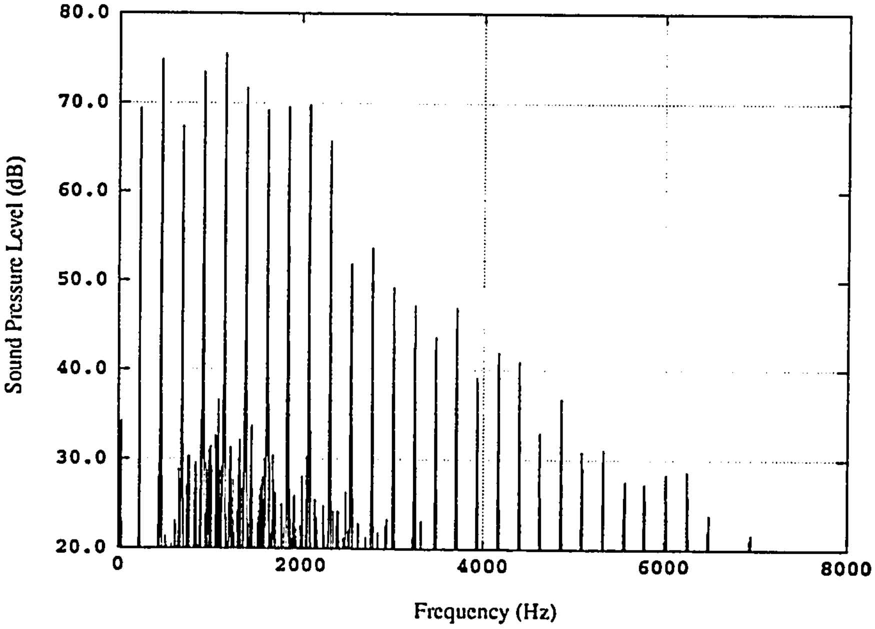


Figure 2.1: Frequency-domain representation of a trombone tone (B♭3, alto trombone).

Paulus and Zwicker's algorithm takes into account the sensitivity of the ear to sounds at different frequencies. This is accomplished by weighting each component according to curves provided in ISO R532B.

Next, the frequency-resolving capabilities of the ear are taken into account through critical bands. A critical band is a range of frequencies over which loudness depends only upon total sound energy. Components that lie in different critical bands contribute more to overall loudness than those that lie in the same (Zwicker and Fastl, 1992). The width of critical bands varies over the audible range; they range from about 100 Hz for frequencies below 500 Hz to about 1/4 of an octave above 1000 Hz. The conversion from frequency to critical bands is given by Zwicker and Terhardt (1980). The unit of critical bands is the *bark*; components that lie within one bark are in the same critical band. This provides a conversion from frequency into barks; a component at 1000 Hz can just as easily be described as a component at 8.5 barks.

Finally, the effects of masking are taken into account. Briefly, a component is masked by other components that lie within a critical bandwidth below it. Masked components do not contribute to overall loudness. This effect is modelled by extending the loudness contribution of each component upwards in the bark dimension with a sloping line.

The result of these three transformations can be seen in Figure 2.2. The absissa represents loudness in sones as a function of critical band rate. The contours of the curve result from adjusting for the sensitivity of the ear, critical bands, and masking. The contribution to total loudness of a particular component at z barks is represented by the absissa of the curve at z. This is called the specific loudness, and is designated by $n'(z)$.

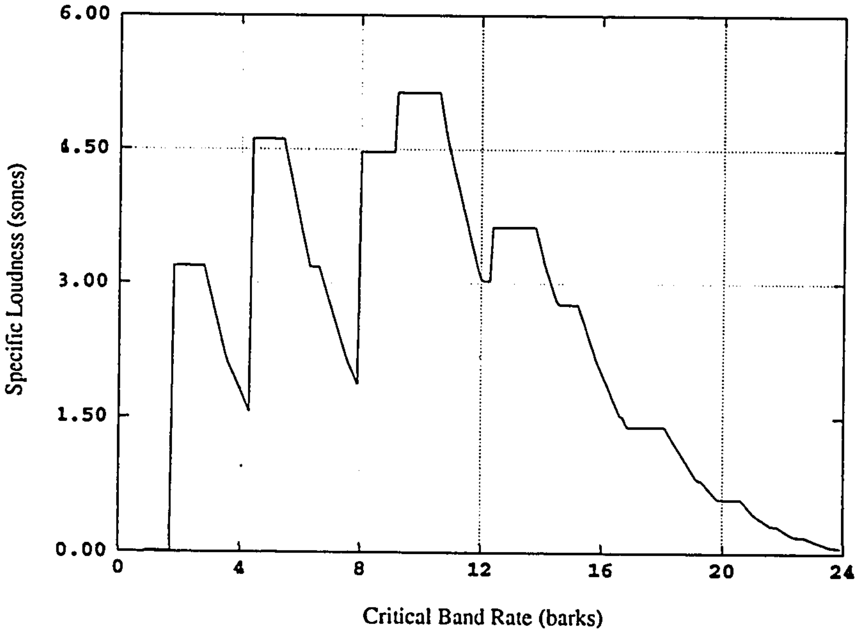


Figure 2.2: Specific loudness vs. critical band rate of a trombone tone.

To compute total loudness, the contribution of each component is summed by calculating the area under the curve in Figure 2.2. Total loudness N is calculated by

$$N = \int_0^{24barks} n'(z)dz. \tag{2.1}$$

The sone scale is designed to double in value with a perceived doubling in loudness; thus, a 20 sone sound is predicted to be twice as loud as a 10 sone sound; a 40 sone sound is twice as loud as one at 20 sones, and so on.

Typical loudnesses encountered in musical listening are given by Campbell and Greated (1987), who compare musical dynamics to loudness in sones. They state that a dynamic of *ppp* is 0.5 sones; *fff* is 64 sones. Goad (1994) found that a violin at *pp* produces between 3 and 7 sones; a single trombone at *ff* produces between 40 and 68 sones.

2.2 Brightness Prediction

Von Bismark (1974a) found that subjectively judged sharpness provides a means by which listeners can discriminate between sounds of similar loudness and pitch. He applied standard psychoacoustical procedures to a variety of sounds to test its consistent measurability, ascertain the factors that most strongly influence it, and produce a quantitative measure (1974b). He found that sharpness is influenced by several factors. There is an approximately exponential relationship between sharpness and upper limiting frequency. Sharpness rises with lower limiting frequency, but is less sensitive to this than to upper limiting frequency. Sharpness is largely insensitive to spectral fine structure. It "is an attribute separable from pitch" (p. 167), but the sharpness of narrow-band sounds increase rapidly with frequency. Subjects are "able to distinguish between sharpness and loudness" (p. 167) but sharpness increases slightly with sound pressure level. He found that a weighted sum of specific loudness as a function of critical band rate yielded values proportional to subjective loudness. Von Bismarck's formula is

$$sharpness = \frac{\int_0^{24barks} n'(z)g(z)dz}{N}, \qquad (2.2)$$

where z and $n'(z)$ are defined as above, and $g(z)$ is a weighting function representing the sharpness of narrow-band noise at constant loudness. The unit of predicted sharpness is the *acum*. Values of r_0 and c are chosen to match calculated to subjectively measured sharpnesses at a single point. The denominator represents total loudness N, reflecting the observation that sharpness is roughly independent of loudness. Von Bismarck concluded that this simple model approximated the subjective estimates "with fair accuracy" (1974b, p. 171).

Von Bismarck utilized stimuli with similar loudnesses. Aures (1985) performed psychophysical scaling experiments on a broad sample of sounds of varying loudness, and found that some adjustments to von Bismarck's formula were required. The weighting function $g(z)$ was modified slightly, becoming

$$g'(z) = e^{(.171z)}. \tag{2.3}$$

Aures also found improved agreement with subjectively measured sharpness if the denominator was replaced by a loudness-dependent function, given by

$$\log \frac{N+20}{20}, \tag{2.4}$$

where N is total loudness in sones as given in Equation 2.1. With these modifications, Aures reported very good agreement between calculated values and psychacoustically measured sharpness for all loudnesses measured. The correlation factor was 0.912.

For present purposes, *brightness* in acums is defined according to Aures' expression; that is,

$$brightness = \frac{\int_0^{24barks} n'(z)g'(z)dz}{log\frac{N+20}{20}} \tag{2.5}$$

Von Bismark's Figure 9 (1974b, p. 170) shows that a ratio between subjectively judged "sharpnesses" is represented with the same ratio of predicted sharpness in acums. This is true of Aures' brightness formulation as well. Thus, a sound with a predicted brightness of 4 acums would in theory be judged as "twice as bright" as a 2 acum sound.

As an indication of of brightnesses encountered in musical listening, the tuba at *pp* produces tones of about 0.5 acums; the E♭ clarinet at *fortissimo* produces nearly 6 acums on the note E♭6 (Goad 1994).

2.3 *Pitch Prediction*

Pitch depends primarily on fundamental frequency, but is also influenced by intensity and spectrum (Ward, 1970). A model for predicting the pitch of simple (sine) tones has been developed by Stevens, Vokmann and Newman (1937). Its units are the *mel.* However, musical tones are nearly always complex rather than simple, and the mel scale is not very satisfactory. For complex tones as produced by the trombone, fundamental frequency is a better predictor. A doubling in fundamental frequency produces the interval of an octave, which is perceived by musicians as "twice as high" in pitch (Campbell and Greated, 1987). Thus, like loudness in sones and brightness in acums, a constant ratio in fundamental frequency results in a constant ratio in perceived pitch (that is, a constant musical interval).

2.4 Summary

For periodic tones like those produced by trombones, perceived brightness, loudness and pitch can be predicted with brightness in acums, loudness in sones, and fundamental frequency in Hertz.

Chapter 3

TROMBONE TONES AND TROMBONE TIMBRE

How are trombone tones perceived? This is a complex problem because knowledge of subjective perceptions comes from descriptions, which differ in form and substance from perceptions. In this chapter, the perception of trombone tones is examined using various methodologies, including adjectival descriptions, identification, similarity ratings, and semantic differential scaling.

3.1 Describing Trombone Timbre

The contrast between timbre perception and timbre description can be illustrated with an example. Trombonist A wishes to relate a tone quality to Trombonist B. The sound passed through Trombonist A's auditory system to the brain, where it was "processed" in many ways. It was compared to memories of other tones, categorized according to attributes and combinations of attributes, and judged according to various criteria. A complex mental representation was formed, comprising not only a model of the tone itself but also how it compares to other models. A verbal description was then attached to the representation. The description may attempt to convey something of the sound itself, or subjective responses to the sound, or perhaps how the sound was produced. When Trombonist B receives this description, its terms are compared to B's own cognitive categories, and a new mental representation of the tone is developed based upon what the *words* mean to Trombonist B. The description is not the sound itself; at best it incompletely represents Trombonist A's subjective responses. Adjectives used in musical description rarely have a one-to-one correspondence to sound characteristics (Sandell, 1989), and are further biased by personal experience and preferences (Edwards, 1978). They are inherently cross-modal, intended to represent a non-verbal experience in a verbal domain. The mental representation that Trombonist B generates, then, reflects a conception of Subject A's reaction to the tone, rather than the tone itself.

This is not to say that verbal descriptions are uninformative. On the contrary, their primary difficulty is that a single term simultaneously carries many kinds of information. A description like "bland" reflects not just tone quality, but the listener's reaction to it, why it is significant to Trombonist A, and why Trombonist A thinks it should be significant to Trombonist B. Verbal descriptions function only to the degree that perceptions are similar and terminology consistent.

3.1.1 Semantic Descriptions: Orchestration Manuals

Semantic description is the most accessible means for dealing with the subjective aspects of trombone tones. It is the norm in theoretical writings on music as well as in musical discourse. How is trombone timbre described by writers on music?

Orchestration manuals provide a convenient source of semantic descriptions, since they are devoted specifically to musical instruments. Furthermore, they are influential in that they both reflect and guide musical practice. An example is the volume by Adler (1989), which has had wide impact. It includes a chapter on the brass section as a whole, and a second chapter with sub-sections detailing each instrument. Nine pages are devoted to describing the range, power, technique and uses of each type of trombone. Curiously, apart from a brief comparison of alto and tenor trombone in the extremes of their ranges there is almost no mention of timbre. Adler seems to simply assume familiarity with trombone timbre. His brief description of alto trombone timbre is typical in its appropriation of adjectives from non-auditory modalities: "the high notes are pure, mellow, yet brilliant, although not as piercing as those of the tenor trombone" (p. 315).

Sandell (1991) surveyed semantic description of timbre in orchestration manuals, and found a consistent approach.

> To characterize these qualities, authors of orchestration manuals ... use one-word semantic descriptors or phrases intended to suggest the aural impression of the instrument with evocative metaphors, simile, onomatopoeia, and comparisons to other modalities. (p. 11)

Descriptors for timbre are regularly taken from visual, tactile, and gustatory modalities, geometric and material analogies, environmental conditions and emotional states. While conceding that "Semantic description appears to play an important role as a tool for making basic qualitative decisions about timbre" (p. 21), Sandell finds them ineffective for describing bases for timbral similarity. For example, Adler's description of alto trombone timbre is indeed evocative; however, it could easily apply to certain clarinet or horn tones.

Sandell identified a subset of descriptors that are used relatively consistently. These include a *dark-bright* continuum, which seems related to spectral centroid (Grey and Gordon, 1978); *rough* and *coarse*, which refer to low tones and may be related to the psychoacoustic quality of *roughness* (Terhardt, 1978); and vocal analogies and *nasality*, which may be related to formant structure (Meyer, 1978). In practice, there is a tendency to specify comparisons inadequately. He notes that "when a tenor trombone is described as 'bright,' the implicit meaning is often 'the brighter member of the trombone family.' In a larger context ... the trombone is relatively dark" (1991, p. 20-21).

Sandell concludes that semantic descriptions are of limited utility for describing timbre. They are not uniquely associated with stimulus properties, and often reflect "complex aesthetic reactions to stimuli rather than direct information about many of the perceptual processes" (p. 20). Furthermore, words simply aren't available for some sounds. For the purpose of the present study, semantic descriptions are unacceptable for timbral description because they are neither consistently applied nor interpreted.

A somewhat different approach is taken by Reitz (1950), who avoids cross-modal descriptions in favor of comparisons with other instrument timbres. Additionally, he is careful to define conditions (e.g., range and dynamics) under which comparisons apply. For the trombone, he says that

> The modern trombonist can produce two great extremes of tone color in this range [B♭2-B♭4]. At fortissimo dynamic levels with sharp articulation, a very biting bass trumpet quality is possible....When played above [B♭4] the tone loses some of its biting edge and assumes a more French-horn like quality ....Writing the direction 'quasi-horn' on a trombonist's part will cause a sensitive player to change his tone even more to simulate that of French horn. (p. 51)
>
> Throughout the natural tessitura of the trombone a very soft singing quality is possible at dynamic levels of *pianissimo* to *mezzo-forte*. The

> sound is almost a vocal quality, resembling the timbre of a baritone voice. (p. 51)

Compared to Adler's description, Reitz's gives a clearer representation of trombone timbre, and may be more useful as an orchestration guide. He identifies range, dynamic, and performer actions that influence perceived timbre. However, his description is predicated on familiarity with other timbres which themselves vary with range, dynamics, and articulation. Furthermore, it lacks precision: how is fortissimo trombone *different* from bass trumpet?

We can conclude that semantic description as commonly used in orchestration manuals forms an inadequate basis for timbral measures. The majority of descriptors in common use lack specificity and accuracy. A few, however, seem widely understood and accepted. These include the *dark* to *bright* continuum, *rough* and *coarse*, and *nasal*.

3.1.2 Subjective Descriptions: Trombonists' Vocabularies

Trombonists have their own vocabulary for describing trombone timbre. Compared to that of orchestration manuals, it is oriented toward characteristics that are susceptible to performer manipulation, and emphasizes the volitional nature of timbre production. The performance literature stresses the importance of developing a personal mental representation or "tone concept" of what one's tones should sound like (Kleinhammer, 1963). This process begins with careful listening and imitation, and is guided by considerations of ensemble blend, performance venue, and musical style. In the early stages, the concept is imitative and derivative, but is expected to eventually become distinct and individual. Kleinhammer assures his readers that "As your concept matures, you will develop a tone of your own with the qualities you have sought out from hearing others" (1963, p. 36). It is expected that the performer will subsequently select instruments and accessories that will assist in the execution of their tone concept; however, the concept combined with the actions and characteristics of the performer are the overriding influences. "It is well known among brass players that a good player can compensate for the shortcomings of a particular instrument and the 'sound' he makes is, to some extent, independent of the instrument" (Elliot and Bowsher, 1982, p. 181).

The terms brass players use to describe timbre may indicate the dimensions of tone concept. While timbre is described by an enormous vocabulary there are a few descriptions that seem more common or more basic than others. The *size* of a tone is often referred to. "Students are often encouraged to 'work on their sound' or 'get a bigger sound' " states Mays (1980, p. 54). The second is the *brightness* of tone, ranging from *dark* to *bright.* The two scales are often combined, in descriptions such as "the biggest, darkest horn tone in the world" (Farkas, 1956, p. 52). It is interesting to note that many of the descriptors used for sounds are applied to instruments as well, implying a causal relationship. *Big* sounds are thought to come from large-bore instruments with large bells; a heavier-gauge bell is thought to produce a heavier tone (Edwards, 1978). Further confusing the issue is the fact that activities required for producing a large and dark tone are themselves large: big breaths, large mouth cavity, open throat. The same adjectives can be applied to various components of the sound-producing system. Thus, brass players refer to a dark sound, or a dark instrument, player, or mouthpiece. This practice suggests the formation of a cognitive model that combines characteristics of the system and expectations for the tones it produces.

The most systematic exploration of trombonists' vocabulary has been that of Edwards (1979), who sent a postal survey to all of the trombonists in the London musicians' guild. Respondents used some 77 different adjectives to describe the sound the sound of their instruments. Of these, Edwards found that 22 related to what he called *responsiveness*, while the remaining 55 referred to tone quality. He reported that

> It is not clear how a trombonist perceives the sound quality of his instrument. Its size is significant to him, since he uses such words as Enormous, Huge, Big, Small, Tiny, Big-Bore and Small-Bore to describe an instrument; but he also uses a plethora of other adjectives, such as Heavy, Wooly, Dull, Bright, Flat, Square, American, Mellow and Beefy, to describe other aspects of the sound. There is also a distinction between the trombone sound which is cutting ... and that which won't carry. (p. 18)

3.2 Timbre Dimensionality: Perceptual Space via Similarity Ratings

From the above, it is clear that some subjective descriptions of trombone tones have more validity than others. A *perceptual space* for trombone tones can be defined by the most important and distinctive of these. The number of dimensions in this timbre space is the minimum number of axes required to describe all points in the space, to within an acceptable accuracy. One way of determining dimensionality is through similarity ratings and multi-dimensional scaling.

As noted above, timbre research is complicated by confusion between perceptions and their descriptions. Similarity ratings offer a means for subjective comparisons without specific verbal categories. Listeners are asked only to evaluate the similarity of tones on a scale ranging from very similar to very dissimilar. The results are subjected to multidimensional scaling (MDS) algorithms to produce a map, in which distance reflects perceived dissimilarity. A perfect embedding of n evaluations can always be accomplished in an $n - 1$ dimensioned space; however, satisfactory results can often be achieved in smaller spaces. A *stress* or *goodness of fit* value reflects how well subjects' ratings can be accommodated. It is up to the researcher to determine the number of dimensions that are necessary (i.e., the stress value that can be tolerated). The usual means for determining this are (1) looking for break points in the plot of stress vs. number of dimensions, and (2) estimating the interpretability of the dimensions. The significance of a stress value can be inferred by comparing it to that produced by random data (for example, see Edwards, 1978).

The difficulty of MDS is that the dimensions' meanings must be inferred *post hoc* by the experimenter. This is commonly accomplished by correlating coordinates from each perceptual dimension with objective characteristics. For example, Grey and Gordon (1978) found that dissimilarity ratings between their tones could be accomodated in three dimensions. Tones which displayed high values in one of their dimensions all had spectral envelopes with a large amount of high-frequency energy. They formulated a spectral measure, and correlated this measure with their coordinates. Similarly, the results of MDS have been compared to ratings on verbal scales (Bloothooft and Plomp, 1988). To summarize, MDS is useful for deducing the dimensionality of timbre perception; however, additional information is needed to establish the meaning of each dimension.

Edwards (1978) applied non-metric multidimensional scaling techniques to adjectives reported in interviews. The limitations of this technique must be noted. Similarity judgements are inferred by the investigator, who tabulates instances when pairs of adjectives appear to have been used with antonymous or synonymous meanings. Edwards admits that the data were "obtained in a dubious and incomplete manner" (p. 417). Because of this, and because many pairings of adjectives are not rated, the technique is relatively inaccurate; nevertheless, the solution is suggestive. Edwards identified "two orthogonal axes—one of size, from small and tiny to big, wide, and big bore, and one of desirability, from British, dull dark, heavy, woofy and anonymous to masculine, clear, and trombone-like" (p. 417).

Edwards' solution is reproduced in Figure 3.1. The vertical dimension reflects size; adjectives used to describe the sound of trombones of similar bore size tend to have similar ratings in this dimension. Edwards reports that session players (trombonists who specialize in popular music and recording) prefer instruments described by adjectives toward the bottom of the diagram. Symphony players prefer those in the upper left corner. Problems arise in positioning points due to scant data, and also because some words have more than one meaning depending upon application. For example, "*dark* is a good symphonic quality, but a bad quality for some session players. *Edgy* can mean rasping and bad, or *bright* and good" ((p. 418). This may explain some of the surprising inconsistencies in the diagram. For example, *huge* and *enormous* appear close to *small centered* and *small bore* and nearly opposite from *big. Euphonium-like* is near *tuba-like* but far from *baritone-like.*

Be that as it may, Edwards' solution seems to suggest that trombone timbre can be adequately described in a small-dimensional space using purely subjective judgements, perhaps well-chosen verbal scales. It will be shown in the following section that attempts to refine and validate these verbal scales have failed.

3.3 Timbre Dimensionality: Descriptive Space via SDS and VAME

A parallel approach attempts to isolate the dimensionality of timbre description. Verbal scales are investigated for applicability to a set of stimuli. Since free-form verbal descriptions, such as those found in orchestration manuals, have been found to lack precision, more specific techniques are utilized. Semantic differential scaling (SDS) and magnitude estimation employ verbal descriptors in a quantitive framework, in hopes of greater precision and validity. In SDS, a listener is asked to rate a stimulus along a scale specified by polar semantic descriptors. Verbal attribute magnitude estimation (VAME) is similar in form, but addresses the problem of establishing antonymous endpoints for the scales. In SDS, a listener may be asked to rate a sound on the scale *dull* to *bright*; in VAME, to rate on the scale *not bright* to *bright* or *0* to *bright*. Finally, correlations and factor analysis are applied to ratings to determine scales that consistently and accurately describe the stimuli.

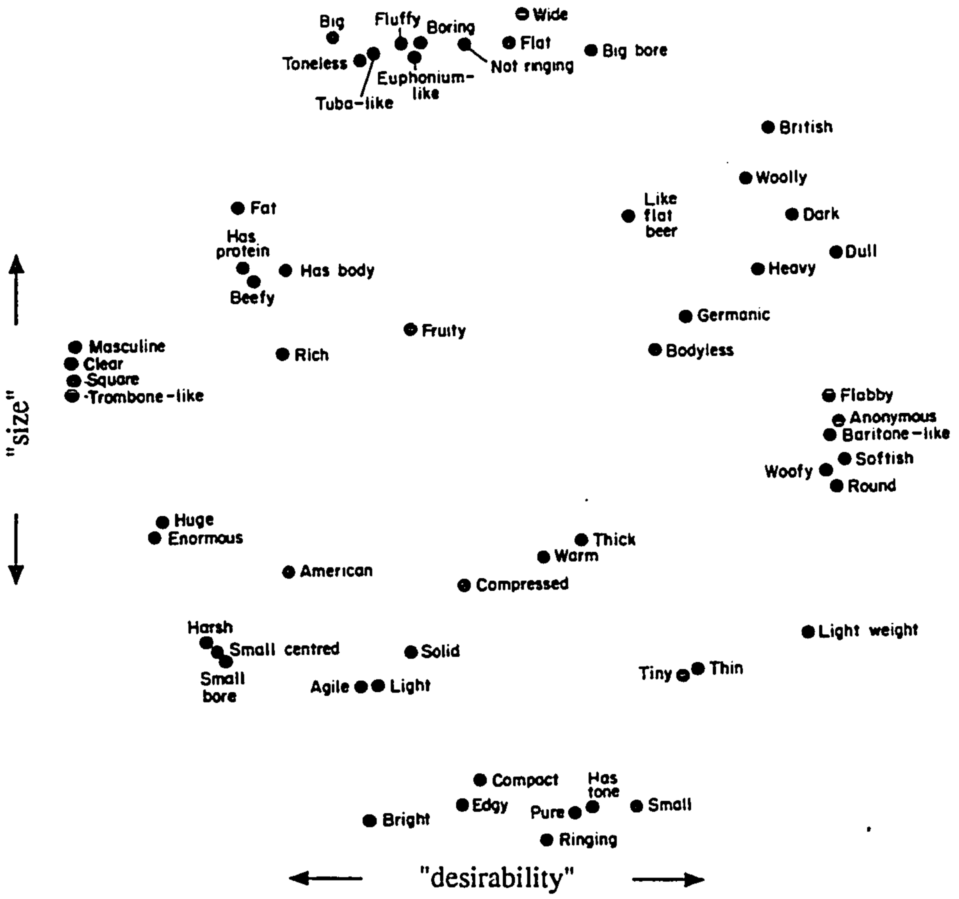


Figure 3.1: Edwards' (1978) two-dimensional scaling of adjectives used to describe trombone timbre.

The success of these techniques depends upon careful selection of descriptors. Stimuli must differ in the dimension that the scale represents, which is to say that scales must be matched to the stimulus. As an example, Kendall and Carterette (1993B) applied VAME to soprano wind dyads, and found that the attribute *reed* was salient. It is doubtful that the same would be found for trombone timbres. Additionally, scales should capture as many salient timbral differences as possible.

Pratt and Bowsher (1978) examined the use of semantic scales for assessing trombone tones. The scales *dull* to *bright*, *not penetrating* to *penetrating*, *compact* to *scattered*, and *unpleasant timbre* to *pleasant timbre* were selected, based upon work by Pratt and Doak (1976), von Bismark (1974a), Grey (1977) and Edwards (1978, 1979). They found that some listeners can reliably discriminate between instruments and performers using the semantic scales *dull* to *bright*, *not penetrating* to *penetrating*, and *unpleasant timbre* to *pleasant timbre*. They further found that "the instrument has the greatest effect on timbre; the player has a measurable effect, but no listener could discriminate between the two different mouthpieces when played on the same instrument" (p. 430). In subsequent experiments they found that pitch and loudness have a greater effect on ratings than either instrument, player or mouthpiece.

They concluded that "the use of SDS to rate the subjective characteristics of trombones quantitatively is clearly feasible" (p. 434), and that the most important acoustical factor governing one professional trombonist's preference for instruments is timbre. However, results indicate that, at least for this choice of stimuli and scales, SDS is of limited utility. Subjects' ratings differed significantly, and had to be analyzed separately. Furthermore, some subjects did not discriminate between trombones or trombonists using the scales. While SDS seems to capture important tonal differences, it does not seem to lend itself to consistent use and interpretation.

3.4 Identification of Trombone Tones

Instrument identification has sometimes been used as a criterion for timbre perception. Clearly, identification indicates perception of at least some characteristics necessary for synthetic listening. Kendall (1984) has made the point that it additionally requires pre-existing verbal categories, and that failure to identify does not necessarily imply that timbral characteristics are not perceived. Furthermore, identification is probably dependent upon a subset of timbral characteristics. Discrimination between trombones and trombonists may well be based on other attributes.

Using identification as the criterion, complete trombone tones are quite easily discriminated from those of other instruments. Clark, Luce, Abrams, Schlossberg and Rome (1963) found that recorded trombone tones were identified with an accuracy of 75%; furthermore, all misattributions remained within the brass family. Clark and Milner (1964), using musically experienced subjects, obtained a 92% identification accuracy, again with all misattributions in the brass family. The most difficult comparisons are between members of the brass family, under certain conditions of range and dynamics (Clark et al., 1963).

Clark et al. (1963) also determined that, for single tones, the attack transient was of preponderant importance. Identification was unaffected by removal of portions of the steady state; tones as short as .06 seconds were named as accurately as complete tones. However, removing the onset caused identification to plummet to 28%.

These findings impact the present study. If the transient truly is essential for identification, then timbre perception cannot be discussed without taking it into account. On the other hand, if the same information can be extracted from a number of quasi-steady states, then this holds greater promise since it is easier to describe a number of steady states than an onset. Transients will be discussed in greater depth in Chapter 6.

3.5 Summary

Perception of trombone tones appears to take place in a low-dimensional space. Description of trombone tones is also low-dimensional. Identification can be used as an indicator of timbre cues. For single tones, the onset transient strongly affects identification.

Verbal descriptions are informative for both listeners and performers. Descriptions which can be used to discriminate between trombone tones are perceived ***size***, ***dull*** or ***dark*** vs. *bright*, and *penetrating* vs. *not penetrating*. However, since they are used idiosyncratically they are ill-suited as a general tool for trombone timbre description.

Chapter 4

OBJECTIVE DESCRIPTIONS OF TROMBONE TIMBRE

In the previous chapter, subjective attributes of trombone timbre were presented. Physical attributes of brass tones will be discussed in this chapter. As noted previously, since timbre is by definition a subjective phenomenon, it is incorrect to regard any physical measurement as a direct measure of timbre. However, aspects of timbre undoubtedly correlate with some combinations of physical attributes. By relating perceived timbral changes with observed physical changes, we can develop an understanding of what constitutes trombone timbre and how it can be described.

4.1 Descriptions of Trombone Tones

Sound results from air pressure fluctuations at rates between 20 to 20,000 Hz. Pressure changes bring about oscillations in the middle and inner ear organs, which are converted into nerve impulses and eventually perceived as sound. A microphone can be substituted for the tympanic membrane, and the electrical signal it produces monitored as an indicator of sound.

Figure 4.1a shows a microphone response to a trombone note. An onset, a relatively steady section, and a decay are apparent. A small portion of the steady-state is shown in greater detail in Figure 4.1b. Axes show air pressure (Pascals) as a function of time (milliseconds; 2,500 ms = 2.5 seconds). A complicated pattern is observed which repeats after approximately 5 ms. Since the ordinate of the plot is time, this is a *time-domain* representation. The description is complete, in that it accurately describes pressure at the microphone for the entire sample. However, perceived sound is not well represented in this manner. It is easy to create pressure series which sound alike but whose time-domain representations differ drastically (Strong and Clark, 1966).

Human ears generally respond to musical tones as frequency analyzers; that is, they represent complex pressure signals as combinations of simple oscillations at

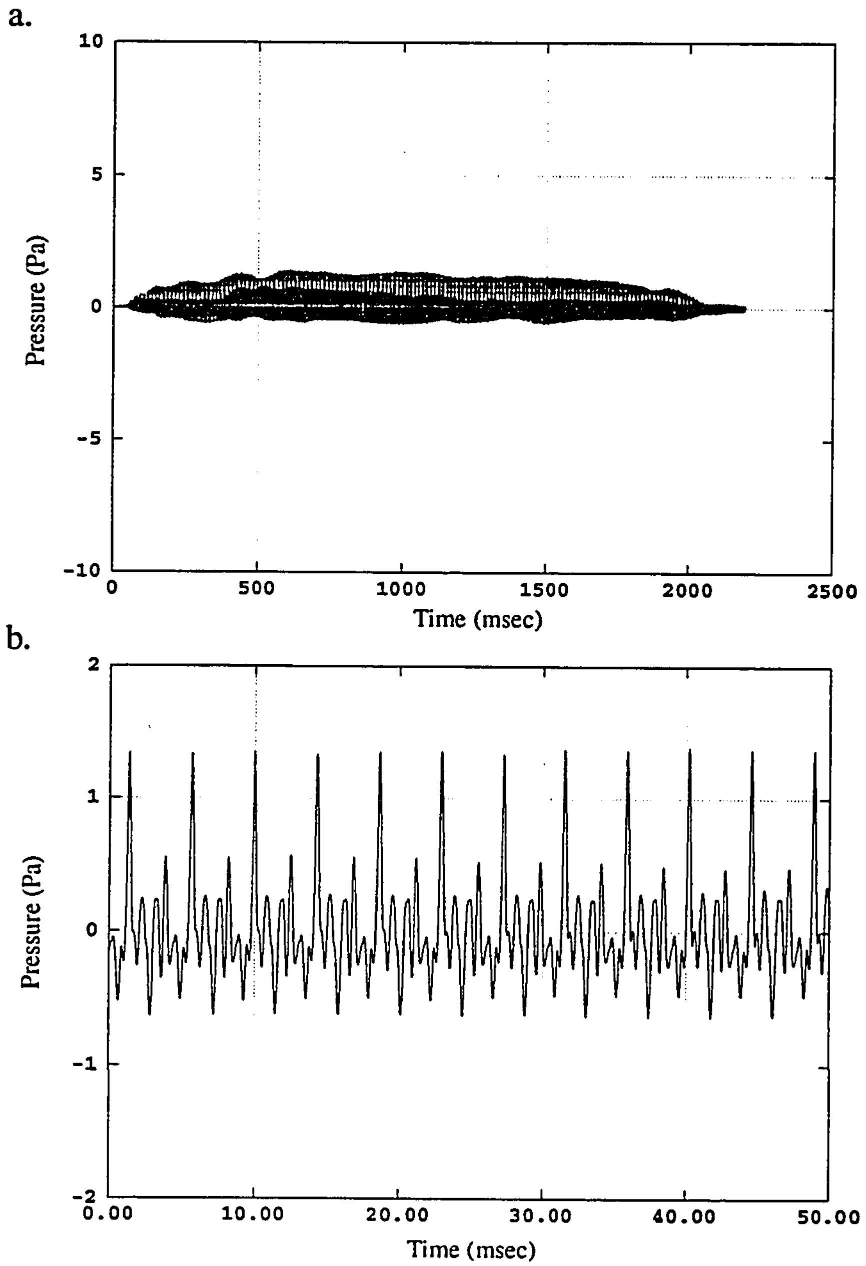


Figure 4.1: Time-domain representations of the note B♭3. a. Complete tone; b. Small portion of steady-state.

various frequencies. A similar effect can be produced mathematically by applying a Fourier transform to a time-domain representation. Figure 4.2a shows the results of a Fourier transform of the steady portion of the trombone tone shown above. The ordinate is now in units of frequency; this can be called a *frequency-domain* representation. It shows peaks at multiples of 207 Hz, which is a good representation of how our hearing system responds to harmonic tones (von Helmholtz, 1885/1954). A Fourier representation (with both magnitude and phase information) is also complete, and the original time series can be reconstructed if so desired.

Musical tones are rarely static; at the very least they have beginnings and ends. A number of means are available for combining time-domain and frequency-domain descriptions to capture these changes. One of these is a short-time Fourier transform (STFT), which can be thought of as a discrete Fourier transform applied successively to small sections of a sound. Figure 4.2b shows the results of an STFT applied to the first 0.25 seconds of a trombone tone. The left-to-right axis shows frequency, the front-to-back shows time, and the vertical axis shows sound pressure level in dB. It can be seen that each harmonic has a complicated amplitude contour in time.

All three of these formulations have been applied to trombone tones. The results will be examined for characteristic features that may contribute to timbre perception.

4.2 Time-Domain Descriptions

From the work of von Helmholtz (1885/1954) to the present, timbre has been described primarily in the frequency domain. One of the reasons for this is that waveforms whose time-domain representation differs markedly can sound the same to auditors. Strong and Clark (1966), for example, synthesized several trumpet-like tones with identical spectral amplitudes but very different waveforms, by manipulating relative phases of components during synthesis. Auditors were unable to distinguish them. Waveforms also change when spectral amplitudes are varied (rather than just phases), but the changes are more likely to be audible.

Conscious manipulation of performance variables can change the waveform. Elliot and Bowsher (1982) observed that the pressure waveform in a trombone mouthpiece varies when a performer attempts to play with a *dark* or *bright* tone quality. Audibility of these changes was not reported.

Studies stressing the timbral importance of transients often examine short-term

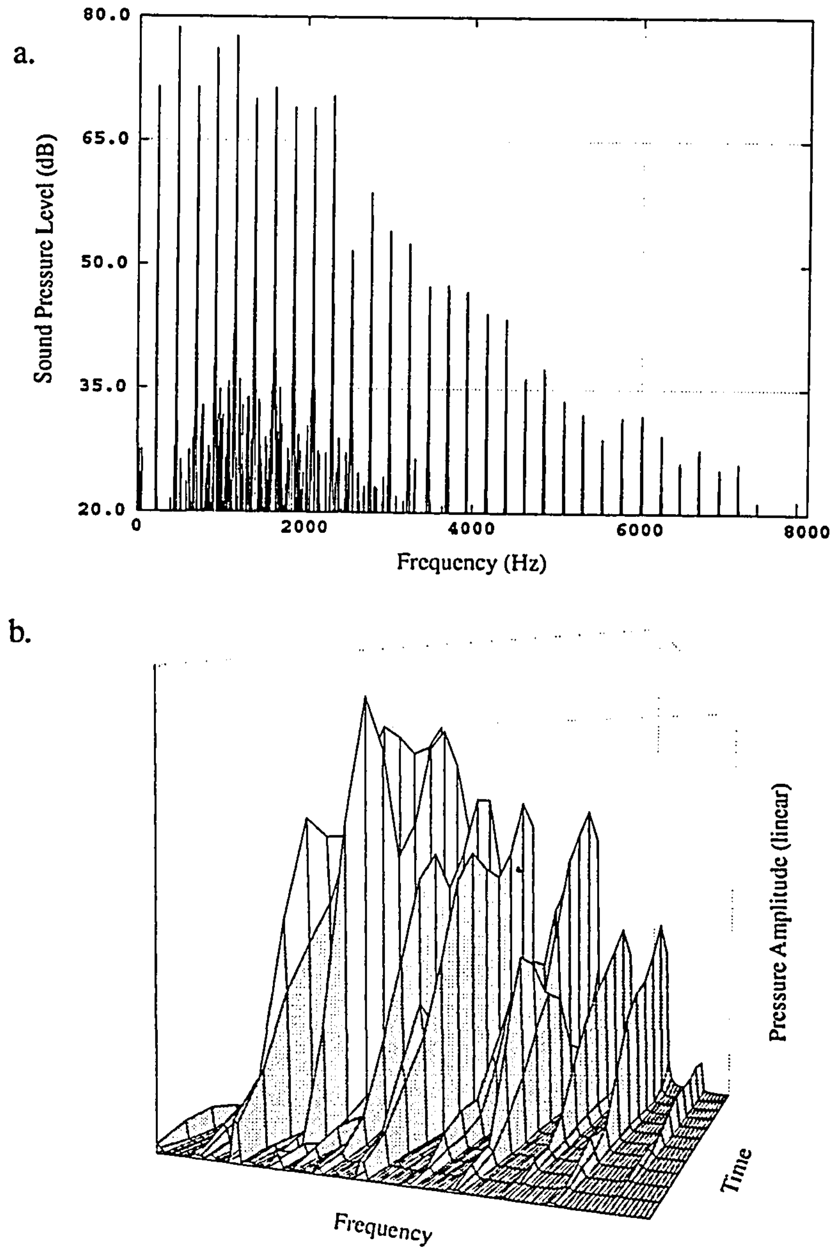


Figure 4.2: Frequency-domain representations of the note B♭3. a. Steady state; b. Attack transient.

variations in frequency-domain descriptions, for example via STFT's (Strong and Clark, 1966a; Luce and Clark, 1967; Beauchamp, 1975, 1982; Grey, 1977). The attack transient for brass instruments is carefully described by Luce and Clark (1967). The transient of the fundamental is about seven cycles, independent of note frequency. The transients of harmonic partials can be divided into those below about 400 Hz and those above it. The lower partials tend to reach their steady-state values at about the same time, and before those of the higher partials. For partials above 400 Hz, the transient period increases with frequency of the partial, tending to about 70 ms. for the highest. The amplitude of these higher partials also decreases with increasing frequency. Short (about 25 ms.) amplitude-modulated "blips" may appear in various partials during the transient. These blips occur throughout the range of the instrument, and increase with intensity of the note.

The perceptual significance of the attack transient will be discussed in Chapter 6. While a great deal of information is contained in combined frequency and time-domain representations, no concise way of describing them has yet been devised.

The release transient appears to have little effect on timbre perception (Clark et al., 1963).

4.3 Frequency-domain Descriptions: Spectral Envelopes

Many authors have reported spectral measurements of steady-state trombone tones, although a comprehensive collection of spectra along the lines of Moorer, Grey and Strawn (1978) has not appeared. As a first approximation, relative spectral amplitudes at can be described in terms reminiscent of a low-pass filter: "The spectrum during the steady state may be simply and approximately described by giving the rate of rise below the cutoff frequency, the cutoff frequency, and a rolloff rate above this frequency" (Luce and Clark, 1967, p. 1240). These are indicated in Figure 4.3 for a representative tone.

Luce and Clark (1967) observed that tenor trombone tones exhibit a cutoff frequency in the range of 500-600 Hz. For a given performer, the rolloff rate above cutoff depends primarily on dynamic level. This permits the calculation of an averaged spectral envelope which applies to tones at a given dynamic.

Table 4.1 shows values for rise below cutoff, cutoff frequency, and rolloff above frequency for a number of instruments, performers, and dynamic levels, as reported

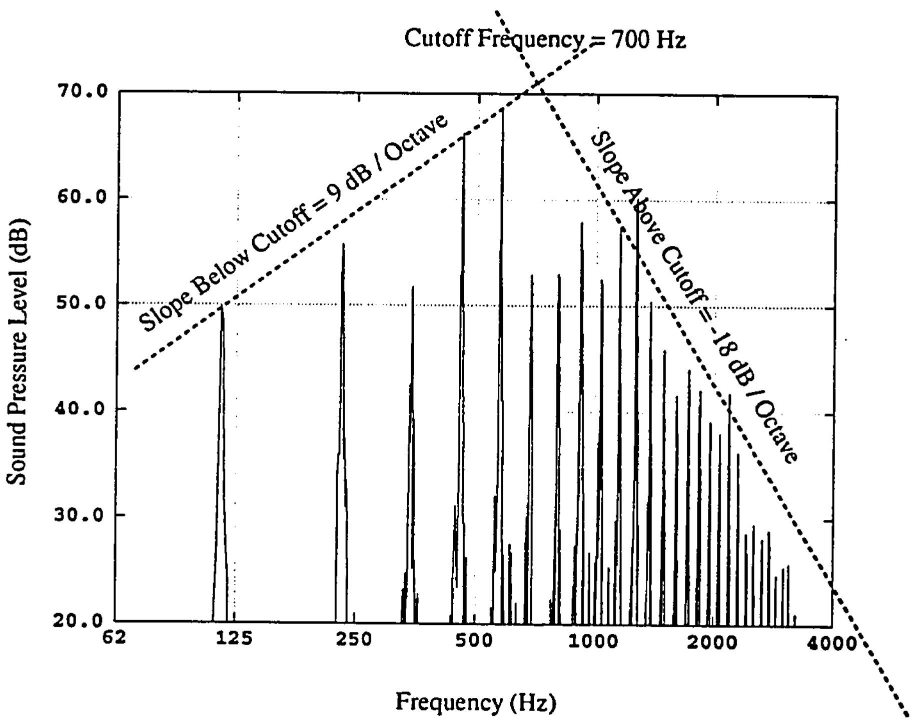


Figure 4.3: Spectrum of a trombone tone, showing spectral cutoff frequency and line segments showing the slope of the spectral envelope above and below cutoff.

by Luce and Clark (1967). Cutoff frequency is similar for orchestral instruments of the same voice type (Benade and Kouzoupis, 1985). Rolloff rates vary with dynamic level, but are related to both instrument and individual performer. Since the envelope represents averaged behavior it omits detail. For example, trombonists are sensitive to "duff" notes found on some instruments (Pratt and Bowsher, 1978). Spectral characteristics of these notes would be smoothed by averaging, and would have minimal effect on a spectral envelope.

Table 4.1: Brass family cutoff frequencies and rolloff rates, at various intensities (After Luce and Clark, 1967).

Instrument & Intensity	Cutoff Frequency	Rolloff Rate Above Cutoff
Trumpet 65 dB SPL	1000 Hz	-18 dB / Octave
Trumpet 75 dB SPL	1000 Hz	-9 dB / Octave
Trombone 55 dB SPL	500-600 Hz	-22 dB / Octave
Trombone 70 dB SPL	500-600 Hz	-17 dB / Octave
French Horn 50 dB SPL	500 Hz	-20 dB / Octave
French Horn 70 dB SPL	500 Hz	-15 dB / Octave
Tuba 60 dB SPL	300 Hz	-25 dB / Octave
Tuba 75 dB SPL	300 Hz	-18 dB / Octave

Luce and Clark (1967) present spectral envelopes for two trombonists using their own instruments, performing at *pp*, *mf*, and *ff* dynamics. These are shown in Figure 4.4. Each curve reflects averaged harmonic amplitudes at a specific intensity. An increase in radiated power is achieved primarily by an increase in the amplitudes of high-frequency partials. Results from the two trombonists differ noticeably, both in rolloff rates and cutoff frequency. Meyer (1978) gives spectral envelopes for tenor and bass trombones at a single dynamic level. The most pronounced difference between them is that bass trombone shows a higher rolloff rate above cutoff.

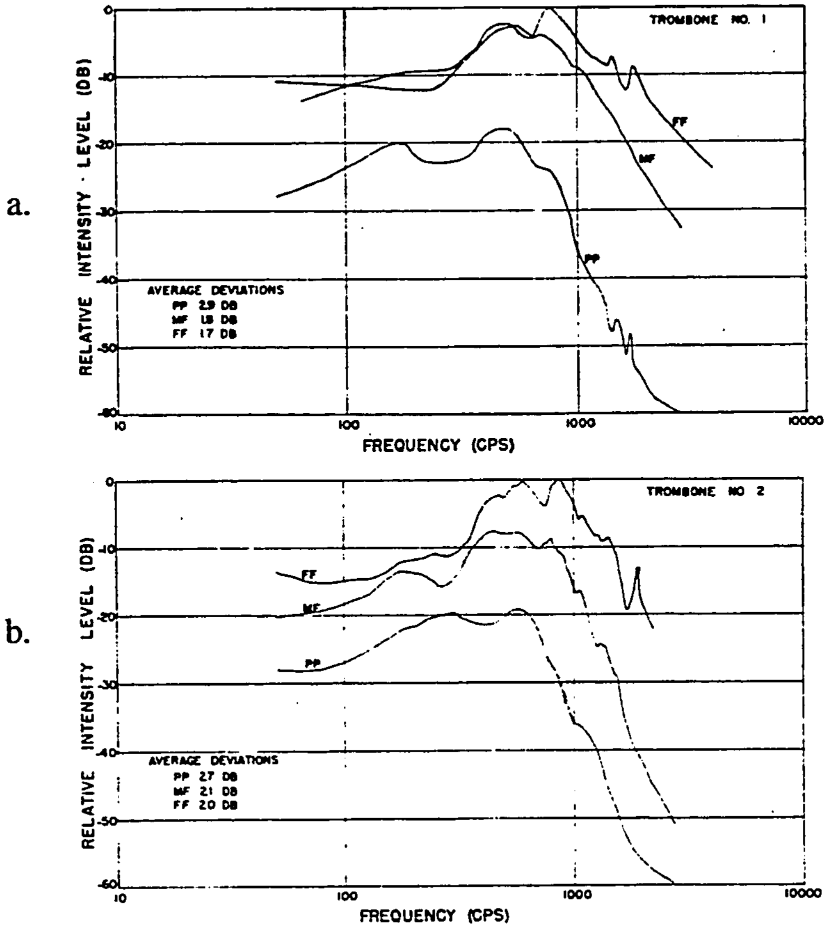


Figure 4.4: Spectral envelopes for two trombonists, at three dynamic levels (Luce and Clark, 1967). a. Perfomer 1; b. Performer 2.

Changes in cutoff frequency produce audible timbral changes. Campbell (1994a) had listeners identify tones with varying spectral cutoff frequencies. She found that just-noticeable differences (*jnd's*) in cutoff frequency depended on spectral slope above cutoff. For a slope of -24 dB per octave (comparable to the trombone at *pp*) jnd's were approximantely 4-5% of the fundamental frequency. For a slope of -6 dB per octave (comparable to trombone at *ff*), jnd's were 24-26% of the same. Jnd's for low-pass filter slopes also depend on level (Zwicker and Fastl, 1992).

Changes in the slope above and below cutoff also produce audible timbral changes. Campbell (1993/1994) reports that spectral slopes above cutoff of -6 to -24 dB/octave corresponded to about 15-20 jnd's. Since the slope above cutoff strongly influences sharpness and brightness predictions via Equation 2, Campbell concludes that her findings "support recent work highlighting the primacy of perceived sharpness in timbral judgements" (p. 58).

To summarize, steady-state trombone tones can be roughly characterized by a spectral envelope reminiscent of a low-pass filter. Differences between performers and instruments are reflected in spectral envelopes. Parameters are cutoff frequency and slopes above and below cutoff; perceived timbre is sensitive to changes in any of these. Slope above cutoff depends strongly upon overall intensity. We may surmise that a measure based on the relative proportion of high-frequency energy as a function of intensity level would capture both general characteristics of trombone tones and individual differences as well. Aures' calculation of brightness reflects changes in this proportion.

4.4 Frequency-domain Descriptions: Formants

In contrast to Luce and Clark, Meyer (1978) emphasizes the presence of peaks and dips in his spectral envelopes, referring to them as *formants.* A formant is a fixed resonance which strongly affects radiated sound. Vocal timbre is often described in terms of formants, and vowel sounds each have a characteristic formant configuration. Meyer's spectral envelopes for tenor and bass trombones are reproduced in Figure 4.5. Formants are indicated by the smooth bumps in the curve below 2 KHz. Meyer locates the first formant at 520 Hz for tenor trombone, and at 370 Hz for bass trombone.

Peaks and dips that resemble formants are indeed present in trombone spectral envelopes. However, vocal tone production differs from trombone playing in that the vocal chords oscillate relatively autonomously, while trombone embouchures are tightly coupled to the air column. Formant terminology is based upon a source-resonator model of tone production which is not easily applied to trombones. An additional difficulty is that while formants describe perceived similarities in vowels, this is not the same as describing vocal timbre. As Benade puts it, "We might not be willing to say that the singer produces the same color when he sings *ah* at the bottom of his range as he does at the top of it" (1976, p. 377).

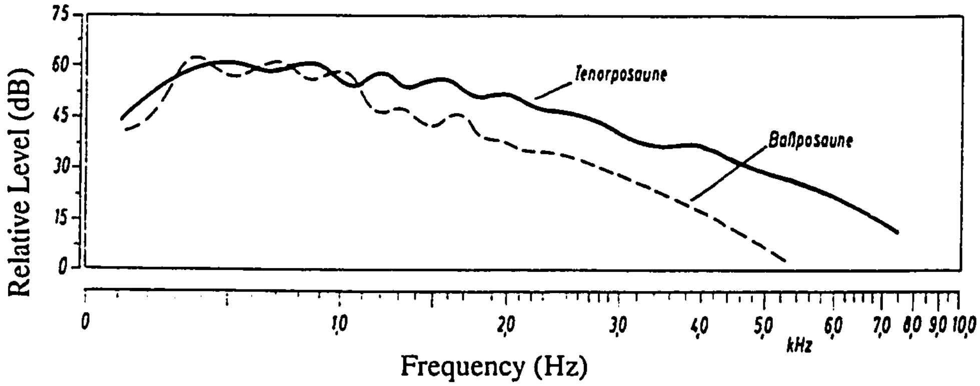


Figure 4.5: Spectral envelopes for tenor and bass trombones, at a single dynamic level (Meyer, 1978).

4.5 Summary

Three types of objective descriptions have been applied to trombone tones. Time-domain descriptions are not appropriate for timbre studies, since changes in waveforms do not always correspond to audible variations. Frequency-domain descriptions include spectral envelopes. These are relatively compact and precise, although they represent averaged behavior. Formant descriptions have been applied to trombone spectra, but are poorly suited to trombone sound production. Additionally, combined time- and frequency-domain descriptions provide a great deal of information about perceptually important transients. However, a compact description of transients has not been developed.

Steady-state trombone spectra can be conveniently described based upon spectral envelope characteristics, namely the cutoff frequency and slopes above and below cutoff. Perceptual correlates of cutoff frequency and slope are summarized in Aures' brightness measure.

Chapter 5

TROMBONE TONES AND PSYCHOACOUSTICS

Since timbre is multidimensional, there may be perceptual attributes of greater subjective importance than brightness. The utility of brightness measurements depends not only upon how well it describes one particular attribute, but upon how important that attribute is compared to others. Three questions must be answered: (1) how many subjective dimensions can be discerned, (2) what is their relative importance, and (3) with what objective features are they associated? Based on this knowledge, brightness measurements can be placed in a broader framework.

Because trombone timbre is of interest almost exclusively to trombonists, the literature dealing with it is not extensive. By placing trombone tones within the larger context of timbre perception, however, additional information is available, and findings may have broader application.

The field of psychoacoustics is dedicated to relating subjective perceptions to objective quantities. The following is an admittedly eclectic review of research into the perception of musical and quasi-musical tones from a psychacoustic perspective.

5.1 Psychoacoustics and Timbre

Psychoacoustics seeks to establish connections between subjective perceptions and objective quantities. Since timbre is acknowledged to be a multi-dimensional attribute (Plomp, 1970) this can be a complicated procedure. How can one establish relationships between sound features and perceptual and cognitive correlates without *a priori* decisions about their number and forms?

A helpful paradigm has been developed by Plomp and his co-workers (Plomp, 1970, 1976; Klein, Plomp and Pols, 1970; Bloothooft and Plomp, 1988). Three kinds of relationships are established and compared. A *perceptual space* for tones is derived from similarity ratings, which are subjected to multi-dimensional scaling. A second kind of perceptual space, not necessarily identical to the first, is derived from semantic differential ratings and called *semantic space*, since it reflects not only how tones are perceived, but how they might be described verbally. Finally, differences between objective qualities of the tones are expressed in a *spectral space.* By a procedure of rotation to maximal congruence, the tones' coordinates in their respective spaces are compared, to determine what relationships are suggested.

The strength of this procedure is its ability to isolate and correlate perceptual, semantic, and objective aspects of tones with few assumptions. One of its drawbacks is that results for one set of sounds may not apply to another. As Plomp (1976) wrote,

> Based upon a specific set of stimuli, three factors alone appeared to be sufficient to describe the differences satisfactorily. This number cannot be generalized. If we had started from a set of tones differing only in the slope of their sound spectra, a single factor would have been sufficient. It is also possible to select nine stimuli which would require, for example, five dimensions to represent their timbres appropriately. (p. 197)

An additional drawback is the daunting amount of data that must be collected. Multi-dimensional scaling algorithms require comparisons whose numbers increase exponentially with the number of stimuli. It is improbable that musical timbre space can be spanned with a practically-sized set of stimuli.

Most of the studies mentioned here only partially ascribe to this paradigm. Several, for example, establish perceptual or semantic spaces for their tones; coordinates in this space are then correlated with pre-selected physical descriptions (e.g., Iverson and Krumhansl, 1993; Grey, 1977; Wessel, 1979). While perhaps less rigorous than Plomp's method, it has the advantage of acknowledging informed intuition as to what may be important features. It also permits consideration of the onset transient, whose complex nature defies easy description.

It should be noted that in practice the number of dimensions and their relative importance are related questions. The number of dimensions is often determined by a ranking of their importance; the experimenter simply truncates the list when the benefit of including more dimensions is outweighed by increased complexity.

5.2 Timbre Dimensionality: Perceptual Space

As described in Chapter 3, similarity ratings of musical tones can be analyzed with multidimensional scaling techniques in order to estimate the number of dimensions necessary to describe them. In view of Plomp's warning about the dependence of findings on stimuli, we will examine results by stimulus type.

A number of investigations have been reported on replicated-period vowel sounds (Bloothooft and Plomp, 1988; Pols, 1970; Klein, Plomp, and Pols, 1970). All of these found that the prediction of timbre with third-octave averaged spectra is valid (indeed, that "spectral analysis with a bandwidth smaller than a critical band [about a quarter of an octave] makes no sense." (Pols, 1970). All agree that between one and five dimensions are needed to characterize vowel sounds, but that three is generally sufficient. The earlier studies in this group found that perceptual dimensions correlated well with spectral dimensions representing formant frequencies and levels (Pols, 1970; Klein, Plomp, and Pols, 1970). This was surprising since the tones did not sound vowel-like and few listeners suspected their derivation. More recently Bloothooft and Plomp (1988) found that the most important spectral dimension of both stationary vowels and song phrases is associated with the perceptual dimension of sharpness. It alone accounted for some 45% of total spectral variance. They also found that the semantic attribute of *Shärfe* is the only dimension on which most listeners could agree in their judgements. The differences between earlier and later studies can be ascribed to the fact that sharpness was formulated as a spectral attribute only in 1974 (von Bismarck, 1974b). To conclude, a few dimensions are sufficient to describe steady-state vowel and vowel-like sounds. Perceptual and semantic attributes are associated with either formant detection or sharpness detection; more recent work indicates the latter.

A second group of studies utilized simplified musical instrument tones developed by Grey (1975). The tones were generated by replacing intensity and pitch envelopes for individual harmonics with line-segment approximations, and equalizing for perceived loudness, pitch, and duration. The tones are short (280 - 400 ms), and it has been noted that the lack of a steady state may have influenced findings (Kendall and Carterette, 1991). Multidimensional scaling of similarity ratings has indicated either two salient dimensions (Wessel, 1979), or three salient dimensions (Grey, 1975, 1977; Grey and Gordon, 1978; Gordon and Grey, 1978). In all cases, one dimension is associated with spectral energy distribution, while the remaining dimensions are associated with transient phenomena. Grey and Gordon (1978) investigated spectral quantities for correspondence with the first dimension, and found that the centroid of the loudness function was most successful in approximating it.

A third group of studies, using musical or quasi-musical tones, largely confirm Grey and Gordon's findings. Iverson and Krumhansl (1993) investigated perception of complete tones, equalized for loudness; they found two important perceptual dimensions. The first of these correlates with the centroid of the loudness function, and the second relates to the nature of the onset transient. Plomp (1976) investigated steady-state tones derived from musical instruments, and found three important perceptual dimensions. The first of these correlates well with sharpness. Kendall and Carterette (1991) used dyads of complete tones, equalized for loudness but much longer than Grey's. They found two stable dimensions, which they labelled *Nasal/Not Nasal* and *Rich/Brilliant.* They suggested that nasality "may be closely related to sharpness." (p. 380). Ehresman and Wessel (1978) used fifteen synthesized tones, and found two dimensions, related to the distribution of spectral energy and characteristics of the onset transient ("Brightness" and "Spectral Variation").

It appears that timbre perception can be modelled in a low-dimensional space. Investigations with both steady-states and complete tones have confirmed that one important dimension correlates with spectral distribution. The correlation has been established with the distribution expressed in terms of spectral slope, centroid of spectral amplitudes, centroid of loudness function, and sharpness.

5.3 Timbre Description: Descriptive (Semantic) Space

Can timbre *description* be similarly modelled in a low-dimensioned semantic space? As discussed previously, the prevailing paradigm for examining this utilizes semantic differential scaling or verbal attribute magnitude estimation. The studies discussed below have all used musical and quasi-musical tones.

Von Bismarck (1974a) applied SDS to a carefully constructed set of tones. Stimuli were synthesized complex tones and noises which varied systematically in spectral slope, formant structure, spectral gaps, and bandwidth. Listeners rated tones on 30 scales. It was found that ratings could be expressed as linear combinations of four factors which together accounted for 91% of the variance. The most important factor alone accounted for 44% of the variance, and was described by the scale *dull–sharp*. The second factor, accounting for 26% of the variance, was described by *compact–scattered* and "was clearly used . . . to differentiate between noise and tonal stimuli" (p. 157).

Other studies have supported the importance of something like a *dull–sharp* axis. Pratt and Doak (1976), for example, used only three carefully-chosen scales, and found that "subjects were able to discriminate between certain sounds of varying harmonic content, using the Dull/Brilliant scale with the greatest reliability" (p. 328). Bloothooft and Plomp (1988) had subjects rate vowel sounds and song phrases on 21 scales. They found that "most semantic scales are used according to sharpness." (p. 854). For all subsets of stationary vowels, "only sharpness turned out to be a verbal attribute of timbre on which most listeners . . . agreed in their judgments . . . . In conformity with von Bismarck, sharpness was found to be acoustically related to the slope of the spectrum" (p. 858).

Only Kendall and Carterette (1991; 1993a,b) question the usefullness of sharpness for describing musical tones. In an extensive set of experiments, they investigated perceptual, semantic, and spectral attributes of simultaneously-sounding instrument dyads. In the perceptual phase (1991) they first found that a two-dimensional solution accomodated their data satisfactorily, and observed that the most consistent dimension "may be closely related to 'sharpness' of von Bismarck" (1991, p. 380). In order to confirm semantic correlates for their perceptual dimensions, they had subjects rate dyads on eight semantic scales selected from von Bismarck (Kendall and Carterette 1993a). Listeners failed to differentiate between instrument dyads using these scales. Additional experiments using semantic scales drawn from a standard music textbook showed that four factors accounted for over 90% of the variance: *power*, *strident*, *plangent*, and *reed* (1993b). In view of this, they concluded that their two-dimensional perceptual space could be interpreted in terms of semantic dimensions of *nasal* (negative) versus *rich* positive and *reedy* (positive) versus *brilliant* (negative).

Kendall and Carterette's conclusion that von Bismarck's scales are inappropriate for soprano wind dyads can be understood in a number of ways. They suggest that von Bismarck's adjectives "lack sufficient connection with musical culture (practice and performance) to correlate highly with timbre space" (1993a, p.466). Their own list of adjectives, drawn from specifically musical sources, obviates this problem. Another explanation is that difficulties were introduced by translation of adjectives (German to English), especially as related to non-antonymous endpoints for bipolar scales. It could also be that perceptual spaces for dyads are not the same as those for single tones. Most likely, this is simply another illustration of Plomp's dictum that perceptual and semantic spaces depend upon stimuli. It is not surprising that ratings of musical tones were inconsistent on the scale *compact-scattered*: von Bismarck found that this scale served to differentiate between noises and tones. On the other hand, it is difficult to imagine *reedy* applied to trombone tones. It should also be noted that in the experiments utilizing von Bismarck's scales, none of the subjects were trained musicians. In all studies that have used listener training as an independent variable, significant differences have been found between musicians and non-musicians; musicians make more reliable judgements (Bloothooft and Plomp, 1988; von Bismarck, 1974; Miller and Carterette, 1975; Kendall and Carterette, 1991; Wedin and Goude, 1972).

To conclude, studies have shown that semantic description of musical tones, like perception, can be accommodated within low-dimensioned spaces. For steady-state tones, the most significant semantic dimension is described by *sharpness* or *brightness* or *nasality*. This dimension appears related to the relative amount of high-frequency energy to low-frequency energy, whether expressed in terms of spectral slope, ratio of energy in selected frequency regions, or formant characteristics.

5.4 Objective Correlates of Timbre

A number of means have been utilized to relate objective characteristics of tones to perceptual and verbal dimensions. Dutch researchers have developed a very general technique, in which steady-state tones are described as vectors in an 18-dimensioned space of third-octave levels (e.g. Bloothooft and Plomp, 1988). In this way, the entire audible spectrum of their stimuli are described to approximately the degree of resolution of the human hearing apparatus. Others have described their stimuli by spectral peak levels (Wedin and Goude, 1972) or critical band levels (Kendall and Carterette, 1993b). In all of these studies, spectral vectors are matched to the tones' coordinates in perceptual and/or semantic spaces using principle components analysis or rotation to maximum congruence.

Coordinates in perceptual spaces can also be correlated with pre-selected physical descriptions. For example, Grey (1977) examined similarity ratings through both multi-dimensional scaling and heirarchical clustering techniques. Plots of spectral amplitudes as a function of time were then examined in light of suggested groupings. This permitted interpretation in terms of combined attributes, such as spectral aspects of onset transients. Von Bismarck (1974b) used tones that varied systematically, permitting him to establish directly that one perceptual dimension was related to both spectral slope and bandwidth.

In a number of studies using complete, naturalistic tones, it was concluded that important perceptual dimensions are related to transient phenomena. The most influential work has been that of Grey and Gordon (Grey, 1977; Grey and Gordon, 1978; Gordon and Grey, 1978), who found that two of three dimensions were connected with transients. The first was synchronicity of attack and decay of higher harmonics, and the second was spectral fluctuation. Wessel (1979) and Iverson and Krumhansl (1993) also described perceptual dimensions related to onset transients. Transients will be discussed more thoroughly in Chapter 7.

All studies have shown the perceptual importance of spectral distribution. However, there is not a consensus as to whether this is best described in terms of spectral slope or absolute frequency location of the amplitude envelope. The latter is prominent in studies that have used vowell-like stimuli, since it facilitates interpretation in terms of formants. However, it is also found in studies influenced by Lichte's (1941) article positing "brightness", defined as "a function of the location on the frequency continuum of the midpoint of the energy distribution" (p. 479), as an independent attribute of complex tones. This formulation, also called the spectral centroid, appears in the work of Grey (1977), Beauchamp (1982), and Iverson and Krumhansl (1993).

Spectral slope or harmonic content may provide a more general description. Bloothooft and Plomp (1988) found that voice classification depends on spectral slope, and suggested that their subjects behaved as "sharpness detectors" rather than "formant detectors." Campbell (1993/1994) examined the absolute identification of tones that differed in fundamental frequency, cutoff frequency, and spectral slope above cutoff (expressed as filter order). She concluded that the consistency of subject's responses "lends strong support to the notion of the primacy of sharpness in judgments of the steady-state timbre of signals with low-pass spectral characteristics" (p. 52). Other proponents of the spectral slope viewpoint are are von Bismarck (1974) and Miller and Carterette (1975).

5.5 Models of Timbre Perception

It is a complicated task to take the suggestive but inexact relationships described above and formulate an objective measure. On the one hand, a calculation scheme should make sense in terms of what is known about hearing. On the other hand, the values it produces must correlate well with appropriate perceptual and verbal dimensions. Finally, while the scheme must apply to the specific tones in question, generality is also desired. A number of attempts have been made to do exactly this, which can be evaluated on how well these criteria are met.

A two-dimensional scheme has been developed for vowel and vowel-like sounds, based upon the first two formant frequencies and levels. Configurations correlate highly with perceptual dimensions (Klein, Plomp and Pols, 1970). As previously mentioned, a recent study of vowel-like timbres has concluded that sharpness or brightness may be a more basic perception (Bloothooft and Plomp, 1988).

It has been observed that the frequency location of the mid-point of spectral energy (*spectral centroid*) correlates with perceived brightness of sounds and may be a useful measure (Lichte, 1941). It has furthermore been shown that this timbral brightness is "not similar to pitch brightness in pure tones", but an independent property of complex tones (op. cit.). A number of schemes for calculating this have been utilized. Beauchamp (1982) derived an expression for a "center of gravity or brightness (BR) of a spectrum vector" which is referenced not to frequency but to harmonic number. It represents "the harmonic number at which the area under the spectral envelope ... is 'balanced' " (p. 397). Beauchamp found that synthesized tones could be made more realistic by matching the time-varying BR of musical instrument tones. While clearly a useful quantity, BR does not incorporate many of the known attributes of hearing, such as the frequency-dependence of loudness and sharpness. Helmholtz, for example, observed that pure tones have different timbres according to tessitura (1954). By Beauchamp's calculation, all pure tones have an identical BR.

Grey and Gordon (1978) investigated a number of different formulations for correlation with their spectral dimension. These included both centroid and median of the amplitude pattern for (1) line-spectrum in linear units (vs. frequency); (2) line-spectrum in decibels (vs. frequency); (3) excitation pattern (vs. critical band rate), and (4) loudness function (vs. critical band rate). They found that "the most successful of the set of models employed here was the centroid of the loudness function" (p. 1496), which correlated at better than 0.9 with their spectral dimension. Grey and Gordon conclude that it is

> an adequate representation of spectral energy distribution, in that it simultaneously takes into consideration the many factors which may be important: overall bandwidth, balance of levels in the lower harmonics, and the existence of strong upper formants. Furthermore, it attempts to model the distribution perceptually. (p. 1498)

The loudness function is a sophisticated model which incorporates many of the known attributes of human hearing. It adjusts levels of the spectral components to reflect hearing sensitivity. It incorporates the concept of critical bands for calculating the contributions of individual harmonics, and models masking. However, there are at least three difficulties that should be examined before applying this measure to musical tones.

The first of these is that, while the loudness function weights spectral features in light of perception, it is not known whether a centroid is the appropriate way to express their respective contributions. The ear may function not as a centroid detector, but as a formant detector (Klein, Plomp and Pols, 1970) or sharpness detector (Bloothooft and Plomp, 1988).

A second problem is found in the interpretation of loudness function centroids. There is no evidence to indicate what values represent perceptual differences, or what kinds of scalings pertain. For example, is the difference between centroids at one and two barks comparable to the difference between centroids at seven and eight barks?

A final difficulty arises when calculating centroids of musical tones. The critical band rate scale, which forms the ordinate of the loudness function, is essentially logarithmic. This means that a complex tone with harmonic overtones will have widely-spaced lower harmonics, and closely spaced higher harmonics. The loudness function calculates contributions to total loudness from each critical band; components that lie in the same critical band contribute less to total loudness than those in other bands (Sharf, 1978). Harmonic tones do not have more than one harmonic per critical band until above the sixth or seventh partial. The calculation of a centroid based on this kind of spacing implies that on the low-frequency side there will be just a few harmonics each of which strongly influences the centroid, while on the high-frequency side there will be many whose individual contribution is less. We also know, from statistical study of room acoustics, that measured levels of individual components vary randomly in a diffuse field (Lindevald, 1987). In the higher frequencies, where there may be several components in a critical band, these variations largely average out. However, in lower bins containing fewer components, variations will strongly affect calculation of the centroid. Centroids calculated from tones recorded in rooms will exhibit large fluctuations based on source and microphone placement, as was documented by Goad (1994, June).

Von Bismarck's sharpness (1974b) addresses the three weaknesses of the centroid described above. Its form is based on psychophysical measurements. It is perceptually scaled, so that equal ratios of values imply equal differences in sensation. Finally, its formulation, based on a weighted sum rather than centroid, is less sensitive to measurement variability.

5.6 Summary

It has been shown that timbre perception can be modelled in a low-dimensioned space. For musical instrument tones, two or three dimensions are sufficient to account for similarity judgments; for vocal tones, a few more appear necessary. It has also been shown that semantic description of musical tones can be reduced to a low-dimensioned space. The most significant semantic dimensions for steady musical tones are described by *sharpness*, *brightness* or *nasality*. The most influential perceptual dimension is related to the relative proportion of high-frequency energy in the spectrum. This is true of the most influential semantic dimensions as well.

Models for relating perception to spectral distribution have been examined in light of the objective characteristics of trombone tones. Formant descriptions are more appropriate for vocal tones than trombone tones. The various formulations of spectral centroid (centroid of linear spectral amplitude, centroid of logarithmic spectral amplitude, centroid of loudness function) cannot be justified on psychoacoustic grounds and are prone to measurement variability.

From convergent lines of inquiry, it has been shown that that the proportion of high- to low-frequency components is a likely candidate for a concise description of trombone timbre. In Chapter 4 it was shown that patterns of variation in this proportion are characteristic of the brasses, and distinctive of individual instruments and performance techniques. It has been shown that important perceptual dimensions are related to this proportion, and that useful semantic scales can also be related to it. In other words, the proportion is likely to vary with trombone tones in ways that are perceptually and cognitively significant. Aures' formulation of brightness appears the most promising predictor of perceived brightness.

Since the relative proportion of high and low frequencies in trombone tones varies with both pitch and loudness, predicted brightness as a function of both pitch and loudness may provide a useful description of trombone timbre.

Chapter 6

THE ATTACK TRANSIENT: CRUCIAL TO TIMBRE PERCEPTION?

A number of researchers have proposed that the onset of a tone, the *attack transient*, is more important to perception than the steady state. For example, Luce and Clark (1967) noted that the relative amplitudes of harmonics in the steady state are affected by many factors, and that it is "unlikely any quantity sensitively dependent upon the performance variables ... could be of first order significance ... in characterizing the timbre" (p. 1233). They concluded that "any adequate representation of orchestral instrument tones must contain information about the attack transient" (p. 1232).

This raises a number of questions. The first is, why is it that trombonists seem less interested in attack transients than steady states? Secondly, what features of attacks are essential, and and what psychological functions do they influence? And finally, is the attack transient necessary for an adequate representation of trombone timbre?

6.1 Performers' Views of Onsets and Tone Quality

All brass players do, of course, attend to the beginnings of their notes. What is remarkable in light of Clark et al. (1963) is that they don't pay more attention. The term *tone quality* is understood by trombonists to be a characteristic of sustained tones, and long-tone exercises are prescribed to remedy deficiencies (Kleinhammer, 1963).

The literature contains many references to the importance of a distinctive, personal tone:

> The quality of a musician's tone shows up in everything he plays. It may be a beautiful, projecting, radiant tone or it may be boisterous, obnoxious, and crude in its personality. The goal of every musician is a tone of the

> purest quality within the realm of his personal concept. (Kleinhammer, 1963, p. 36)

> One must have a mental conception of his own "ideal" tone. This is acquired by constant listening ... Decide for yourself if this player's tone isn't too bright, that one's a little too muffled, etc. Gradually you will come to have very definite ideas about which tone is just right. (Farkas, 1956, p. 52)

The beginnings of notes are usually dealt with under the rubric of articulation, which for trombonists is primarily a problem of tonguing. Treatment of articulation tends to be perfunctory and practical:

> There are only two general types of articulation — slurring and tonguing — for any series of notes. (Farkas, 1956, p. 49)

> The use of the tongue is generally given too much emphasis by teachers. ... The function of the tongue in detached playing is to assist the air and embouchure in producing a clean attack (beginning of a tone). (Kleinhammer, 1963, p. 63)

Trombonists' descriptions of trombone tones contain fewer references to attacks than to steady-state timbre. Edwards' study (1978) discriminated between *sound quality* which "is the quality of sound produced by the instrument when it is playing a steady tone", and *responsiveness* which is "how easily an instrument starts to produce a note, and how easily it changes from one note to another, or from one sound to another" (p. 415). Edwards found 55 adjectives for sound quality but only 22 for responsiveness.

6.2 Possible Reasons for Lessened Interest in Attacks

It is difficult to reconcile the apparent importance of the attack for listeners with performers' opinions that a distinctive and personalized steady-state is paramount. While resolving this inconsistency is outside the scope of the present study, a number of explanations can be ventured.

It may be that transients receive less attention because they are so brief. Luce and Clark (1967) found that the transient lasts from 5 to 10 cycles of the fundamental in the lower octave, and tends to 70 ms for the highest notes. This is too brief for conscious intervention by the performer. So, while the attack is influenced by the performer and important for the listener, brass pedagogy may simply de-emphasize it in favor of aspects that are more directly controllable.

Alternatively, it may be that transients receive less attention from performers because their features are determined by the instrument. Fletcher and Rossing (1991) outline calculations indicating that the initial transient extends over "something like 10 cycles of the horn fundamental, or about 50-100 ms, depending on instrument size" (p. 387). Elliot, Bowsher, and Watkinson (1982) showed that features of the impulse response of the instrument can be directly related to reflections from the bell and internal features of the trombone. An impulse response is conceptually the response of the instrument to an idealized impulse, and determines the "shape" of the pressure signal that drives the lips; it can be expected to strongly influence attack transients.

Finally, it may be that transients are important to performers, but that they respond to the same performer interventions as steady states. The literature emphasizes, for example, that tone quality responds to changes in embouchure and air supply: "The quality of a musician's tone. ... involves embouchure and breath control in the case of a brass player" (Kleinhammer, 1963 p. 36). Attack transients depend on similar quantities: "A fine attack consists of the preparation of the embouchure, the metering of the air, and assistance of the tongue, all in correct timing" (Kleinhammer, 1963 p. 63). A fine attack, then, may be produced by the same actions as a fine tone quality; but it is easier to discuss and work with tone quality.

6.3 Psychological Functions of Attack Transients

We turn now to perceptually important features of the attack, and the psychological functions that they influence. The topic has been investigated from a number of perspectives, and a consensus does not yet exist on many aspects.

At least two paradigms have been used to study relationships between attack transients and timbre. In both, the independent variable consists of stimulus characteristics, such as which portion of a natural or synthetic tone is presented, or what parameters determine how a synthetic sound is generated. In one paradigm, the dependent variable is instrument identification. It is assumed that correct identification by musically literate subjects implies the presence of critical timbral attributes. Another approach utilizes similarity ratings as the dependent variable, applying multidimensional scaling techniques to the results. A search for physical correlates of resulting dimensions may then be undertaken.

As Iverson and Krumhansl (1993) have pointed out, identification and similarity judgements may be based on different information.

> Similarity judgements rely on the comparison of gross acoustic attributes, but identification judgments may rely on acoustic attributes that are more informative of their source. Acoustic characteristics of the onset may be uniquely important for determining if an instrument was struck, blown, or bowed, but gross acoustic differences may be less informative. (p. 2602)

In Western European art traditions, instrument identification is a valued skill, as can be surmised from any American elementary-school music curriculum. However, identification is rarely crucial to musical experience. Timbre discrimination is usually more important, since it facilitates tracking of a voice in a complex texture. Luce and Clark (1967) seem to believe that failure to name a sound implies failure to discriminate it, which may be incorrect.

Iverson and Krumhansl (1993) had subjects rate similarity for portions of tones. They found that

> Ratings for complete tones corresponded to those for onsets, indicating that the salient acoustic attributes for complete tones are present at the outset. Ratings for complete tones also corresponded to those for remainders [tones with the onsets removed], indicating that the salient attributes for complete tones are present also in the absence of onsets. (p. 2595)

They attributed the pattern of similarity to spectral centroid frequency and to the amplitude envelopes of the onsets. They concluded that "the dynamic attributes of timbre are not only present at the onset, but also throughout, and that multiple acoustic attributes may contribute to the same perceptual dimensions" (p. 2602).

Beauchamp (1982) found that his calculation of spectral centroid BR (referenced to harmonic number rather than frequency) was useful in synthesis of brass-instrument tones. Realistic syntheses resulted from matching instantaneous BR. In earlier work, Beauchamp (1975) showed that it was possible to synthesize realistic-sounding cornet tones, independent of dynamic and including transients, by using the original first-harmonic envelope to generate other harmonics.

Since Clark et al. (1963), a number of studies have investigated parts of tones with instrument identification as the dependent variable, but have only partially supported their conclusions. For example, Strong and Clark (1966a, 1966b) synthesized tones with interchanged temporal and spectral envelopes. They found that identification accuracy for altered tones depended on specific spectral envelope characteristics. When envelopes were unique, they played a predominant role in identification. When they were not unique, for example, for trombone, flute and French horn, "the spectral envelope is equal or subordinate to the temporal one in aural significance" (1966a, p. 277). Thayer (1972) exchanged attacks and steady-states between trumpet, oboe, flute and clarinet. He found that removal of transients adversely affected identification, but the effect was relatively weak. For example, music majors' identification of trumpet tones was 59.6 percent for normal tones, 50.61 percent for tones without transients, and 45.13 percent for tones with transients from a different instrument. He concluded that

> The trumpet attack ...does reinforce the information provided by the the trumpet steady-state in identification ... The steady-state by itself or in combination with other attacks is easily confused or identified as the attacking instrument, and the trumpet attack with other steady-states does not provide as much information as one might anticipate. (p. 94)

Instead of relying on identification of tones, which is a learned skill, Kendall (1984) had listeners match unedited recordings of instrument tones with manipulated examples. Stimuli consisted of single tones as well as entire folk song phrases. His edit conditions included normal (complete) tones, tones with onset and release transients omitted ("time-varying steady state"), transients alone, and static steady-state with and without transients. To obtain the "static steady-state" Kendall spliced together copies of a single waveform from the steady state. He found that, for whole musical phrases, transients are neither sufficient or necessary for categorization. For trumpet, the time-variant steady-state is sufficient and necessary for categorization. However, for single-note stimuli, statistically equivalent means were produced with complete tones, time-varying steady-states, and transients alone. Categorization is better in whole-phrase contexts than in single-note contexts.

It appears that attack transients contribute significantly to instrumental timbre. For single notes on some instruments, including the trombone, they provide as much or more information pertinent to identification than the steady state. However, Kendall's work suggests that time-varying steady-states are nearly as informative. When presented in a number of notes, time-varying steady states allows listeners to categorize sounds about as well as normal, complete tones.

6.4 Summary

One can dispute Luce and Clark's claim that any adequate representation of orchestral instrument tones must contain information about the attack transient. Transients and steady-states may influence different kinds of judgements, and one is not always more important than the other. It is especially interesting to note lessened importance of transients in Kendall's study, since his time-varying steady-states include pitch and loudness changes that are ubiquitous in music.

The attack transient may be of unique importance only for determining the excitation of a given instrument, for example, whether it is struck, blown, or bowed (Iverson and Krumhansl, 1993). For comparisons of trombones and trombonists, it may be no more informative than other measures. The difficulty of concisely describing the onset further makes it a less promising measure than brightness.

Chapter 7

MECHANISMS FOR HARMONIC GENERATION IN THE TROMBONE

Trombone tones, like those of other wind instruments, are characterized by partials in harmonic relationships to the fundamental. The distribution of energy in these harmonics determines measured brightness. This chapter is devoted to describing physical mechanisms for the production of harmonics.

The mechanisms take the form of theories or models that relate physical characteristics to spectral content. Such models are based on simplifications and approximations that may or may not be valid under performance conditions. Since a rich performance-oriented literature already exists that describes, in pragmatic terms, the factors that influence brass timbre, it may be asked what contributions can a physical perspective make to the understanding of trombone timbre?

As mentioned previously, one of the primary goals of trombonists is developing a distinctive personal "sound." For this reason their descriptions are likely to emphasize unusual and distinctive aspects of sounds, rather than commonalities. The approach is more analytic than synthetic, to use terminology introduced earlier. A different perspective may be better suited for discerning elements that contribute to synthetic listening. Balzano has suggested that synthetic listening for timbre is "a matter of perceiving underlying dynamics of physical processes" (1986, p. 312). Thus, a physical perspective may suggest what the essential elements of trombone timbre are, and how performer and instrument influence them.

In this chapter the dynamics of tone production in the trombone will be outlined, along with their spectral results. It will be shown that an increase in high frequency spectral energy with loudness is characteristic of brass instruments. Brightness is appropriate for documenting this. It will also be shown that some kinds of differences between instruments are likely to be reflected in brightness. Some performer actions also impact brightness.

7.1 Physical Measures: Input Impedance and Pressure Transfer Function

Before dynamics can be discussed, some basic quantities and relationships must be defined. Input impedance is an important measure of a trombone's response characteristics. It is defined as the complex ratio of pressure to volume velocity, when both are measured at the input plane. It can be thought of as the pressure response to a unit flow, as a function of frequency. Input impedance is useful for describing how the instrument presents itself to the performer at the mouthpiece, and excels as a "single measure of the acoustic response of an instrument because it plays a crucial role in the regeneration of sound at the lips." (Elliot, Bowsher and Watkinson, 1982, p. 1748).

Input impedance does not directly provide information about radiated sound. It describes energy that is reflected back to the measurement plane, and by inference how much is lost, but it cannot indicate whether losses are to sound production or to some other mechanism. For this, a pressure transfer function measurement from mouthpiece to radiated field is useful. The pressure transfer function, like input impedance, reflects the ratio of two measured quantities. They are both pressures, however, and one is measured at the input (mouthpiece) end, whereas the other is measured at the bell. Together, input impedance and pressure transfer function give a useful description of a brass instrument.

Input impedance is a linear measure, since it assumes that flow is proportional to pressure. The pressure tranfer function is also a linear measure. When flow becomes turbulent, linearity can no longer be assumed. Likewise, at very high pressures the tranfer function may not be linear. In general, the models described below are designed to explain typical behavior, rather than behavior at extremes of range and loudness.

7.2 Outline of Tone Production in the Brasses

Tone production in the brasses is a good example of self-sustaining oscillation. In any cycle, part of the energy supplied by the performer is retained to determine how subsequent cycles occur. This can be clarified by considering alternative models.

In brass performance literature, one sometimes finds instruments described as a passive resonators. For example, Leuba (1980) states that the relationship between embouchure and French horn is analogous to that between a horn and undamped strings of a piano into which the horn is played. The instrument resonates sympathetically to vibrations introduced by the embouchure, just as piano strings resonate when components of a tone coincide with their resonant frequencies. The embouchure and instrument behave autonomously: the embouchure produces sound, which the instrument selectively transmits. This model neglects influences of the instrument on the embouchure itself; for example, it cannot explain the difficulty of maintaining an oscillation at any but a few favored frequencies.

More accurate models account for the influence of the air column on the operation of the embouchure. Not only are lip vibration frequencies determined by air column resonances, but the airflow spectrum produced by the embouchure and admitted into the instrument is closely determined; hence, the radiated sound is affected. As Benade (1976) stated,

> The player's lips have a motion that is very strongly influenced by the acoustical properties of the air column to which they are connected. That is, the mechanical shape of a trumpet ... has a direct influence on the shapes of the puffs of air which enter its mouthpiece. The brass instrument's air column has a dual function, then: not only does it (like the vocal tract) transmit sound components selectively from the flow source to the room, it also plays a large role in determining the nature of the incoming flow pattern itself. (p. 391)

The embouchure functions as a valve that admits air flow in such a way that air column resonances are "nourished." These resonances, in turn, drive the valve so that its operation is stable and consistent. Large pressure peaks are developed in the mouthpiece corresponding to resonance frequencies. These resonant frequencies appear as peaks in input impedance plots. A small portion of the energy trapped in these resonances is transmitted to the room as a characteristic brass instrument tone. Another portion is used to drive lip motion and feed vocal tract resonances. The remainder—by far the greatest part—is lost to viscous and thermal wall losses, turbulence, and extraneous mechanical vibrations. Benade and Gans (1968) estimate that 75% of total energy dissipation is due to viscous and thermal damping, while the remaining 25% is divided between vocal tract, reed, turbulence, and radiation losses. Seen in this light, the role of the brass instrument appears to be more to trap energy than transmit it. The retained energy ensures that the portion that is radiated makes up a musically useful sound—one that has stable pitch, consistent tone color, and wide dynamic range.

The timbre produced by a performer/instrument system is influenced by some variables that depend upon only the instrument, and other variables relating to just the performer. In general, however, it is the interaction of performer and instrument that determines timbre. A simplified diagram showing the elements of tone production and their relationships is shown in Figure 7.1. The arrows can be thought of as representing the flow of energy, in terms of air pressure and flow. At the left-hand side of the figure is the box labelled "Player's Windway," which encompasses the lungs as well as the vocal tract. The embouchure is depicted simply as a valve; it is driven both by the windway as well as the instrument, by way of the box labelled "Embouchure Dynamics." The box marked "Instrument" receives energy from the embouchure, some of which it passes along to be radiated and some of which is returned to the embouchure. This outline will be used to organize the discussion of timbre production. In deference to the interactive nature of this system, discussion will begin near the middle, at the box labelled "Embouchure Dynamics."

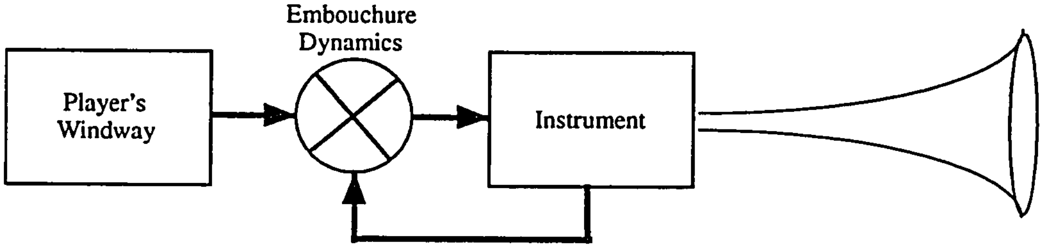


Figure 7.1: Elements of tone production in trombone.

7.3 The Role of the Embouchure

Since the work of Helmholtz in the 19th century, brass players' lips are thought to form a reed of the *outward-blowing* type; that is, they open and permit flow when the pressure gradient across them is positive (Helmholtz, 1885/1954). This is in contrast to clarinet and saxophone reeds, which are *inward-blowing* and close with increased pressure. If one begins with an assumption of a higher pressure in the mouth than in the mouthpiece, it is difficult to see how an outward blowing reed can sustain oscillation. Resonance is achieved when additional flow is introduced into an already high pressure; the defining characteristic of an outward-blowing reed is that it admits flow when the mouthpiece pressure is lower than the mouth. This apparent difficulty is resolved by taking into account the resonant characteristics of the embouchure. The lips can be modelled as a damped spring and mass system which is driven into motion by the oscillating pressure in the mouthpiece. By properly arranging the driving frequency (for example, by adjusting the length of the instrument) and resonant frequency of the embouchure (by tensing facial muscles or other manipulations), the necessary phase relationship between mouthpiece pressure and incoming flow can be established (Campbell and Greated, 1987). The conditions for oscillation are quite strict; the resonant frequency for the embouchure must be closely set and there is only a small frequency range over which oscillation is possible.

The lips have been observed to vibrate smoothly, just closing once per cycle, and presenting an aperture whose area varies smoothly with time (Martin, 1942). If pressure in the mouth is greater than in the mouthpiece (a reasonable assumption), it would seem that the sinusoidal motion of the lips would result in a sinusoidal variation in mouthpiece pressure. This would result in the radiation of a pure tone, with no harmonics. Brass tones are generally not at all like pure tones. The embouchure in fact is a highly non-linear valve, producing energy not only at its oscillation frequency but also at harmonic multiples. Components at harmonically-related frequencies are thus a by-product of embouchure dynamics. The relative amplitude of these components largely determines timbre.

Elliot and Bowsher (1982) have modelled the interaction between embouchure and air column, and how it influences the mouthpiece pressure spectrum. They identified two regimes. The first is marked by a lip opening that is large for a significant part of each cycle, so that the pressure in the mouthpiece rises to be comparable with that in the mouthpiece. In theoretical terms, this corresponds to the case where Z_n, the instrument's input impedance at the nth harmonic, is much greater than R_0, which is the average resistance of the aperture. In this case pressure as a function of time

$P(t)$ is approximated by

$$P(t) = \sum_{n}[-P_1/4P_s]^{(n-1)}P_1 \cos n\omega t \tag{7.1}$$

where n is the harmonic number, P_1 is the pressure magnitude at the fundamental and P_s the pressure from the lungs, and ω is the radian frequency of the fundamental. Equation 7.1 shows pressure in the n^{th} harmonic growing as the pressure at the fundamental to the n^{th} power, leading to a rapid development of the spectrum. This regime is called the *low-pitch case*, since it is chiefly in the lower register that the aperture is large enough for this model to apply. Interestingly, the spectrum does not depend upon the details of the instrument's input impedance. In the words of Elliot and Bowsher,

> For notes of low pitch, where the lip displacement is large, the mouthpiece pressure has the characteristic harmonic structure given in Equation 7.1, which is independent of the impedance of the instrument, provided it is large at all the harmonic frequencies ($Z_n \gg 2R_0$). (p. 195-196)

The *high-pitch case* is encountered when the lip aperture is relatively small, so that ($Z_n \gg 2R_0$) cannot be assumed. The high-pitch case applies for high notes, especially at soft dynamics. While the pressure spectrum cannot be as conveniently described as in the low-pitch case, the levels of harmonics are generally lower and the harmonic structure depends upon details of the input impedance.

To summarize, then, the lip can be modelled as a pressure-operated valve functioning as a simple harmonic oscillator. Regeneration is possible when the resonant frequency of the embouchure is placed in a closely determined relationship with an air column resonance. Once this condition has been satisfied, the non-linear flow characteristics of the lip-valve itself determines harmonic structure in the low-pitch case. In the high-pitch case, harmonic structure develops more slowly with pressure, and structural details are dependent not only on the valve but on the instrument by way of its input impedance. Note that, since aperture size is directly related to blowing pressure and hence loudness (Martin, 1942), the loudness of a tone is correlated with the amount of influence the instrument can exert on the radiated spectrum.

7.4 The Role of the Instrument

Performers are convinced that an instrument influences timbre. As mentioned previously, trombonists are often able to aurally discriminate between notes played by the same performer but using a different instrument (Pratt and Bowsher, 1978). Sound quality has also been found to be one of the principle criteria used by trombonists when evaluating an instrument (Edwards, 1978).

A trombone influences the mouthpiece pressure spectrum through a number of mechanisms. It influences the mouthpiece spectrum by way of its input impedance. When the average resistance of the aperture is greater than the impedance of the instrument (the high-pitch case), "the amplitude of each sound-pressure component within the mouthpiece is proportional to the height of the air-column input impedance curve measured at the frequency of the component" (Benade, 1976, p. 442).

Radiated sound is also influenced by the instrument, through the pressure transfer function. External pressure can be calculated by multiplying mouthpiece pressure by the transfer function (Benade, 1976).

Finally, it is likely that turbulence within the instrument affects radiated sound, by selectively damping certain frequencies. No accepted measure currently exists for this non-linear property.

7.4.1 Influence of the Instrument on the Mouthpiece Spectrum

The nonlinear flow characteristics of the embouchure produce a mouthpiece spectrum with precisely harmonic components, locked together in magnitude and phase through their common source. Benade (1976) has convincingly argued that a satisfactory musical instrument has input impedance peaks arranged in such a way that a number of harmonics are supported by impedance peaks in a cooperative "regime of oscillation." The optimal arrangement places the impedance peaks in a harmonic relationship. Modern trombones have input impedance peaks that are very nearly harmonic, excluding the fundamental which is too flat. Figures 7.2, 7.3, and 7.4 show input impedances of representative alto, tenor, and bass trombones.

The plots are similar in many ways. Each shows a series of sharply defined peaks and dips. If the frequency axis for alto trombone is rescaled to reflect its shorter length, peaks for all trombones lie at nearly the same frequencies. With increasing

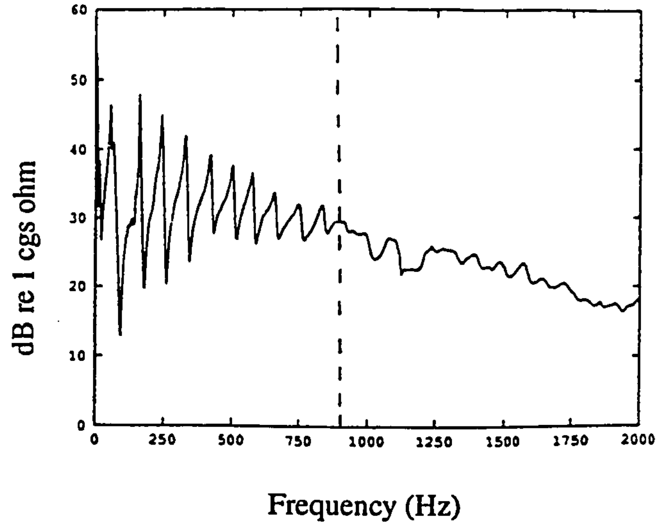


Figure 7.2: Input impedance level of an alto trombone. Dashed line gives approximate input impedance cutoff frequency.

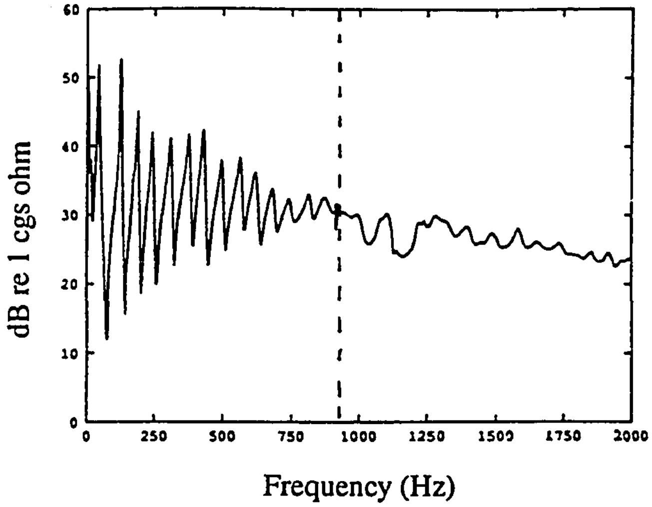


Figure 7.3: Input impedance level of a tenor trombone. Dashed line gives approximate input impedance cutoff frequency.

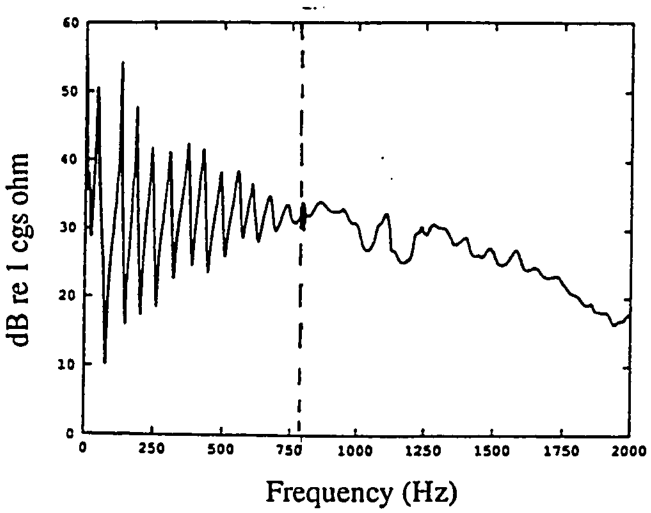


Figure 7.4: Input impedance level of a bass trombone. Dashed line gives approximate input impedance cutoff frequency. [1]

frequency, peaks generally decrease in magnitude, and dips increase, until at about the twelfth harmonic the peaked pattern largely disappears. This is called the *cutoff frequency* of the input impdance (Benade, 1976). Below cutoff frequency, energy is largely trapped in the instrument to assist in regeneration. Components above cutoff are not retained by the instrument, and are radiated more strongly. The height of each impedance peak reflects the degree to which the instrument retains energy at that frequency, although the extent of retention depends upon whether the high-pitch or low-pitch case is dominant. The mouthpiece pressure spectrum is affected by height and alignment of impedance peaks, as well as by cutoff frequency. It has been shown that altering the position of even a single impedance peak on a trumpet can affect the timbre of some notes (Smith and Daniell, 1976), and that a small change in cutoff frequency (2-3%) is readily audible in clarinets (Benade, 1976).

One way of describing both height and degree of alignment of impedance peaks is through the *sum function* (Pratt and Bowsher, 1979). Wogram (1983) found that

---

[1] Figures 7.2, 7.3, and 7.4 were produced using an adaptation of the calibration procedure described in Keefe, Ling and Bulen (1992). The calibration failed in the immediate vicinity of 500 Hz, due to choice of calibration tubes. Peak levels near 500 Hz are suspect; the *frequencies* of the peaks, however, are accurate.

sum function peaks are good predictors of playing frequency.

7.4.2 Influence of the Instrument on Radiated Spectrum

As stated above, input impedance does not provide all necessary information for predicting the radiated spectrum. It deals only with quantities at the input plane, and provides no information about how transmitted energy is allocated between various loss mechanisms. The pressure transfer function provides a means for transforming mouthpiece pressure into radiated sound. Figure 7.5 shows the measured transfer function for a tenor trombone, along with its input impedance (Campbell and Greated, 1987). Regeneration takes place at frequencies near to input impedance peaks; furthermore, harmonics will be enhanced close to impedance peaks. This means that the mouthpiece pressure spectrum will consist almost entirely of components that are close to impedance peaks. However, dips of the transfer function lie close to peaks of the input impedance. Not surprisingly, the transfer function shows that energy in the mouthpiece spectrum is largely trapped by the instrument, especially at lower frequencies. The new information it provides is a measure of the proportion radiated.

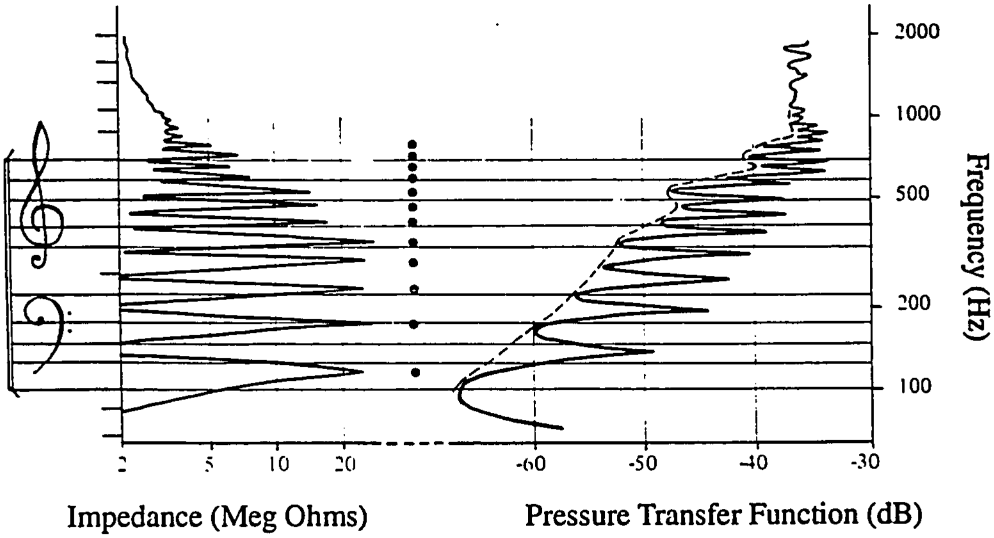


Figure 7.5: Pressure transfer function and input impedance for a tenor trombone (from Campbell and Greated, 1987).

The pressure transfer function's pattern of sharp peaks and dips gives performers another means for influencing timbre. A player can alter playing frequencies by embouchure manipulation ("lipping" a note). This affects mouthpiece spectra by way of input impedance; it also affects radiated spectra through the transfer function. Recall that levels of mouthpiece harmonics decrease as a note is lipped away from its optimal position, since harmonics are not supported by impedance peaks. However, since dips of the transfer function correspond to peaks of the input impedance, harmonics of the lipped note will be transmitted more efficiently. Thus, the two effects act in opposite directions. Benade (1976) asserts that the levels of upper harmonics generally increase when a note is lipped away from its optimal position; May (1980) contends that they are decreased when the a note is lipped down in pitch. Benade and May agree, however, that the effect on timbre is clearly audible.

It is known that playing frequency varies slightly but consistently from player to player (Elliot and Bowsher, 1982; Wogram, 1983). This may account for some of the perceived dependence of timbre upon the individual performer (Pratt and Bowsher, 1978).

To summarize, a trombone influences radiated spectra through several mechanisms. It influences internal spectra through input impedance. The transfer function indicates the proportion of energy from internal spectra that is radiated. The performer has a certain amount of flexibility in determining the nature of the interaction.

7.5 *Effects of the Player's Windway*

While acousticians sometimes dispute the importance of windway resonances to instrument timbre (Backus, 1985), performers "have always insisted upon the importance of getting the proper shapes in a wind-player's air passages" (Benade, 1985). Certainly mouth, tongue, and throat configurations can markedly alter the timbre of trombone tones (Dempster, 1979/1994). However, conscious manipulation of windway resonances for timbral purposes is rarely called for in trombone repertoire. Reasons for this are outside the range of the current investigation. The fact remains that windway resonances can and do influence trombone timbre, and that although rarely fully exploited, they are recognized by traditional pedagogy (Hall, 1955a, 1955b).

Windway resonances vary with shape, dimensions, and wall characteristics of the vocal tract. They are most easily changed by altering tongue position, a technique that is often discussed in brass pedagogy, although as often in relation to obtaining a clean attack as a particular timbre (Fox, 1981). Of course, windway resonances are used in speech to produce vowel formants. They differ somewhat from those used in brass performance, because of different termination (mouth vs. trombone) and because they are driven from the mouth rather than larynx. Fletcher (1983) suggests that formants used when performing the didjeridu (a short, straight, lip-reed instrument with a bore roughly the diameter of a trombone mouthpiece's rim) do not involve the entire vocal tract, but instead "seem to be more analogous to the mouth resonances determining pitch during whistling". The mouth and lip opening make up a kind of Helmholtz resonator, in which the resonant frequency "varies inversely with the square root of the mouth volume and directly with the linear dimension of the opening between the lips or teeth." (p. 37).

Benade (1985) found that windway resonances in clarinet performances had a greater effect on the pressure spectrum in the mouth than on radiated sound, although their effect was perceptible to listeners. The large impedance of the clarinet reed serves to limit the influence of one on the other. Greated and Campbell (1988), on the other hand, have suggested that relatively low impedance of trombone embouchures in the low-pitch case favors transmission of components from mouth to radiated sound. This is congruent with performers' observations, since vocal tract effects are most pronounced in the lower register of the trombone (Dempster, 1979/1994).

7.6 Inferences for Measured Brightness of Trombone Tones

It has been shown that harmonic generation in the trombone, and thus timbre of trombone tones, results from the non-linear valving action of the embouchure. The general consequence is increasing production of upper harmonics with blowing pressure. Thus, a perceptual model based upon the relative proportion of high to low frequency energy as a function of loudness would describe important aspects of trombone tones. Brightness is one possible formulation. Insofar as it reflects underlying physical dynamics of tone production in the brasses, it is a possible cue for synthetic listening of trombone timbre.

Harmonic generation is influenced by interaction between embouchure and air column. The instrument acts through its input impedance. The input impedance cutoff frequency determines the spectral envelope cutoff frequency described in Chapter 4;

it thus directly affects the proportion of high to low frequency energy and hence brightness. Cutoff frequencies of alto, tenor and bass trombones differ by more than the 2-3% which Benade (1976) says is audible for clarinets.

In the high-pitch case, details of the input impedance curve below cutoff are reflected in the radiated spectrum. However, since brightness represents average spectral contour, these will be not strongly influence measurements.

A performer has several mechanisms available for influencing tone color: embouchure impedance, which determines whether high-pitch or low-pitch mechanisms apply; intonation of the tone relative to sum-function peaks; and vocal tract resonances. The performer's ability to alter embouchure impedance is not well understood. Lipping a tone away from its optimal intonation is thought to alter levels of all harmonics, thus affecting brightness. Lipping a tone becomes increasingly difficult at higher loudnesses (Benade, 1976), so these effects may be noticable only at lower loudnesses. Finally, vocal tract resonances may increase brightness by increasing levels of selected harmonics. Since vocal tract resonances are higher in frequency than most trombone tones they rarely affect the fundamental. Benade (1983) asserts that vocal tract resonances are found between 450 and 1500 Hz. However, resonances strongly influence timbre only in the low-pitch case: in the low register and at high loudnesses.

7.7 Summary

It has been shown that the dynamic processes of trombone tone production result in a characteristic relationship between spectral content and blowing pressure. The relationship is influenced by playing frequency. It can be modelled in the perceptual domain by predicted brightness as a function of predicted loudness and pitch. Because of the inherent nature of this relationship, it may form a cue for synthetic listening for trombone timbre.

Differences in cutoff frequency between alto, tenor and bass trombones are likely to result in differences in predicted brightness. Performer differences in intonation and use of vocal tract resonances may bring about variations in predicted brightness, but these are likely to be prominent only in certain regions of pitch and loudness.

Chapter 8

ACCURACY OF LOUDNESS AND BRIGHTNESS MEASUREMENTS

A distinction is sometimes made between measurement precision and accuracy. Precision refers to how exactly a quantity is determined, and is primarily determined by instrumentation and technique. In loudness and brightness measurements as implemented in this study, precision is high. Microphones and recorders are accurately calibrated, and instrument resolution is high.

Accuracy, on the other hand, refers to how closely a measurement approximates a "true" or "correct" value. In many experiments, this involves a consideration of the statistical nature of what is being measured. For example, if one is interested in average human lifespan, one experimental issue is how closely ages are determined: by year, day, minute, etc. This is measurement precision. Accuracy involves the additional issues of error and bias, as well as the degree to which an experimental sample represents a population. Even if individual lifespans are determined precisely, their relationship to human lifespan is statistical, and the resulting lifespan estimate must reflect this.

Both brightness and loudness are calculated from broadband spectra of recorded tones. As is well-known in acoustics, recordings made in rooms are affected by many factors, and any particular spectral measurement may differ from others. It will be shown that spectral measurements have a statistical aspect. In this chapter techniques for spectral measurements will be discussed, along with their influence on accuracy and, therefore, validity.

8.1 Performance Specifications

Since trombones are musical instruments, it is important to capture the tone quality reaching a listener's ear as the result of a musical performance. The role of context must not be overlooked. In brass instrument performance the player has control over nearly all aspects of tone production and "can adjust the regeneration conditions to obtain practically any pitch or tone colour he wants" (Bowsher, 1983, p. 345). Without a musically familiar environment a trombonist is unlikely to perform in a representative fashion. Benade and Larson (1985) have underscored this:

> We wish ... to emphasize in the strongest possible terms that unsuitable specification of the player's task in a laboratory experiment will deprive him of the highly developed functioning of his musically trained auditory and proprioceptive feedback systems. This deprivation inevitably increases the variability of his performance, and thus increases the instability of the resulting analyzed data. (p. 1477)

This precludes, for example, performance in an anechoic chamber, since the sound returned to a musician from an ordinary reverberant space is both influential and customary. Any musician who has recorded in extremely dry acoustical environments can verify the difficulty of performing under such conditions. A suitable environment is moderately reverberant and of appropriate size.

Furthermore, performance specifications must be familiar, and must allow the performer to take advantage of learned musical skills. For example, tones must be of sufficient duration that the performer has time to "zero-in" on optimal adjustments. Benade and Larson suggest that tones are most stably produced "either as single long tones, or as members of a sequence of articulated segments having only the briefest interruptions between them" (1985, p. 1477). Tasks are best specified in musical terminology, such as note names, rhythmic durations and and dynamic markings, rather than frequencies, timings, and sound pressure levels.

The quantitive impact of these considerations is difficult to estimate, and no study has attempted to measure them. Benade and Larson suggest that unmusical specifications contributed to large standard deviations reported by Anastasio and Bussard (1971).

To conclude, playing tasks should be specified in terms that are familiar to a musician. An environment must be provided in which a musician can take advantage of learned adjustment skills, and tones should be long enough that adjustments are possible. When these conditions are met, musicians perform in a consistent manner enabling highly accurate measurements. For clarinet, spectral variability is on the order of 1-2 dB at spectral peak frequencies, reproducible over a period of years (Benade and Larson, 1985).

8.2 Recording Conditions

When sound travels from source to receiver it is modified by its path and the environment in which it travels. In *free field* measurements, reflections do not contribute appreciably to the measured signal, and sound power decreases exponentially with source-receiver separation. Free field conditions exist at moderate distances from the source in very large or very dry spaces, outdoors, and in anechoic chambers. Accurate recording of musical tones is difficult in this kind of environment, quite apart from the previously noted difficulty of obtaining consistent performances. Musical instruments (including trombones) have pronounced frequency-dependent radiation patterns. Free field measurements are highly dependent upon microphone placement.

Musical listening normally takes place in a moderately reverberant space. Suitable rooms are larger in each dimension than several wavelengths of the lowest frequency of interest, and have a reverberation time of 0.5 - 2.5 seconds. In such rooms, reflections follow complicated paths from many surfaces before they die away. At some point, so many reflections are encountered that a modal description becomes appropriate.

In a modal description, stable locations where pressure waves arrive in phase are seen as pressure antinodes, and where they are out of phase as nodes. At very low frequencies, modes are widely separated and nodal regions are large. With increasing frequency the number of room modes increases rapidly and nodal patterns become smaller. It is easy to conceptualize the effects of source and microphone placement in such a field. An intricate three-dimensional system of nodes and antinodes is present throughout the room. If a source is located near a pressure antinode, maximal energy transfer to the room will take place; if near a node, minimal. Likewise, if a microphone is placed near an antinode, measured pressure response will be high, but if near a node, response will be small. Complicating the description is the fact that any number of modes are simultaneously superimposed. Thus, a particular source and microphone placement may exhibit transmission characteristics very different from nearby locations. Benade and Larson (1985) mention that differences of 30 dB are not uncommon, and that displacement of a microphone by a few centimeters can markedly alter response.

A *diffuse field* is a special kind of idealized sound field, in which sound intensity is the same in all directions. This suggests closely spaced resonances, since it results from standing waves involving each wall at nearly all angels of incidence. One measure of modal spacing is the *Schroeder frequency* f_s, which represents the lowest frequency for which, on average, three room resonances lie within the half-power bandwidth of a single resonance. This is given by (Pierce, 1989)

$$f_s = 2000(T_{60}/V), \tag{8.1}$$

where T_{60} is reverberation time (in seconds) and V is volume (in cubic meters). Any component above the Schroeder frequency will excite on average at least three modes, and pressure measured at any point will contain contributions from a number of modes. The sound field can be considered diffuse, when measured sufficiently far away from the source.

At small and moderate distances from a source, direct sound may be more intense than reverberant sound. Recordings may be inordinately sensitive to position and orientation. A relevant measure is the critical radius r_c, defined as the distance from the source at which direct and reverberant fields contribute equivalent sound pressures. It is given by (Benade and Larson, 1985)

$$r_c = .0565 D_\theta (V/T_{60}), \tag{8.2}$$

where D_θ is normalized directivity. For an omnidirectional source, D_θ is 1; for a source that radiates into half-space, it is 2. For brass instruments, D_θ is strongly dependent on frequency. Based on Meyer's (1978) directionality measurements, D_θ for trombone is 1 below 400 Hz; at 500 Hz, it is 5; at 1000 Hz it is 7; at 2000 Hz it is 27.

8.3 Statistical Aspects of Diffuse Fields

In a diffuse field, pressure measurements are dependent upon source and microphone positions. It can be shown theoretically that pressure amplitudes exhibit a Rayleigh distribution with a standard deviation of 5.6 dB (Lindevald, 1987). The advantage of measuring in a diffuse field is not increased accuracy of any measurement; rather, it is that there exists an extended region in which transfer varies with a known distribution. Averaging can be used to reduce variability, as long as measurements are statistically independent. This is the case when individual source and/or microphone positions are separated by more than $\lambda/2$, where λ = wavelength (Lindevald, 1987).

Standard deviation scales as $1/\sqrt{n}$ with the number n of independent measurements. For example, Benade and Larson (1985) used 32 independent spectral measurements (both source and microphone displaced) to produce an expected standard deviation for the room-averaged spectrum of $5.6/\sqrt{32} \approx 1$ dB for clarinet tones. Keefe, Bulen, Arehart and Burns (1994) used 384 measurements (6 source positions x 64 microphone positions) to produce an expected standard deviation of about 0.3 dB for broadband noise measurements. By varying both source and microphone positions, accurate (that is, reproducible and representative) room-averaged measurements can be made to any desired precision.

In addition to obtaining statistically independent measurements, the experimenter must ensure that a diffuse field is present at each measurement position. Lubman (1974) provides guidelines by dividing the sound field into three frequency regions:

- The region of low modal overlap, which he calls the ***volume control region*** since modal overlap depends strongly on volume. Spatial variability is large in this region. It can be avoided by ensuring that all frequencies of interest lie above the Schroeder frequency f_s.

- The *statistical control region*, in which modal overlap is high and the reverberant component of sound energy predominates over direct sound. This corresponds to diffuse field conditions. Lubman says that it begins at about f_s and "extends for several octaves, ending gradually as the field becomes semi-reverberant" and the direct field cannot be neglected (p. 525).

- The region of *direct field control*, in which spatial variance is significantly influenced by the direct field. The frequency at which this region begins depends on both source directionality and room characteristics. Additionally, air absorbtion decreases reberberant field energy at very high frequencies, causing the direct field to predominate. For musical measurements this effect can be neglected, since it is important only above about 10 KHz.

In volume control regions and direct field control regions, spatial variance is high and it is difficult to predict measurement accuracy. Statistical control regions are preferred for sound power measurements. Provided that microphones are sufficiently separated from the source, variance is predictable and averaging can be used. Benade and

Larson (1985) suggest recording at a minimum of 2.5 times the critical radius r_c (calculated for an omnidirectional source). Lubman (1972) notes that ANSI S1.21 gives the minimum permitted distance between source and microphone as $0.08\sqrt{V/T_{60}}$, which is only 1.4 times r_c calculated for an omnidirectional source.

An alternative approach is to position a microphone close to the source, so that the direct field predominates. This is practical for brass instruments, since all musically useful sound comes from the bell. Benade (1976) has suggested that a microphone placed "just in front of a bell (about one bell radius away) receives a sound signal that is in reasonably good agreement with what one expects from ... room averaging procedures" (p. 421). Pressure amplitudes at this point, however, are far in excess of normal listening levels. Both loudness and brightness calculations are referenced to absolute sound pressures, and are neither designed nor validated for extreme levels. Although a close microphone position gives reproducible spectral measurements, it is not appropriate for loudness and brightness measurements.

8.4 Effect of Spectral Fluctuations on Loudness and Brightness Calculations

It has been shown that the accuracy of spectral measurements can be inferred from sound field characteristics. It remains to be shown how measurement variability influences loudness and brightness calculations. Unfortunately, this depends not only upon the sound field and recording conditions, but upon the signal itself.

Loudness and brightness formulas weight spectral regions differently, and measurement variability depends upon both the fluctuation of individual components and on how they are distributed. Fluctuation of a low-frequency component influences loudness more than brightness; fluctuations of high-frequency components affect brightness more than loudness. Furthermore, when components are located in different critical bands they contribute more to both sharpness and loudness than if they are in the same critical band.

8.5 Summary

Measurements of predicted trombone loudness and brightness are not possible with a close microphone placement due to high levels. Recordings are best made in a diffuse field (Lubman's statistical control region), since measurement variability can be predicted. If, however, a diffuse field is not present at all frequencies of interest, measurement variability cannot be confidently predicted and must instead be established empirically.

Chapter 9

METHODS

It has been shown that brightness is an important dimension for timbre perception. It has also been shown that for trombone tones, brightness varies with loudness and pitch in ways determined by the physics of tone production. Brightness is thus an attribute likely to contribute to synthetic listening for trombone timbre, since it is (1) relevant to trombone tones, (2) salient, and (3) essential. Predicted brightness as a function of predicted loudness and pitch will be employed as a description of trombone timbre.

Pratt and Bowsher (1978) found that listeners could discriminate between different instruments played by the same performer as well as different performers using the same instrument, when presented with complete tones. A test of the usefulness of predicted brightness vs. loudness and pitch is therefore whether it captures differences between alto, tenor, and bass trombones, and whether it captures differences between individual performers when using the same instrument.

To examine this, three trombonists each performed on alto, tenor and bass trombones trombones. Notes were recorded covering the entire practical range of pitches and loudness. Comparisons of trombones (same performer) and performers (same trombone) could then be made.

9.1 Materials

Three trombones were utilizied in this study.

- An E♭ alto trombone, Minnick brand, performed with its stock Minnick mouthpiece.

- A B♭/F tenor/bass trombone, Vincent Bach model 42BO, with a Bach 5G mouthpiece.

- A B♭/F/G♭ bass trombone, Vincent Bach model B503O, with a Bach $1\frac{1}{2}$G mouthpiece.

All instruments are of professional quality, and are representative of models used in the United States. Likewise, mouthpieces reflect common usage for the particular instruments. All three instruments belonged to one of the performers (C); however, this performer normally uses slightly different mouthpieces on each of them. Thus, the instrument/mouthpiece combinations are representative of common practice by American trombonists, although no performer was using exactly his accustomed configuration.

9.2 Participants

The participating performers are highly skilled and accomplished trombonists with extensive amateur and professional experience. One is Professor of Trombone at a major university, and the others are doctoral students in trombone performance at the same university.

9.3 Apparatus

Tones were recorded in stereo, with the exception of Performer C on tenor and bass trombones, who was recorded monaurally. A high-quality Larson-Davis reference microphone system was used, consisting of $\frac{1}{2}$-inch condenser microphones model 2445 with pre-amplifiers model 900B, powered with pre-amp/power supply model 2200c. A bias of 200 volts and a gain of 30 dB was used throughout. The pre-amp output was recorded on a Panasonic Digital-Audio Tape Recoder model SU-3700 at a sampling rate of 44.1 kHz. The recorder has a resolution of 16 bits.

All recordings were made in a large rehearsal hall. The volume V of the hall is approximately 2200 cubic meters, and it has a reverberation time T_{60} of approximately .8 seconds. The Schroeder frequency f_s of this room is therefore, according to Equation 8.1, about 38 Hz. The critical radius r_c for an omnidirectional source is, by Equation 8.2, about 3 meters.

Microphones were placed about 6 meters from trombonists and about 6 meters apart, at different heights (1.5 - 2.5 m), and about 30 degrees off-axis. They were separated from each other and from a reflecting surface by more than a half wavelength for frequencies above 115 Hz. While they were closer to the source than the 2.5 times r_c recommended by Benade and Larson (1985), they were farther than the ANSI minimum distance for sound power measurements (Lubman 1974). Above 500 Hz, normalized directivities for trombone tones are greater than 2. This implies that at some frequencies the microphones were beyond r_c, while for others they were not.

Microphones were placed in roughly the same position for each performer. Likewise, performers were located in approximately the same area and were oriented in the same direction.

9.4 Procedure

Participants were asked to play notes from their entire range of loudness and over a specified range of pitches. According to Benade and Larson (1985), a performer produces consistent tones when given "a task whose nature and whose specifications are completely familiar in the terms of his musical experience" (p. 1478). In view of this, performers were allowed to prepare for each tone according to their own practice habits. Thus, Performer C prepared to play each note by performing a lip slur prior to the note, as in many common trombone warm-up exercises (e.g. Hunsberger 1980); Performer A played the notes of the same exercises as isolated tones; and Performer B performed converging scale patterns. Instead of artificially biasing results, accepting accustomed means of preparation enhances the opportunity for each trombonist to produce his or her most characteristic tone quality.

Specified pitch ranges were:

- alto trombone: A1 to F♯5, omitting the notes A♭2 to E2
- tenor/bass trombone: G1 to E♭5, omitting the note B1
- bass trombone: F1 to C♯5

These ranges encompass the normal playing range of each of the instruments (Read 1969). Tones were approximately 1 to 4 seconds in length.

Loudnesses were specified using musical terminology. Performer C performed the above notes at levels that would be understood from classical music notation: *pp*, *p*, *mp*, *mf*, *f*, and *ff*. Following suggestions from performer A, a 7-point scale was used by the other players: *ppp*, *pp*, *p*, *m*, *f*, *ff*, and *fff*. It was felt that this was both consistent with common notational practice (Romantic period and later) and easier to perform, since it reduced confusion between *mezzo-piano* and *mezzo-forte*. While each trombonist performed at loudnesses spanning the normal performance range, it should be noted that this range may not be the same for the three players, due in part to their performance emphases. Performers A and B often perform contemporary solo and ensemble repertoire, which employs dynamics that are considered inappropriate in more traditional contexts. Performer C, on the other hand, performs mostly from the standard repertoire.

9.5 Calibration

Calibration of microphone systems (microphone cartridge, power-supply microphone casing, and pre-amp) was verified with a Bruel and Kjaer model type 4220 pistonphone connected to a Bruel and Kjaer model 2209 sound-level meter. The measured sensitivity of each system was within 0.5 dB of the value supplied by Larson-Davis. The manufacturer's calibration was therefore considered verified and their sensitivities were used in subsequent calculations.

Calibration of the recording system (voltage to bits) was accomplished by disconnecting the microphones from the DAT recorder and substituting a Krohn-Hite model 4300 Signal Generator producing a nominal 0.316 volt RMS sine wave at 1 kHz. The voltage was verified with a voltmeter. The peak-to-peak variation in bits was then compared to the peak-to-peak voltage fluctuation.

The calibration was verified computationally by calculating the sound pressure level required to produce a .316 volt RMS output from the microphone and pre-amplifier. The loudness of the digitized 1-KHz sine tone was calculated using the calibration (bits to volts) of the recording system. Finally, the two were compared. For example, for performer B on alto trombone, the sine wave was measured at .309 volts RMS. Taking microphone sensitivity and gain into account, this corresponds to a sound pressure level of 93 dB. According to Scharf (1978) the loudness of a 93 dB, 1-kHz sine tone is 39.4 sones. Measured loudnesses were 41.60 sones for Channel 1 and 40.75 sones for Channel 2, or within 5% of published values.

Additionally, the predicted brightness of a 14 sone, 8 KHz sine tone was calculated.

According to Aures, the brightness of this should be 9.6 acums. However, using the calibration above, predicted brightness was 9.94 acums. Goad (1994) also noted a slight deviation from Aures' values. This does not affect comparisons described in this study.

9.6 Data Processing

Digital recordings were converted to sound files, and 32,768-point sections of each tone were selected for analysis based on subjective steadiness of the tone throughout approximately $\frac{3}{4}$ second duration. Sharpness, loudness and pitch were calculated using a specialized signal processing program (Ling 1991). Sharpness was calculated based upon Aures' formulation (1985); loudness was calculated using an implementation of Zwicker and Paulus' (1972) algorithm, and fundamental frequency was calculated using a pattern-matching algorithm developed by Keefe. This algorithm was found to be accurate within a few cents, except when very large (generally harmonic) errors occurred (See Appendix B). In this case, fundamental frequency estimates were noticeably incorrect, and an alternative method was used. Fundamental frequency was instead calculated from the frequency of the tone's 10th harmonic. The mean fundamental frequency error from the pattern matching algorithm as compared to the 10th harmonic was 7.4 cents, with a standard deviation of 0.85 cents.

Again, since both sharpness and loudness measures are not designed (and have not been validated) for extremely loud sounds, all tones with a loudness of more than 120 sones were omitted.

9.7 Measurement Accuracy

Using the terminology developed by Lubman (1974), the soundfield encountered by the microphones is best described as "semi-reverberant." This means that at some frequencies microphones were in the statistical control region, while at others they were in the direct field control region. In its low register the trombone radiates nearly isotropically, and microphones were beyond the critical radius. At higher frequencies (certainly above about 500 Hz) directionality becomes more pronounced, and the microphones were affected by the direct field. Since relative intensities of the fields are not known, there is no *a priori* way to determine measurement accuracy. Instead, a number of measurements must be compared and accuracy established empirically.

Considerations of accuracy suggests how the resulting data should be handled in order to facilitate comparisons. The various trombonists each recorded with slightly different microphone locations as well as slightly different playing positions. It is important, therefore, to describe the loudness and brightness as encountered at *representative* locations rather than *specific* locations. For most performers, stereo recordings were made. The microphones were placed symmetrically and at roughly the same distance from the trombonist, and the trombonists were oriented along an axis of symmetry of the room. For music listening, these locations would be regarded as equivalent. This suggests that pooling the data from the separate positions permits the description of a characteristic rather than specific response, and facilitates comparisons between performers. Comparison between channels can be used as an indication of measurement accuracy.

Detailed consideration of these issues is found in Appendix A. To summarize, a pilot study was performed with synthesized tones rather than trombone tones, in order to investigate the effect of microphone and speaker positions for loudness and brightness measurements. It was found that means and standard deviations for brightness and loudness do not vary systematically between multiple microphone positions for each speaker position (n=8), and multiple microphone and speaker positions (n = 24). This suggests that slight changes in source position do not strongly bias measured brightness and loudness.

The effect of microphone position can best be investigated by comparing data from separate channels. Details are again found in Appendix A. To summarize, brightness varies with a standard deviation of about 7% between microphone placements, and loudness varies with a standard deviation of about 15%. Specific microphone locations bias measured brightness and loudness, and can be seen in brightness and loudness ratios between microphones, when averaged for the same sounding pitch. However, the bias is less than either differences between instruments or performers. This suggests that data from both microphone positions can be pooled.

Chapter 10

RESULTS

As discussed earlier, brightness, loudness and pitch are all subjective quantities. Measured loudness in sones, brightness in acums, and pitch in notes are properly described as *predictions*, whose correlations with subjective perceptions can only be investigated with human listeners. In this chapter predicted loudness, brightness and pitch are measured and discussed at length. In the interest of readability, I have omitted the "predicted" qualifier. The reader is asked to remember that measurements reported in this chapter are of tones, and not directly of their perception.

Since loudness predictions in sones and brightness predictions in acums are perceptual measures whose values double with a perceived doubling in sensation levels, it is appropriate to view them in a logarithmic space. In this way, a given displacement along an axis reflects a constant ratio in sensation level. Likewise, a logarithmic frequency axis is used in all plots, labelled with note names ("C4" = middle C @ 261.6 Hz.). While this is not the same as using a perceptual measure for pitch (the mel), it is justifiable since the mel scale does not deviate significantly from logarithmic frequency for notes in the range of the trombone. Furthermore, trombone tones are all complex, and the mel strictly applies only to pure tones.

10.1 Loudness Ranges

The range of loudnesses varied widely between performers. In Figures 10.1, 10.2, and 10.3 the loudness ranges produced by each performer on each trombone is shown. Each point denotes the loudness of a single tone at a given pitch. Vertical lines join notes of the same pitch class; some jitter is seen due to inaccuracies in intonation. On each instrument, performer B employed the widest range and Performer C the narrowest. On the note B♭3 on bass trombone, for example, Performer B produced a range of 9 to 190 sones, while Performer A performed from 13 to 120 and Performer C only from 29 to 99 sones. Thus, on this note Performer B had the ability to double loudness 4.4 times, Performer A 3.3 times, and Performer C 1.8 times, calculated by taking the logarithm base 2 of the ratio of maximum to minimum loudnesses. The loudness range seems more strongly associated with the performer than with the instrument. For example, loudnesses produced by Performer A on alto, tenor and bass trombones (middle plots of Figures 10.1, 10.2, and 10.3 respectively) are quite similar; they differ, however, from those produced by Performers B or C on any trombone. The only exception to this trend is Performer A on bass trombone, for whom the minumum loudness for each pitch is considerably higher than on alto or tenor.

10.2 Brightness Ranges

A similar pattern is evident in the ranges of brightnesses produced by each performer, shown in Figures 10.4, 10.5, and 10.6 for alto, tenor, and bass trombones respectively. As was the case for loudness, for each instrument Performer B produced the widest range of brightnesses and Performer C the narrowest. As with loudness, the ranges produced are more closely associated with performer than with instrument. Also of interest is the relative uniformity of each performer's brightness ranges, as compared with loudness. For example, for Performer B on alto trombone, the lower limits of loudness vary widely across the range of pitches (Figure 10.1b); the lower limits of brightness are more uniform (Figure 10.4b). The explanation for this relative uniformity awaits further investigation. It may simply be an artifact of the brightness model, which preferentially weights high-frequency contributions. Soft trombone tones have few harmonics, and will not be affected by such weighting. Alternatively, it may be a direct result of performers' actions. When attempting to perform a scale at a musical dynamic of *ppp*, performers may strive for uniform brightness instead of uniform loudness.

The majority of loudness and brightness ranges is common to all players. However, for notes of the same pitch and loudness, brightness is not necessarily the same.

10.3 Brightness as a Function of Loudness and Pitch

The relationships between loudness, brightness, and pitch can be explored using a number of plotting techniques. The most intuitive is simply a three-dimensional plot of log frequency vs. log loudness vs. log brightness, as in Figures 10.7-10.9. Lines

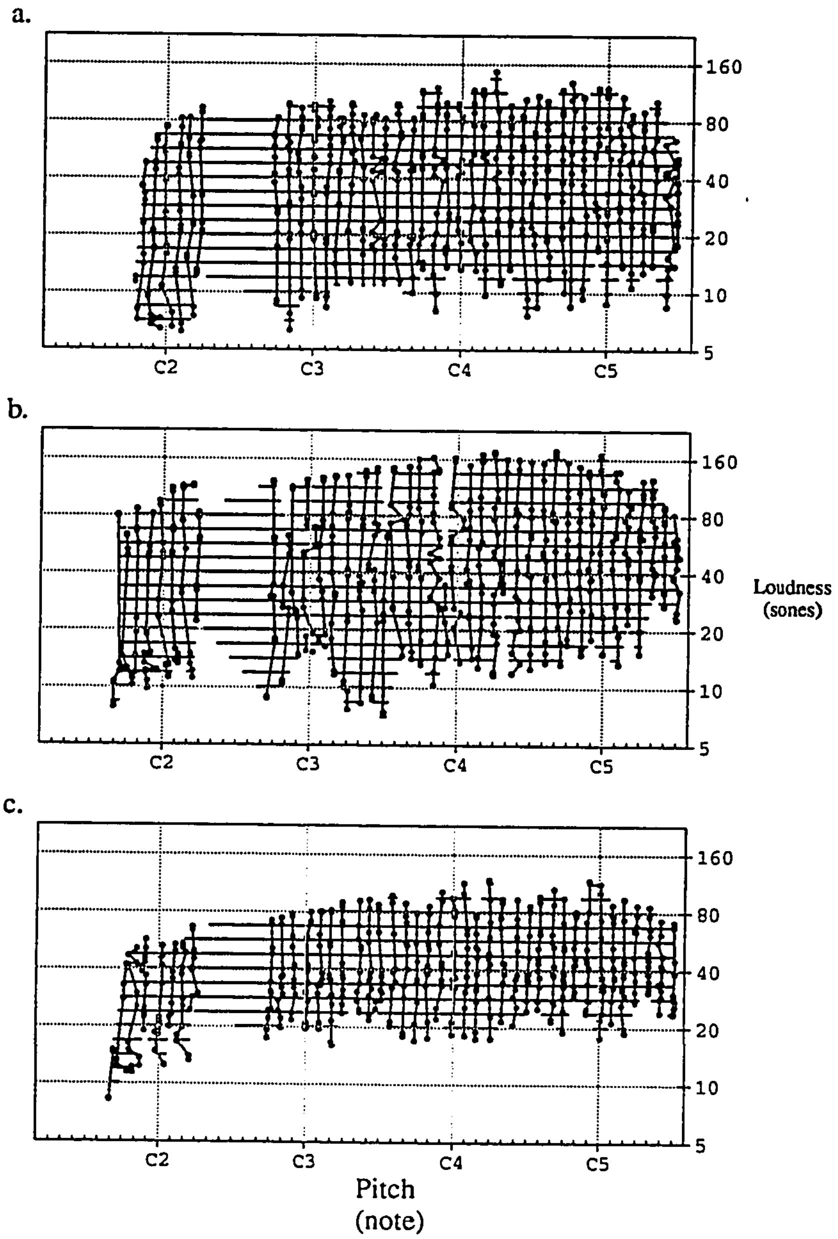


Figure 10.1: Loudness produced by each performer on alto trombone. a. Performer A; b. Performer B; c. Performer C.

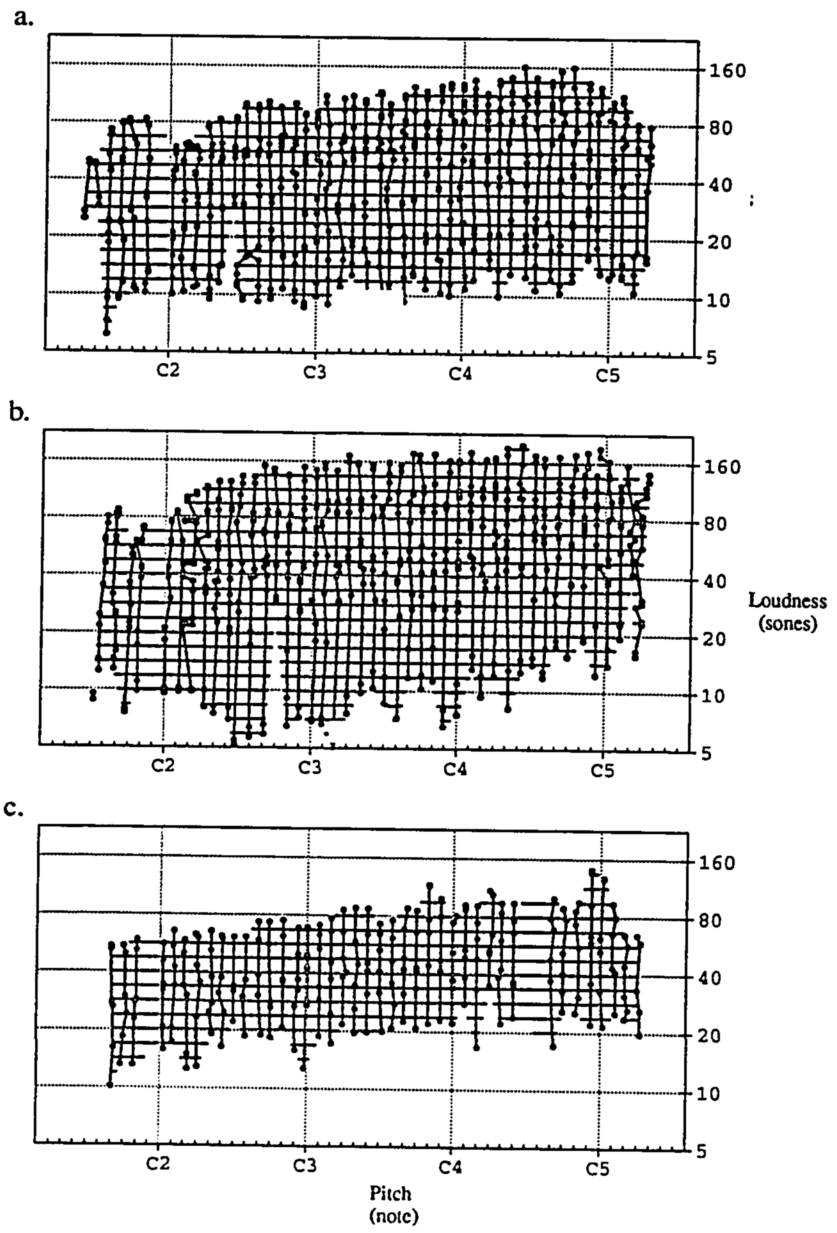


Figure 10.2: Loudness produced by each performer on tenor trombone. a. Performer A; b. Performer B; c. Performer C.

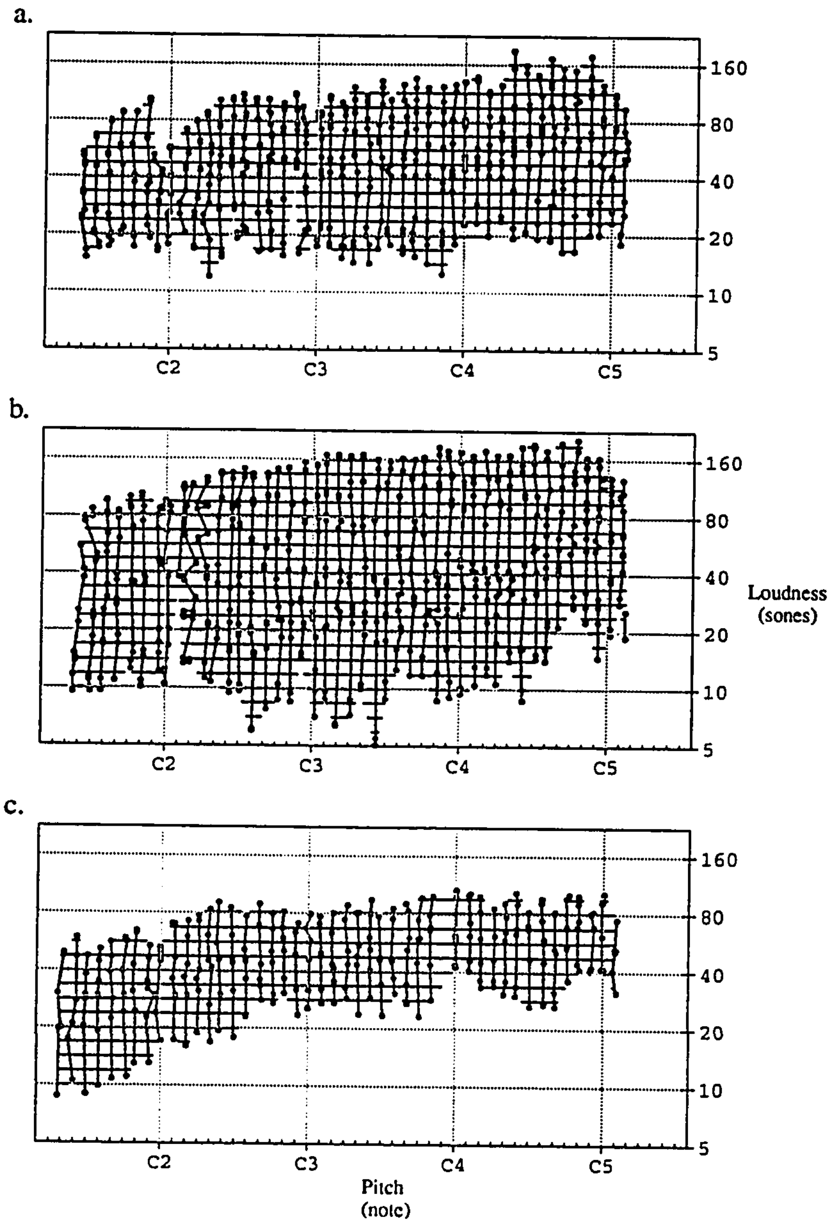


Figure 10.3: Loudness produced by each performer on bass trombone. a. Performer A; b. Performer B; c. Performer C.

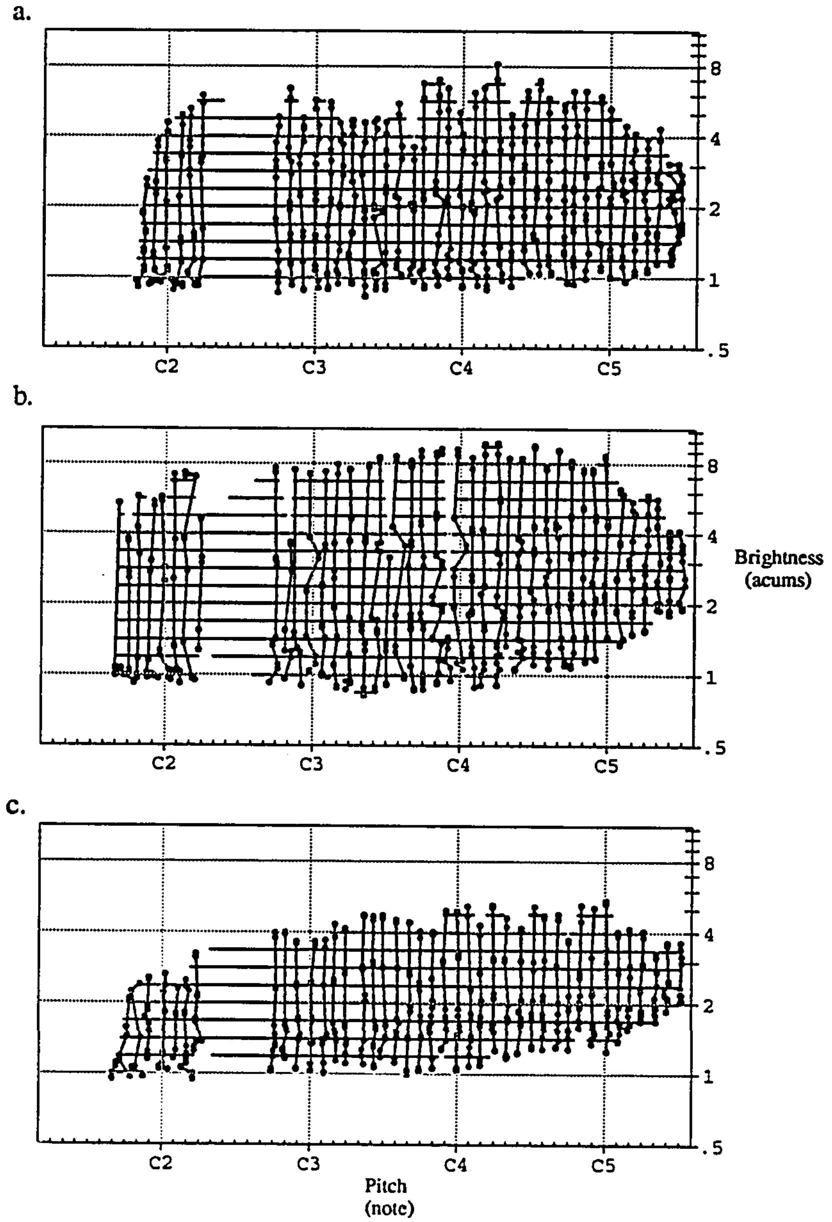


Figure 10.4: Brightness produced by each performer on alto trombone. a. Performer A; b. Performer B; c. Performer C.

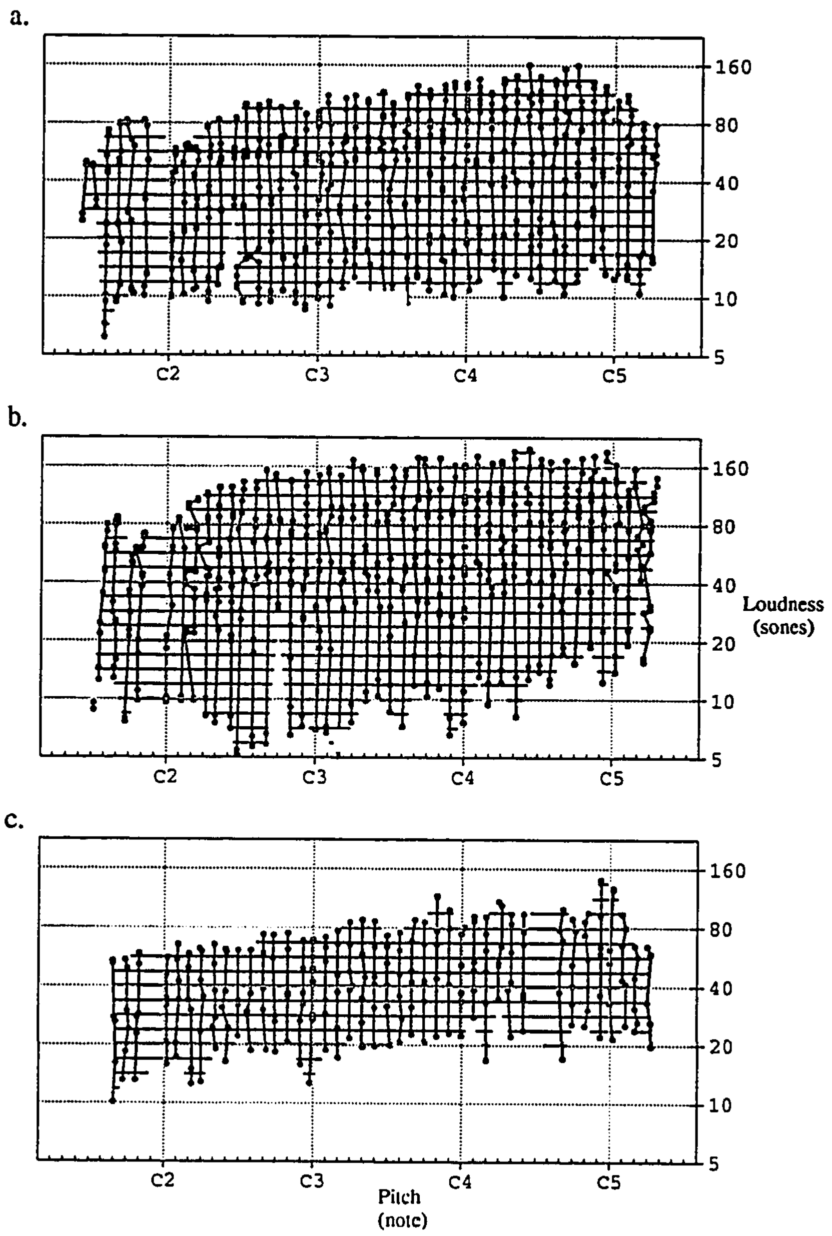


Figure 10.5: Brightness produced by each performer on tenor trombone. a. Performer A; b. Performer B; c. Performer C.

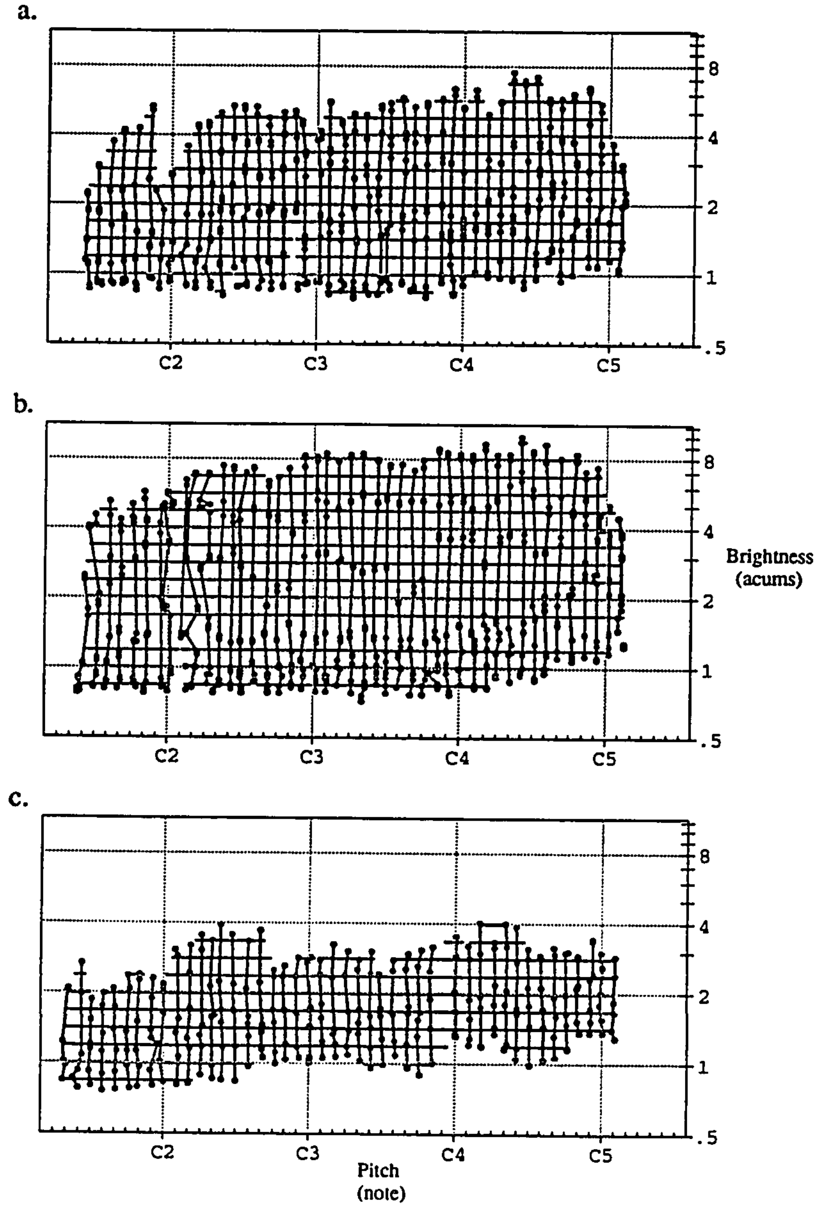


Figure 10.6: Brightness produced by each performer on bass trombone. a. Performer A; b. Performer B; c. Performer C.

parallel to the loudness-brightness plane are made up of line segments connecting data points. Lines parallel to the loudness-brightness plane are linearly interpolated equal-loudness lines, at values of 5 sones x $2^{n/4}$, where n is an integer. In this way every fourth line portrays a doubling of loudness.

A number of observations can be readily made. To begin with, similar contours characterize all trombone/performer combinations. Brightness appears to increase monotonically with loudness, excluding minor fluctuations that may be due to measurement variability. At low loudnesses, brightness increases slowly with loudness. At high loudnesses the increase is much quicker. The relationship appears to be exponential. In all plots, the minimum loudness and brightness is greater in the highest registers; that is, the lower right-hand corner of each surface appears to turn upwards. The degree to which this occurs depends on the trombonist. It is least evident for Performer A (Figures 10.7a, 10.8a, 10.9a), and most apparent for Performer B (Figures 10.7b, 10.8b, 10.9b). The effect is not pronounced for Performer C, perhaps because the minimum loudness level is generally higher than for either A or B.

Visual perspective makes it impossible to estimate coordinates from this kind of plot. A less intuitive but more informative plot is obtained by plotting brightness as a function of note, at a number of loudness levels. This is equivalent to collapsing the above plots along the loudness axes. Brightness on equal-loudness lines is shown in Figures 10.10-10.12. Lines are generated at loudnesses of 10 sones x $2^{n/2}$, so that every other equal-loudness line corresponds to a doubling in loudness.

In each plot, the lowest loudness lines (10-20 sones) are nearly flat. This means that for these loudnesses, brightness varies little with pitch. At medium and high loudnesses, however, most lines show a negative slope from low to middle ranges. Above the middle range, many lines assume a positive slope. The breakpoint is about C4 for tenor and bass trombones and E♭4 for alto trombones. For loudnesses above 20 sones, brightness is greater in the low register than in the middle register, and often increases again in the upper register.

A pattern is also seen in the vertical spacing of equal-loudness lines. Low loudness lines are close together; separation increases at higher loudnesses. For example, 14 sone lines are only about 15% brighter than 10 sone lines, but 80 sone lines are between between 50 and 100% brighter than 57 sone lines. This reveals a strong influence of loudness on brightness. Since equal-loudness lines are separated by a constant factor

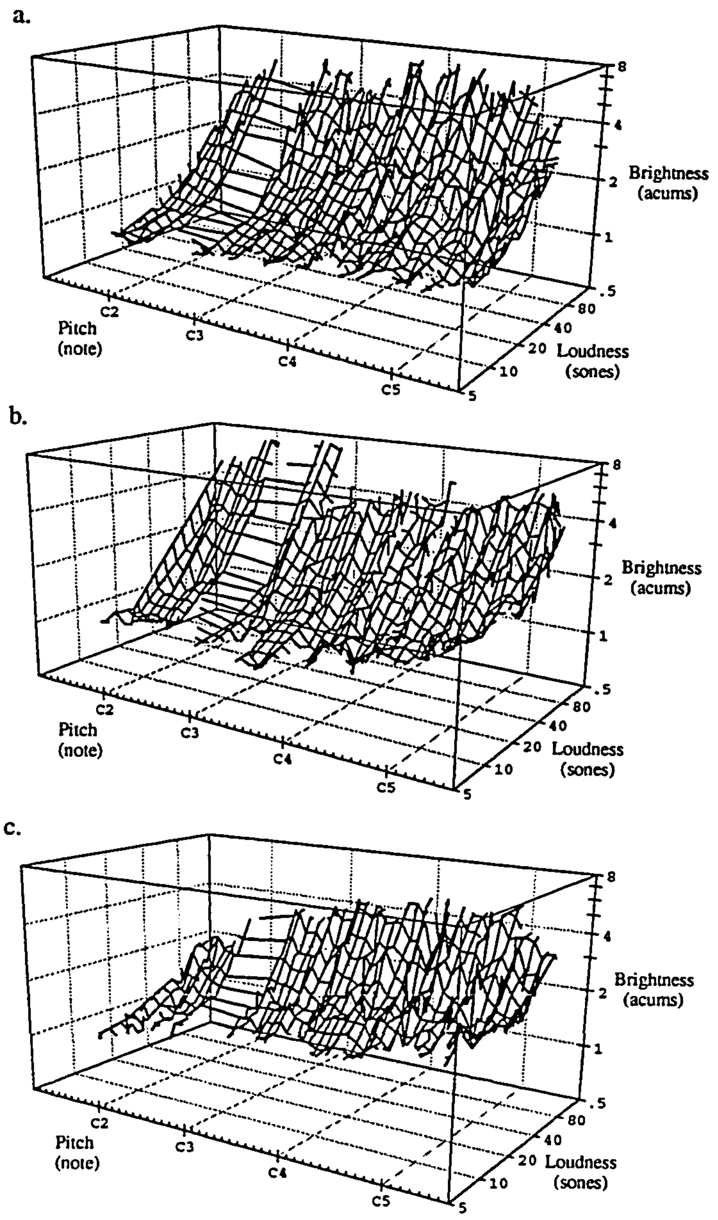


Figure 10.7: Brightness vs. loudness and pitch, for alto trombones. a. Performer A; b. Performer B; c. Performer C.

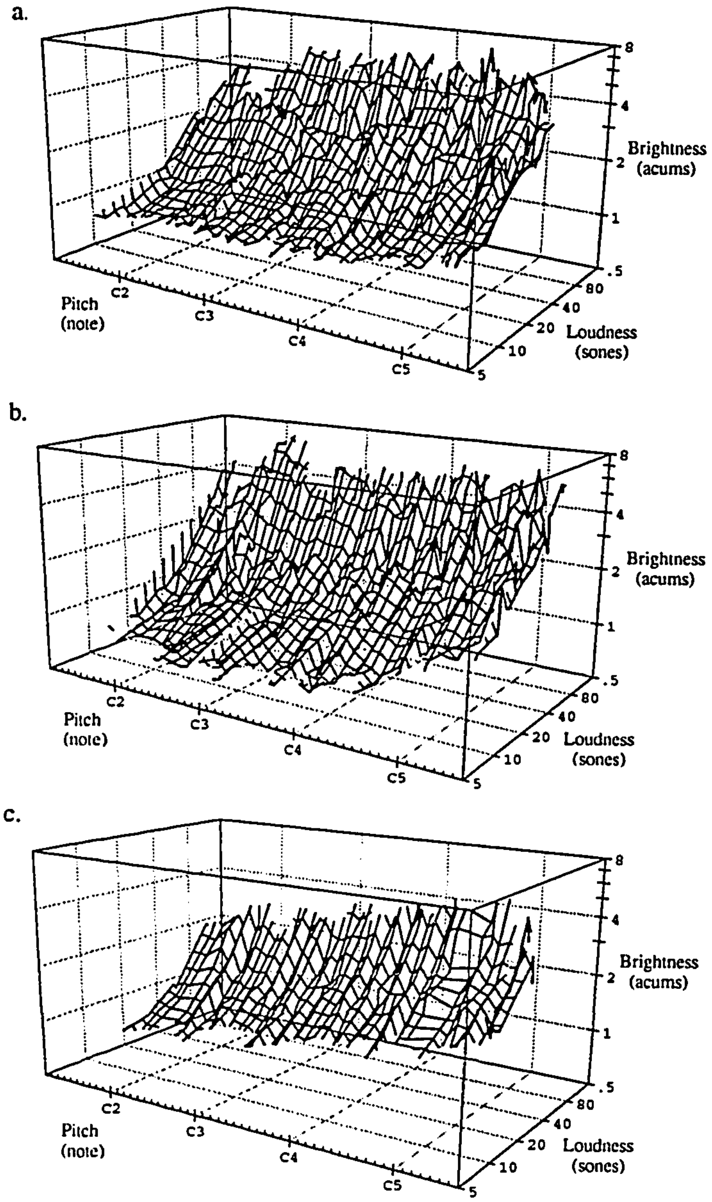


Figure 10.8: Brightness vs. loudness and pitch, for tenor trombones. a. Performer A; b. Performer B; c. Performer C.

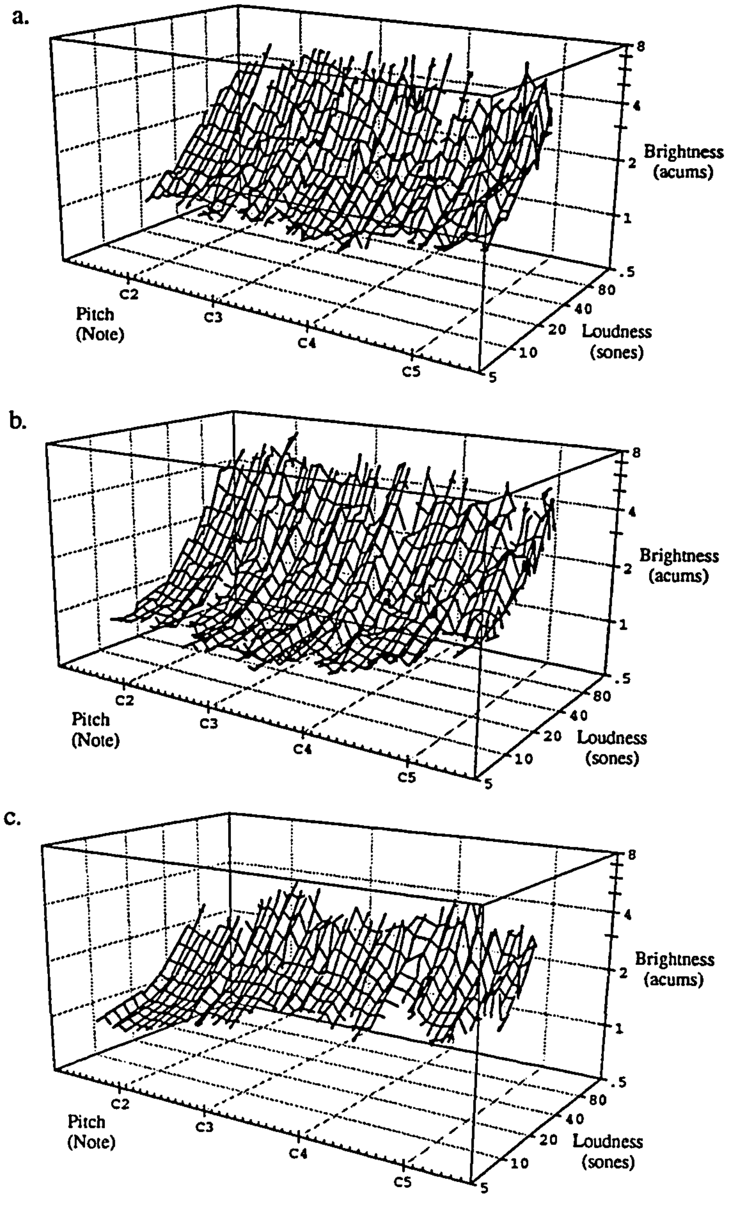


Figure 10.9: Brightness vs. loudness and pitch, for bass trombones. a. Performer A; b. Performer B; c. Performer C.

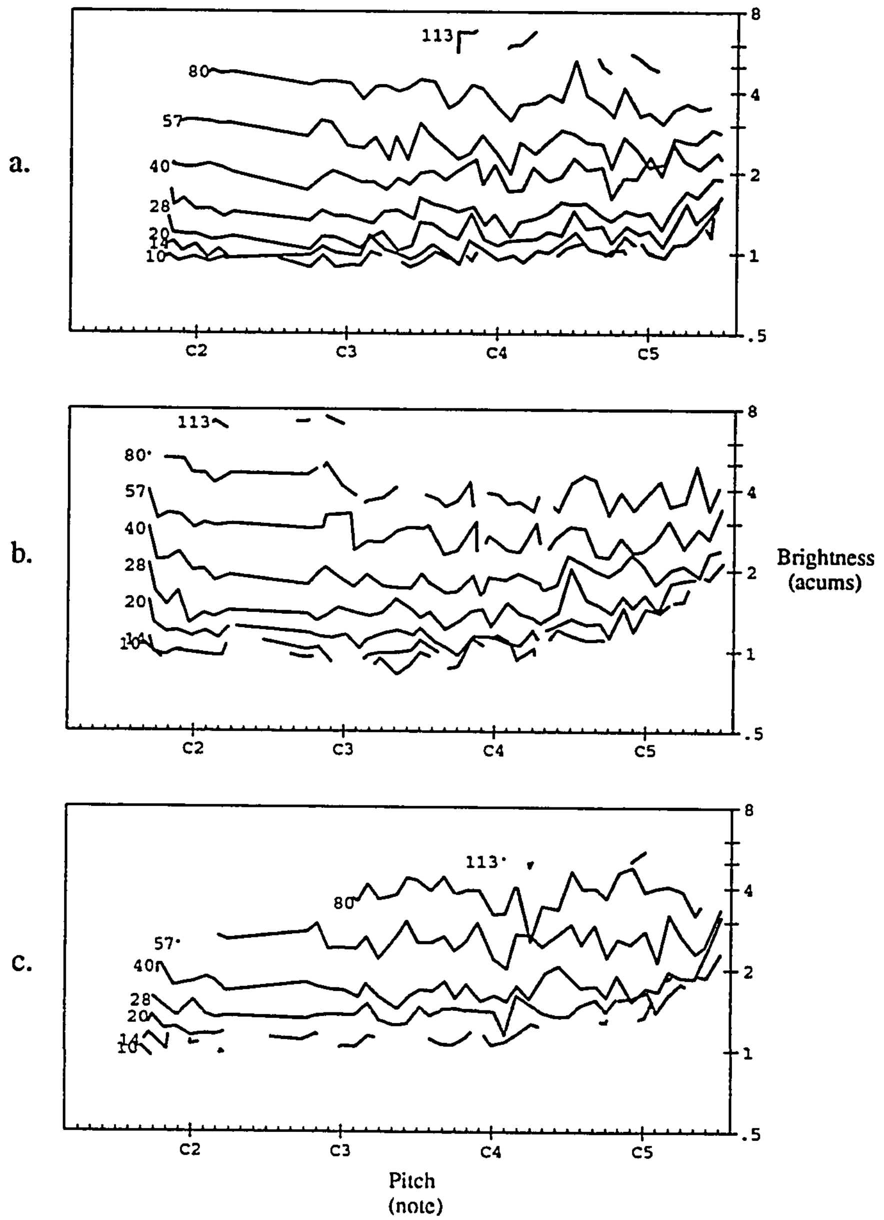


Figure 10.10: Brightness along interpolated equal-loudness lines, for alto trombones. Loudnesses are indicated to the left of each line. a. Performer A; b. Performer B; c. Performer C.

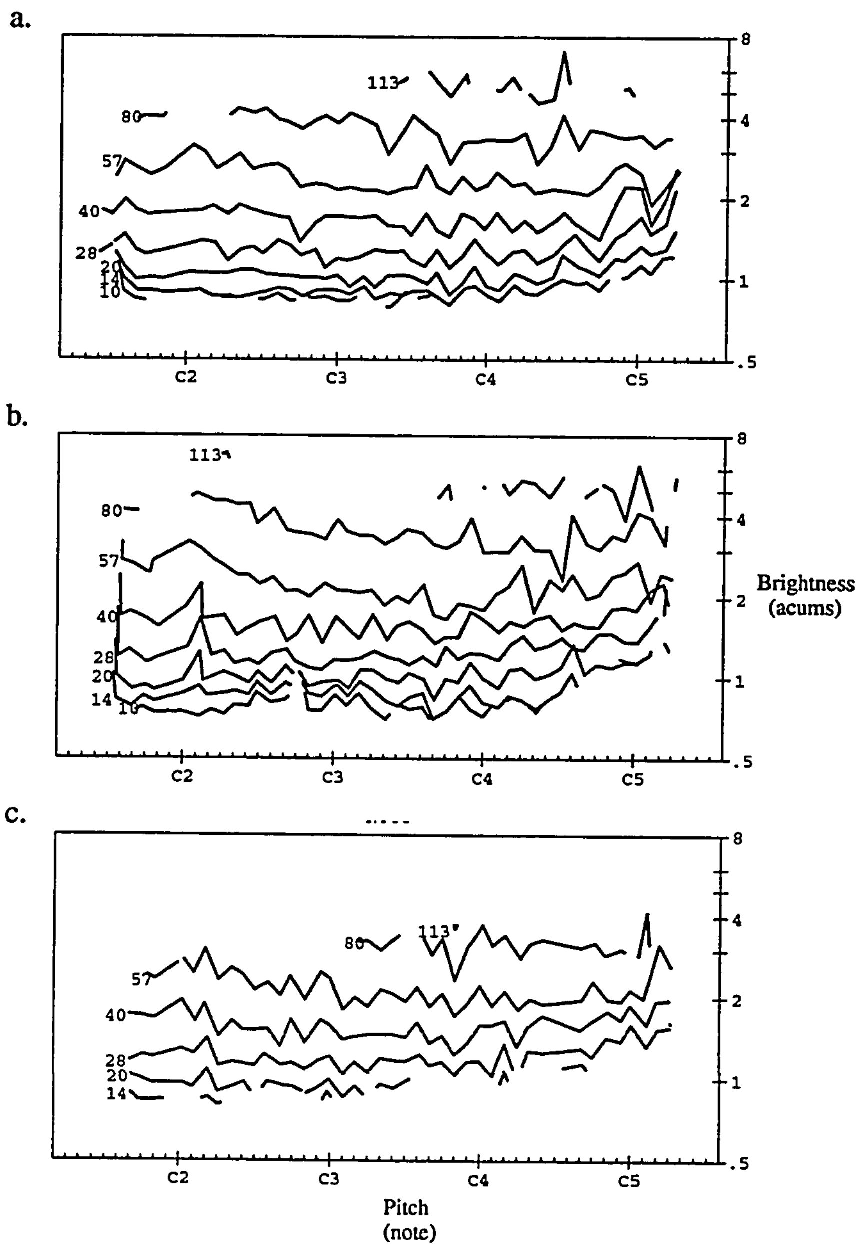


Figure 10.11: Brightness along interpolated equal-loudness lines, for tenor trombones. Loudnesses are indicated to the left of each line. a. Performer A; b. Performer B; c. Performer C.

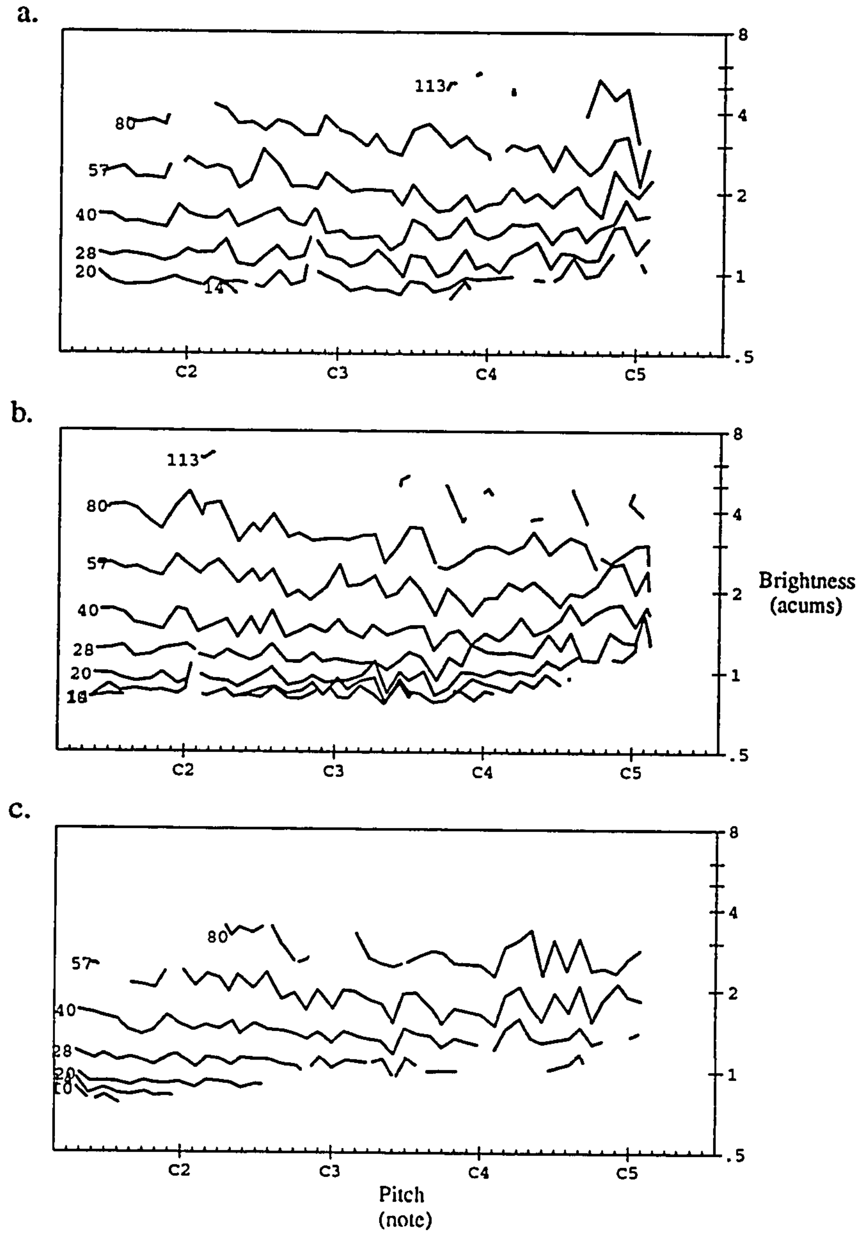


Figure 10.12: Brightness along interpolated equal-loudness lines, for bass trombones. Loudnesses are indicated to the left of each line. a. Performer A; b. Performer B; c. Performer C.

$(\sqrt{2})$, the spacing between them reflects the rate of change of brightness with loudness. The rate of change of brightness with loudness will be called the *brightness gradient.* In addition to being related to loudness, the brightness gradient is influenced by pitch. As can be seen in the plots, equal-loudness lines are more widely separated in low registers than in high registers. This means that the brightness gradient is greater in the low register.

Some general comparisons are possible based on Figures 10.10–10.12. The trombone has a noticeable effect on brightness levels. For each performer, brightness at a given loudness on the alto trombone is higher than those of the tenor by a factor of about 0.1 to 0.2. Above C5 this trend is not clear. For each performer, brightnesses on tenor are greater than those on bass by a factor of perhaps 0.1. At the same pitch and loudness, tenor trombone brightness is generally between those of alto and bass.

Performers seem to influence brightness in subtle ways. Differences are most pronounced in the middle pitch range, and for medium loudnesses (28–57 sones). For example, on alto trombone the most consistent difference between performers is that Performer C has lower brightnesses at 40 sones for all but the highest and lowest notes (compare Figure 10.10 a, b, and c). On tenor trombones, Performer A is slightly brighter than B and C through most of the range; B and C are practically indistinguishable (Figure 10.11). On bass trombone, A and B are very similar, but C is less bright through the middle portion of the pitch range and above 40 sones (Figure 10.12).

10.4 Summary

Measurements of loudness as a function of pitch show that each performer produced similar loudness ranges regardless of instrument. These ranges differ, however, between performers. Furthermore, each performer produced similar ranges of brightness regardless of instrument.

Brightness plotted as a function of loudness and pitch produced similar general contours for all performer/combinations. Brightness is influenced strongly by loudness, and to a lesser degree by pitch. The brightness gradient is useful for quantifying relationships between loudness and brightness.

Chapter 11

ANALYSIS

In the previous chapter it was seen that alto, tenor and bass trombones produce similar patterns of brightness as a function of loudness and pitch. In this chapter, brightness and brightness gradient are examined for patterns distinctive of instrument and performer, so that comparisons can be made. Since measurement variability complicates some comparisons, averaging and smoothing will be used when necessary.

11.1 Brightness gradient

The brightness gradient can be used for describing differences between instruments and performers. Since gradients are sensitive to measurement inaccuracies, smoothing in the loudness-brightness plane is advisable. Frequency measurements are more accurate than loudness or brightness, so smoothing in this dimension is not required.

Smoothing in the loudness-brightness plane has been implemented in the following fashion. Data for one note, over its range of loudnesses and brightnesses, is extracted for analysis. An exponential curve is fitted to the loudness and brightness values by a least-means square method, producing an expression for brightness as a function of loudness for the particular note in question. It was found empirically that exponential curves describe measured data better than polynomial expressions up to order 5. From the fitted expression, brightness can be calculated from loudness, and the brightness gradient can be obtained by taking the partial derivative with respect to loudness.

Figures 11.1, 11.2, and 11.3 show brightness gradients for alto, tenor, and bass trombones respectively. To obtain each of the curves, the brightness gradient has been calculated for each of the single-pitch exponential expressions described above, at levels of 20, 40, 60, 80 and 100 sones. Lines have been drawn connecting points at the same loudness. Thus, each curve represents the brightness gradient at a single loudness. Smoothing has been performed in the loudness-brightness plane, but no smoothing has been performed in the pitch dimension.

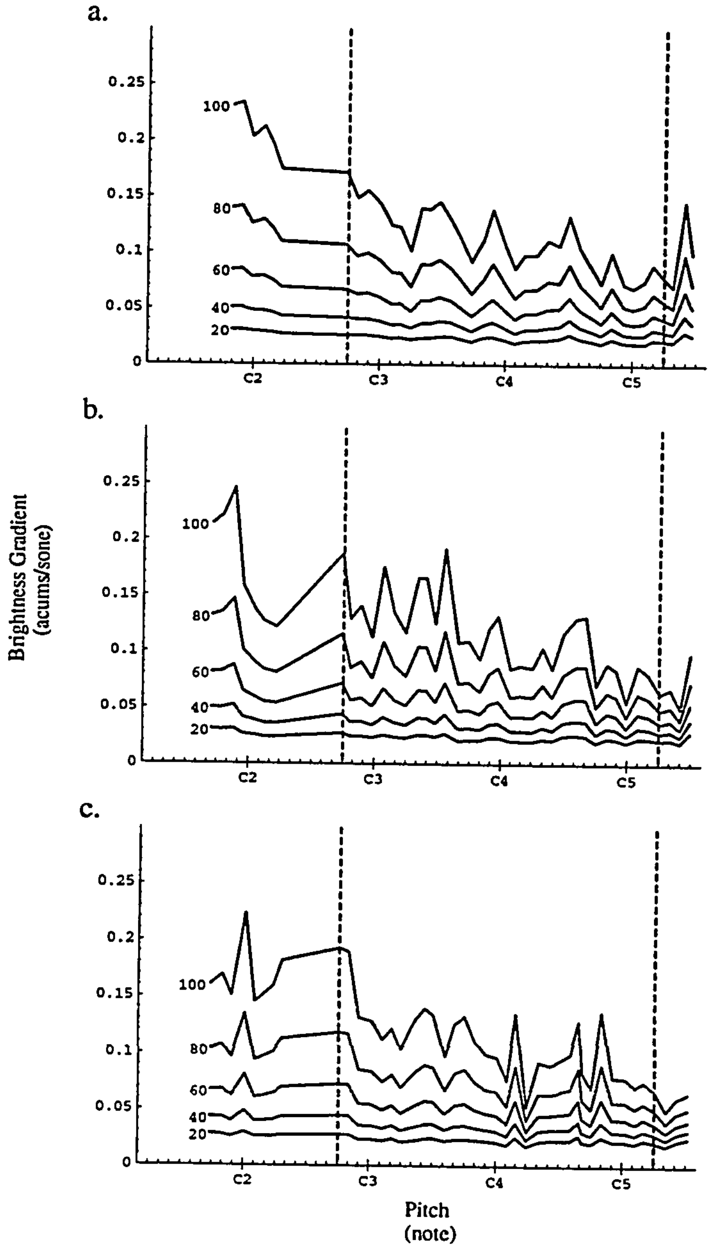


Figure 11.1: Brightness gradients at five loudness levels, for alto trombones. a. Performer A; b. Performer B; c. Performer C.

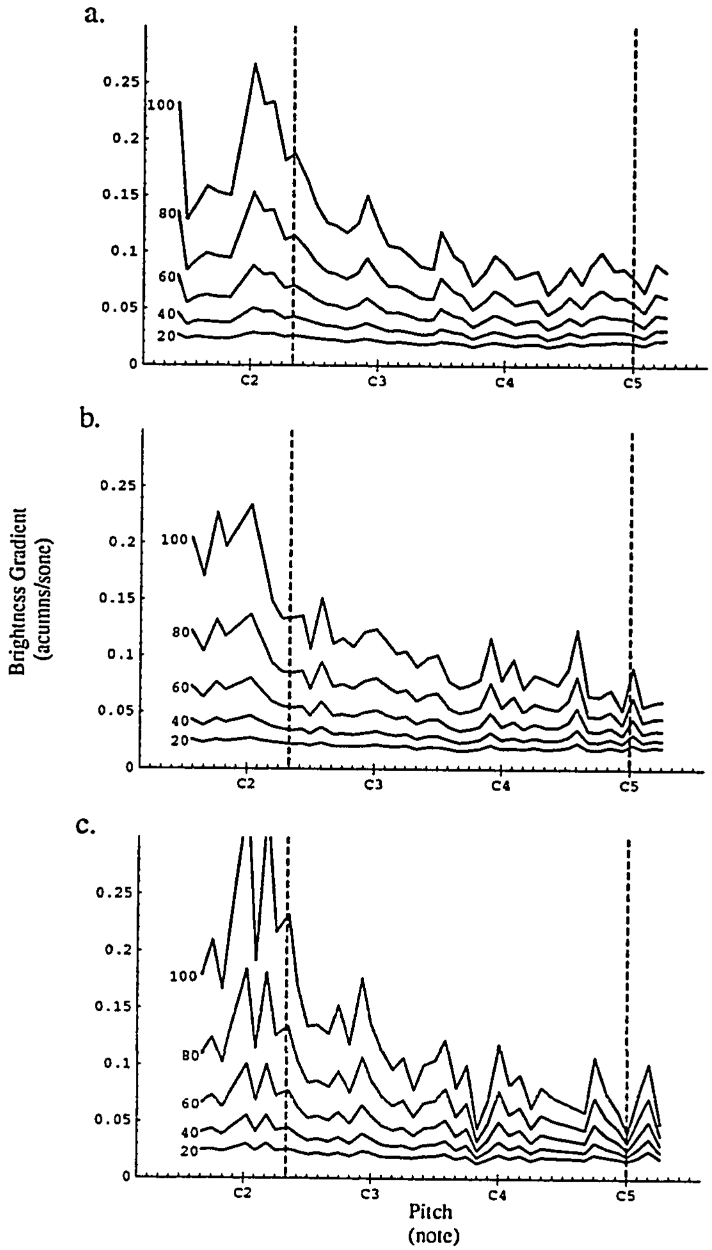


Figure 11.2: Brightness gradients at five loudness levels, for tenor trombones. a. Performer A; b. Performer B; c. Performer C.

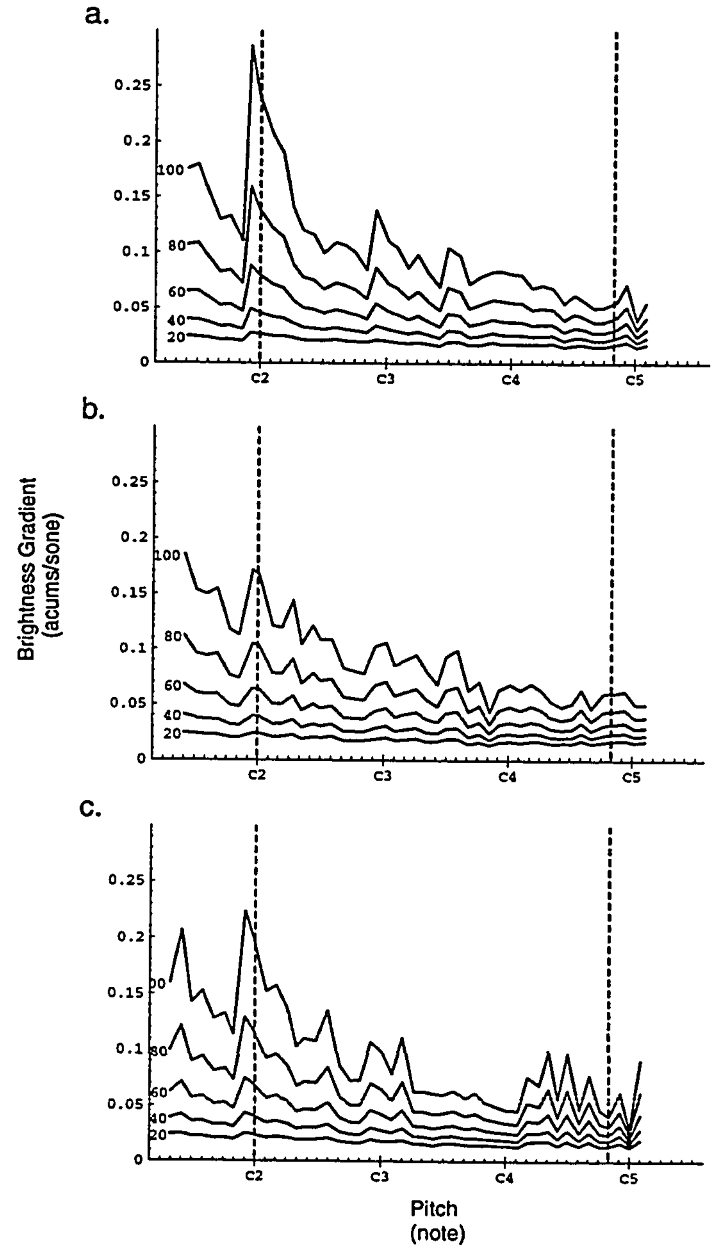


Figure 11.3: Brightness gradients at five loudness levels, for bass trombones. a. Performer A; b. Performer B; c. Performer C.

For all performers on all trombones, the brightness gradient is greater in the low register than in the high register. A pattern emerges especially clearly in the plot from Performer A on bass trombone (Figure 11.3a), but can be seen in all plots. At each loudness level, the brightness gradient follows a series of peaks and dips whose frequencies are consistent across loudnesses. Peaks occur at B1, B2, and F♯3; dips occur precisely one half-step below these, at B♭1, B♭2, and F3. These are significant notes for a trombonist. B♭1, B♭2 and F3 are first position notes, so that the trombone is being played in its shortest configuration (about 8 feet). B1, B2, and F♯3 are all long-position notes; F♯3 is in fifth position (about an 10.7 foot tube length), B2 in seventh (12 feet), and B1 in approximately fifth position with both F and G♭ valves engaged (15 feet).

It appears that brightness gradient is affected by both register and slide position. The relative contribution of each can be assessed by examining the brightness gradient for notes played in the same positions, as in Figure 11.1. To obtain these plots, the brightness gradient was first determined for each note at a loudness of 80 sones. This value was chosen because it represents a high loudness that is within the range of all performers. The notes were then separated according to sounding length (slide position). No discrimination was made between tubing provided by the slide and tubing provided by the valve; thus, sixth position and F-attachment first position were considered equivalent. Finally, notes with the same sounding length were connected by lines. Solid lines connect first-position notes; dotted lines connect fourth-position notes and dashed lines seventh position notes.

Three trends are apparent in these plots. First, the brightness gradient for each sounding length is greater in the lower register than in the higher register. This is indicated by the negative slope of nearly all line segments. The first position notes show that as pitch drops by three octaves (B♭4 to B♭1 for tenor and bass trombones; E♭5 to E♭2 for alto trombones) the brightness gradient increases by between 50 and 100 percent. The only consistent exception to this trend is seventh position notes for bass trombone. Secondly, notes in longer positions generally have higher brightness gradients. The first position (solid) line is usually lowest, fourth position (dotted) intermediate, and seventh position (dashed) is usually highest. Exceptions are fourth position on tenor trombones, which shows about the same gradient as first position, and seventh position on bass trombone, which has a lower gradient than fourth posi-

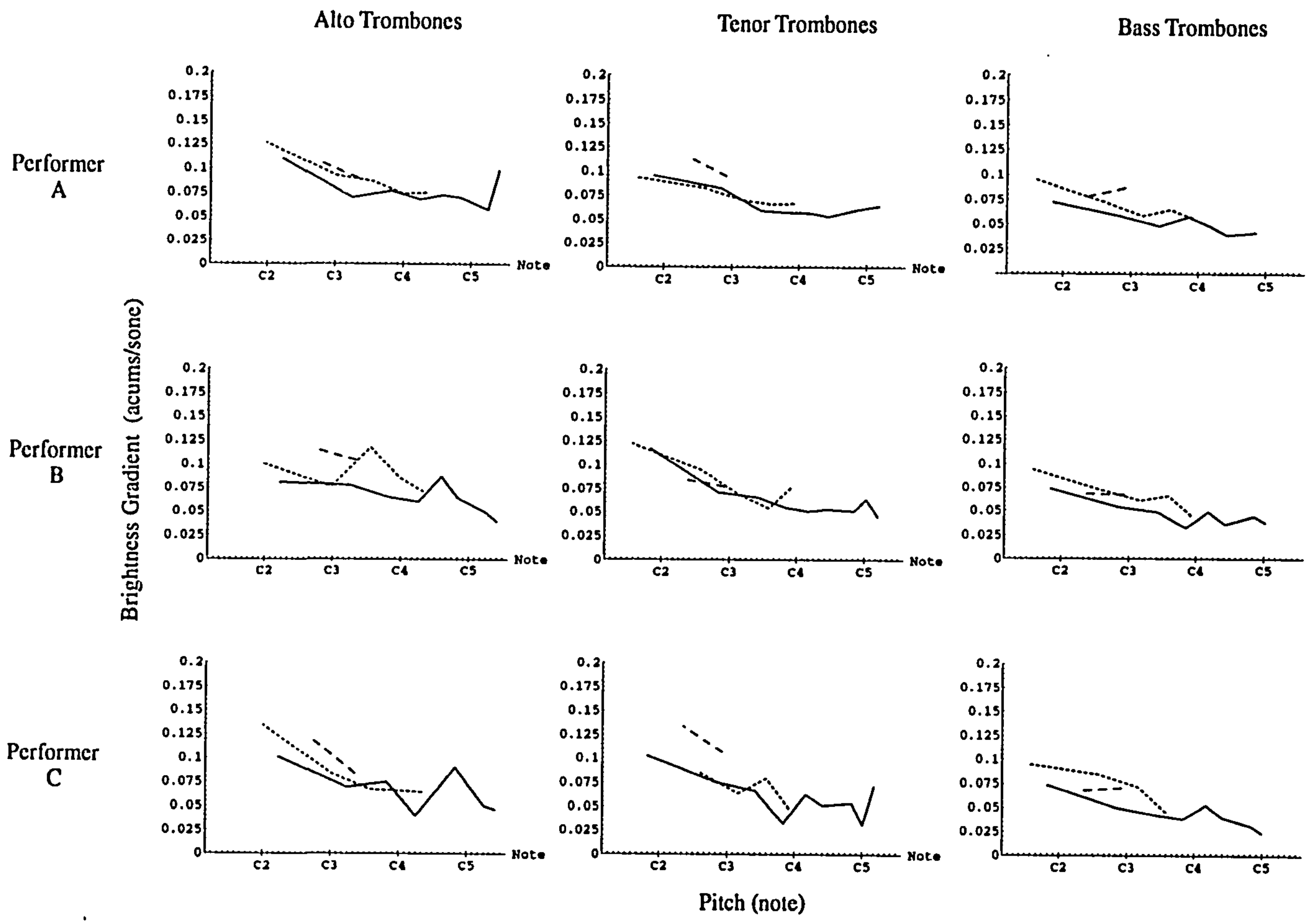


Figure 11.4: Brightness gradient as a function of slide position, at 80 sones. Solid lines are first position notes; dotted notes are fourth position, and dashed notes are seventh position.

tion for Performers B and C. Between first and seventh position, the gradient changes by about 25%. Finally, while alto and tenor trombones show comparable brightness gradients, bass trombone gradients tend to be lower by about 25%.

11.2 Brightness gradient and alternate positions

Slide position is sometimes a performance variable, since many trombone notes can be performed in more than one position. The dependence of brightness gradient on tubing length suggests that brightness would increase more rapidly with loudness for notes in longer positions than in shorter positions.

Data were collected for a number of notes in both regular and extended ("alternate") positions. For each tone, log brightness vs. log loudness is plotted. This is like looking down the pitch axis of the three-dimensional plots in Figures 10.7-10.9, but for only a single pitch. As previously discussed, logarithmic axes reflect perceptions better than linear axes for loudness and brightness. Figure 11.5 compares measurements of a number of notes in their normal positions with others in alternate positions. Each sets of points is shown with a fitted exponential curve. The expected pattern is found for all combinations except G4 on alto trombone. At low and moderate loudness levels, brightnesses are similar; at higher loudnesses brightnesses are greater in the alternate positions than in the regular positions. For the loudest tones, brightness in an extended position may exceed that of the regular position by as much as 20%.

A performer can also select whether to use tubing actuated by a valve. Many trombonists believe that sharp bends within the valve and the tubing it engages adversely affects timbre (e.g., Yeo 1992). On bass and tenor-bass trombones, the note F2 can be played in either sixth position with the valve not engaged, or first position with the valve engaged. The sounding length of the instrument is the same. Figure 11.6 shows log brightness as a function of log loudness for four instances of these notes. Solid lines indicate sixth position notes; dashed lines indicate valve with first position. A consistent pattern is not apparent.

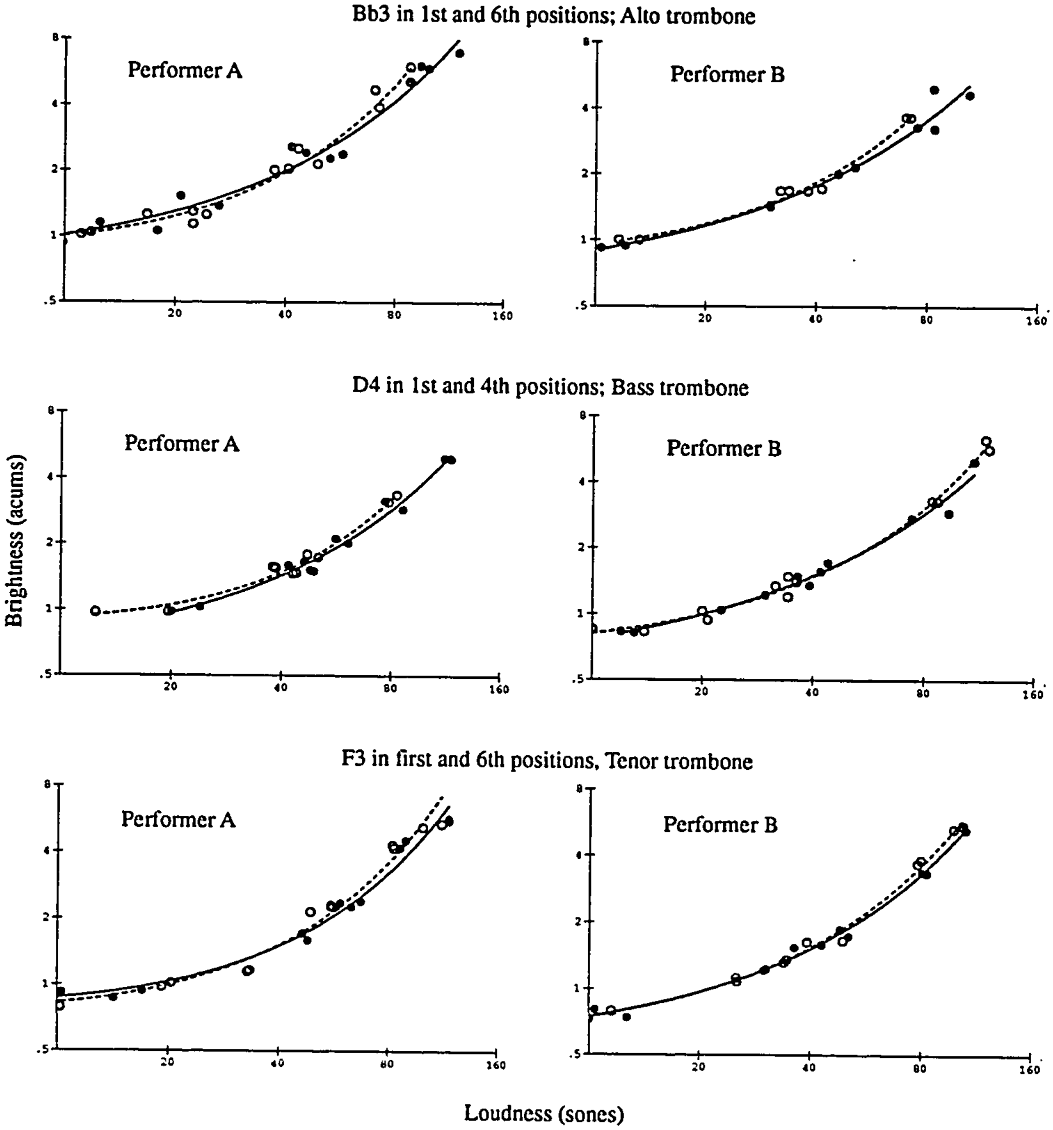


Figure 11.5: Brightness as a function of loudness for alternate and primary positions. Solid lines and filled circles are primary positions; alternate (extended) positions are shown with dashed lines and open circles.

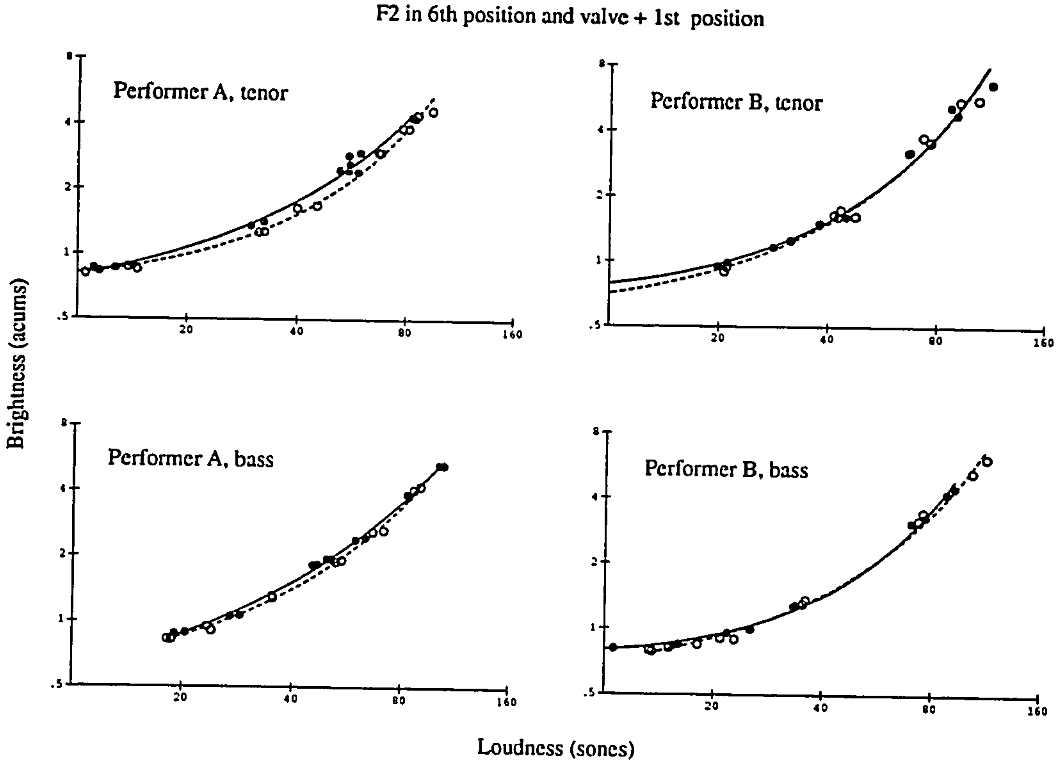


Figure 11.6: Brightness as a function of loudness for sixth-position and F-attachment notes. Solid lines with filled circles are sixth position; valve notes are shown with dashed lines and open circles.

11.3 Comparisons: Average Brightness Differences

Does an alto trombone sound brighter than a tenor trombone, at the same loudness and pitch? Does one performer sound brighter than another? For a gross comparison, the mean difference in log brightness can be calculated for tones at the same pitch and loudness. If one performance is generally brighter than another, this will be reflected in the mean difference. The standard deviation of these differences indicates distribution spread. However, if a performance is brighter in some regions and less bright in others, then the mean will be uninformative. Likewise, a standard deviation cannot indicate whether variance is distributed throughout ranges of loudness and pitch, or confined to specific regions. To answer these questions, graphical representations must be employed.

Since data points are not equally spaced, interpolation must be made in order to compare them. The scheme that has been adopted is shown graphically in Figure 11.7, which shows brightness vs. loudness for two notes. Solid dots represent measured loudnesses and brightnesses. Each measured point is matched to a point linearly interpolated between the other set's measured points (open circles). No points are extrapolated beyond measured extremes in loudness. Each pair of points therefore represents brightnesses at a given loudness, for a certain note. The difference in log brightness is then calculated. Besides compensating for unevenly spaced data, this scheme has the advantage of being symmetric; that is $A - B = -(B - A)$. This eliminates the need for comparing data sets twice.

Performer/instrument combinations are compared by calculating differences in log brightness for all notes, and then calculating means and standard deviations. Figures 11.8a, b, and c show the mean difference in log brightness between all combinations of trombones and performers, along with error bars showing plus and minus one standard deviation. In each cluster of three points, the leftmost point reflects Performer A, the center point is Performer B, and the rightmost point is Performer C. The same information, along with 95% confidence intervals for the means, is presented in tabular form in Tables 11.1, 11.2, and 11.3. A 95% confidence interval indicates the range about a sample mean for which the probability is 0.95 that the population mean is within it. A normal distribution is assumed; from histograms of differences (not shown) this assumption appears justified.

Starting from the left hand side of Figure 11.8a, the first cluster of points compares

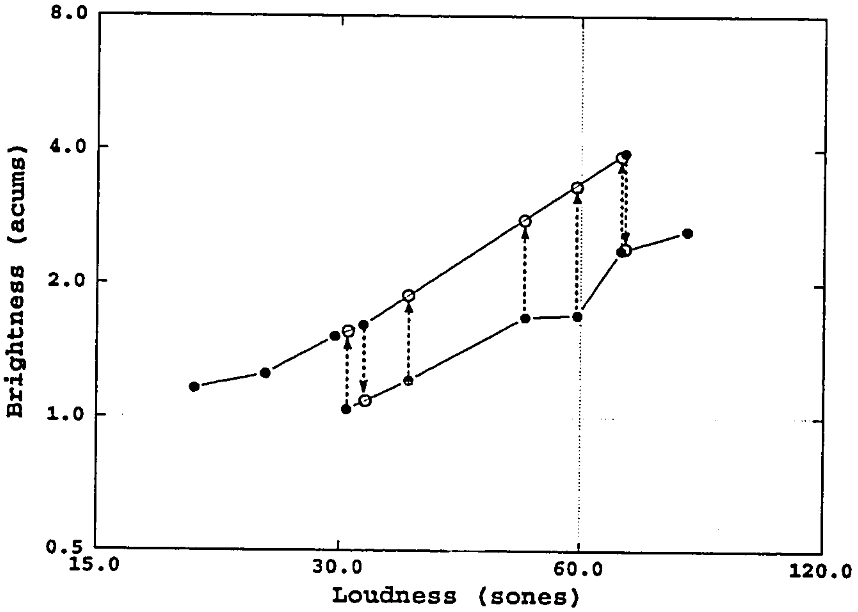


Figure 11.7: Interpolation scheme for comparing brightness at unequal loudnesses. Solid dots are measured data points; open circles are interpolated.

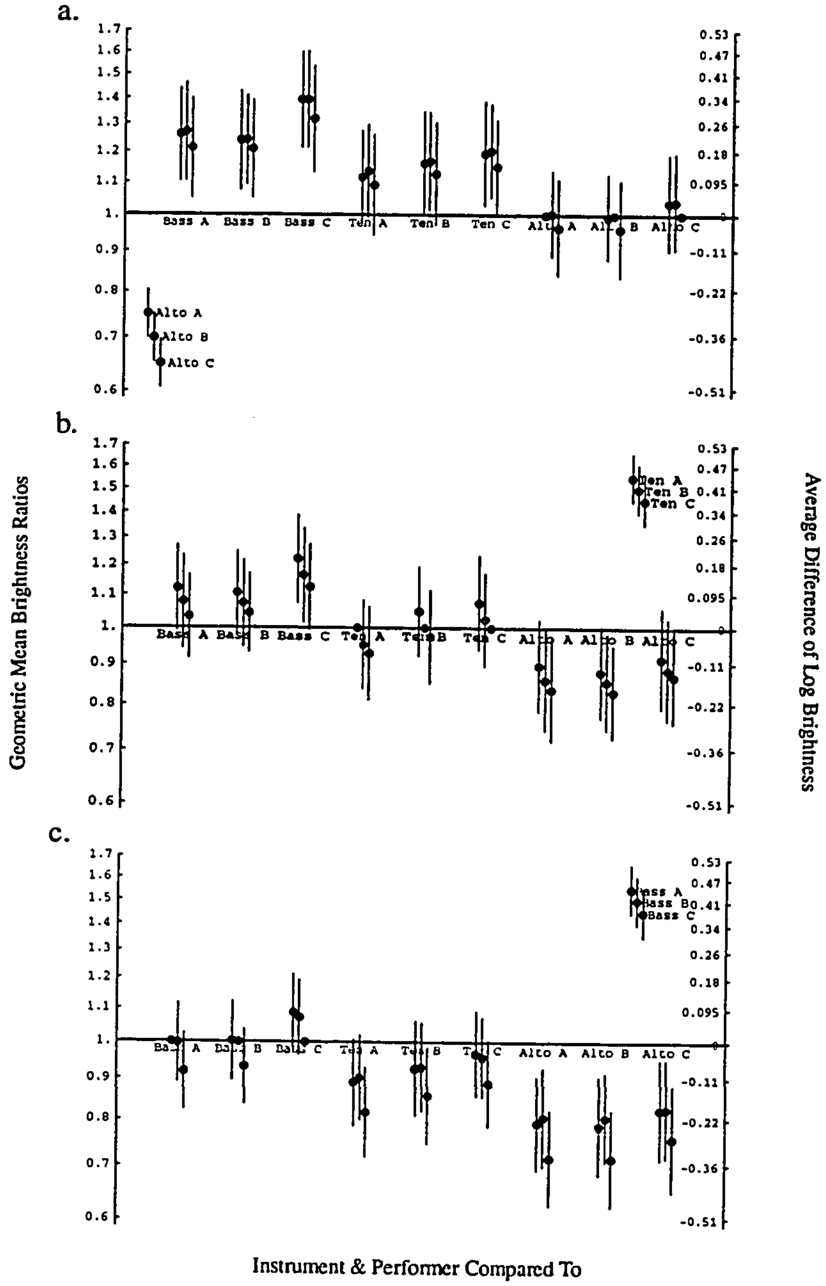


Figure 11.8: Mean differences in log brightness and geometric mean ratios between all combinations of instruments and performers. a. Alto vs. all other trombones; b. Tenor vs. all other trombones; c. Bass vs. all other trombones.

Table 11.1: Alto trombones vs. all others: mean difference of log brightnesses, 95% confidence interval for the mean, and standard deviation.

	Performer A alto	Performer B alto	Performer C alto
Performer A Bass	-0.232 ± 0.0098 0.134	-0.24 ± 0.0107 0.141	-0.194 ± 0.0109 0.143
Performer B Bass	-0.215 ± 0.0102 0.142	-0.216 ± 0.00946 0.128	-0.191 ± 0.0107 0.14
Performer C Bass	-0.332 ± 0.013 0.138	-0.333 ± 0.0132 0.139	-0.277 ± 0.0138 0.153
Performer A Tenor	-0.109 ± 0.00906 0.133	-0.128 ± 0.0095 0.133	-0.0875 ± 0.0108 0.145
Performer B Tenor	-0.15 ± 0.0104 0.149	-0.156 ± 0.0103 0.141	-0.12 ± 0.011 0.147
Performer C Tenor	-0.179 ± 0.0136 0.15	-0.186 ± 0.0124 0.135	-0.14 ± 0.0117 0.135
Performer A Alto	0. ± 0. 0.	-0.00437 ± 0.00836 0.124	0.0357 ± 0.00957 0.139
Performer B Alto	0.00437 ± 0.00836 0.124	0. ± 0. 0.	0.0395 ± 0.00986 0.14
Performer C Alto	-0.0357 ± 0.00957 0.139	-0.0395 ± 0.00986 0.14	0. ± 0. 0.

Table 11.2: Tenor trombones vs. all others: mean difference of log brightnesses, 95% confidence interval for the mean, and standard deviation.

	Performer A tenor	Performer B tenor	Performer C tenor
Performer A Bass	-0.115 ± 0.00825 0.125	-0.0753 ± 0.00916 0.135	-0.0332 ± 0.0106 0.122
Performer B Bass	-0.102 ± 0.00801 0.121	-0.0716 ± 0.00858 0.126	-0.0438 ± 0.0103 0.114
Performer C Bass	-0.2 ± 0.011 0.128	-0.153 ± 0.0119 0.138	-0.119 ± 0.0131 0.124
Performer A Tenor	0. ± 0. 0.	0.0484 ± 0.00832 0.129	0.0738 ± 0.0117 0.136
Performer B Tenor	-0.0484 ± 0.00832 0.129	0. ± 0. 0.	0.0245 ± 0.0119 0.135
Performer C Tenor	-0.0738 ± 0.0117 0.136	-0.0245 ± 0.0119 0.135	0. ± 0. 0.
Performer A Alto	0.109 ± 0.00906 0.133	0.15 ± 0.0104 0.149	0.179 ± 0.0136 0.15
Performer B Alto	0.128 ± 0.0095 0.133	0.156 ± 0.0103 0.141	0.186 ± 0.0124 0.135
Performer C Alto	0.0875 ± 0.0108 0.145	0.12 ± 0.011 0.147	0.14 ± 0.0117 0.135

Table 11.3: Bass trombones vs. all others: mean difference of log brightnesses, 95% confidence interval for the mean, and standard deviation.

	Performer A bass	Performer B bass	Performer C bass
Performer A Bass	0. ± 0. 0.	0.0026 ± 0.00757 0.112	0.0843 ± 0.00899 0.11
Performer B Bass	-0.0026 ± 0.00757 0.112	0. ± 0.0 0.	0.0702 ± 0.00929 0.107
Performer C Bass	-0.0843 ± 0.00899 0.11	-0.0702 ± 0.00929 0.107	0. ± 0. 0.
Performer A Tenor	0.115 ± 0.00825 0.125	0.102 ± 0.00801 0.121	0.2 ± 0.011 0.128
Performer B Tenor	0.0753 ± 0.00916 0.135	0.0716 ± 0.00858 0.126	0.153 ± 0.0119 0.138
Performer C Tenor	0.0332 ± 0.0106 0.122	0.0438 ± 0.0103 0.114	0.119 ± 0.0131 0.124
Performer A Alto	0.232 ± 0.0098 0.134	0.215 ± 0.0102 0.142	0.332 ± 0.013 0.138
Performer B Alto	0.24 ± 0.0107 0.141	0.216 ± 0.00946 0.128	0.333 ± 0.0132 0.139
Performer C Alto	0.194 ± 0.0109 0.143	0.191 ± 0.0107 0.14	0.277 ± 0.0138 0.153

Performer A on alto trombone, Performer B on alto trombone, and Performer C on alto trombone each to Performer A on bass trombone; the next cluster compares Performers A, B, and C on alto trombone to Performer B on bass trombone, and so on. Two scalings of the absissa are provided. On the right hand side of the plot, the absissa is given in units of log brightness. A more intuitive scaling is found on the left hand side of the plot. Since

$$1/n \sum_{n} (\log a_n - \log b_n) = \log[\sqrt[n]{\frac{a_1 a_2 \cdots a_n}{b_1 b_2 \cdots b_n}}], \tag{11.1}$$

the mean difference in log brightness can be thought of as the log of the geometric mean of brightness ratios. Thus, a mean difference in log brightness of 0 is the same as a (geometric) mean ratio of 1; that is, the instruments compared would have on average the same brightness.

Figure 11.8a compares performances on alto trombone to all others; Figure 11.8b shows tenor trombones vs. all others and Figure 11.8c shows bass trombones vs. all others. A number of observations can be made. First, for each performer the alto is on average brighter than the tenor, and the tenor is brighter than the bass. Furthermore, altos, tenors and basses as groups are distinct; in no instance, for example, does any tenor appear brighter than an alto or less bright than a bass. Consistent differences between performers show up as well. Performer C is consistently less bright than either A or B. Performers A and B are on average very similar in brightness for alto and bass trombones; on tenor trombone, however, B is roughly intermediate between A and C.

Notice that standard deviations are on the order of differences between instruments, and considerably larger than differences between players. As mentioned above, a single-valued standard deviation cannot indicate whether variance is randomly distributed through loudness and pitch spaces, or concentrated in regions where differences are predominantly positive or negative. Graphical representations will be presented in the following section which can address this question.

11.4 Comparisons: Graphical Representations

As mentioned above, differences in log brightness can be expressed as brightness ratios. Description in terms of ratios has been employed in the following discussions, since it can be more directly related to perception.

Three-dimensional plots are not satisfactory for comparing performer/instrument combinations, since visual perspective makes it impossible to estimate coordinates. Scatter plots are preferable. Brightness ratios for each note are calculated as described above, and plotted on loudness vs. pitch axes. Each ratio is presented as a dot at its pitch and loudness; the size of the dot is scaled (by radius) to the brightness ratio. For clarity, ratios greater than 1.0 are presented in separate plots from ratios less than 1.0.

11.4.1 Comparisons of Different Trombones, Same Player

To examine how brightness as a function of loudness and pitch can be used to differentiate between trombones, brightness ratios are presented comparing alto, tenor and bass trombones as performed by the same player. Figure 11.9 compares Performer A on alto, tenor, and bass trombones. Figure 11.9a shows that, for all but a few notes in the upper range, ratios are greater than 1, indicating that alto is brighter than tenor. Ratios tend to be greatest between 40 and 80 sones and in the middle range. Ratios less than 1 are consistently present only above the note A♭4, at both high and low loudnesses, but not at moderate loudnesses.

Figures 11.9b compares Performer A on tenor vs. bass trombones. For all but a few scattered points, tenor is brighter than bass. Ratios tend to be largest for pitches above F4 and for loudnesses in excess of 40 sones.

Finally, Performer A's alto trombone is compared to bass trombone, in Figures 11.9c. The alto is considerably brighter than the bass except for the very highest notes B4, C5 and C♯5, for which they are of comparable brightness.

Similar comparisons are presented for Performer B on alto, tenor, and bass trombones, as shown in Figures 11.10a, b, and c.

The following observations can be made:

- Alto trombone is brighter than tenor in general. The differences between alto and tenor as played by Performer B are similar to those shown by Performer A, with two exceptions. First, Performer B does not exhibit as many high range tones with ratios less than 1.0. By referring back to Figure 10.10, it can be seen that above the note B4 equal-loudness contours for Performer B tend to turn upwards. For Performer A this tendency is less pronounced. Secondly, some extremely bright tones are produced by Performer B in the lowest register. Two notes–C♯2 and E4–are on average less bright than on tenor.

- Tenor trombone is on average brighter than bass trombone. However, compared to Performer A differences are smaller and more evenly distributed. There is a marked tendency for tenor to be less bright than bass below 14 sones.
- Alto trombone is consistently brighter than bass trombone. Differences are similar to those of Performer A, except for the very highest and very lowest notes.

Finally, the following observations can be made when comparing alto, tenor, and bass trombones for Performer C, in Figures 11.11a, b, and c.

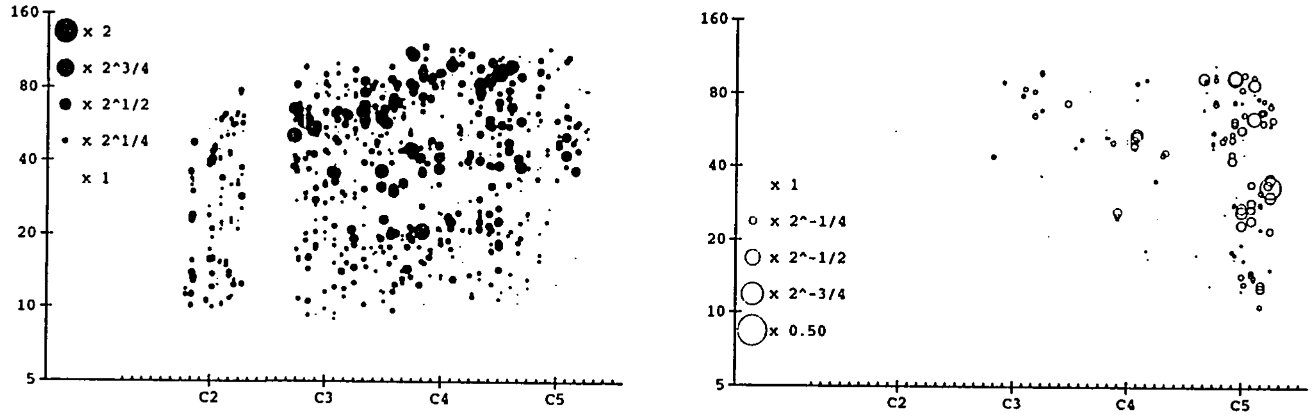
- Alto is brighter than tenor. Ratios tend to be larger at loudnesses in excess of 40 sones.
- Tenor is brighter than bass, especially for high pitches and loudnesses.
- Alto is much brighter than bass. Ratios tend to be higher for Performer C than either A or B.

To summarize, brightness ratios as a function of pitch and loudness can be used to differentiate between alto, tenor, and bass trombones. For all performers alto is brighter than tenor which is brighter than bass. Relatively small individual differences appear in the degree to which one is brighter than the other, and the ranges over which ratios are comparable.

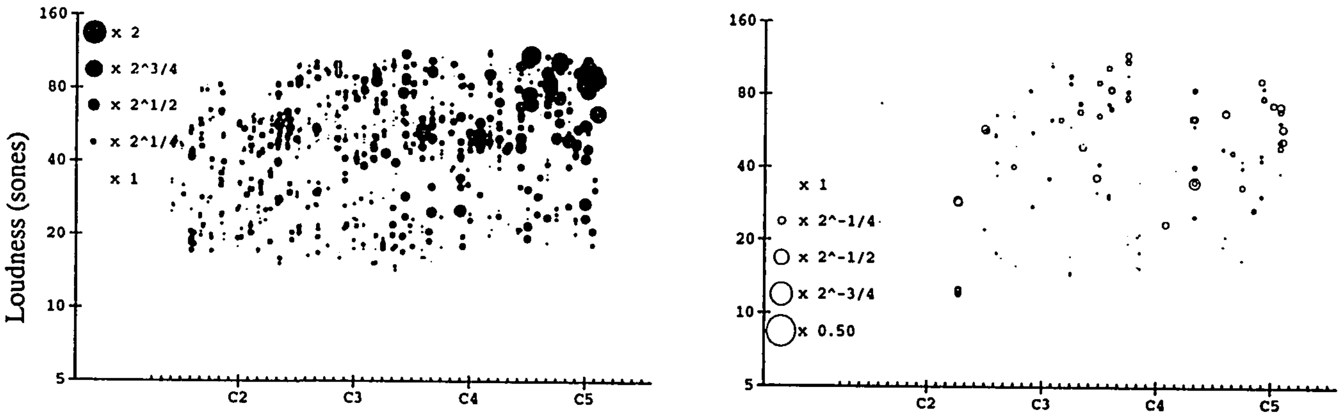
11.4.2 Comparisons of Different Performers, Same Instrument

It remains to be shown whether brightness as a function of loudness and pitch can be used to differentiate between performers using the same instrument. As shown in average ratios (Figure 11.8), differences between performers are much smaller than those between instruments. In fact, they are on the order of differences in brightness measured at the two microphone positions used in this experiment. The implications of this will be discussed below. Some sort of averaging or smoothing appears necessary.

a. Alto vs. Tenor Trombone



b. Tenor vs. Bass Trombone



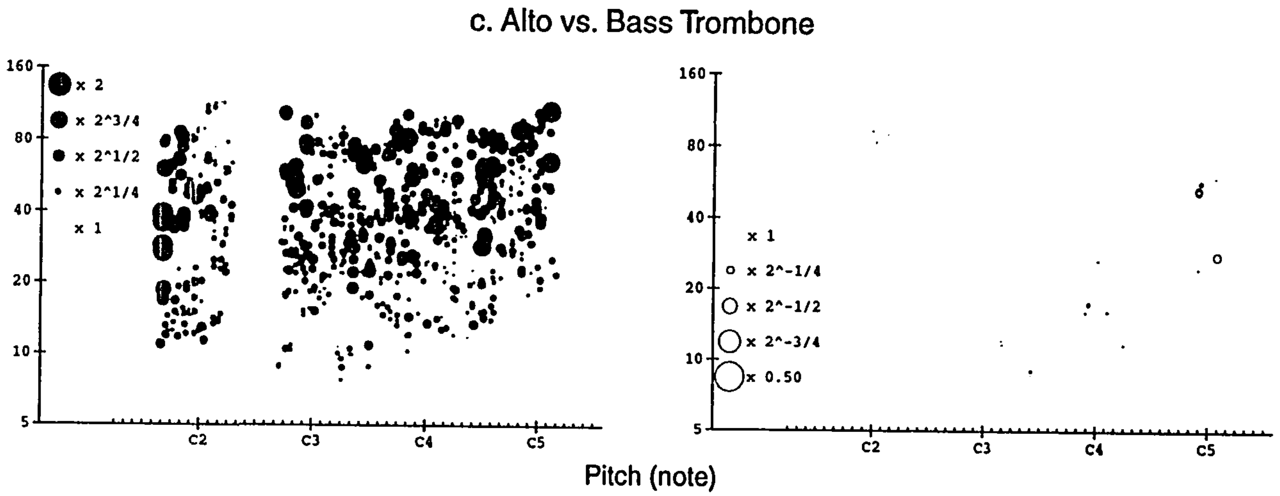
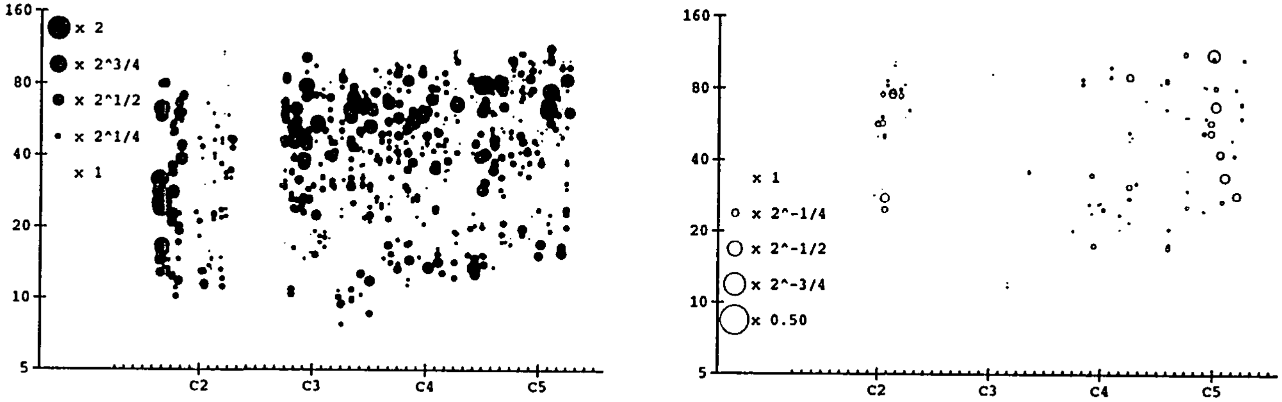
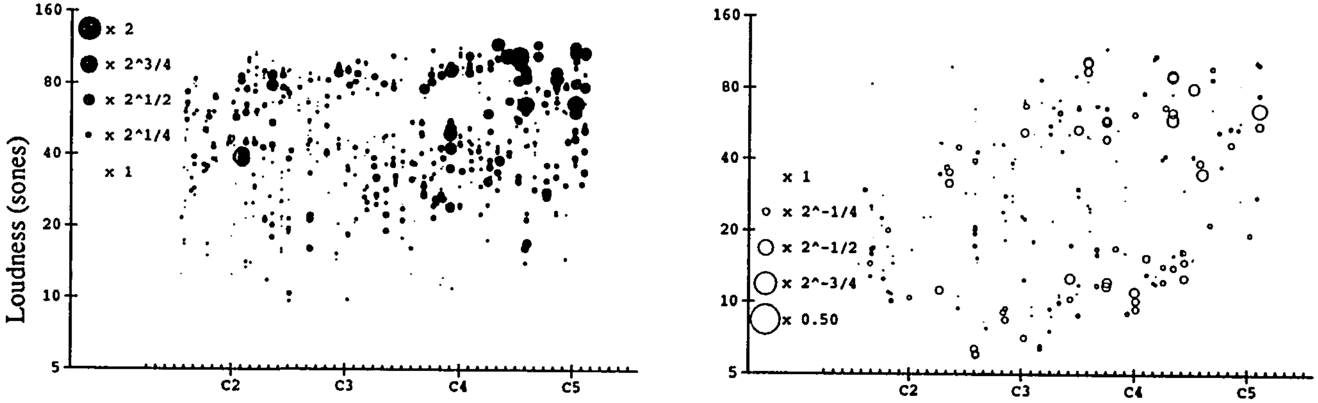


Figure 11.9: Note by note ratios comparing Performer A on alto, tenor and bass trombones. Ratios greater than 1 appear on the left plots; ratios less than 1 are on the right. a. Alto vs. tenor trombone; b. Tenor vs. bass trombone; c. Alto vs. bass trombone.

a. Alto vs. Tenor Trombones



b. Tenor vs. Bass Trombones



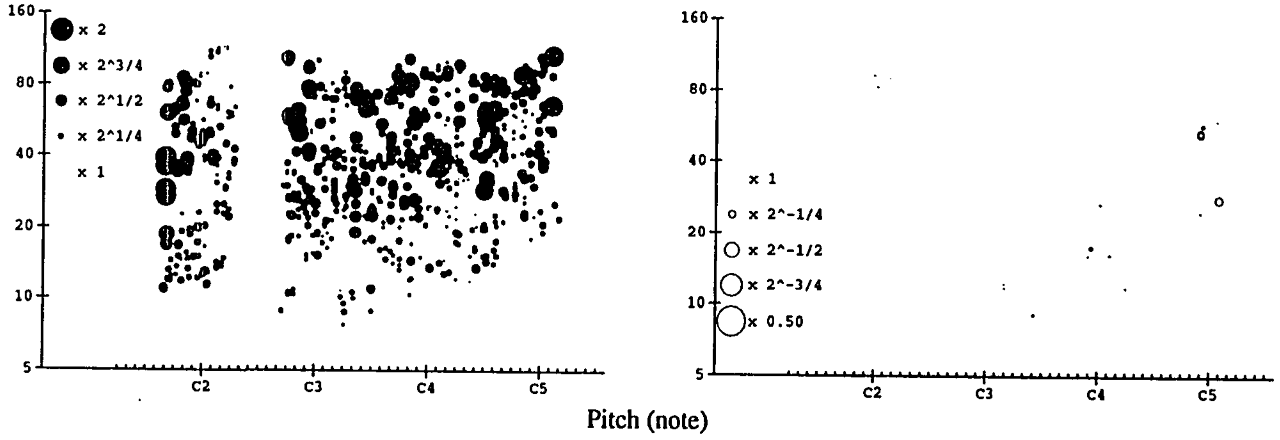
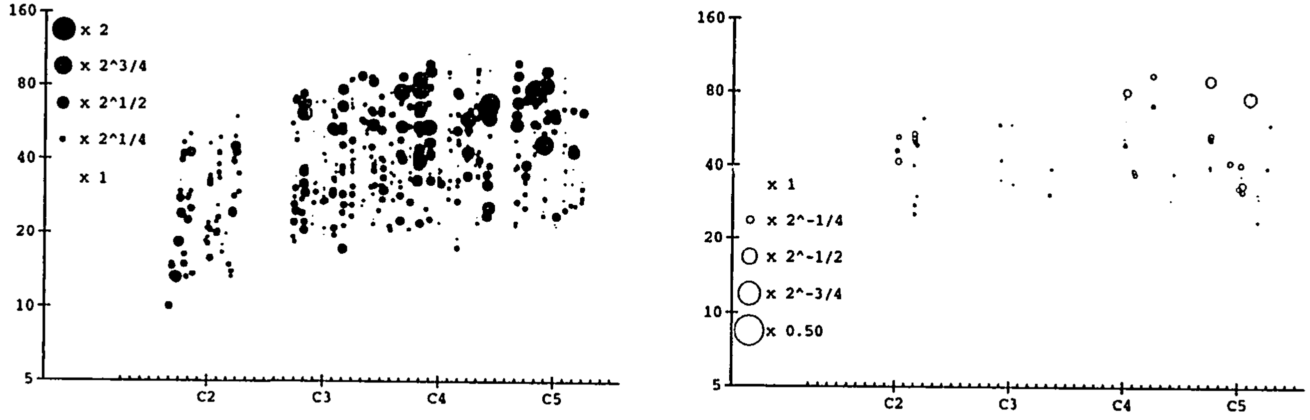
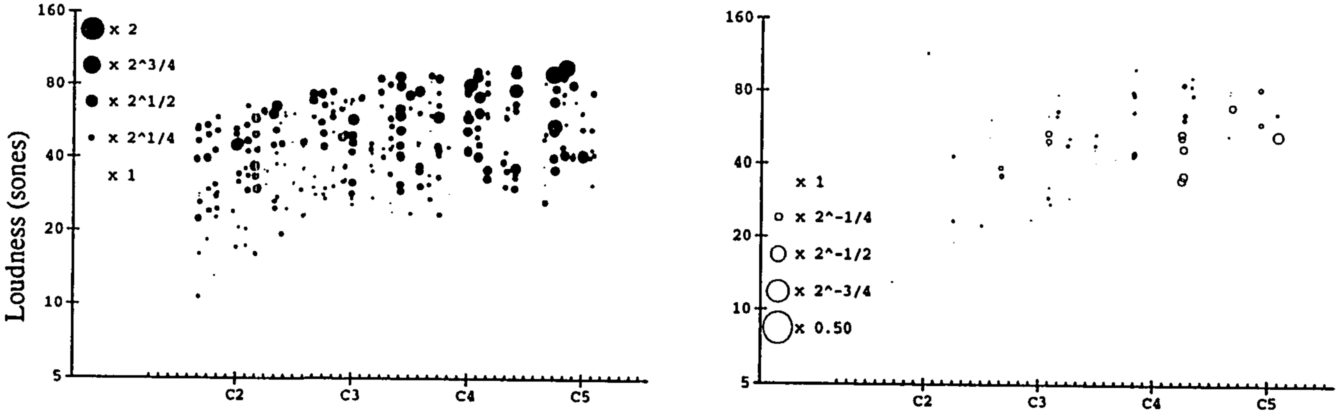

c. Alto vs. Bass Trombone

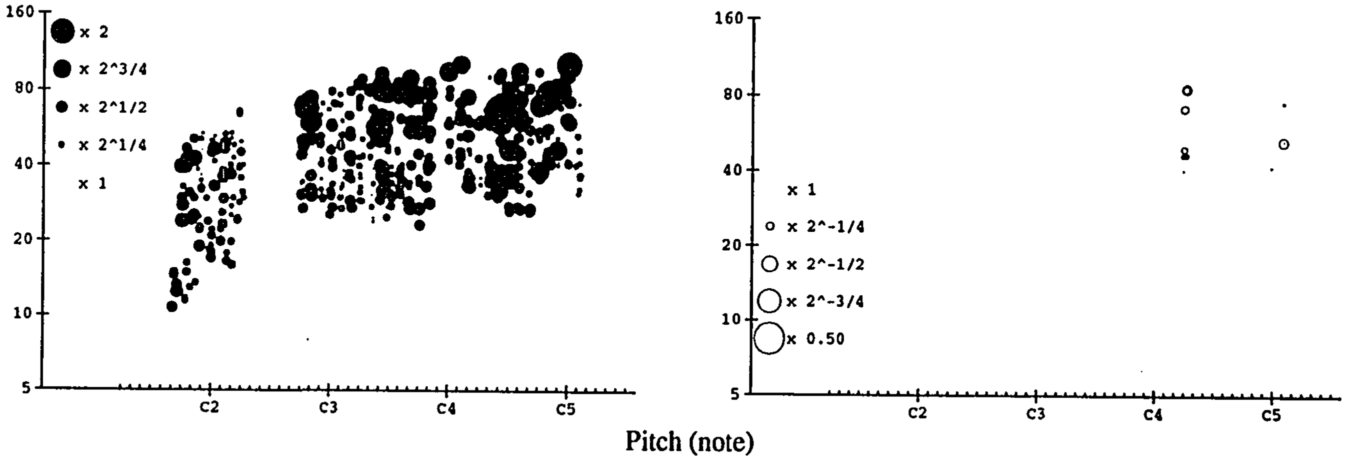
Figure 11.10: Note by note ratios comparing Performer B on alto, tenor and bass trombones. Ratios greater than 1 appear on the left plots; ratios less than 1 are on the right. a. Alto trombone vs. tenor; b. Tenor trombone vs. bass; c. Alto trombone vs. bass.

a. Alto vs. Tenor Trombone



b. Tenor vs. Bass Trombone





c. Alto vs. Bass Trombone

Figure 11.11: Note by note ratios comparing Performer C on alto, tenor and bass trombones. Ratios greater than 1 appear on the left plots; ratios less than 1 are on the right. a. Alto trombone vs. tenor; b. Tenor trombone vs. bass; c. Alto trombone vs. bass.

Figure 11.12 illustrates the problem for the specific case of comparing Performer A and C on alto trombone. Brightness ratios as a function of pitch and loudness

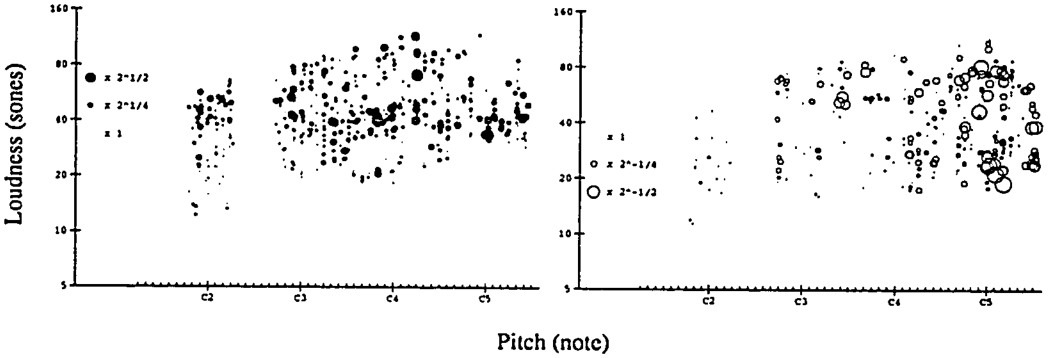


Figure 11.12: Note by note brightness ratios comparing Performer A on alto trombone to Performer C on alto trombone. Ratios greater than 1 appear on the left plots; ratios less than 1 are on the right.

are presented as before. Note that ratios are in general smaller than when comparing different instruments. Note also that ratios appear to vary not just with pitch, but with loudness as well. For example, in the range from A♭4 to F5, brightnesses tend to be greater for Performer A in the vicinity of 30 sones; for louder and softer notes, however, Performer C tends to be brighter.

In the three-dimensional grid plots of Figures 10.7-10.9, it was seen that log brightness vs. log loudness vs. log frequency values appear to cluster about a surface whose form is similar for all trombones and trombonists. This suggests that appropriate surfaces can be fit to the data, and the surfaces compared rather than the raw data. In this way, global trends can be isolated that are difficult to observe in the raw data. Of course, a surface simple enough to be manageable will necessarily eliminate detail as well as smooth measurement variability. The advantage of fitting to a surface is that it effectively smooths in all three dimensions, and the surface is constrained by the largest number of points. For example, in the calculation of the brightness gradient an exponential curve was fitted to measurements for a single pitch. Anomalous values near the endpoints disproportionately influence the fit. When fitting to a surface, however, neighboring pitches as well as loudnesses help to determine appropriate parameters.

Measured log brightness can be closely approximated by a surface that varies quadratically with both log loudness and log frequency; that is, an equation of the form

$$\log brightness = (a_1 * \log loudness^2 + a_2 * \log loudness + a_3)(b_1 * note^2 + b_2 * note + b_3) \tag{11.2}$$

where the a's and b's are coefficients determined by a least means square fit, and *note* is a scaled version of log frequency, given by

$$note = 12 \log_2 (f / f_0) \tag{11.3}$$

where f is fundamental frequency and f_0 is the equal-tempered frequency of C1 (32.7 Hz). Thus, C1 is scaled to 0, C♯1 is scaled to 1, and so on.

Figure 11.13 shows representative data (Performer A on tenor trombone), along with a surface of the form above. Data points are shown with filled-in dots; black line segments show the difference in log brightness between data and fit. Two views are presented, one from upper front right, and one from lower back left. As can be seen, there is little evidence of systematic errors between data and fit; points are equally distributed above and below the surface. Table 11.4.2 lists the coefficients of each surface's equation.

Table 11.4.2 gives two indications of error between each data set and surfaces fit to each data set. The *total error* is simply the sum of differences between measured log brightness and the surface, normalized by the total number of data points. For an ideal fit this should be 0, signifying that systematic errors either are not present or are compensated for by offsetting errors in another region. *Absolute error* is the sum of the absolute value of the same, also normalized. This indicates the total differences between log data and fit, regardless of sign, and should be a small number. Three observations can be made: (1) errors between surfaces fit to a data set are much smaller than between surfaces fit to other data sets; (2) errors between surfaces fit to other trombones *of the same type* (e.g., a surface fit to one bass trombone performance compared to data from other bass trombone performances) are smaller than those comparing a surface fit to one kind of trombone to data from a different kind. This suggests that each surface represents its respective data set very well; that surfaces represent data from the same kind of trombone but other performers less well; and that they do a poor job of representing data from different trombones.

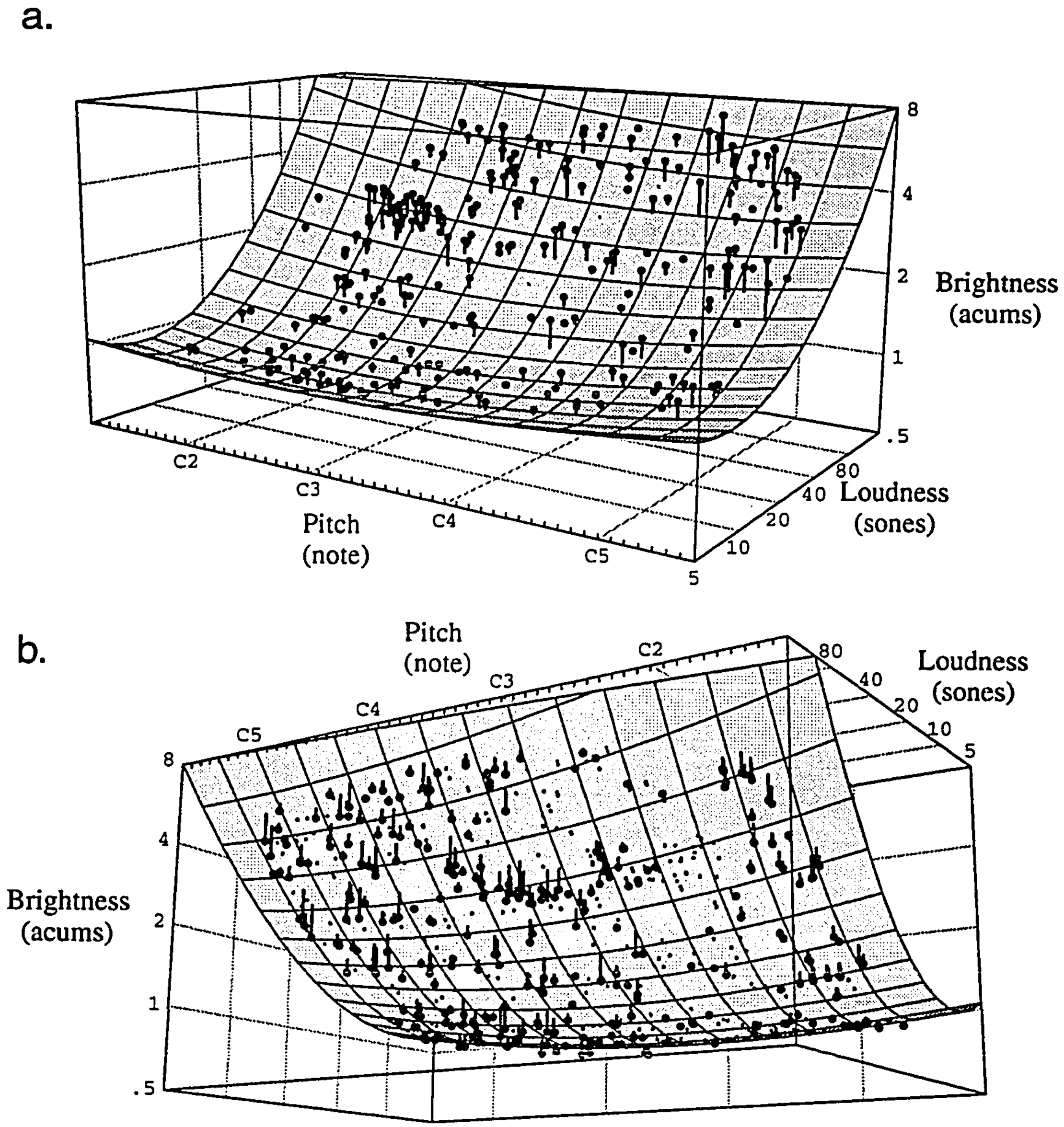


Figure 11.13: Brightness vs. loudness and pitch for Performer A, tenor trombone, with fitted surface. Dots are data points; line segments show distance from surface. a. Front view; b. Back view.

Table 11.4: Smoothing surface coefficients for each trombone/trombonist combination.

		x	x^2	y	xy	x^2y	y^2	xy^2	x^2y^2
Performer A Bass	1.28E-1	1.4E-1	-2.55E-3	-7.77E-1	-8.28E-2	1.64E-3	2.48E-1	1.09E-2	-2.38E-4
Performer B Bass	2.09E	-1.77E-3	-3.24E-4	-1.9E	-3.22E-3	4.18E-4	4.08E-1	-4.84E-4	-6.35E-5
Performer C Bass	2.22E	2.69E-2	-1.74E-3	-2.13E	-2.2E-3	8.46E-4	4.55E-1	-2.74E-3	-7.94E-5
Performer A Tenor	1.81E	-1.24E-2	-3.14E-5	-1.69E	3.23E-3	2.E-4	3.79E-1	-1.3E-3	-2.91E-5
Performer B Tenor	1.81E	-1.53E-2	-2.28E-4	-1.79E	3.1E-3	4.04E-4	4.08E-1	-1.56E-3	-5.89E-5
Performer C Tenor	2.66E	-9.59E-2	2.31E-3	-2.27E	5.18E-2	-1.05E-3	4.8E-1	-9.16E-3	1.5E-4
Performer A Alto	3.35E-1	1.15E-1	-2.19E-3	-6.79E-1	-7.84E-2	1.58E-3	2.3E-1	1.11E-2	-2.39E-4
Performer B Alto	1.23E	2.31E-2	-6.21E-4	-1.14E	-2.92E-2	7.55E-4	2.92E-1	4.21E-3	-1.25E-4
Performer C Alto	1.88E	-4.22E-2	1.5E-3	-1.39E	-2.67E-3	-2.34E-4	3.03E-1	2.11E-3	-1.72E-5

Table 11.5: Two measures of error between measured brightness and fitted surface: total error, and absolute error.

		Fitted Surface								
Measured Data		Perf. A Bass	Perf. B Bass	Perf. C Bass	Perf. A Tenor	Perf. B Tenor	Perf. C Tenor	Perf. A Alto	Perf. B Alto	Perf. C Alto
Perf. A Bass	Total Error	-3.63E-12	3.96E-3	-7.97E-2	1.08E-1	6.07E-2	4.96E-2	2.25E-1	2.39E-1	2.12E-1
	Absolute Error	7.54E-2	7.35E-2	1.04E-1	1.2E-1	1.18E-1	8.9E-2	2.25E-1	2.39E-1	2.14E-1
Perf. B Bass	Total Error	3.68E-3	8.94E-12	-8.E-2	1.06E-1	5.76E-2	4.87E-2	2.11E-1	2.23E-1	2.E-1
	Absolute Error	7.78E-2	7.03E-2	1.01E-1	1.19E-1	1.13E-1	8.68E-2	2.14E-1	2.23E-1	2.02E-1
Perf. C Bass	Total Error	8.53E-2	7.19E-2	3.1E-12	1.94E-1	1.43E-1	1.27E-1	3.11E-1	3.05E-1	2.85E-1
	Absolute Error	1.07E-1	9.5E-2	7.09E-2	1.95E-1	1.74E-1	1.37E-1	3.11E-1	3.05E-1	2.85E-1
Perf. A Tenor	Total Error	-1.12E-1	-1.01E-1	-2.01E-1	4.33E-12	-5.29E-2	-6.74E-2	1.11E-1	1.16E-1	8.87E-2
	Absolute Error	1.26E-1	1.15E-1	2.01E-1	7.51E-2	1.04E-1	1.01E-1	1.29E-1	1.29E-1	1.15E-1
Perf. B Tenor	Total Error	-6.08E-2	-4.53E-2	-1.5E-1	5.7E-2	3.12E-12	-1.41E-2	1.6E-1	1.64E-1	1.32E-1
	Absolute Error	1.E-1	9.64E-2	1.55E-1	9.85E-2	8.34E-2	8.42E-2	1.73E-1	1.72E-1	1.5E-1
Perf. C Tenor	Total Error	-5.12E-2	-6.01E-2	-1.33E-1	4.45E-2	-4.09E-3	-4.94E-12	1.43E-1	1.59E-1	1.48E-1
	Absolute Error	9.47E-2	9.2E-2	1.4E-1	1.07E-1	1.04E-1	7.67E-2	1.81E-1	1.69E-1	1.57E-1
Perf. A Alto	Total Error	-2.44E-1	-2.26E-1	-3.31E-1	-1.24E-1	-1.76E-1	-1.89E-1	3.52E-13	8.11E-3	-2.45E-2
	Absolute Error	2.45E-1	2.27E-1	3.31E-1	1.39E-1	1.88E-1	1.98E-1	7.73E-2	8.34E-2	9.69E-2
Perf. B Alto	Total Error	-2.33E-1	-2.19E-1	-3.26E-1	-1.16E-1	-1.66E-1	-1.87E-1	-6.74E-3	5.59E-12	-3.57E-2
	Absolute Error	2.35E-1	2.2E-1	3.26E-1	1.3E-1	1.78E-1	1.93E-1	8.74E-2	7.84E-2	9.44E-2
Perf. C Alto	Total Error	-1.92E-1	-1.87E-1	-2.81E-1	-9.7E-2	-1.4E-1	-1.49E-1	2.65E-4	2.52E-2	-2.45E-12
	Absolute Error	1.96E-1	1.91E-1	2.81E-1	1.25E-1	1.63E-1	1.6E-1	1.13E-1	9.06E-2	8.54E-2

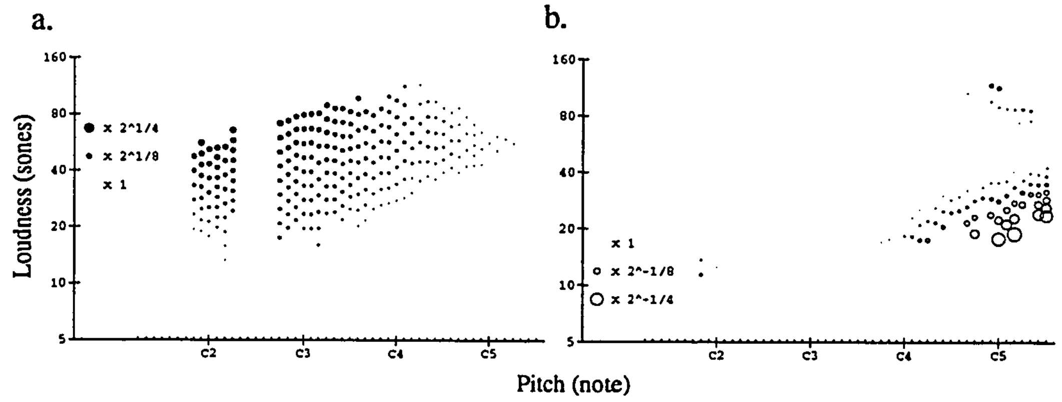


Figure 11.14: Brightness ratios of Performer A to Performer C on alto trombone, by means of fitted surfaces. a. Ratios greater than 1.0; b. Ratios less than 1.0.

Since these equations appear to represent the data exceptionally well, it may be possible to use them for comparisons. We can return to the comparison of Performers A and C on alto trombone. The scatter plot for Performer A presented previously in Figure 11.12 shows that ratios vary widely and a pattern is not obvious. Closer inspection, however, shows that at moderate loudnesses (about 40 sones) ratios are consistently greater than 1.0, indicating that Performer A is brighter. At high pitches, however, numerous ratios less than 1.0 are seen, especially at low and high loudnesses. Are these regularities captured in a comparison of fitted surfaces?

Figure 11.14 shows a comparison of the fitted surfaces representing Performers A and C on alto trombone. To obtain these plots, surfaces were first fitted to the respective data sets. Then, for each pitch, the loudness range that the sets have *in common* was found. Ten evenly spaced loudnesses were then interpolated over this range, and disks drawn with radii proportional to the ratio of brightness between the surfaces. No dots are shown beyond what is common to both sets. It can be seen that patterns apparent in the scatter plot of Figure 11.12 also appear when fitted surfaces are compared. That is, for high pitches A tends to be less bright than C, especially at both high and low loudnesses; and for lower pitches A is brighter than C, especially at high loudnesses. This suggests that the surfaces maintain sufficient detail to provide interesting comparisons, while eliminating measurement variability.

Figures 11.15-11.17 show brightness ratios for each combination of performer, on alto, tenor, and bass trombones, calculated via fitted surfaces as above. Figure 11.15, for example, shows brightness ratios of Performer A to Performer B. Figure 11.15a shows brightness ratios for alto trombone, Figure 11.15b for tenor, and Figure 11.15c for bass trombone.

Comparisons of like-instrument brightnesses by means of fitted surfaces reveals patterns of differences between performers that extend across instrument types. For example, Performers A and B are compared for alto, tenor, and bass trombones in Figure 11.15. For all trombones, Performer A is slightly brighter than Performer B over the majority of pitches and loudnesses. However, in the high register brightness for Performer A falls below that of Performer B. This always begins in the middle loudness region (30-40 sones), but with increasing pitch soon encompasses the entire loudness range. The surface in this region is concave upwards. In the lowest pitch regions Performer A's brightness also dips below that of Performer B, but only at the extremes of loudness. The congruence of the top and bottom plots of Figure 11.15 (alto and bass, respectively) is remarkable. The middle plot of the same figures, comparing A and B on tenor trombone, differs in detail but not form; overall ratios of brightness are slightly larger throughout.

Similarities of form are clear when comparing Performers A and C (Figure 11.16) across trombone types as well. In all cases brightnesses are most similar in the middle and lower register, at low loudnesses. Ratios are largest at high loudnesses. In the upper range brightness ratios begin to decrease, and for alto and tenor trombones Performer A becomes less bright than C over a limited range. However, in contrast with the pattern observed between Performers A and B, this region is concave downwards.

Comparisons between Performers B and C across instruments (Figure 11.17) show a similar pattern as between A and C, with two differences. First, for all trombones the decrease in brightness ratios in the upper pitch range is less than that observed between Performers A and C. Secondly, Performer B brightness dips below that of Performer C in the lower register for the tenor trombone, but not for alto or bass.

While these patterns are striking, it is important to note that the ratios between players are on the order of $2^{(1/8)}$, or about 9%. Thus, they are on the order of the standard deviation for brightness measurements. Could the observed patterns of differences between performers be an artifact of the fitting procedure?

11.5 Measurement Accuracy for Smoothed Surfaces

As discussed in Appendix A, for an estimate of accuracy for the surfaces it is necessary to return to the raw data. This is because errors in measured loudness and brightness vary with dynamic and register. Errors near the extremes will affect contours more than errors in the middle; thus, a knowledge of the overall standard deviation does not necessarily indicate confidence in a particular region. The simplest means of estimating accuracy in a particular region is by comparing left and right channel recordings of the same tone.

Confidence in each of the circles in Figures 11.15-11.17 has been assessed in the following manner. For each performer, measurements from the right channel were separated from the left. Surfaces were then fitted to each channel according to Equation 11.2. For each point of comparison, the brightness ratio between right and left channel $\frac{br_r}{br_l}$ was calculated, for each performer. An estimate of error was calculated according to

$$error = \sqrt{[\frac{br_r}{br_l}(p)]^2 + [\frac{br_r}{br_l}(q)]^2}, \tag{11.4}$$

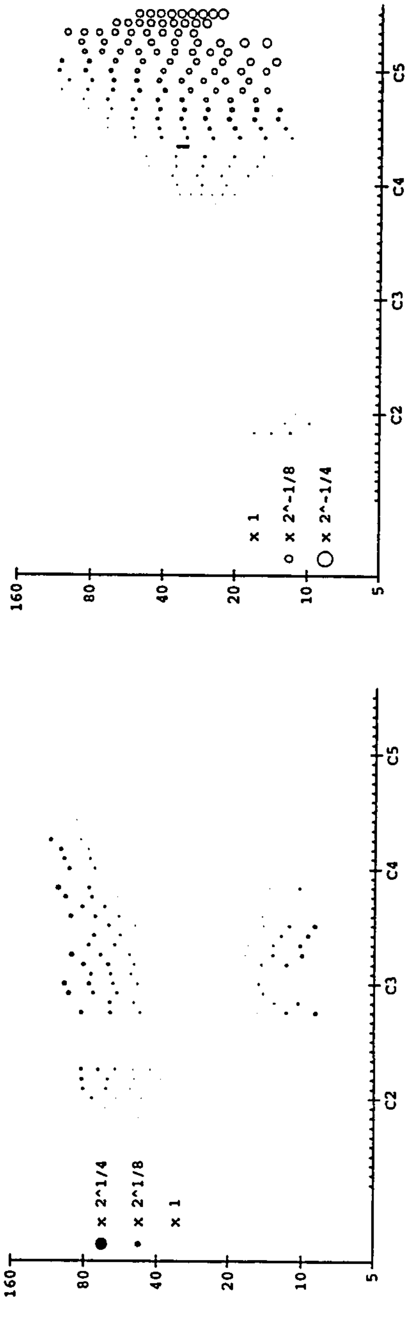
where $\frac{br_r}{br_l}(p)$ corresponds to the brightness ratio between channels for performer p, and $\frac{br_r}{br_l}(q)$ for performer q. For Performer C on tenor and bass trombones, only one channel was recorded; in this case, the error for the other performer (for whom there were two channels) was simply doubled.

Figures 11.15-11.17 were then reprocessed, and ratios of smaller magnitude than the error calculated by Equation 11.4 omitted. The results are shown in Figures 11.18-11.20. Figures 11.15 can be compared to Figures 11.18; Figures 11.16 can be compared to Figures 11.19, and so on. In general, ratios greater than $2^{\frac{1}{8}}$ or less than $2^{\frac{-1}{8}}$ have survived. This means that many measured differences between performers are greater than the error suggested by interchannel variability.

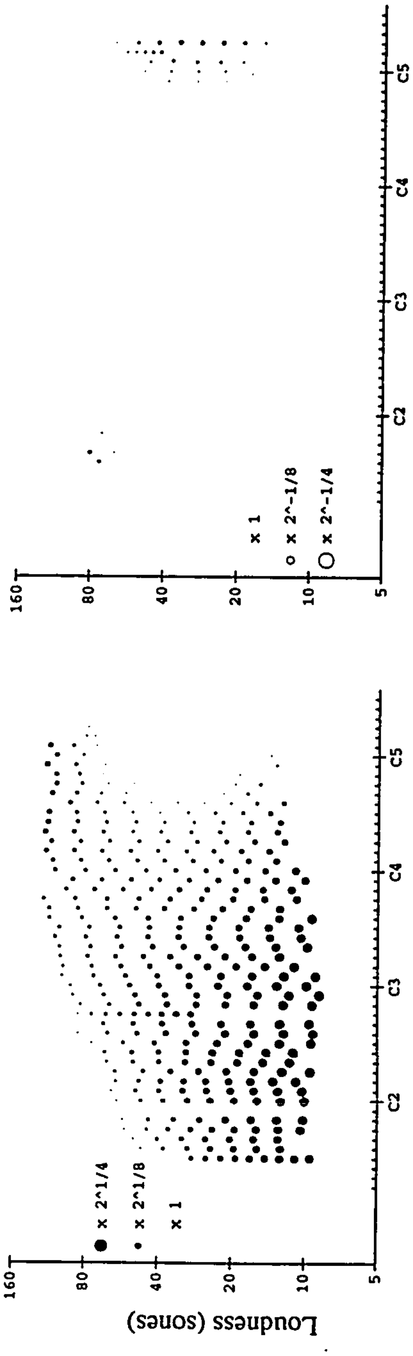
To summarize, differences between performers on the same instrument are not apparent from the raw data due to measurement variability. Some kind of smoothing is required. By fitting measured loudness, brightness and pitch values to a surface of the form of Equation 11.2, patterns in brightness ratios emerge that are distinctive of the performers being compared, and which generally carry across alto, tenor, and bass trombones. The ratios are close to measurement resolution; however, they are robust. When individual channels are compared, the same patterns emerge.

Performer A vs. Performer B

a. Alto Trombone



b. Tenor Trombone



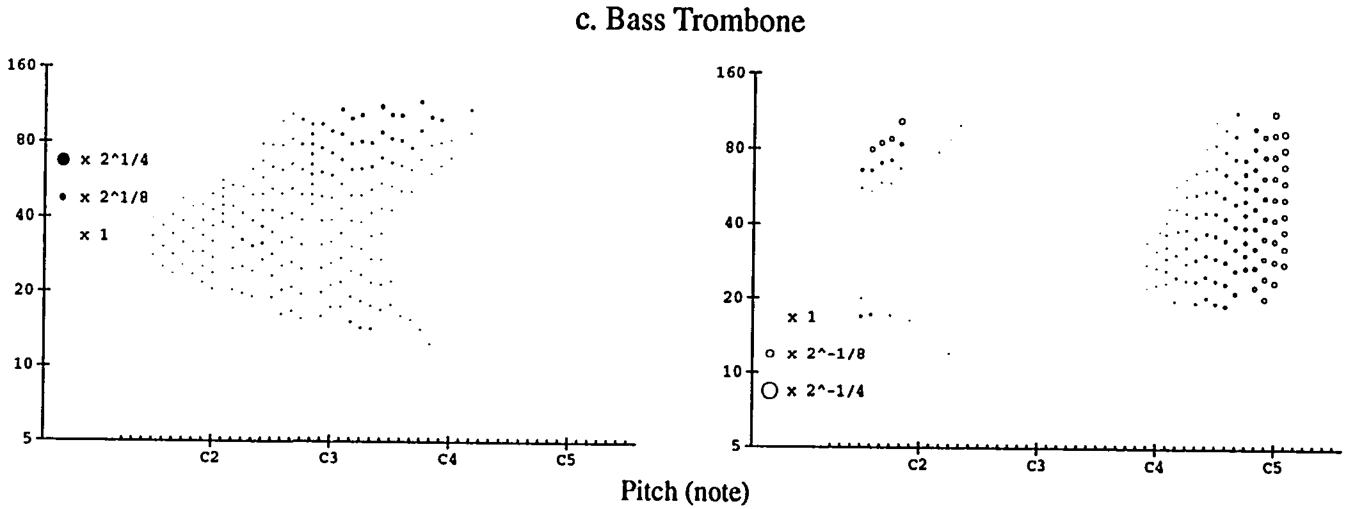
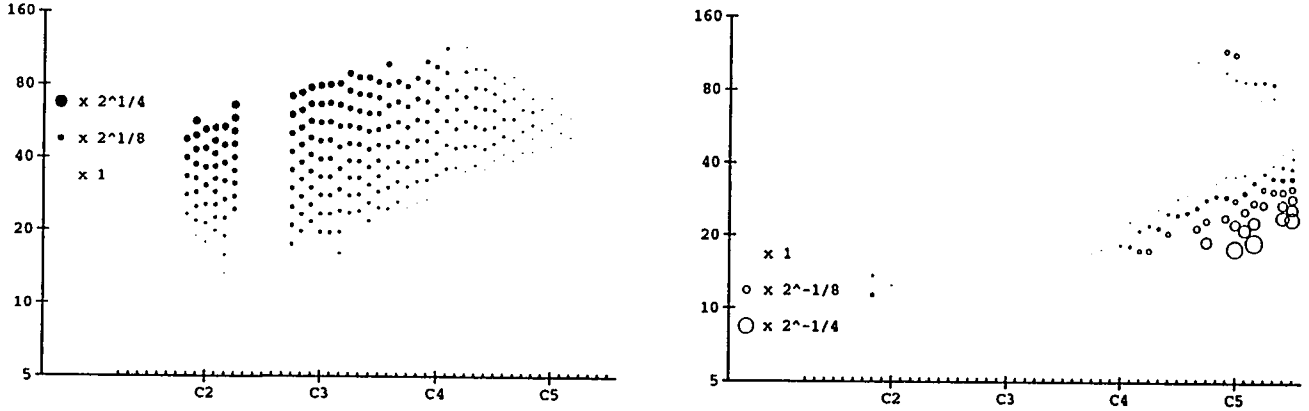

c. Bass Trombone

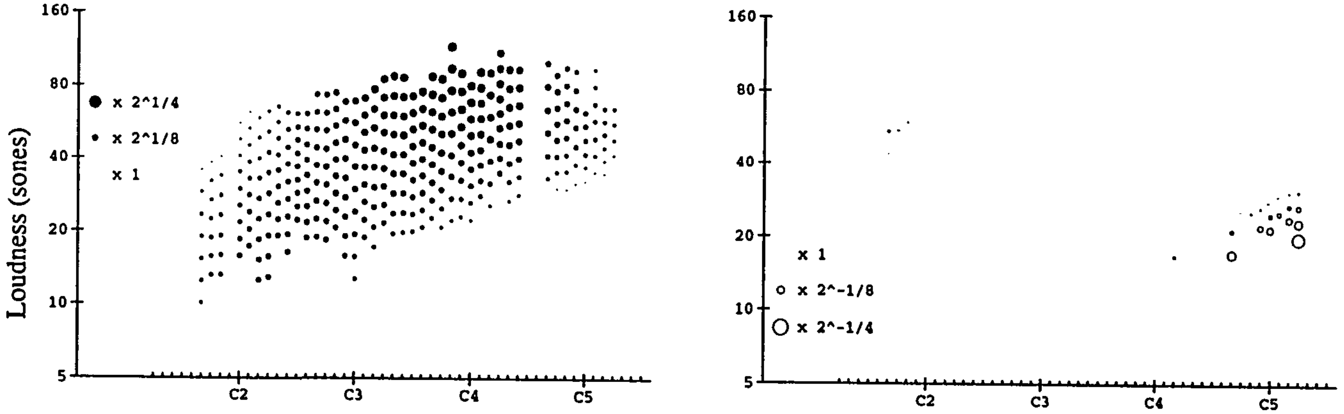
Figure 11.15: Brightness ratios of Performer A to B, by means of fitted surfaces. a. Alto trombone; b. Tenor trombone; c. Bass trombone.

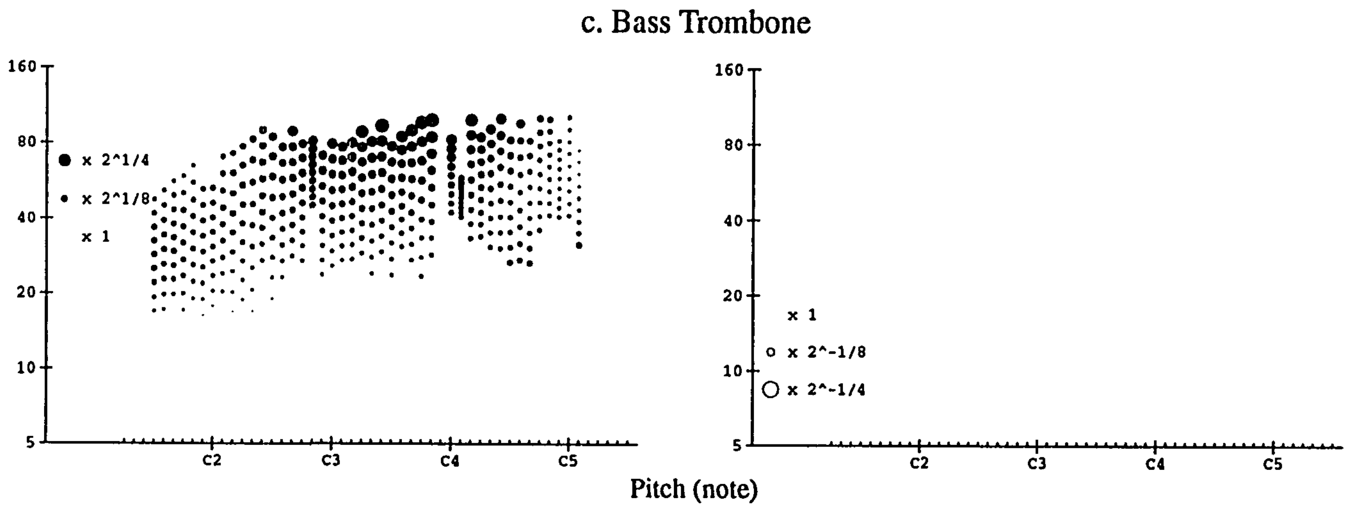
Performer A vs. Performer C

a. Alto Trombone



b. Tenor Trombone



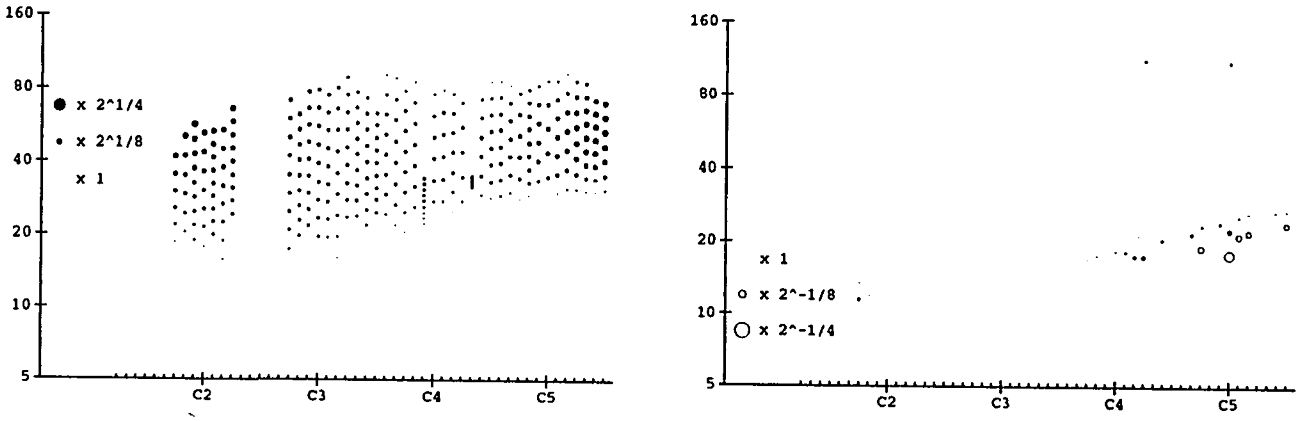


c. Bass Trombone

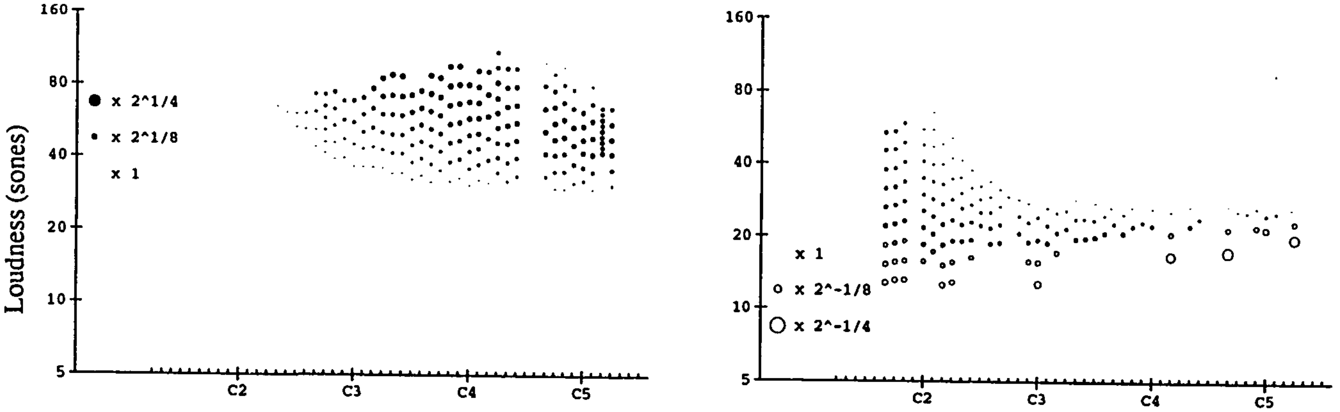
Figure 11.16: Brightness ratios of Performer A to C, by means of fitted surfaces. a. Alto trombone; b. Tenor trombone; c. Bass trombone.

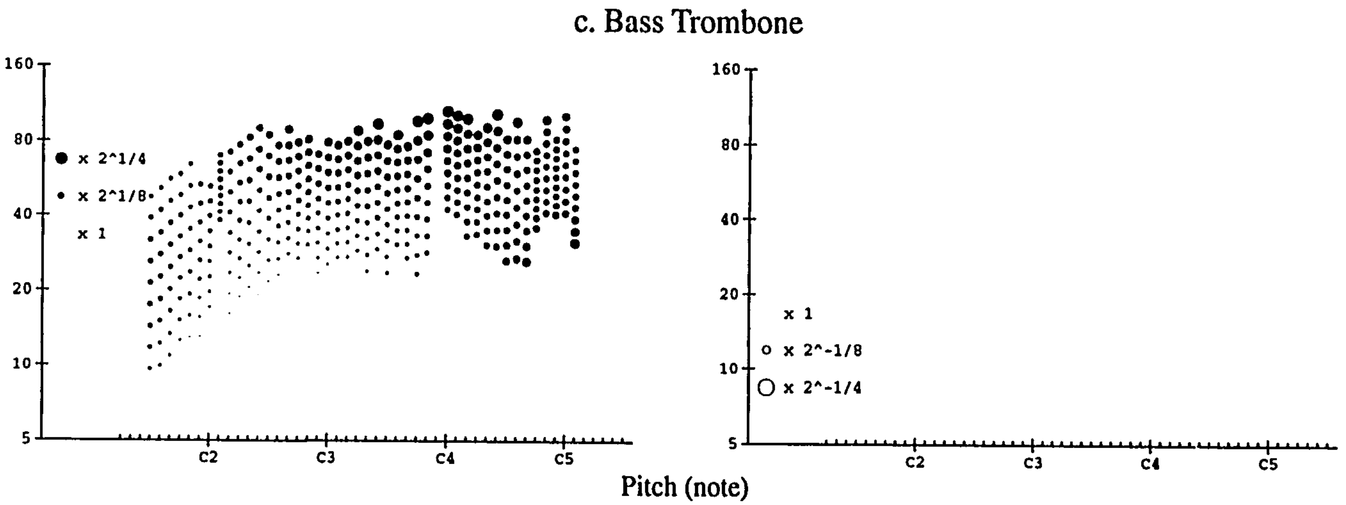
Performer B vs. Performer C

a. Alto Trombone



b. Tenor Trombone



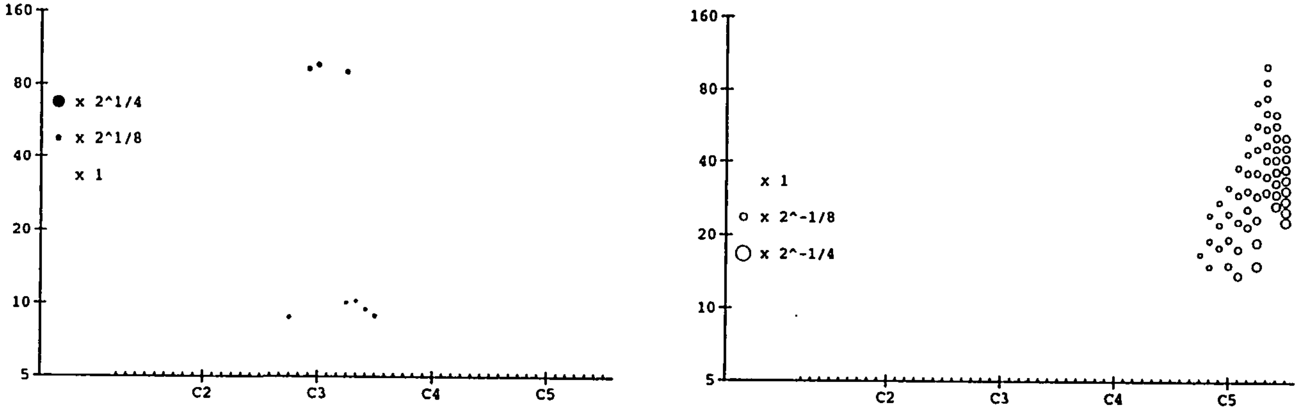


c. Bass Trombone

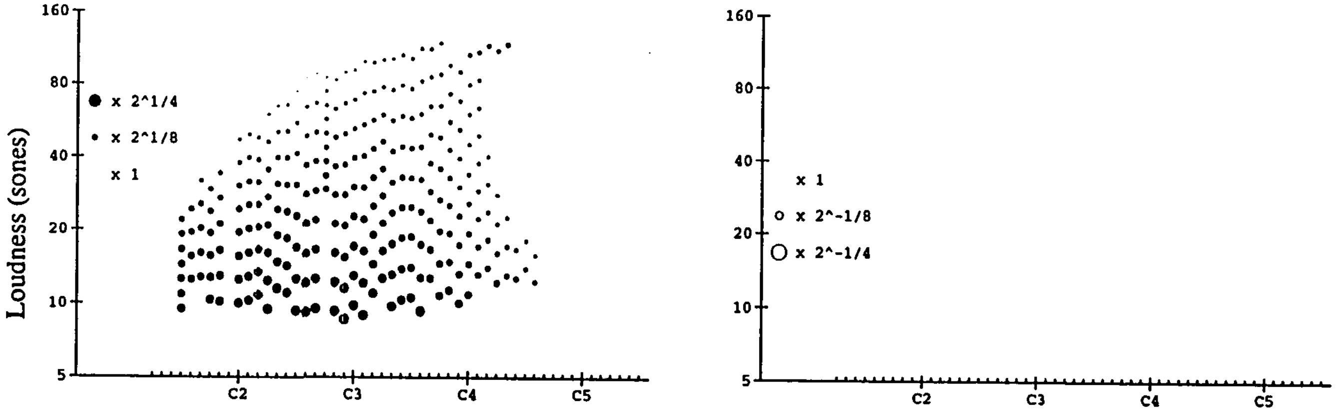
Figure 11.17: Brightness ratios of Performer B to C, by means of fitted surfaces. a. Alto trombone; b. Tenor trombone; c. Bass trombone.

Performer A vs. Performer B

a. Alto Trombone



b. Tenor Trombone



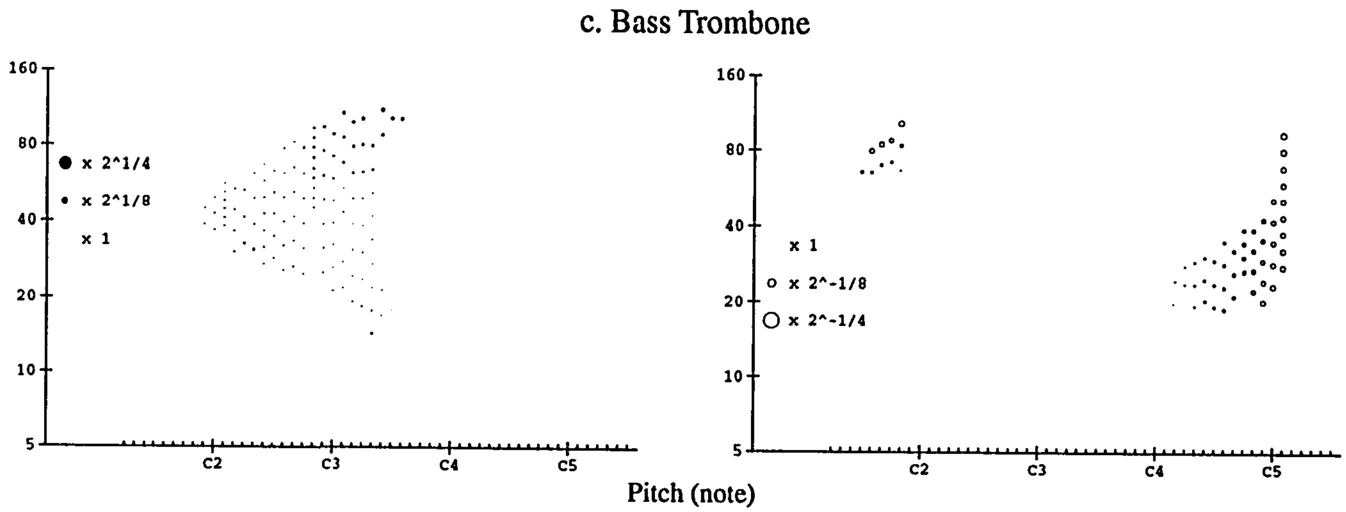
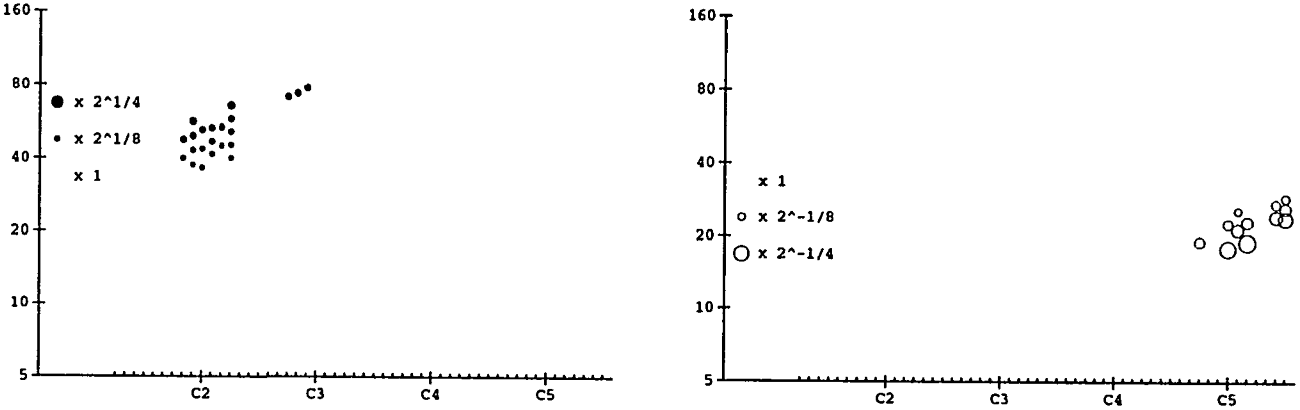


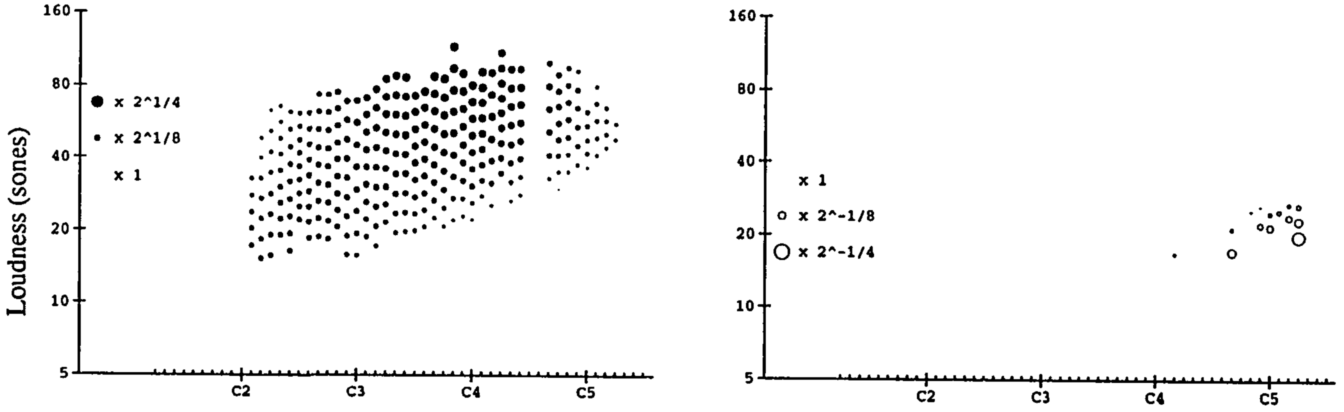
Figure 11.18: Brightness ratios of Performer A to B, by means of fitted surfaces. Only ratios exceeding cross-channel measurement variability are shown. a. Alto trombone; b. Tenor trombone; c. Bass trombone.

Performer A vs. Performer C

a. Alto Trombone



b. Tenor Trombone



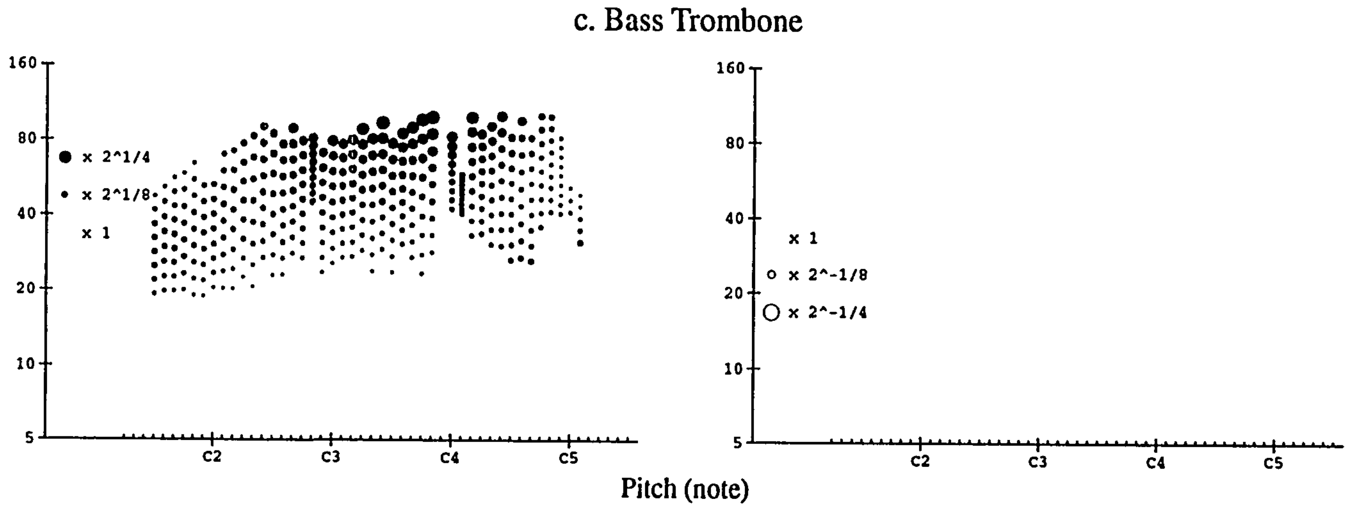
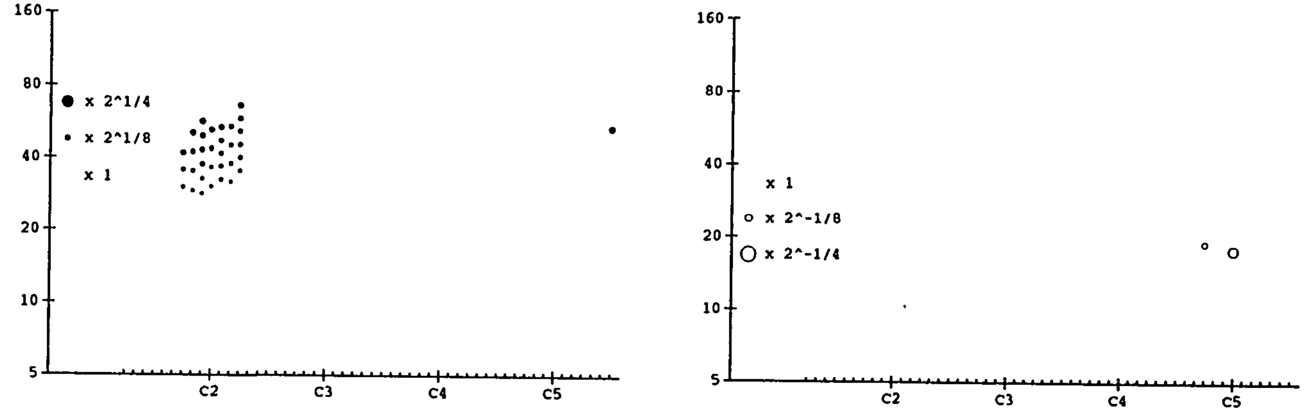

c. Bass Trombone

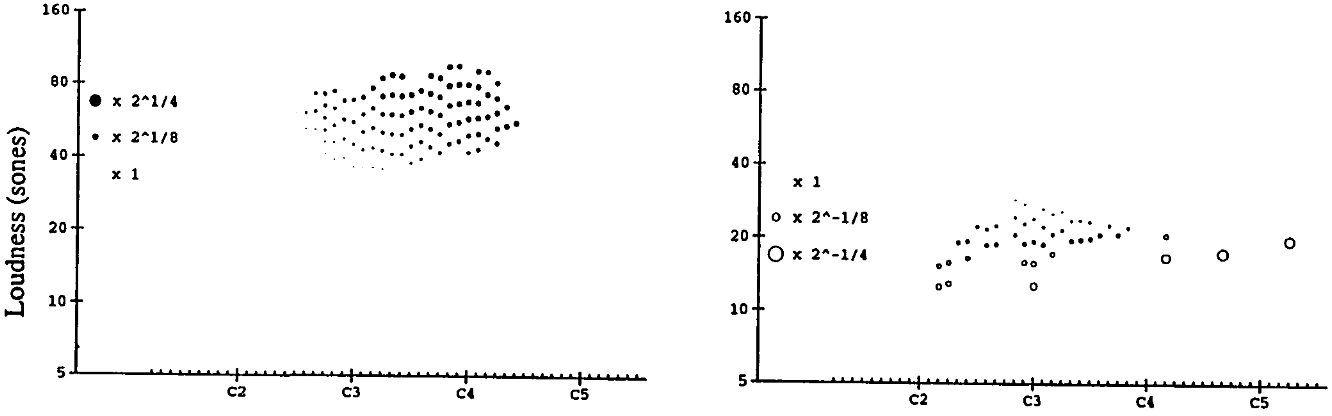
Figure 11.19: Brightness ratios of Performer A to C, by means of fitted surfaces. Only ratios exceeding cross-channel measurement variability are shown. a. Alto trombone; b. Tenor trombone; c. Bass trombone.

Performer B vs. Performer C

a. Alto Trombone



b. Tenor Trombone



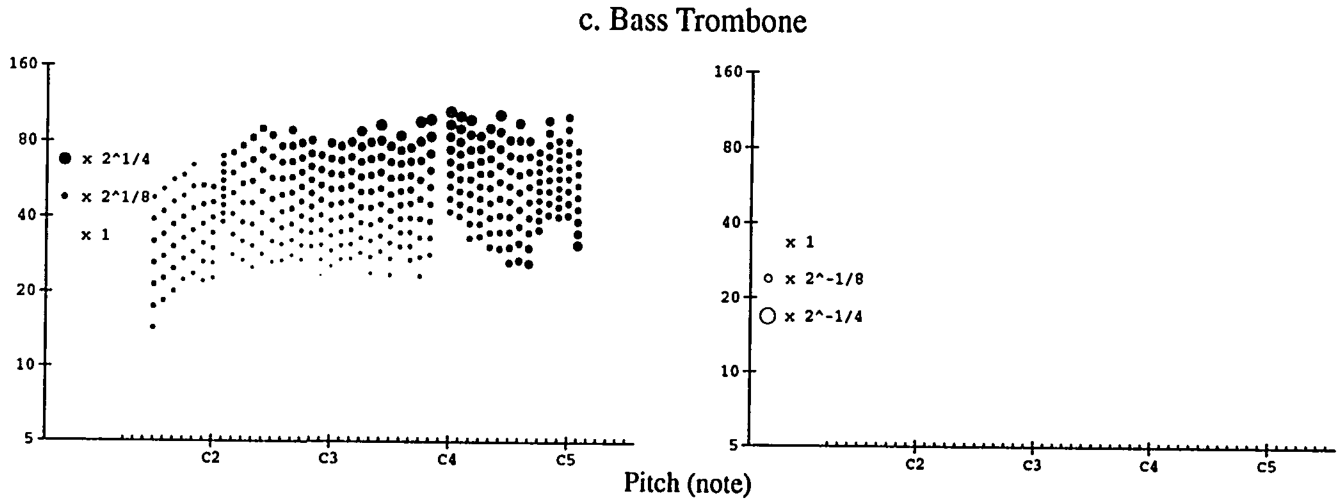


Figure 11.20: Brightness ratios of Performer B to C, by means of fitted surfaces. Only ratios exceeding cross-channel measurement variability are shown. a. Alto trombone; b. Tenor trombone; c. Bass trombone.

Surfaces as calculated by Equation 11.2 preserve general characteristics of trombone brightness as a function of loudness and pitch. Some details are lost; an example is the relationship between rate of change of brightness with loudness on instrument length discussed above. On the positive side, measurement variability is smoothed over. General behavior across performers can be examined using these surfaces.

11.6 Summary and Discussion

Figure 11.21 shows a compact form of the information presented in the previous sections. Brightness as a function of pitch has been calculated at three loudnesses for each instrument and performer combination, based on equations for the fitted surfaces (see Table 11.4.2).

As with any smoothing function, confidence is highest when there are many nearby points to constrain the fit, and lowest where there are few points and near extremes. Unfortunately, the players did not utilize the same ranges of loudness in their performances. As a result, confidence in the fitted functions for Performers A and B at, both 80 and 20 sones is very high, since they performed a large number of notes near that loudness. For Performer C, however, these are close to the minimum and maximum loudness for each note, and confidence is lower.

Figure 11.21a shows brightness as a function of pitch for alto trombones. Vertical lines indicate conventional playing ranges (Read 1969). It can be seen that in the lower and middle pitch ranges Performer A is brighter than Performers B or C. However, in the upper range Performers B and C become as bright or brighter than Performer A, especially at lower loudnesses.

Figure 11.21b shows the same for tenor trombones. Again, Performer A is brighter for most of the pitch and loudness range. Performer C is the least bright, except at 20 sones. Differences tend to be largest in the middle pitch range.

Figure 11.21c shows the same for bass trombones. In this plot, Performers A and B are virtually identical, while Performer C is markedly lower in brightness, especially at 80 sones. The low values for Performer C at 20 sones and above the note C3 are to be understood as an extrapolation, since there are no recorded tones for this Performer below about 24 sones in this pitch range.

Figure 11.22 similarly shows the brightness gradient for each combination. Differences between players are difficult to characterize. Performers A and B tend to be

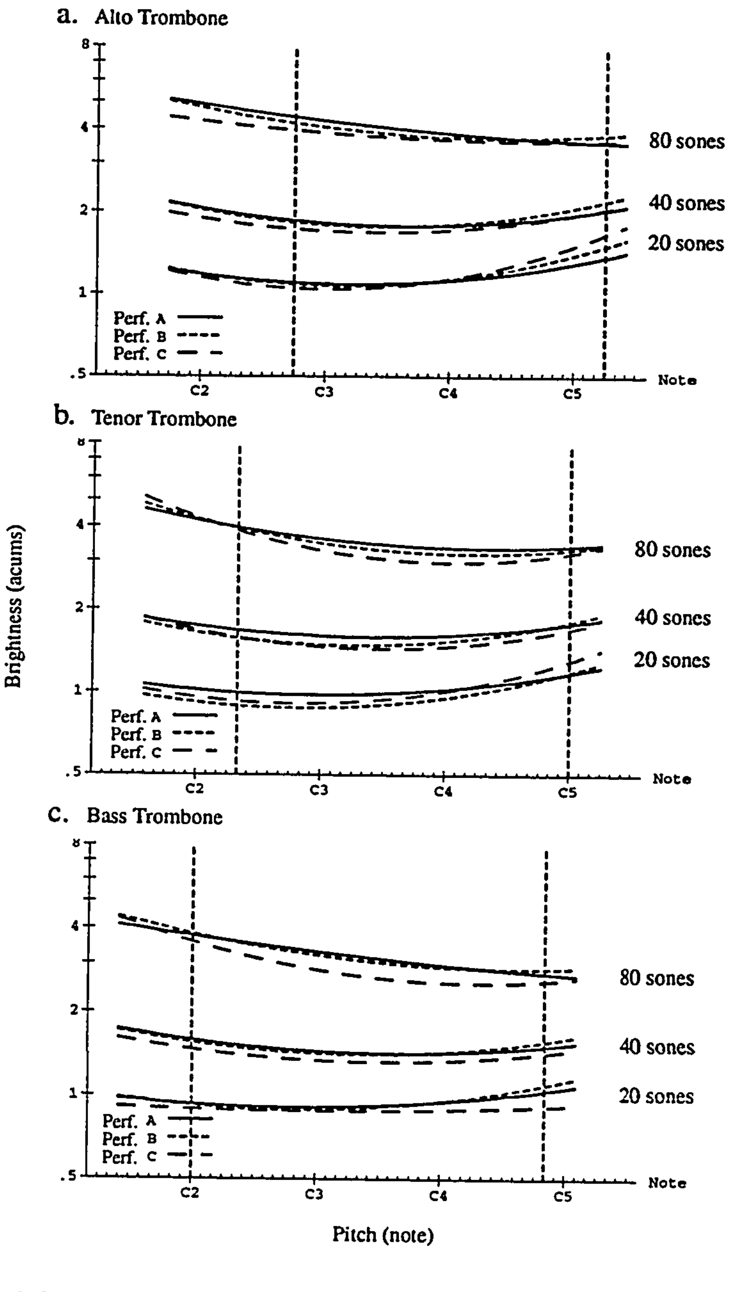


Figure 11.21: Brightness as a function of pitch, at loudnesses of 20, 40, and 80 sones. a. Alto trombones; b. Tenor trombones; c. Bass trombones.

similar across instruments and loudnesses, while Performer C diverges. For alto and tenor trombones, Performers A and B show a gradual decrease in brightness gradient with increasing pitch, along with a convergence, so that at the highest pitches slope varies less with loudness. Performer C, on the other hand, displays a wide range of slopes up to the highest pitches. On bass trombone, in contrast, Performer C shows a tendency for gradients to converge in the highest register. As noted before, the fit is not well constrained at low loudnesses for this performer.

Since differences between performers are much smaller than differences between instruments, a general characterization of each instrument's brightness behavior may be extracted by pooling like-instrument data, and then fitting with an surface of the form of Equation 11.2. Table 11.6 gives coefficients for these equations. Table 11.7 shows total differences and absolute differences between data and fit. Errors between pooled data and fit are on the order of those for individual performers (c.f. Table 11.4.2), suggesting that these equations represent the pooled data very well.

Figure 11.23a, shows brightness at 20, 40, and 80 sones taken from the equations given in Table 11.6. Pitch ranges have been restricted to the traditional limits mentioned above. The following general observations can be made:

- For all loudnesses, alto trombone is consistently brighter than tenor or bass trombones.
- Tenor and bass trombones are similar at low loudnesses, but diverge at higher loudnesses. The greatest differences are seen above the note B♭4 at 40 sones and louder.
- At each loudness, brightness as a function of pitch follows approximately parallel paths for the three instruments, when viewed over their traditional ranges.

Figure 11.23b, shows brightness gradients at 20, 40, and 80 sones taken from the equations in Table 11.6. Pitch ranges have again been restricted to traditional limits. The following general observations can be made:

- All instruments show higher gradients at high loudnesses, and lower gradients at low loudnesses.

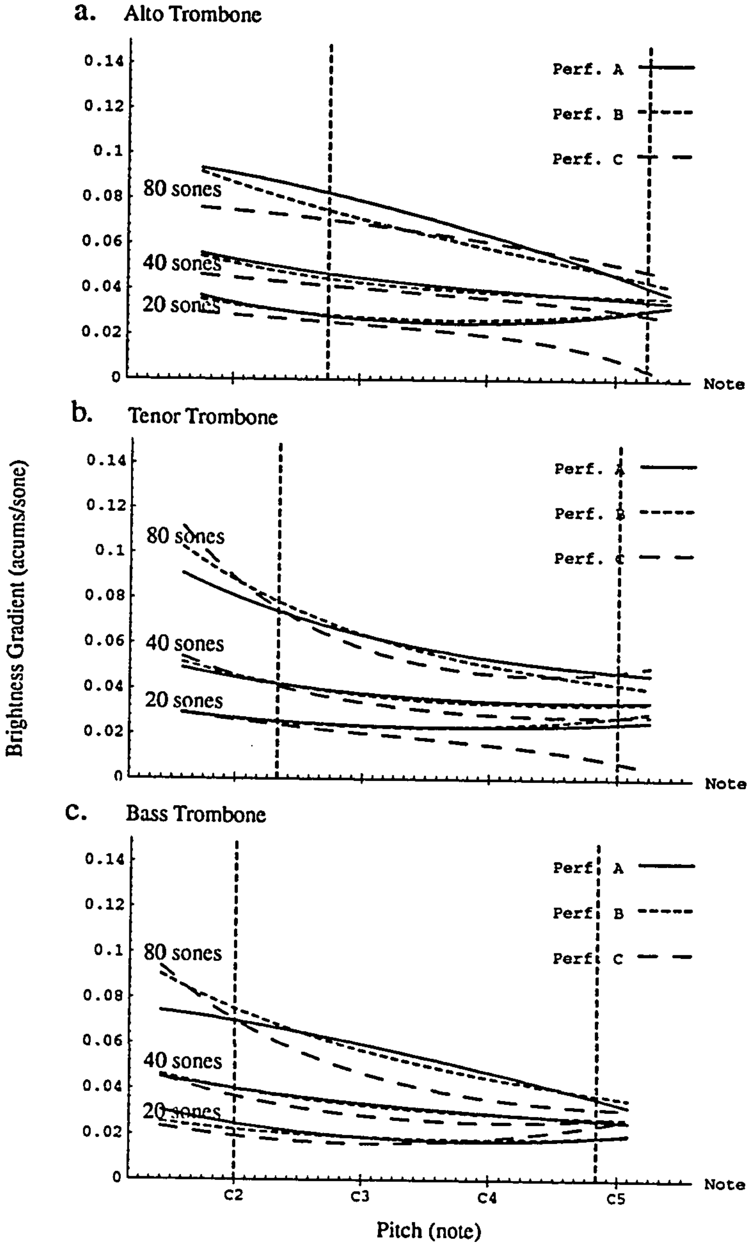


Figure 11.22: Brightness gradient as a function of pitch, at loudnesses of 20, 40, and 80 sones. a. Alto trombones; b. Tenor trombones; c. Bass trombones.

Table 11.6: Coefficients for surface fitted to data pooled across performers.

		x	x^2	y	xy	x^2y	y^2	xy^2	x^2y^2
Pooled Altos	9.58E-1	7.52E-2	-1.62E-3	-1.02E	-5.69E-2	1.28E-3	2.76E-1	8.02E-3	-1.97E-4
Pooled Tenors	1.93E	-2.65E-2	1.63E-4	-1.8E	9.52E-3	1.52E-4	4.01E-1	-2.32E-3	-2.38E-5
Pooled Basses	2.02E	4.73E-3	-3.45E-4	-1.88E	-5.27E-3	3.99E-4	4.06E-1	-2.58E-4	-6.08E-5

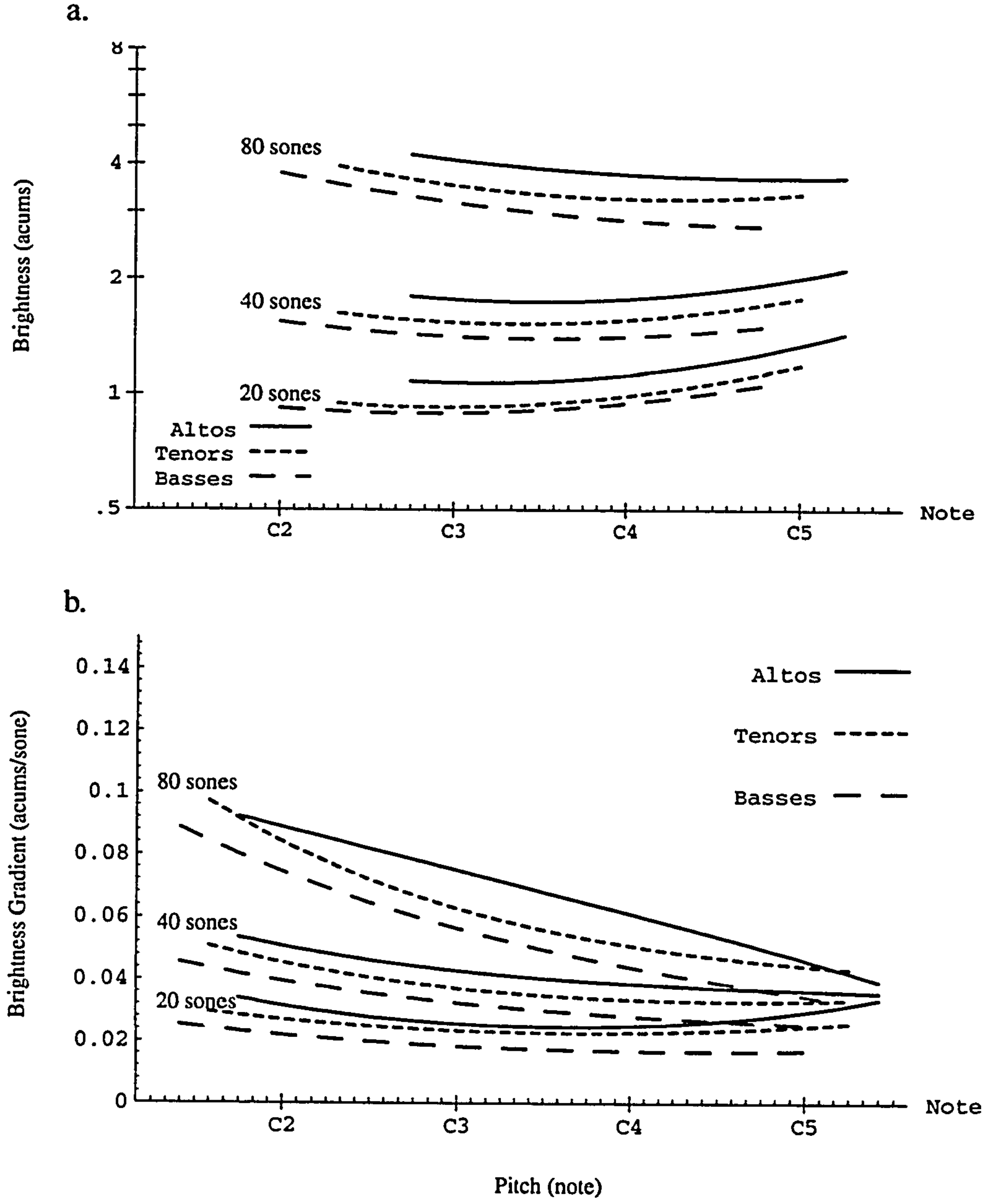


Figure 11.23: Brightness and brightness gradient at loudnesses of 20,40, and 80 sones, taken from surfaces describing data pooled across performers. a. Brightness as a function of pitch; b. Brightness gradient as a function of pitch.

Table 11.7: Two measures of error between measured log brightness and fitted surfaces, for data pooled across trombones: total error, and absolute error.

		Fitted Surface		
Measured Data		Pooled Altos	Pooled Tenors	Pooled Basses
Pooled Altos	Total Error	-1.76E-12	-1.41E-1	-2.37E-1
	Absolute Error	8.38E-2	1.55E-1	2.38E-1
Pooled Tenors	Total Error	1.34E-1	-5.68E-12	-9.45E-2
	Absolute Error	1.47E-1	8.54E-2	1.16E-1
Pooled Basses	Total Error	2.32E-1	9.21E-2	-4.31E-13
	Absolute Error	2.33E-1	1.21E-1	7.79E-2

- All instruments show convergence toward a single gradient in the upper register, and divergence in the lower registers.

- Brightness gradients are similar for tenor and bass at each loudness; alto trombone differs at high loudnesses.

11.6.1 Accuracy: General Applicability of Findings

Microphones were placed at distances and orientations representative of normal listening listening to live music; that is, about 6 meters from the sound source, off axis. In this particular room, the critical distance at which direct and diffuse sound fields produced by an omnidirectional source contribute equivalent sound pressure levels is about 3 meters. However, the trombone radiates high-frequency sounds in an increasingly directional manner (Meyer, 1978); for some frequencies, r_c was certainly more than 6 meters. Thus, these recordings were made under *semi-reverberant* conditions (Lubman, 1972). Compared to diffuse-field measurements, it is likely that high-frequency contributions were somewhat over-represented at higher loudnesses. This biases brightness measurements more than loudness. The comparisons described in this study are not likely to be affected by this bias, since radiation patterns are presumably similar for alto, tenor, and bass trombones. However, it could influence comparisons to measurements made with other microphone placements, or made in different rooms.

Predicted loudness, brightness and pitch for trombone tones have been reported by Goad (1994). Subject C participated in her study as well, recording pianissimo and fortissimo scales on alto, tenor, and bass trombones. Loudness and brightness predictions differ systematically between the studies. Since predictions differ in both loudness and brightness, it is difficult to directly compare measurements; however, the relationship between brightness, pitch and loudness can be described for my study with a fitted surface as given by Equation 11.2. At fortissimo levels, Goad's brightnesses are consistently above the surface. At pianissimo levels, they tend to be below the surface, but there are numerous exceptions. Also, Goad's loudnesses tend to be lower than what I measured.

A number of factors may have contributed to this. The first is that slightly different instruments and mouthpieces were employed. For bass trombone, the same mouthpiece was used in both studies, but the trombones were slightly different. An older Bach B50 with a customized lead-pipe was used in Goad's study; in the present study a newer Bach B50 3O was used. For alto and tenor trombones, the instruments were the same, but the mouthpieces were different. In Goad's study a Bach 14D was used on the alto, and a Bach 4G on tenor; in my study, a Minnick alto trombone mouthpiece and a Bach 5G tenor mouthpiece were used. The effect on brightness of these minor changes is not known.

A second contributing factor is microphone placement and orientation in the room. The same rehearsal hall was used in both studies, but performers were located in different places and oriented differently. Microphones were closer in Goad's study, but placed at a greater angle from the bell axis. It is possible that radiation patterns, or direct reflections from a nearby wall, influenced measurements.

A further possibility is an error in system calibration. Goad used a different calibration procedure than I did. Repeated reviews have not revealed procedural errors in either method. Since a known reference tone was not recorded in Goad's study, it is not possible to compare calibrations *post hoc.* I have experimented with varying the calibration to examine the effect on brightness as a function of loudness. As compared to my original calibration, a calibration value 3 dB lower leads to higher brightness at high loudnesses, similar brightness at low loudnesses, and lower loudnesses overall.

To conclude, the only published measurements of predicted loudness and brightness for tromone tones are those of Goad (1994). Systematic differences are observed between her measurements and mine. It is not possible at this time to positively identify the reasons for these differences; some explanations include differences in instruments, in performing instructions, in recording technique and in calibration. This does not invalidate the comparisons made in the present study; however, until a plausible explanation can be found specific values should be treated with caution.

11.6.2 Comparison to Other Instruments

Instruments other than trombones were not recorded in this study, and comprehensive measurements like those reported here have not been made for other instruments. However, Goad (1993) has reported simultaneous loudness, brightness and pitch predictions for selected orchestral instruments. Her measurements were performed on tones from a two-octave diatonic scale, at two loudnesses.

To compare trombone brightness to that of other instruments, it is useful to establish a representative expression for brightness as a function of loudness and pitch. This has been done by pooling like-instrument data across subjects, and fitting a surface as described by Equation 11.2. The shaded surfaces shown in Figures 11.24 and 11.25 were obtained in this way, using data from tenor trombones.

Values for brightness as a function of loudness and pitch for French horn and tenor saxophone have been taken from Goad (1994) and compared to tenor trombone. Figure 11.24 compares French horn brightness to tenor trombone (horn is considered to be a tenor instrument). It can be seen that predicted brightnesses are greater for the horn than the trombone at fortissimo, and less at pianissimo. However, these observations must be treated with caution, since they are reminiscent of the differences between Goad's and Bulen's measurements of the same trombonist. With this in mind, the most striking difference between horn and trombone tones is the reduced range of loudnesses utilized by this hornist, especially in the lower register. So, while horn brightness is similar to that of the tenor trombone at the same loudness and pitch, it doesn't appear to have the same loudness or brightness range.

11.25 compares tenor saxophone to tenor trombone. It can be seen that the saxophone is uniformly brighter than the trombone.

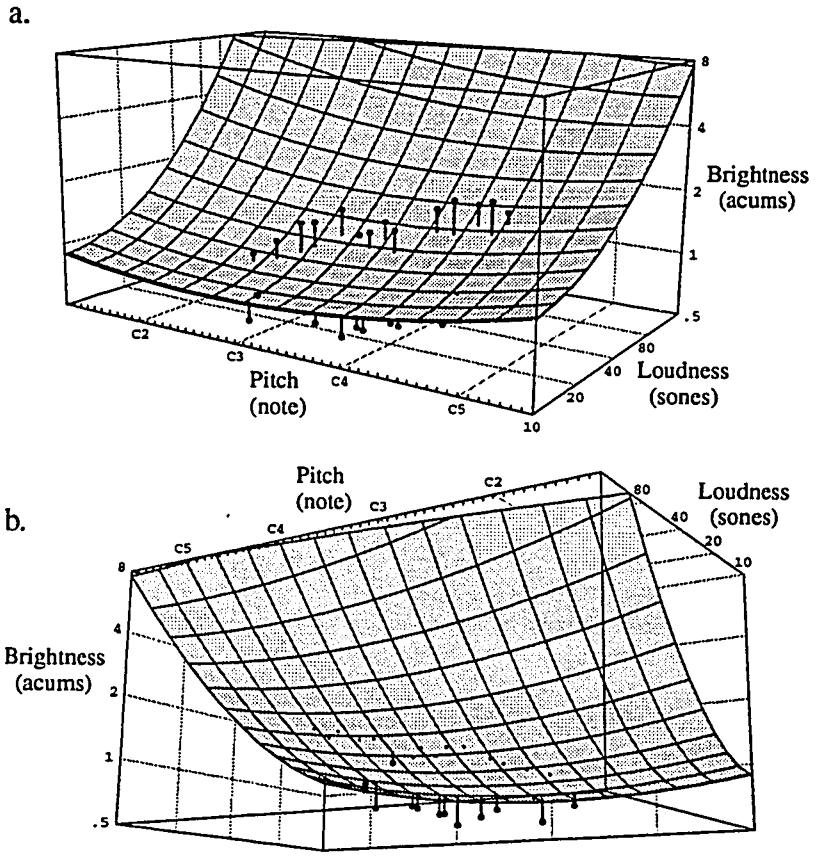


Figure 11.24: Brightness of French Horn tones (dots) compared to those of tenor trombone (shaded surface). a. Front view; b. Back view.

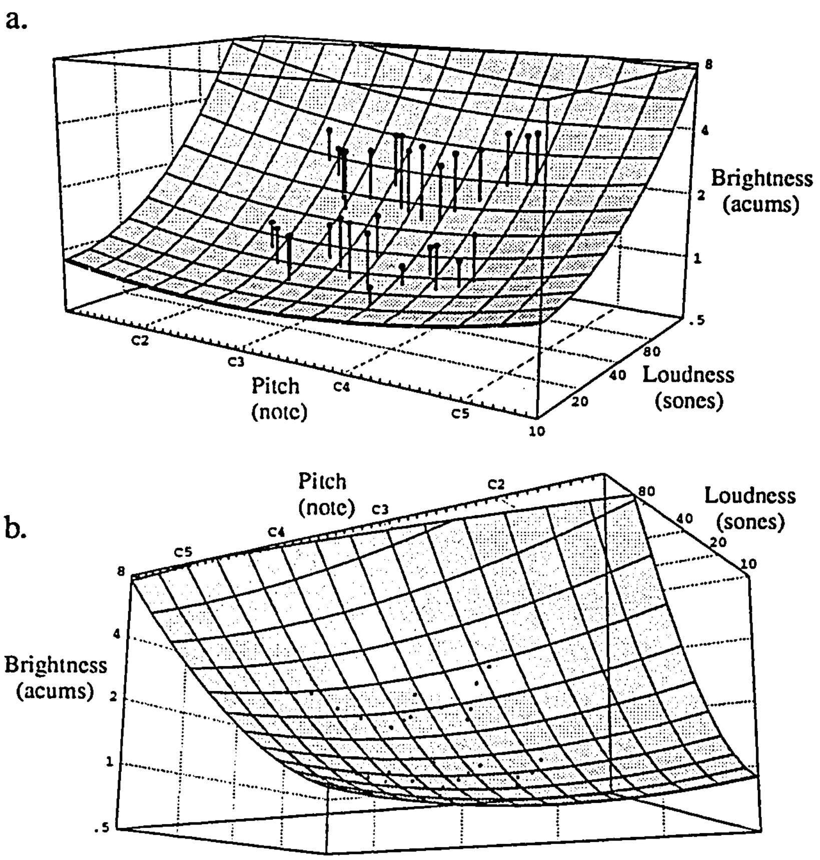


Figure 11.25: Brightness of tenor saxophone tones (dots) compared to those of tenor trombone (shaded surface). a. Front view; b. Back view.

It appears that brightness as a function of loudness and pitch is useful for discriminating trombones from other instruments.

Chapter 12

SUMMARY, CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

To summarize, at the outset it was suggested that an objective description of trombone timbre would be useful to trombonists. Brightness is a prominent dimension of timbre perception in general and is perceived by trombonists to be an important component of trombone timbre. For trombone tones, brightness is both salient and distinctive.

Since brightness is influenced by both loudness and pitch, a three-dimensional representation of trombone tones (brightness vs. loudness vs. pitch) is indicated. Brightness, loudness and pitch are subjective attributes which cannot be directly measured. However, objective measures can be used to *predict* brightness in acums, loudness in sones, and pitch in musical note names.

Predicted brightness, loudness, and pitch have been calculated for tones produced by three performers each playing alto, tenor and bass trombones. Findings are described below. These will be related to the research objectives of differentiating between tones from alto, tenor and bass trombones, and between tones from different performers using the same instrument.

12.1 Findings

- Three-dimensional plots of predicted brightness vs. loudness vs. pitch display a consistent, distinctive shape for all performers and all instruments. Brightness increases monotonically with loudness. Variation of brightness with pitch tends to be dependent upon loudness. For example, at low loudnesses brightness is fairly constant across the pitch range; at high loudnesses, on the other hand, one finds increased brightnesses at both very low and very high pitches.

- Log brightness can be closely approximated with a quadratic function of log loudness and log frequency. The log brightness surface increases rapidly with increasing loudness, is fairly constant across frequency, and is slightly concave upwards in both loudness and brightness dimensions.

- While displaying similar contours, clear brightness differences exist between alto, tenor and bass trombones. Comparisons can be made on the basis of average difference in log brightness, or of geometric mean ratios of tones at the same pitch and (interpolated) loudness. On average, alto trombone is brighter than tenor by about 15 percent. Tenor is brighter than bass by about 10 percent.

- Average differences in log brightness sometimes distinguish between performers when using the same instrument. On alto trombone, the average difference between subjects A and B is smaller than the 95% confidence interval for the mean; however, subject C is on average about 4% less bright than either subject A or subject B. On tenor trombones, A is consistently brightest, with B about 5% less bright and and C about 8% less bright. On bass trombone, A and B are again very similar, but C is less bright than A by about 9%, and less bright than B by about 7%.

- When comparing one subject to another on the same instrument, subtle but consistent patterns are seen regardless of instrument. For example, a similar pattern characterizes brightness differences between subjects A and B, whether compared on alto, tenor, or bass trombones.

- The rate of change of brightness with loudness is related to instrument length, that is, to slide position. The rate is greater for longer positions. It is also related to register, and is greater for low notes than high notes.

- In the rehearsal room where measurements were conducted, measurement variability between two microphone positions was about 7 percent for brightness and 15 percent for loudness. Increased variability in loudness measurements is found at low loudnesses and high pitches. Frequency measurement errors were negligible.

These results indicate that brightness as a function of loudness and pitch holds promise as a representation of one important aspect of trombone timbre.

12.2 Relationship of Findings to Research Objectives

12.2.1 Differences between Alto, Tenor, and Bass Trombones

Alto, tenor and bass trombones (same player) can be discriminated on the basis of brightness as a function of loudness and pitch. Differences are sizable and consistent, and can be described in terms of average difference in log brightness. This is shown in Figure 11.8 and Tables 11.1a, b, and c. For each player, alto is brighter than tenor, which is brighter than bass. Across players the same pattern applies, but the degree may vary somewhat. For example, tenor trombone as played by Performer C is only slightly brighter than bass trombone as played by Performer A.

12.2.2 Differences between Performers

Differences between performers are smaller than differences between alto, tenor and bass trombones. They are sometimes reflected in average differences in log brightness. As seen in Table 11.1a, b, and c, for all but one comparison the means differ by more than the 95% confidence interval for the mean. The exception was Performer A to B on alto trombone. In no instances, however, did means differ by more than one standard deviation. More generally, however, it is the *pattern* of differences across pitches and loudnesses that is distinctive. That is, a similar pattern of brightness differences is observed between Subject A and B, whether playing alto, tenor, or bass trombones.

12.3 Implications for Trombone Performance and Pedagogy

The relationships documented here between brightness of alto, tenor and bass trombones are congruent with musical descriptions and instrument usage. The alto trombone is consistently brighter than tenor or bass, even at low loudnesses. Its use in orchestras largely reflects this: one uses it to ease the effort of playing in the high register, perhaps because a given brightness can be achieved at a lower dynamic. As Sauer has observed, the alto "sounds bright and alive at ... reduced dynamics" where the tenor trombone sounds "dull and lacks impact and intensity" (1984, p. 41). The bass trombone, on the other hand, behaves somewhat differently in spite of the fact that, until the second valve is depressed, it is the same length as the tenor.

The bass trombone "serves as a bridge between the tenor trombone and the tuba" (Mays, 1980, p. 55). It is less bright than the tenor over its entire range, although considerably brighter than the tuba (Goad, 1994). However, it makes extensive use of long-position notes and notes that require the valve(s) to be engaged. The higher brightness gradient under these circumstances may permit the bass trombonist to approach, within a reasonable dynamic range, both the very low brightnesses of the tuba and the higher brightnesses of the tenor trombone.

A number of suggestions for trombone pedagogy present themselves; however, a word of caution is in order. Like other objective tools that a musician uses, predicted brightness is appropriate for directing and developing listening and thinking skills, rather than for indicating what is right and wrong. A metronome, for example, does not confer a musical sense of time; instead, it presents a convenient reference from which to begin to understand it. An electronic tuner is productive when used to study tendencies in comparison to a reference, rather than for simply determining what is in-tune and what isn't. Likewise, brightness is not in itself musical or unmusical; measurements of predicted brightness are best used to understand how it can be used in a musical context.

Since embarking on this study, I listen more carefully for brightness and how it is manipulated in performance. The key finding of this study is that trombone brightness is primarily determined by loudness; influences of the instrument type are secondary, and differences between performers are considerably smaller in magnitude. This heirarchy can be exploited as a pedagogical tool. If a student's brightness is unsatisfactory in a given situation, one can first focus on loudness. Following that, contributions of the instrument and performer can be considered.

A recent experience suggests one way in which this can be applied. A student was having difficulty achieving an acceptable tone quality in the low register of the tenor trombone. The piece that he was studying—the Hindemith *Sonata*—uses melodic phrases covering a wide range, including some seventh-position notes, all at forte levels. As played by this student, the lower notes "stuck out" and disturbed the melodic line; the problem was perceived to be one of timbre.

As a first approach, we concentrated on the loudnesses of the notes in each phrase. We found that, when played slowly with emphasis on breath support, loudnesses were fairly consistent and timbre difficulties were lessened. At tempo, however, the problems re-emerged. This guided our listening to the attacks. We found that the attack he was using throughout was quite hard, and that the first part of the tone was louder than the sustained portion. While this was audible in the middle and upper register, it was only disturbing in the lower register. We directed our attention to providing a steady, well-supported air column, intended to produce consistent loudness. As a result, the timbral problems were greatly lessened. In this case, loudness provided a means for understanding and addressing timbral problems.

In another recent experience, timbre provided a means for addressing loudness problems. A student was working on the second movement of the Halsey-Stevens *Sonata*, a movement which is generally subdued, with subtle *crescendi* and *decrescendi.* The student was having problems establishing a *piano* dynamic that permitted dynamic changes without distorting the melodic line. By concentrating on tone quality rather than loudness, he was able to consistently produce appropriate dynamics.

The production of a consistent brightness in melodic passages is often important for trombonists. This was brought out in a recent performance of the Nehleybel *Concerto for Bass Trombone and Wind Ensemble* featuring Douglas Yeo as bass trombone soloist. In tutti passages, the trombone tones were partially masked by the ensemble; what projected through was the bright "edge" of the tone. It was remarkable to listen for the consistency of this aspect as Yeo traversed from the pedal register through the valve register.

12.4 Brightness and Sound Concept

As discussed previously, trombonists value timbre as an individual characteristic. It is traditionally approached by the development of a personal sound concept (Kleinhammer, 1963) which is then implemented by the performer.

If brightness is an important component of sound concept, then the measurements reported here, along with the physics of sound production as discussed in Chapter 7, can give some indications of how sound concept might be realized. Besides loudness, the performer can manipulate embouchure resonant frequency ("lipping" the intonation of a note) and vocal tract resonances (vowels). While these parameters were not part of this study, it is interesting to examine their influences on brightness and conjecture how they may have been used.

Benade (1976) has suggested that aligning vocal tract resonances with components of a tone help stabilize the regime of oscillation, producing a tone quality that is both recognizable and disirable. By selectively enhancing upper harmonics, this also increases brightness. Of the three performers in this study, the one that is most skilled at manipulating vocal tract resonances is Performer A; performer B is a close second, and Performer C is less adept. It is suggestive that Performer A was consistently brighter than C, and generally brighter than Performer B.

Performer C consistently had the lowest brightness levels. It is interesting that he usually performs with his tuning slide perhaps a half-inch shorter than either A or B. All things being equal, this would suggest a fundamental in first position about 18 cents sharper than A or B. Since he generally manages to play in tune, he may be lipping many tones down. May (1980) suggested that lipping a note down diminishes upper harmonics, which would lower brightness.

12.5 Implications for Future Research

12.5.1 Predicted vs. Perceived Brightness, Loudness and Pitch

The models used in this study to predict loudness, brightness and pitch are scaled to reflect perceived doublings in stimulus attribute with doublings in predicted value. It reasonable to assume that a predicted doubling is perceptible; however, none of the differences reported here have been that large. Are differences in predicted attribute of just a few percent actually perceptible? And if so, under what conditions?

A difference limen for brightness has not been established, and its determination is an important goal for future research. It will not be an easy task. Risset and Matthews (1982) have noted that psychometric scalings may be inherently inaccurate, since judgements of "twice as bright" or "half as bright" are difficult. Many subjects and many trials are required, since standard deviations and inter-subject variability are high.

A difference limen for trombone brightness may be elusive for another reason. Trombone brightness and loudness are intimately connected, and are likely to be confounded in experimental situations. This may be the explanation for Clark and Milner's (1964) finding that the performance loudness of a trombone tone can be accurately identified from a recording, even when playback levels are manipulated. In fact, Clark and Milner's subjects could do this better for trombone tones than any other instrument. I imagine that deciding whether two similar tones differ in loudness or in brightness would be a difficult task. To summarize, predicted brightness as a function of loudness and pitch can be used to discriminate between trombones and trombonists; however, it remains to be shown how closely predicted differences correspond to perceived differences.

12.5.2 Methodological Enhancements

Loudness and brightness prediction is based upon accurate spectral measurements. As mentioned previously, it is difficult to make accurate spectral measurements of musical instruments tones. Instruments radiate directionally, making near-field measurements difficult. Far-field measurements are complicated by room resonances. Benade and Larson (1985) showed that consistent results can be obtained by varying both source and microphone positions and performing spectral averaging. However, this is impractical for the present study, since many trombone tones are of necessity quite short.

Benade and Larson suggest that, for brass instruments, a useful approximation to the room-averaged diffuse field spectrum can be obtained with a microphone located one radius from the bell and on-axis. However, loudness and brightness measures are referenced to absolute sound pressures, and are perceptually validated only for levels tolerated by human subjects. Sound pressure levels at the bell of a trombone are far in excess of this.

An adaptation of brightness and loudness measures might be made to permit near-field recording of trombone tones. If Benade and Larson's assertion about brass spectra at the bell is correct, the adaptation will be a frequency-independent attenuation. Experimental determination and validation of this requires accurate room-averaged trombone spectra over a range of performance conditions. Once accomplished, it would simplify measurements by greatly reducing measurement variability. Given the regularity of the brightness vs. loudness and pitch contours measured so far, it is likely that they could be specified with fewer, but more accurate, measurements.

12.5.3 Trombone Performance Investigations

A faster method for data acquisiation would facilitate a larger data base of trombones and trombonists. While the performers in this study have broad performing experience, they form a relatively homogeneous sampling of trombonists. All perform primarily in the Western art-music tradition; all use a symphonic tenor-bass trombone as their primary instrument, although of different makes. Performers B and C have studied trombone with Performer A. A broader sample is needed to be able to discuss trombone timbre in general. Another area of great interest is in the comparison of similar instruments. For example, subject A normally performs on a model of tenor-bass trombone favored by many soloists. Subject C normally plays a tenor-bass trombone that is popular among symphonic players. The instruments are virtually identical in dimensions, but subject A's instrument is currently considered "lighter" or "brighter." Are these kinds of differences reflected in brightness vs. loudness and pitch, and can they be related to performing style?

The ways in which loudness and brightness are manipulated in performance should also be investigated. One possibility is suggested by Goad's measurements of a two-octave major scale, performed at fortissimo on alto and tenor trombones (1994, p. 52). The general trend of these plots is a slow increase in brightness with increasing pitch; however, the trend is broken twice in each plot, at the same scale degrees. In all four cases, the note that is brighter than suggested by the general curve is located a half-step below its neighbor. Do performers use brightness to inflect leading tones, in the same way that pitch is sometimes used (Ward, 1970)?

Formal relationships can be investigated as well as harmonic. Do performers alter loudness and brightness in response to structural function? In a homophonic texture, is a melodic line treated differently from an accompanying line in terms of loudness and brightness? In a polyphonic texture, are loudness and brightness treated similarly in all parts? Finally, how are loudness and brightness manipulated to balance the various pitches in chords?

Another interesting topic concerns differences between tones produced in short positions and long positions. At high loudnesses, notes produced in long positions are brighter than nearby pitches produced in short positions. Extreme examples are the notes B1 on bass trombone, a second-partial note for which both valve sections are engaged to extend the instrument to nearly twice its normal length, and B♭1, which is a first-position fundamental note. One would expect that B1 would be regarded as unusual or unrepresentative. The opposite seems to be true, however. For example, Nelson Riddle asserts that "pedal notes [such as B♭1] . . . are used mainly for effect and contrast" (1985, p. 63). He adds that the bass trombone "provides the firm low register not attainable through the use of pedal tones." (op. cit.), presumably by supplying long-position second harmonic notes. An interesting experiment would be to measure brightness, loudness and pitch as a trombonist played a melodic section that alternated B1 and B♭1, to determine what kinds of adjustments (if any) are made as compared to the same tones in isolated contexts.

Timbral blend could also be examined in terms of brightness. Sandell (1991) investigated the blend of re-synthesized musical instrument tones, and related findings to measured spectral centroids. Blend was described as ranging from segregation into distinct timbral entities to complete fusion into a single composite timbre. He found that

> Ratings of blend showed primary effects for centroid (the location of the midpoint of the spectral energy distribution) and duration of the onset for the tones. Lower average values of both centroid and onset duration for a pair of tones led to increased blends. (p. iii)

As previously noted, brightness and spectral centroid are both attempts to quantify the relative distribution of spectral energy. However, their results are not easily comparable, and conclusions concerning spectral centroid cannot be directly applied to brightness measurements.

We can calculate the brightness of notes in a chord. For tenor trombone, a B♭2 - F3 - B♭3 chord at 80 sones would have brightnesses of 3.69, 3.48, and 3.39 acums, calculated from a fit to Subject A's data. If the same chord was performed using bass, tenor, and alto trombones (all by Subject A), brightnesses would be 3.36, 3.48, and 3.91. These figures are plotted in Figure 12.1. In the chord with tenor trombones only, brightness decreases with increasing pitch. When using alto, tenor, and bass trombones, on the other hand, brightness increases. This ascending pattern also appears in a a fortissimo chord on violins, taken from Goad (1994). Of course, in actual performance musicians may manipulate brightness by varying loudness or other means. Furthermore, different kinds of blend are musically acceptable; examples are found in choral and solo singing. In trombone performance, the French have long favored homogeneous groupings of trombones, while Americans and Germans favor heterogeneous groupings (Guion, 1988).

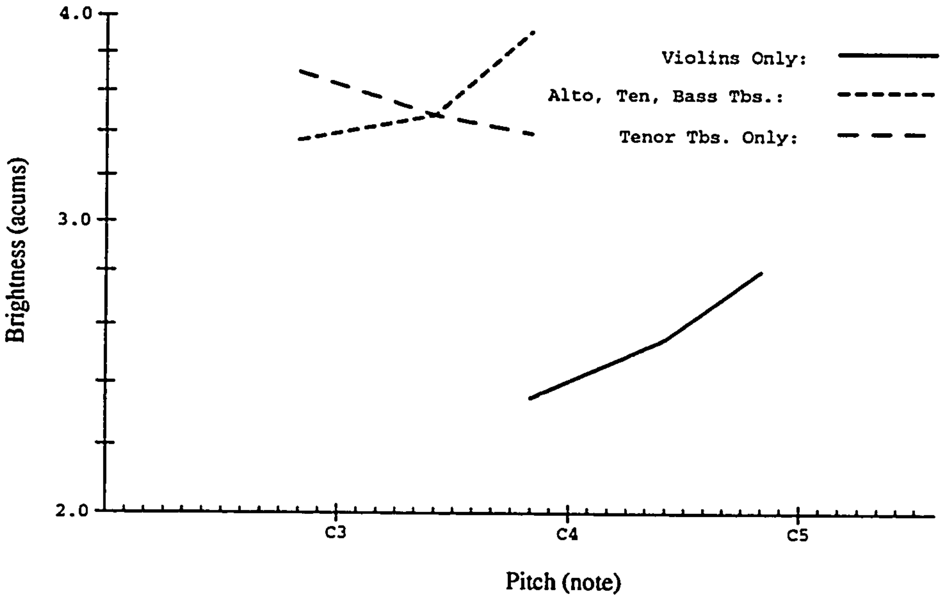


Figure 12.1: Brightness of notes in an 80-sone B♭ Major chord, as performed with three tenor trombones, or with alto, tenor, and bass trombones. For comparison, a fortissimo chord on violins is also shown.

12.6 Inferences for Music Perception

The perception of musical instrument timbre has presented a number of difficult problems. For example, each instrument of the orchestra is valued in part for its distinctive timbre, and not just for register and power. Yet it is easy to establish conditions under which their tones cannot be reliably distinguished (e.g. Luce and Clarke, 1967). Similarly, trombonists put considerable effort into refining a personal sound quality, and spare no expense in acquiring instruments and accessories to help them make that sound. But Pratt and Bowsher (1978) found that under experimental conditions it was difficult for listeners or players to distinguish the sounds of instruments or players.

Certainly part of the difficulty has been caused by barren, unmusical contexts. An anecdote can illustrate this. I recently spent several weeks working with a new mouthpiece in an effort to improve a particular aspect of my own "sound". However, playing alone—even in a familiar practice room—I could not so much as describe the differences between old and new. It was suggested that I try playing with piano accompaniment. Suddenly the differences were obvious, and I could quickly judge the relative merits of each mouthpiece.

As the present measurements document, even a room specifically designed for music performance gives rise to spectra that vary noticeably between listening locations. Fortunately, it appears that humans are able to compensate for the vagaries of transmission paths. In an elegant set of experiments, Watkins and co-workers (Watkins, 1991; Watkins and Makin, 1994) have identified a *negative auditory after image*, which functions as a variable filter to compensate for distortion. One experiment presented listeners with a synthesized voice that intoned "The next vowel is," followed by a burst of white noise. The voice, however, had been modified with a special filter. Subjects consistently heard the subsequent noise as a particular vowel, which resulted from applying to the noise the inverse of the filter that had been applied to the voice. In other words, in the course of listening to the short phrase subjects generated a corrective filter and applied it to subsequent sounds. For my mouthpiece evaluation, the sound of the piano seemed to cause the same effect: suddenly I had a yardstick by which to judge my sound.

I suspect, however, that this is not the only mechanism that improves listeners' perceptual abilities in musical situations. It seems likely that listeners internalize the principled ways in which related sounds change, and from that information learn how to connect them. Balzano's point that "Perceiving timbre is more a matter of perceiving underlying dynamics of physical processes than perceiving the places of things in abstract ordered structures" (1986, p. 312) means exactly this: sounds are unified by a cognitive picture of how they are produced. This is an example of synthetic listening for timbre, as described in the Introduction.

Trombone and trombonist are elements of a dynamic system which produces sound. Like most dynamic systems, its characteristics are not apparent from its behavior when operating in a steady state. By observing the motion of a driven spring, for example, one cannot deduce its stiffness or mass or resonant frequencies. However, if driving force is changed in a known direction (faster frequency or greater amplitude, for example) then the response of the spring reveals its characteristics. The start-up transient contains a great deal of information about both the system and its relationship to driving forces (Fletcher and Rossing, 1991). One cannot develop a model of a system without information about either its start-up transients or its response to well-defined changes in driving force.

This analogy helps explain observed peculiarities in tone identification. It suggest reasons for the difficulty of identifying a single steady-state tone, and provides plausible reasons for identification improvement when a transient is included. It also explains why a number of distinct steady-states may be as useful for identifying the source as a transient, as was found by Kendall (1986). Different pitches, loudnesses, vibrato, and note transitions are all important sources of information.

How do plots of predicted brightness vs. loudness vs. pitch relate to this? I would argue that they encapsulate the most important perceptual results of brass tone production, and thus are a particularly worthwhile representation. Most components of brass tone production influence spectral development as a function of frequency and intensity. The instrument acts through cutoff frequency, loss mechanisms, and details of the impedance curve; the performer through playing frequency, vocal tract resonances, and impedance of the embouchure (Elliot and Bowsher's high pitch and low pitch regimes). Thus, spectral changes over a range of pitches and loudnesses provide information about the system and how it is being driven. Expressing them in perceptual quantities allows an estimate of the audibility of changes.

Do relationships between loudness, brightness and pitch have any kind of cognitive reality? Recent research by Malara and Marks (1990) has suggested that they do. Using a full speeded paradigm, they examined response timing as a function of orientation of these dimensions. Redundancy gains and interference inhibitions were found indicating preferred combinations of loudness, timbre, and pitch. Of course, it is a leap to apply this directly to musical studies. Malara and Marks' stimuli were few and not especially musical; their unit of timbre was the duty-cycle of a square wave. Their study does not specifically show that very subtle predicted brightness differences, such as those found between trombones and trombonists, have perceptual or cognitive reality.

Other investigators have discussed timbre perception in terms suggestive of a cognitive model of tone production. For example, in an exhaustive examination of replicated-period sung vowels, Bloothooft and Plomp (1985) correlated perceptual, cognitive, and spectral spaces. They found that timbral differences could be well-represented in a two-dimensional space. The first dimension was referred to as the spectral effect of vocal effort, that is, "entirely the consequence of variation in the glottal sound source" (p. 862). The second dimension, on the other hand, reflected the "spectral effect of pharyngeal volume" (p. 863). In other words, cognitive and perceptual spaces correlate well with dimensions representing physical characteristics of the system and how it is being driven. As discussed in Chapter 7, differences in trombone tones can also be related to the physical system (via cutoff frequency and transfer function) as well as blowing pressure. Brightness as a function of loudness and pitch encapsulates these relationships.

12.7 Conclusion

As previously mentioned, the *bright-dark* continuum is already part of trombonists' vernacular. This is true although brightness is neither consistently judged (Pratt and Bowsher, 1978) nor uniformly desired (Edwards, 1978). It is hoped that, by describing the characteristics of trombone brightness, and its dependencies upon loudness, register, instrument and performer, an improved understanding will result that benefits both performance and pedagogy.

REFERENCES

Adler, S. (1989). *The Study of Orchestration* (2nd ed.). New York: W. W. Norton & Co.

American Standards Association (1960). *American Standard Acoustical Terminology.* New York: American Standards Association.

Anastasio, A. & N. Bussard (1971). Mouth air pressure and intensity profiles of the oboe. *Journal of Research in Music Education, 19*, 62-76.

Appert, D. L. (1980). The alto trombone – its uses, problems, and solutions. *Journal of the International Trombone Association, 8*, (March 1980), 13-14.

von Aures, W. (1985). Berechnungsverfahren fur den sensorischen Wohlklang beliebiger Schallsignale [Procedure for calculating the sensory euphony of arbitrary sound signals]. *Acustica, 59*, 130-141.

Backus, J. (1985). The effect of the player's vocal tract on woodwind instrument tone. *Journal of the Acoustical Society of America, 78*, 17-20.

Backus, J., & Hundley, T. C. (1971). Harmonic generation in the trumpet. *Journal of the Acoustical Society of America, 49*, 509-519.

Balzano, G. J. (1986). What are musical pitch and timbre? *Music Perception, 3*, 297-314.

Beauchamp, J. W. (1975). Analysis and synthesis of cornet tones using nonlinear interharmonic relationships. *Journal of the Audio Engineering Society, 23*, 778-795.

Beauchamp, J. W. (1982). Synthesis by spectral amplitude and "brightness" matching of analyzed musical instrument tones. *Journal of the Audio Engineering Society, 30*, 396-406.

Benade, A. H. (1976). *Fundamentals of Musical Acoustics.* New York: Oxford University Press.

Benade, A. H. (1983). Air column, reed, and player's windway interactions in musical instruments. In I. R. Titze & R. C. Scherer (Eds.), *Vocal Fold Physiology, Biomechanics, Acoustics and Phonatory Control* (pp. 425-452). Denver, CO: The Denver Center for the Performing Arts.

Benade, A. H., & Gans, D. J. (1968). Sound production in wind instruments. In A. Bouhuys (Ed.), Sound production in man [Special issue]. *Annals of the New York Academy of Sciences, 155*, 247-263.

Benade, A. H. & Kouzoupis, S. N. (1985). Spectral envelopes of orchestral instruments. *Journal of the Acoustical Society of America, 78*(Supp. 1), S75.

Benade, A. H. & Kouzoupis, S. N. (1988). The clarinet spectrum, theory and experiment. *Journal of the Acoustical Society of America, 83*, 292-304.

Benade, A. H. & Larson, C. O. (1985). Requirements and techniques for measuring the musical spectrum of the clarinet. *Journal of the Acoustical Society of America, 78*, 1475-1498.

Benade, A. H., & Lutgen, S. J. (1988). The saxophone spectrum. *Journal of the Acoustical Society of America, 83*, 1900-1907.

von Bismarck, G. (1974a). Timbre of steady sounds, a factorial investigation of its verbal attributes. *Acustica, 30*, 146-159.

von Bismarck, G. (1974b). Sharpness as an attribute of the timbre of steady sounds. *Acustica, 30*, 159-172.

Bloothooft, G., & Plomp, R. (1988). The timbre of sung vowels. *Journal of the Acoustical Society of America, 84*, 847-860.

Bloothooft, G., & Plomp, R. (1986). Spectral analysis of sung vowels. III. Characteristics of singers and modes of singing. *Journal of the Acoustical Society of America, 79*, 852-864.

Bouhuys, A. (1965). Sound power production in wind instruments. *Journal of the Acoustical Society of America, 37*, 453-456.

Bowsher, J. M. (1978). Experiments on the subjective assessment of trombones. *Journal of the Institute of Musical Instrument Technology, 3*(3), 96-102.

Bowsher, J. M. (1983). Thoughts on the difficulty of assessing brass instruments. *Proceedings, 11th International Conference on Acoustics* (Paris).

Campbell, M., & Greated, C. (1987). *The Musician's Guide to Acoustics.* New York: Schirmer Books.

Campbell, S. L. (1994a). Uni- and Multidimensional Identification of Rise Time, Spectral Slope, and Bandwidth. (Doctoral dissertation, University of Washington, 1993). *Dissertation Abstracts International B, 55*(2), 617.

Campbell, S. L. (1994b). Uni- and multidimensional identification of rise time, spectral slope, and cutoff frequency. *Journal of the Acoustical Society of America, 96*, 1380-1387.

Chowning, J., Grey, J., Moorer, J. A., & Rush, L. (1982). *Instrumental Timbre and Related Acoustical Phenomena in the Perception of Music.* (Report No. STAN-M-11). Stanford, CA: Stanford University, Center for Computer Research in Music and Acoustics.

Clarke, M. E., Luce, D., Abrams, R., Schlossberg, H., & Rome, J. (1963). Preliminary experiments on the aural significance of parts of tones of orchestral instruments and on choral tones. *Journal of the Audio Engineering Society, 11*, 45-54.

Clarke, M. E., & Milner, P. (1964). Dependence of timbre on the tonal loudness produced by musical instruments. *Journal of the Audio Engineering Society, 12*(1), 28-31.

Dempster, S. (1994). *The Modern Trombone, a Definition of its Idioms.* Athens, OH: Accura Music. (Original work published 1979).

Devors, J. L. (1987). *Probability and Statistics for Engineering and the Sciences.* New York: Brooks/Cole Publishing Company.

Dunnwald, H. (1990). Ein erweiteres Verfahren zur objektiven Bestimmung der Klangqualität von Violinen [An extended method of objectively determining the sound quality of violins]. *Acustica, 71*, 269-276.

Edwards, R. M. (1978). The perception of trombones. *Journal of Sound and Vibration, 58*, 407-424.

Edwards, R. M. (1979). A postal survey of trombone players in the London area. *Journal of the International Trombone Association, 9* (Jan. 1979), 15-19.

Ehresman, D. E., & Wessel, D. L. (1978). *Perception of timbral analogies* (Rapports IRCAM 13). Paris: IRCAM.

Elliot, C. A. (1975). Attacks and releases as factors in instrument identification. *Journal of Research in Music Education, 23*, 35-40.

Elliot, S. J., & Bowsher, J. M. (1982). Regeneration in brass wind instruments. *Journal of Sound and Vibrationm 83*, 181-217.

Elliot, S., Bowsher, J.M., & Watkinson, P. (1982). Input and transfer response of brass wind instruments. *Journal of the Acoustical Society of America, 72*, 1747-1760.

Evans, E. F. (1992). Auditory processing of complex sounds, an overview. *Philosophical Transactions of the Royal Society of London, B 336*, 295-306.

Fletcher, N. H. (1983). Acoustics of the Australian didjeridu. *Australian Aboriginal Studies, 1*(1), 28-37.

Fletcher, N. H., & Rossing, T. D. (1991). *The Physics of Musical Instruments.* New York: Springer-Verlag.

Fox, F. (1981). Bull's eye!. *The Horn Call, 11*(2), 65-66.

Gartner, J. (1981). *The Vibrato* (E. W. Anderson, Transl.). Regensburg, Germany: Gustav Bosse Verlag.

Goad, P. J. (1992). *Loudness and Sharpness Measurements Relevant to Musical Instrument Timbres.* (University of Washington Systematic Musicology Technical Report No. 9202). Seattle, WA: University of Washington, Systematic Musicology Program.

Goad, P. J. & Keefe, D. H. (1992). Timbre discrimination of musical instruments in a concert hall. *Music Perception, 10*, 43-62.

Goad, P. J. (1994). Timbral sharpness and modulations in frequency and amplitude, implications for the fusion of musical sounds. Unpublished doctoral dissertation, University of Washington.

Goad, P. J. (1994, June). *Sharpness, a Perceptually Based Measure of the Spectral Dimension of Musical Timbre.* Paper presented at the 127th Meeting of the Acoustical Society of America, Cambridge, MA.

Gordon, J. W., & Grey, J. M. (1978). Perception of spectral modifications on orchestral instrument tones. *Computer Music Journal, 2*(1), 24-31.

Gove, P. B., (Ed. in Chief) (1981). *Webster's Third New International Dictionary of the English Language, unabridged.* Springfield, MA: G. & C. Merriam.

Green, D. M. (1988). *Profile Analysis, Auditory Intensity Discrimination.* New York: Oxford University Press.

Grey, J. M. (1975). An Exploration of Musical Timbre. (Rep. No. STAN-M-2). Stanford, CA: Stanford University, Center for Computer Research in Music and Acoustics.

Grey, J. M. (1977). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America, 61*, 1270-1277.

Grey, J. M. (1978). The role of psychoacoustics in computer music. (Rep. No. STAN-M-38). Stanford, CA: Stanford University, Center for Computer Research in Music and Acoustics.

Grey, J. M. & Gordon, J. W. (1978). Perceptual effects of spectral modifications on musical timbres. *Journal of the Acoustical Society of America, 63*, 1493-1499.

Guion, D. M. (1988). *The Trombone, Its History and Music 1697-1811.* New York: Gordon and Breach.

Hall, J. C. (1955a). To 'ah-ee' or not to 'ah-ee'. *The Instrumentalist, 9* (April 1955), 36-39.

Hall, J. C. (1955b). Effect of oral and pharyngeal cavities on trumpet tone quality. *Journal of the Acoustical Society of America, 27*, 996.

Helmholtz, H. L. F. (1954). *On the Sensations of Tone* (A. J. Ellis, Transl.). New York: Dover. (Original work published 1885)

Hunsberger, D. (1980). *The Remington Warm-Up Studies.* Athens, OH: Accura Music.

International Organization for Standardization (1975). *Acoustics–Method for calculating loudness level* 1st ed. (International Standard ISO 532). [Geneva, Switzerland]: International Organization for Standardization.

Iverson, P., & Krumhansl, C. L. (1993). Isolating the dynamic attributes of musical timbre. *Journal of the Acoustical Society of America, 94*, 2595-2603.

Keefe, D. H., Bulen, J. C., Campbell, S. L., & Burns, E. M. (1994). Pressure transfer function and absorption cross section from the diffuse field to the human infant ear canal. *Journal of the Acoustical Society of America, 95*, 355-371.

Kendall, R. A. (1984). The role of transients in listener categorization of musical instruments, an investigation using digitally recorded and edited musical phrases.

Dissertation Abstracts International, 45(8), 2297. (University Microfilms No. AAC-8425754)

Kendall, R. A. (1986). The role of acoustic signal partitions in listener categorization of musical phrases. *Music Perception, 4*, 185-214.

Kendall, R. A., & Carterette, E. C. (1991). Perceptual scaling of simultaneous wind instrument timbres. *Music Perception, 8*, 369-404.

Kendall, R. A., & Carterette, E. C. (1993). Verbal attributes of simultanous wind instrument timbres, I. von Bismarck's adjectives. *Music Perception, 10*, 445-468.

Kendall, R. A., & Carterette, E. C. (1993). Verbal attributes of simultaneous wind instrument timbres, II. Adjectives induced from Piston's Orchestration. *Music Perception, 10*, 469-502.

Kennan, K. (1970). *The Technique of Orchestration* (2nd ed). Englewood Cliffs, N.J., Prentice-Hall.

Klein, W., Plomp, R., & Pols, L. C. W. (1970). Vowel spectra, vowel spaces, and vowel identification. *Journal of the Acoustical Society of America, 48*, 999-1009.

Kleinhammer, E. (1963). *The Art of Trombone Playing.* Evanston, IL: Summy-Birchard Co.

Klippel, W. (1990). Multidimensional relationship between subjective listening impression and objective loudspeaker parameters. *Acustica, 70*, 45-54.

Krumhansl, C. L., & Iverson, P. (1992). Perceptual interactions between musical pitch and timbre. *Journal of Experimental Psychology, Human Perception and Performance, 18*, 739-751.

Kuttruff, H. (1973). *Room Acoustics.* New York: John Wiley & Sons.

Lichte, W. H. (1941). Attributes of complex tones. *Journal of Experimental Psychology, 28*, 455-480.

Lindevald, I. (1988). The perception of musical tones in the statistical sound fields of rooms. (Doctoral Dissertation, Case Western Reserve University, 1987) *Dissertation Abstracts International B, 48*(7), 2002.

Ling, R. (1991). Ndsp, NeXT Digital Signal Processing [Computer software]. Seattle, WA: University of Washington.

Lubman, D. (1974). Precision of reverberant sound power measurements. *Journal of the Acoustical Society of America, 56*, 523-533.

Luce, D., & Clarke, M. (1967). Physical correlates of brass instrument tones. *Journal of the Acoustical Society of America, 42*, 1232-1243.

Malara, R. D., & Marks, L. E. (1990). Interaction among auditory dimensions: timbre, pitch, and loudness. *Perception and Psychophysics, 48*, 169-178.

Martin, D. W. (1942a). Lip vibrations in a cornet mouthpiece. *Journal of the Acoustical Society of America, 13*, 171-174.

Martin, D. W. (1942b). Directivity and the acoustic spectra of brass wind instruments. *Journal of the Acoustical Society of America, 13*, 309-313.

Mays, R. (1980). Bass trombone sound concept. *The Instrumentalist, 35* (November 1980), 54-58.

McIntyre, M. E., Schumacher, R. T., & Woodhouse, J. (1983). On the oscillations of musical instruments. *Journal of the Acoustical Society of America, 74*, 1325-1345.

Miller, J. R., & Carterette, E. C. (1975). Perceptual space for musical structures. *Journal of the Acoustical Society of America, 58*, 711-720.

Moorer, J. A., Grey, J. M., & Strawn, J. (1978). Lexicon of analyzed tones. Part 3, the trumpet. *Computer Music Journal, 2*(2), 23-29.

Padgham, C. (1986). The scaling of the timbre of the pipe organ. *Acustica, 60*, 189-204.

Paulus, E., & Zwicker, E. (1972). Programme zur automatischen Bestimmung der Lautheit aus Terzpegeln oder Frequenzgruppenpegeln [Computer programs for calculating loudness from third-octave band levels or from critical band levels]. *Acustica, 27*, 253-266.

Pierce, A. D. (1989). *Acoustics, An Introduction to its Physical Principles and Applications.* New York: Acoustical Society of America.

Pitt, M. A., & Crowder, R. G. (1992). The role of spectral and dynamic cues in imagery for musical timbre. *Journal of Experimental Psychology, Human Perception and Performance, 18*, 728-738.

Plomp, R. (1970). Timbre as a multidimensional attribute of complex tones. In R. Plomp and G. F. Smoorenburg (Eds.), *Frequency Analysis and Periodicity Detection in Hearing.* Leiden, Netherlands: Sijthoff.

Plomp, R. (1976). *Aspects of Tone Sensation.* New York: Academic Press.

Pollard, H. F. (1988). Feature analysis of musical sounds. *Acustica, 65*, 232-244.

Pols, L. C. W., van der Kamp, L. J. Th., & Plomp, R. (1969). Perceptual and physical space of vowel sounds. *Journal of the Acoustical Society of America, 46*, 458-467.

Pols, L. C. W. (1970). Perceptual space of vowel-like sounds and its correlation with frequency spectrum. In R. Plomp and G. F. Smoorenburg (Eds.), *Frequency Analysis and Periodicity Detection in Hearing.* Leiden, Netherlands: Sijthoff, p.463-473.

Pratt, R. L. & Bowsher, J. M. (1978). The subjective assessment of trombone quality. *Journal of Sound and Vibration, 57*, 425-435.

Pratt, R. L. & Bowsher, J. M. (1979). The objective assessment of trombone quality. *Journal of Sound and Vibration, 65*, 521-547.

Pratt, R. L. & Doak, P. E. (1976). A subjective rating scale for timbre. *Journal of Sound and Vibration, 45*, 317-328.

Preis, A. (1984). An attempt to describe the parameter determining the timbre of steady-state harmonic complex tones. *Acustica, 55*, 1-13.

Rasch, R. A., & Plomp, R. (1982). The perception of musical tones. In D. Deutsch (Ed.), *The Psychology of Music.* New York: Academic Press.

Read, G. (1969). *Thesaurus of Orchestral Devices.* New York: Greenwood Press.

Reitz, J. V. (1950). Brass instruments in the contemporary orchestral situation. Unpublished doctoral dissertation, University of Washington, 1950.

Riddle, N. (1985). *Arranged by Nelson Riddle.* Secaucus, NJ: Warner Bros. Publications.

Risset, J. D., & Matthews, M. V. (1969). Analysis of musical-instrument tones. *Physics Today, 22*(2), 23-30.

Risset, J. D., & Wessel, D. L. (1982). Exploration of timbre by analysis and synthesis. In D. Deutsch (Ed.), *The Psychology of Music.* New York: Academic Press.

Runyan, W. E. (1979). The alto trombone and contemporary concepts of trombone timbre. *Brass Bulletin-International Brass Chronicle, 28*, 43-50.

Sandell, G. J. (1992). Concurrent timbres in orchestration, a perceptual study of factors determining 'blend'. (Doctoral Dissertation, Northwestern University, 1991). *Dissertation Abstracts International, 53*(3), 660.

Sauer, R. (1984). The alto trombone in the symphony orchestra. *Journal of the International Trombone Association, 12*(3), 41.

Scharf, B. (1978). Loudness. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of Perception, Vol. IV* (pp. 187-233). New York: Academic Press.

Smith, R. A., & Daniell, G. J. (1976). Systematic approach to the correction of intonation in wind instruments. *Nature, 262*, 761-765.

Stevens, S. S., Volkmann, J., & Newman, E. B. (1937). A scale for the measurement of the psychological magnitude pitch. *Journal of the Acoustical Society of America, 8*, 95-190.

Strong, W., & Clark, M. (1966a). Synthesis of wind-instrument tones. *Journal of the Acoustical Society of America, 41*, 39-52.

Strong, W., & Clark, M. (1966b). Perturbations of synthetic orchestral wind-instrument tones. *Journal of the Acoustical Society of America, 41*, 277-285.

Strong, W. J. (1990). *Computer Simulation of a Player-Trumpet System.* Paper CCC3 presented at the Acoustical Society of America Conference, Penn State, PA.

Terhardt, E. (1978). Psychoacoustic evaluation of musical sounds. *Perception and Psychophysics, 23*, 483-492.

Ueda, K., & Akagi, M. (1990). Sharpness and amplitude envelopes of broadband noise. *Journal of the Acoustical Society of America, 87*, 814-819.

Ward, W. D. (1970). Musical perception. In J. Tobias (Ed.), *Foundations of Modern Auditory Theory.* New York: Academic Press.

Warren, W. H., Jr., & Verbrugge, R. R. (1984). Auditory perception of breaking and bouncing events, a case study in ecological acoustics. *Journal of Experimental Psychology, Human Perception & Performance, 10*, 704-712.

Watkins, A. J. (1991). Central, auditory mechanisms of perceptual compensation for spectral-envelope distortion. *Journal of the Acoustical Society of America, 90*, 2942-2955.

Watkins, A. J., & Makin, S. J. (1994). Perceptual compensation for speaker differences and for spectral-envelope distortion. *Journal of the Acoustical Society of America, 96*, 1263-1282.

Watkinson, P. S., Shepherd, R., & Bowsher, J. M. (1982). Acoustic energy losses in brass instruments. *Acustica, 51*, 213-221.

Wedin, L., & Goude, G. (1972). Dimension analysis of the perception of musical timbre. *Scandinavian Journal of Psychology, 13*, 228-240.

Wessel, D. L. (1979). Timbre space as a musical control structure. *Computer Music Journal, 3*(2), 45-52.

Wick, D. (1971). *Trombone Technique.* London: Oxford University Press.

Wogram, K. (1983). The acoustical properties of the brass instruments. *The Horn Call, 13*(2), 19-31.

Yeo, D. (1985). The bass trombone, innovations on a misunderstood instrument. *Journal of the International Trombone Association, 13*(3), 22-28.

Zwicker, E., and Fastl, H. (1990). *Psychoacoustics, Facts and Models.* Berlin: Springer-Verlag.

Appendix A

MEASUREMENT ACCURACY AND STEREO RECORDINGS

When recording sounds in a room appropriate for music, the reverberant sound field causes spectral measurements from any combination of source and microphone positions to differ from others. In this study, trombonists performed from positions that were equivalent in terms of musical listening, although not physically identical. Likewise, microphones were place in similiar but not identical positions. Spectral measurements, and hence measured brightness and loudness, can be expected to vary based on differences in position and orientations. The questions to be addressed in this Appendix are the following:

- How much will brightness and loudness measurements vary over the range of source and microphone positions used in this study?
- What is the accuracy of (1) individual measurements of brightness and loudness, (2) fitted surfaces as described in Equation 11.2?
- How should the data be utilized in light of this? For example, when stereo recordings are available, should measurements from each microphone be analyzed separately or pooled together?

A.1 Pilot Study: Measurement Variability with Displacement of Source and Microphones

The accuracy of loudness and brightness measurements can be predicted theoretically if the sound field is predominantly diffuse. In the current measurement, the sound field is not diffuse, but "semi-reverberant" (Lubman, 1974). At low frequencies the diffuse field predominates, but at high frequencies the direct field may be as intense as the diffuse. Thus, it is not possible to confidently predict measurement variability from a theoretical perspective. It can be investigated empirically, however.

A pilot study was undertaken to examine the effect of source and microphone placement on loudness and brightness measurements. Since it is unreasonable to expect a trombonist to repeatedly produce identical tones or to sustain them through microphone repositioning, a synthesized tone was used that resembles a loudly-played trombone.

A.1.1 Materials

Tones were produced with an Ensonic model VFX synthesizer, using a trombone-like present ("synth-horns"), modified to eliminate AM and FM modulation. Each tone was produced by striking firmly to produce a rich, *mezzo-forte*-quality tone; the key was then taped down. The signal was amplified and played through a single JBL model LE-8 8-inch broad-range speaker, mounted in a Morgan Sound enclosure. The sound source was therefore comparable in size to a trombone bell, although directional characteristics are not comparable. Levels were used that corresponded roughly to sound pressure levels generated by trombone at a *mezzo-forte* dynamic. The tones D2, D3, D4 and D5 were used, which span the range of the composite trombone family.

Microphones and source were positioned as in trombone recordings, with microphones about 6 meters from the source and symmetrically placed approximately 30 degrees from midline. The speaker was placed on a chair, about .5 meters from the floor.

A.1.2 Procedure

For each tone, stereo recordings were made (about 2 seconds each) with the speaker in three distinct positions. For each speaker position, each microphone took four positions. Thus, there were eight microphone placments for each of three speaker placements, for a total of 24 recordings for each tone.

Microphone positions were varied within a sphere of approximately 1 meter in radius. Speaker position was varied within a disk of approximately 1 meter in radius. Although height from the floor was not varied, horizontal and azimuthal positions were varied by turning and tilting the speaker.

A.1.3 Results and Discussion

Since brightness in acums and loudness in sones are perceptual measures (i.e., a doubling in sensation level corresponds to a doubling of the predicted value), results were converted to logarithmic form. In this way, appropriate means and standard deviations could be calculated. As discussed in Chapter 10, average difference in log brightness is the same as the log of the geometric mean of brightness ratios (see Equation 11.1). The standard deviation gives an appropriate indication of variability when displayed on log axes.

To examine measurement variability caused by source and microphone displacement, the standard deviation of log brightness and loudness measurements has been calculated for each note across three speaker positions and eight microphone locations (n=24). The results are presented graphically in Figure A.1. For example, the mean loudness for the note D2 is 54 sones, and the mean brightness is 2.7 acums. Vertical bars show one standard deviation above and below the mean.

Standard deviations of log brightness ranges from .08 to .12 log acums. In linear terms, these correspond to a standard deviation of about 8 and 13% respectively. For log loudness, standard deviations ranged from .08 to 0.17 log sones, or 8-19 %.

A second issue concerns the interpretation of recordings made with a single source position but multiple microphone positions. In a diffuse field, energy transfer depends on the nodal system at both source and receiver locations. To what degree can recordings made with one source location be compared to recordings with another source location? This can be examined by separating out tones recorded with the same source position. Means and standard deviations for these (n=8) are shown in Figure A.2. It can be seen that, in the case of loudness measurements for the note D5, means can differ by more than one standard deviation. This does not occur with brightness measurements.

To summarize, measurement variability due to source and microphone displacement is from 8 to 13% for brightness and from 8 to 19% for loudness, for these synthesized tones. The amount of variability appears to depend on fundamental frequency. Using a single source position but multiple microphone positions sometimes produces mean brightness and loudness values that differ by more than one standard deviation from measurements using another source position.

Strictly speaking, these figures apply only to the synthesized tones that were used.

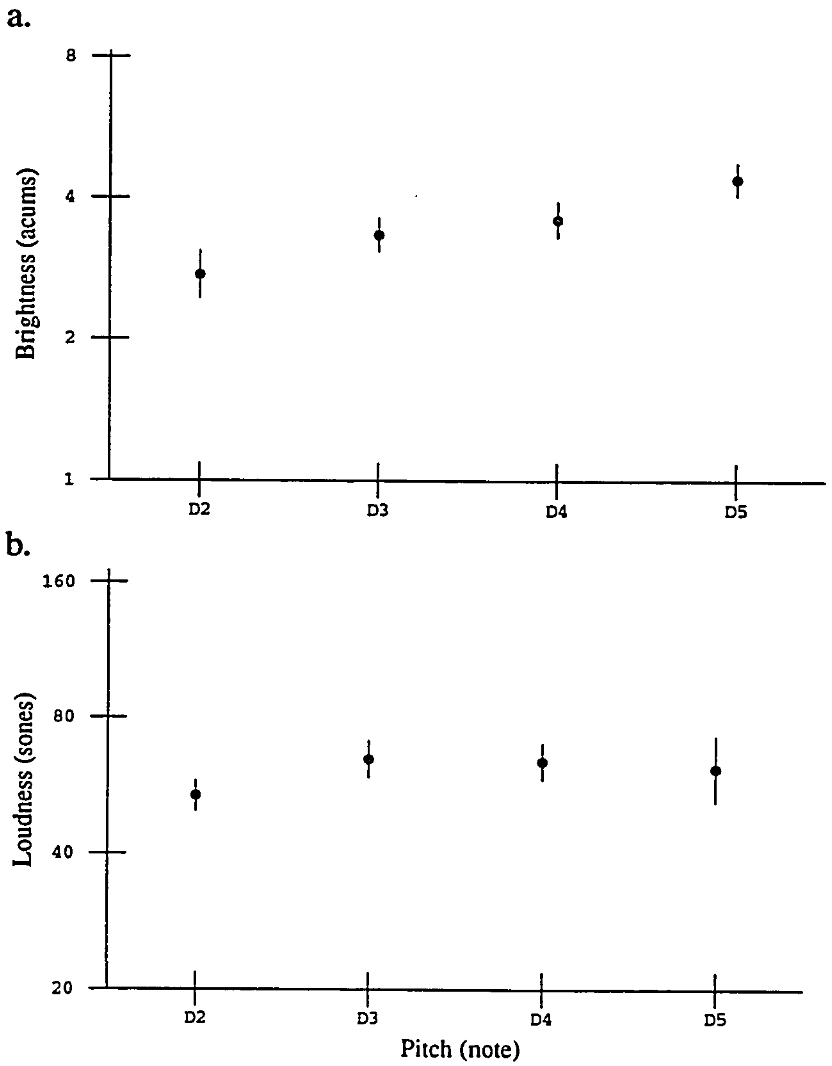


Figure A.1: Mean loudness and brightness of a synthesized tone, averaged over three source positions with eight microphone positions each. a. Brightness; b. Loudness.

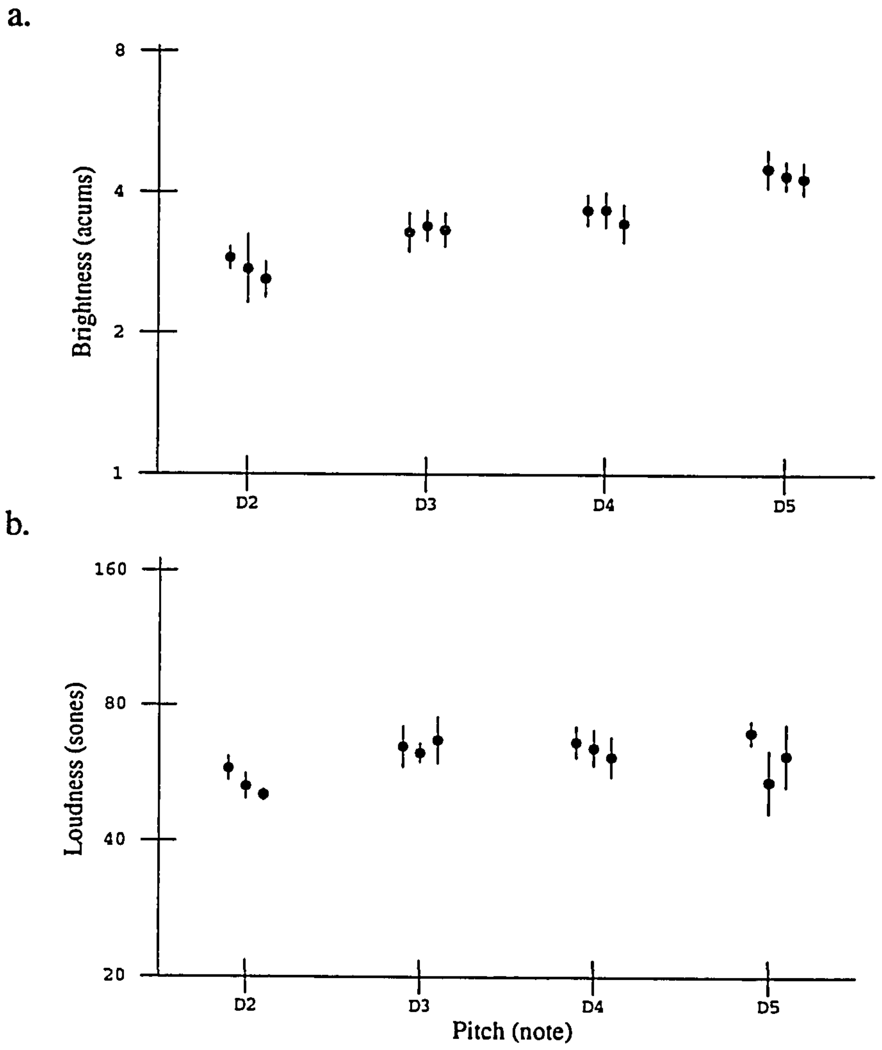


Figure A.2: Mean loudness and brightness of a synthesized tone produced at three distince source positions. For each source, measurements are averaged over eight microphone positions each. a. Brightness. b. Loudness.

This is because both loudness and brightness calculations involve averaging within critical bands. For example, some tones with high-frequency harmonics of a low fundamental will have many components within a critical band, and measurement variability will be low. Others may have a small number and be strongly affected by fluctuations of individual components. As described in Chapter 4, the harmonic structure of trombone tones vary with register and intensity; thus, measurement variability may vary over ranges of pitch and loudness.

A.2 Measurement Variability for Trombone Tones

There are 1,709 two-channel recordings of trombone tones, providing information about variability of loudness and brightness measurements for a single source but with two microphone positions. A two-channel estimate of standard deviation for loudness and brightness can be calculated for each tone; these can then be averaged. For log brightness, differences averaged to 0.011, reflecting a geometric mean of 1.011. The standard deviation of the log differences was 0.060, corresponding to a ratio of 1.062. Thus, tones within one standard deviation in mean brightness lay between .95 and 1.07 of the mean. For log loudness, differences averaged to -0.0143, corresponding to a geometric mean of 0.986. The standard deviation was 0.0796, so loudnesses within one standard deviation of the mean fell between and 0.85 and 1.15 times the mean. As expected, this is a somewhat smaller range than that suggested by synthesized tones. Thus, log brightnesses differences averaged to within about 1% of 0, with a standard deviation of about 7%. Log loudnesses differences averaged to within 1.5% of 0, with a standard deviation of about 15%.

A.2.1 Influence of Stationary Microphones on Measurements

As seen in Figure A.2, source location affects measured loudness and brightness, at least for some tones. Due to reciprocity, it is likely that a single microphone position will likewise bias measurements, even if the source is displaced. By examining measurement variability between microphones, it is possible to investigate by how much the recordings from one single microphone position may vary from those at another, as well as how variability changes with loudness and pitch.

This can be examined empirically using trombone tones. Performer B happened to perform all three trombones with exactly the same microphone position. Comparing measurements from these positions the positions gives an idea of the size of the bias, and whether it consistently appears across position changes by the trombonists. Since room resonances remain fixed in frequency, it is appropriate to perform averaging across frequencies.

A.2.2 Bias Due to a Single Microphone Position

Bias was investigated in the following manner. For each trombone, stereo channels were first separated. From these, data for one pitch at all its loudness levels were extracted. The difference in log loudness and log brightness between notes at each microphone position were then calculated. Means and standard deviations of these differences were then calculated for each note. Finally, these were converted to geometric mean ratios for interpretability, and are shown in Figure A.3 and Figure A.4.

Three features are noticeable in these plots. First, while for most notes one channel does not differ markedly from the other, there are a number of notes in each plot where differences are greater than one standard deviation. Secondly, the magnitude of the bias is on the order of measurement variability. Mean ratios for brightness are all less than $2^{1/8}$ (about 9%), and for loudness they are less than 19%. Third, the contours of the plots are similar. For example, all show a dip at B♭3, a peak at B3, etc. This can be quantified by calculating a cross-correlation of the lists, after dropping points that aren't held in common. The results of this are shown in Table 11.2 for brightness, and Table 11.3 for loudness. It can be seen that the curves are moderately correlated, suggesting a consistent bias introduced by microphone positioning. Finally, ratios cluster about 1.0; any smoothing function that operates across frequency (such as Equation 11.2) will be likely to eliminate the bias.

A.2.3 Measurement Bias and Fitted Surfaces

Since brightness differences between alto, tenor and bass trombones are greater than this bias, it can be disregarded for these comparisons. However, differences between players using the same instrument are sometimes small. As shown in Chapter 11, to compare players it is necessary to smooth data by fitting to a surface as in Equation 11.2. To what degree do biases apparent in Figures A.3 and A.4 affect fitted

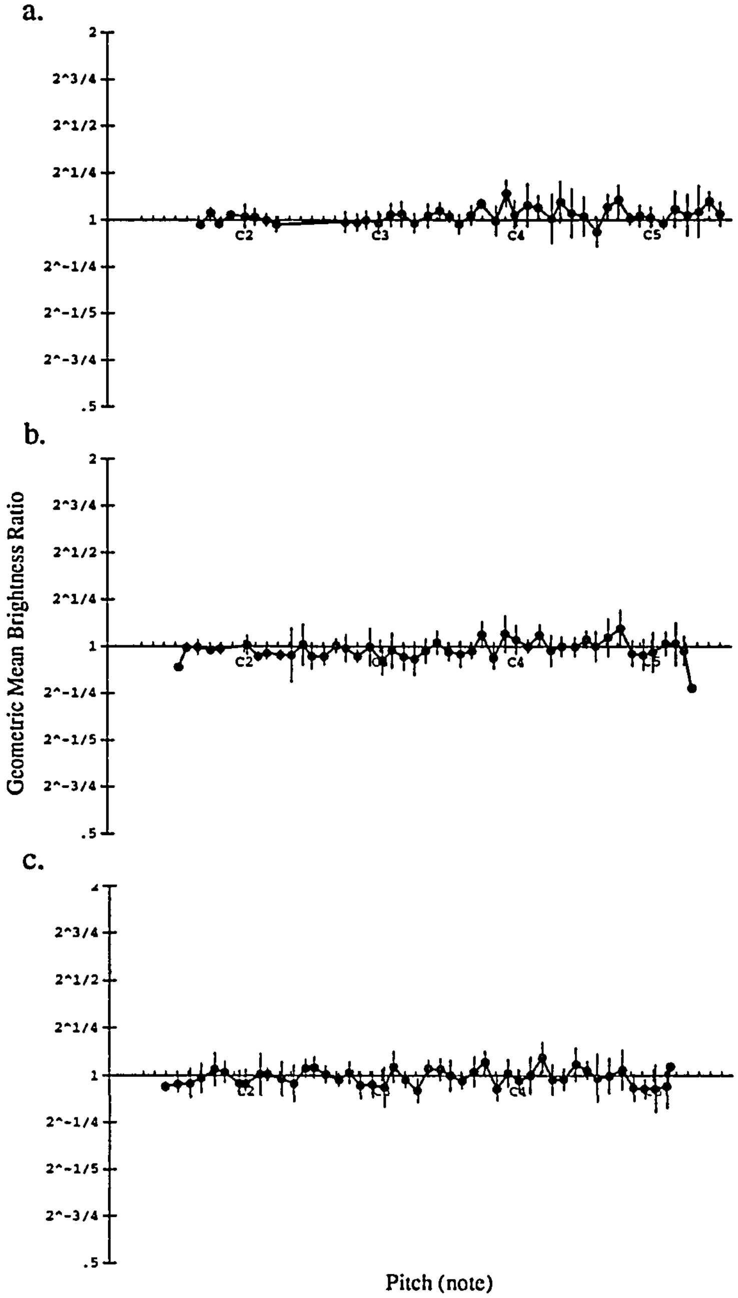


Figure A.3: Geometric mean brightness ratios and standard deviations for notes recorded at two microphone positions. a. Performer B, alto trombone; b. Tenor trombone; c. Bass trombone.

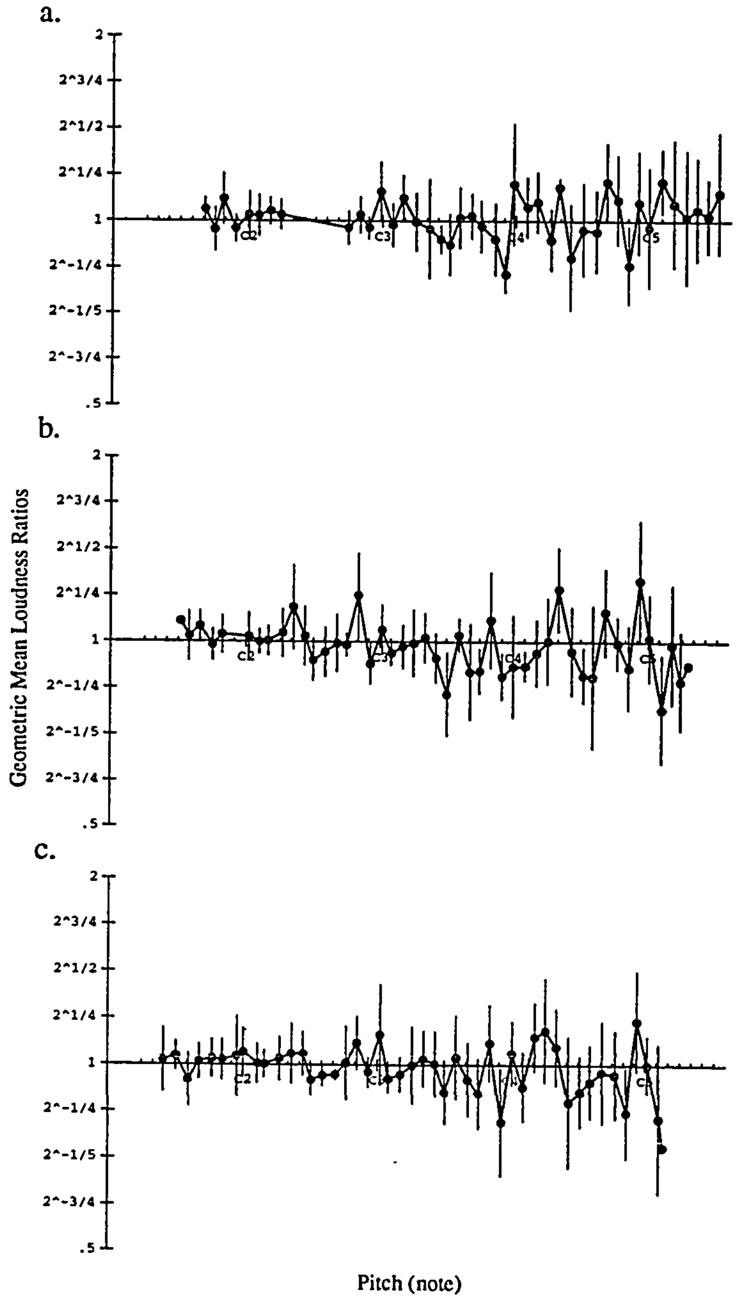


Figure A.4: Geometric mean loudness ratios and standard deviations for notes recorded at two microphone positions. a. Performer B, alto trombone; b. Tenor trombone; c. Bass trombone.

Table A.1: Cross correlations of geometric mean brightness ratios between microphones, for three different trombones. Strong correlations indicate a consistent measurement bias.

	Performer B Alto	Performer B Tenor	Performer B Bass
Performer B Alto	1		
Performer B Tenor	.66	1	
Performer B Bass	.47	.57	1

Table A.2: Cross correlations of geometric mean loudness ratios between microphones, for three different trombones. Strong correlations indicate a consistent measurement bias.

	Performer B Alto	Performer B Tenor	Performer B Bass
Performer B Alto	1		
Performer B Tenor	.31	1	
Performer B Bass	.40	.74	1

surfaces? How different are surfaces produced by each microphone? Finally, how do differences arising from microphone placement compare to differences between performers?

To investigate this, right and left microphone channels were separated, and surfaces of the form of Equation 11.2 fit to each. Differences in log brightness were then calculated at evenly-spaced loudness increments. Across all trombones and performers, the mean difference was .03 log acums, or about 3%. The standard deviation was .037 log acums, or about 4%. Differences between players are greater than this in certain regions. The difficulty is that differences between channels are not constant across pitch and loudness.

The most conservative (and laborious) approach is that taken in Figures 11.18, 11.19, and 11.20 of Chapter 10. Differences between channels for each player were first calculated for each note; an error function was then calculated for that one note by Equation 11.4. Differences between players were reported only if they exceeded the error function.

Differences between channels were large only for Performer C on alto trombone. The reasons for this are not clear. In general, the patterns of differences that appear in Figures 11.15, 11.16, and 11.17 are also apparent when comparing individual channels.

A.3 Summary

In light of the above, comparisons were made on the basis of pooled data. When smoothed by fitting to a surface, differences between recordings of the same tones at two loctaions were on average smaller than differences between trombones and trombonists. More to the point, since microphone positions differed slightly for each performer, data from two microphones gives a more general description of room response than does a single response.

Appendix B

FUNDAMENTAL FREQUENCY ESTIMATION

The trombone tones measured in this experiment are both periodic and harmonically complex. To the extent that both of these conditions are true, fundamental frequency can be established with a resolution surpassing that of the Fourier transform binwidth.

The simplest way to accomplish this is simply to determine the frequency of an upper harmonic spectral peak, and then divide by the harmonic number. This was done for all of the notes for one trombone (subject C, alto trombone). The tenth harmonic was utilized except for a very few notes in which the tenth harmonic was weak; in this case the eighth or ninth harmonic was substituted.

For a 44.1 KHz sampling rate and a 32,768-point FFT size, the analysis binwidth is 1.35 Hz. If the tenth harmonic is used, then fundamental frequency is determined to within .135 Hz. For the note Bb1 at 51.9 Hz this translates to an accuracy of 4.5 cents; for Bb4 at 415.3 Hz accuracy is 0.6 cents. Unfortunately, this procedure is too time-consuming for general use.

The signal-processing program *NDSP* (Ling, 1991) includes two algorithms for fundamental frequency estimation. One of these, based on a peak-frequency fitting algorithm by Keefe, is appropriate for the present study.

Figure B.1a shows fundamental frequency as estimated from the tenth harmonic. The ordinate is arbitrary; each diagonal line of notes represents notes from one dynamic level. Figure B.1b shows fundamental frequency estimates from the peak-frequency fitting algorithm. As can be seen, results are similar except for loud, low-pitched tones. Where differences occur, they tend to be large.

Figure B.2 shows a comparison between the two estimates in cents (1/100 of a semitone). Errors clearly fall into two categories: very small errors (on the order of a few cents), and very large errors (greater than 400 cents). Large errors are easily spotted during data processing. The small errors all lie within 16 cents; the mean error is 0.85 +/- 0.50 cents and the standard deviation is 7.44 cents (n = 225).

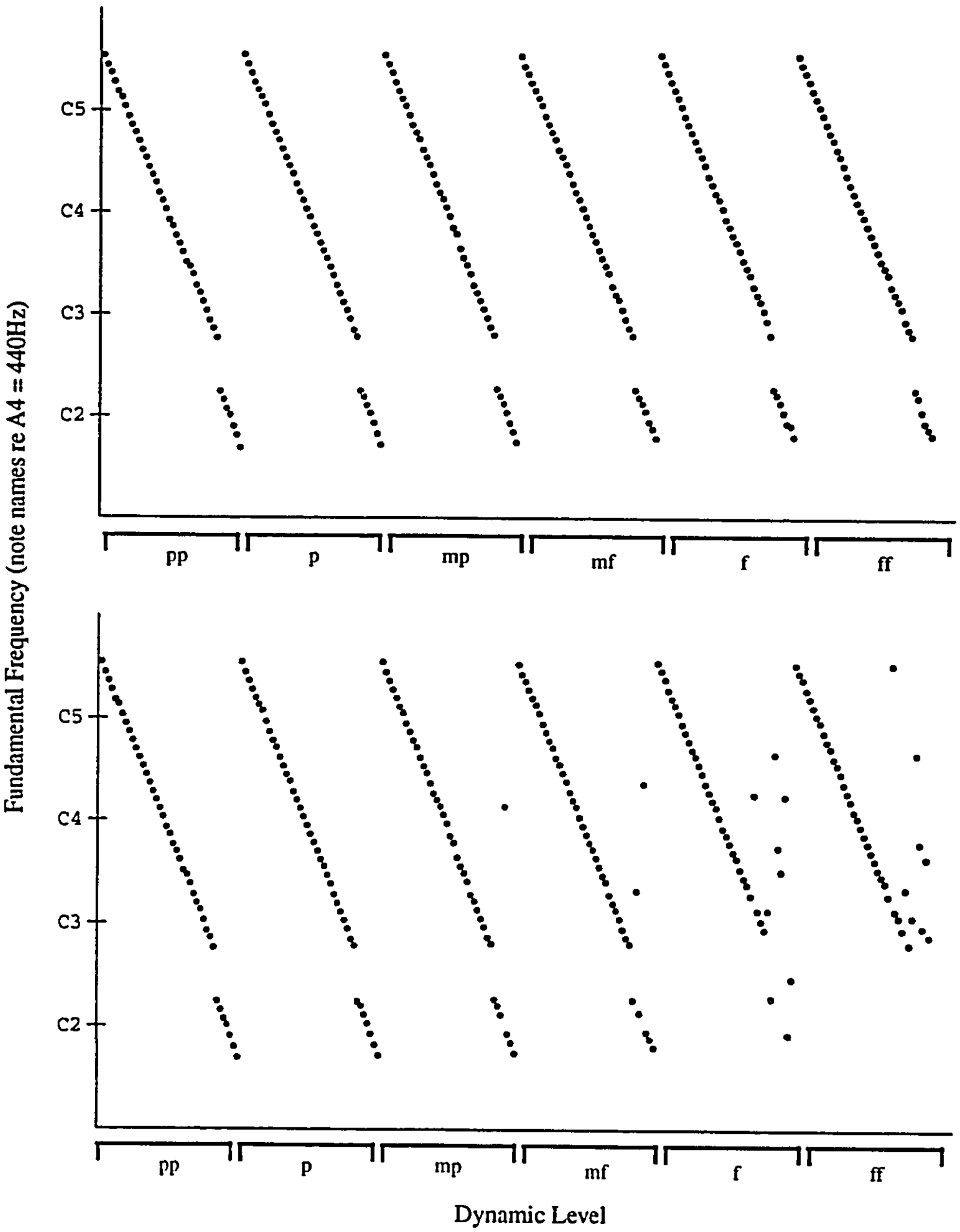


Figure B.1: Fundamental frequency estimates for notes of various dynamics levels and pitches, using two different methods. a. Estimate based on frequency of tenth harmonic; b. Estimate based on peak-frequency fitting algorithm.

Thus, 95% of the small-error estimates will be between -11.4 and + 13.1 cents of the estimate generated from the tenth harmonic. This is on the order of the tuning ability of musicians: "when tuning an instrument to the same pitch as a standard, the typical musician will show a standard deviation, in repeated settings, of about 10 cents" (Ward, 1970). Peak frequency fitting accurately determined fundamental frequency estimate for 93% of the tones, and all errors were with loud, low tones.

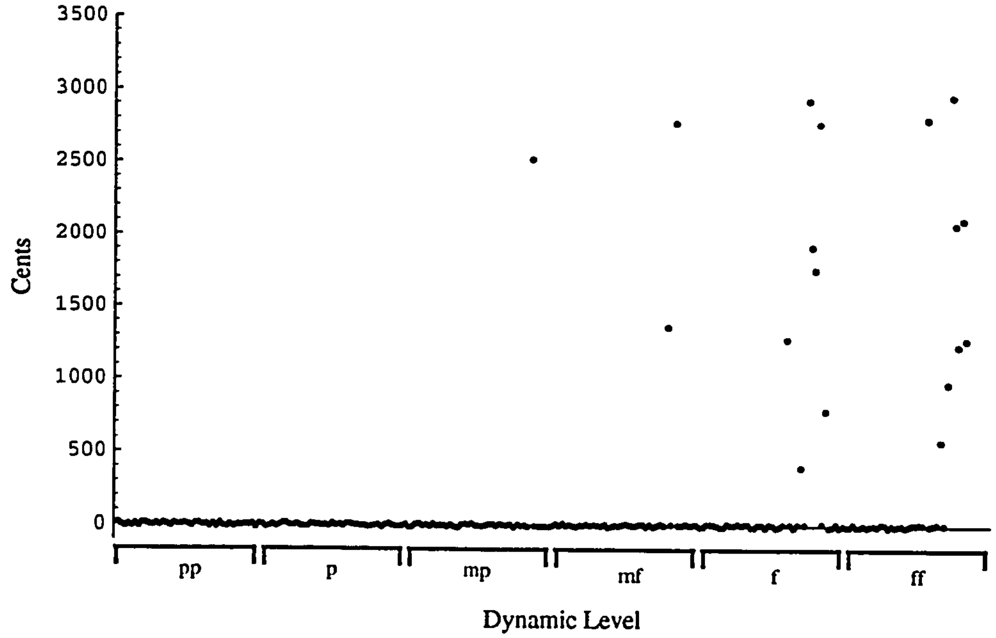


Figure B.2: Difference between fundamental frequency estimates produced by the 10th harmonic and spectral peak fitting methods, for notes of various dynamics levels and pitches.

It was concluded that fundamental frequency can be calculated using the peak-frequency fitting algorithm. Large errors are obvious, and can be resolved by estimating fundamental frequency with an upper harmonic. For situations where a large error is not indicated, a 95% confidence interval is from -11.4 to +13.1 cents.

Biographical Note: Jay C. Bulen

Education

Doctor of Musical Arts (Trombone Performance) University of Washington, Seattle, June 1995

Ph.D. Candidate in Systematic Musicology (Musical Acoustics) University of Washington, Seattle, in progress

Master of Music (Trombone Performance) Arizona State University, Tempe, June 1982

Bachelor of Fine Arts (Trombone Performance) California Institute of the Arts, Valencia, June 1980

Publications and Recordings

Keefe, D.H., Bulen, J.C., Campbell, S.L., and Burns, E.M. (1994). Pressure transfer function and absorption cross-section from the diffuse field to the infant ear canal. *Journal of the Acoustical Society of America 95*(1), p. 365-371.

Keefe, D.H., Bulen, J.C., Arehart, K.H., and Burns, E.M. (1993). Ear canal impedance and reflection coefficient in human infants and adults. *Journal of the Acoustical Society of America 94*(5), p. 2617-2638.

Keefe, D.H., Ling, R. and Bulen, J.C. (1992). Method to measure acoustic impedance and reflection coefficient. *Journal of the Acoustical Society of America 91*(1), p. 470-485.

Burns, E.M., Harrison, W.A., Bulen, J.C., and Keefe, D.H. (1993). Voluntary contraction of middle ear muscles: effects on input impedance, energy reflectance and spontaneous otoacoustic emissions. *Hearing Research 67*, p. 117-127.

Harrison, W.A., Burns, E.M., Bulen, J.C., Campbell, S.L., Keefe, D.H., and Culpepper, N.B. (1991). Voluntary control of the middle ear reflex. Abstracts of the Fourteenth Midwinter Research Meeting, Association for Research in Otolaryngology.

Morton Subotnick: *After the Butterfly.* Mario Guarneri, soloist, 20th Century Players. Nonesuch 78001 (D).