



The Capacity for Simulation by Physical Modeling to Elicit Perceptual Differences Between Trumpet Sounds

Robin Tournemenne, Joël Gilbert, Jean-François Petiot

► To cite this version:

Robin Tournemenne, Joël Gilbert, Jean-François Petiot. The Capacity for Simulation by Physical Modeling to Elicit Perceptual Differences Between Trumpet Sounds. *Acta Acustica united with Acustica*, 2016, 102 (6), pp.1072-1081. 10.3813/AAA.919020 . hal-01504199

HAL Id: hal-01504199

<https://hal.science/hal-01504199>

Submitted on 27 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The Capacity for Simulation by Physical Modeling to Elicit Perceptual Differences Between Trumpet Sounds

Robin Tournemenne¹⁾, Jean-François Petiot¹⁾, Joël Gilbert²⁾

¹⁾ IRCCyN (Institut de Recherche en Communications et Cybernétique de Nantes UMR CNRS 6597),
École Centrale Nantes, 1 rue de la Noë, Nantes, France.
[robin.tournemenne, jean-francois.petiot]@irccyn.ec-nantes.fr

²⁾ LAUM (Laboratoire d'Acoustique de l'Université du Maine UMR CNRS 6613), Université du Maine,
Le Mans, France. joel.gilbert@univ-lemans.fr

Summary

This paper investigates the subjects' ability to perceive trumpet sound changes, the sounds being either simulated by physical modeling or played by a real musician. The objective of the paper is to study to what extent simulations are able to create perceptually distinct sounds and to compare the discrimination on simulated sounds to the discrimination obtained when the instrument is played. Two experimental factors are considered to study perceptual discriminations by a panel of subjects: changes in the resonator geometry (instrument discrimination) and changes in the sound dynamics of the same instrument (dynamics discrimination). The changes in the resonator geometry are made with tiny geometrical variations of a parameterized leadpipe, whereas changes in dynamics are generated modifying control parameters for simulated sounds and playing dynamic instructions for real musicians. The same note (Bb4) is simulated using the harmonic balance technique and played by a real musician. To assess the discrimination, a "same-different" test has been carried out involving a panel of 24 listeners and pairs of sounds coming from the simulations or the recordings (for the sake of the experiment, the recorded sounds are resynthesized). For the different experimental conditions (differences between trumpets or between dynamics), the abilities of the listener to discriminate different sounds and to identify identical sounds are analyzed. Using the Receiver Operating Characteristics plots, the results show that even if the differences are more noticeable with sounds played by the musician, simulations are able to produce noticeable differences between the instruments, mainly due to playing frequency differences. They confirm the capacity of the physical model to represent realistic playing conditions for the trumpet.

1. Introduction

Investigating physical models of musical instruments is an interesting way of raising knowledge about their functioning, and may provide better assistance in their design. Yet, assistance in instrument design using physical modeling simulations is still at an early stage [1] and further work is still necessary to enhance its reliability. An interesting comparison of the sounds of reed instruments of different types (single or double reed) is made using time domain simulations in [2]. With huge variations in the bore of the instrument (two instruments of different families), the results show that simulated sounds and real sounds exhibit the same differences. Time domain simulations are also described in [3] to investigate perceptual differences between guitar sounds with psychoacoustical

testing. These experiments investigate just noticeable differences between simulated sounds from different guitars, even if the "virtual" modifications of the instrument are not necessarily realistic and cannot be manufactured. The ability for sound simulations to predict the sound characteristics of a real instrument through physical modeling is an interesting challenge for the computer-aided design of instruments. It needs to integrate different phases to confirm that simulations can represent the sounds of real instruments. The first phase consists in choosing a relevant physical model to represent the functioning of the instrument [4, 5]. Then, once a physical model simulating sounds is chosen, a second phase consists in defining to what extent the simulated sounds are in agreement with their equivalent real sounds played by "real" musicians on real instruments [2]. In [6], the authors focused their work on how simulations by physical modeling could be used to predict certain characteristics of trumpet sounds like the playing frequency, the spectral centroid and the spectral

irregularity. These objective measurements are interesting for instrument comparisons, but perceptual aspects of the sounds must be addressed to consider simulations as an aided design tool. A third phase then concerns the study of the agreement between simulations and real sounds from a perceptual point of view. [7] presents a first study on the noticeable differences between sounds, whether they be simulated or real. The present paper is in the continuum of this work and focuses on perceptual aspects. While based on the same experimental data as [7], the present paper undertakes a more complete analysis of the perceptual differences between sounds. Like in previous papers, the harmonic balance technique [8] is used to simulate the sounds, with different control parameters for the simulation. Likewise, small variations in the internal geometry of the trumpet are made by controlling the geometry of the leadpipe, an important part of the bore connecting the mouthpiece to the tuning slide. The general aim of this paper is to study whether simulations through physical modeling are able to produce perceptible differences between instruments. An additional objective is to compare the perceptive differences obtained with the simulations to those obtained with a real trumpeter playing the instruments. The present study does not focus on the nature of the perceptive differences, but focuses only on a first step in their magnitude. For this reason, the perceptual test considered in this study is a discrimination test that aims to study whether differences between sounds are detected. Different protocols can be proposed for a discrimination test, like the ABX test [9] or the 2 AFC (2 alternative forced choice) protocol [10]. The latter is generally used when the underlying dimension can be specified. Given that the differences between the sounds proposed in the experiments are extremely subtle, according to timbre and pitch, the present study considers a simple protocol called the same-different test, or AX test [11]: a pair of sounds is presented, the participant is simply asked to indicate if the two sounds are different or the same. This protocol, simpler than the ABX test because only two sounds are played before the decision of the participant, is retained for the simplicity of its implementation. The paper focuses on two aspects of discrimination: (1) discrimination between different instruments, obtained by small variations of the geometry of the resonator. In this case, pairs of sounds include two sounds generated by different instruments; (2) discrimination between different dynamics of the same instrument, obtained by varying the control parameters of the simulations, or the instructions given to the trumpeter. In this case, pairs of sounds include two sounds generated by the same instrument. Two categories of sounds are proposed for the discrimination tests: simulated sounds, obtained by simulations by physical modeling, and real sound, obtained by the recording of a real trumpeter. It is important to note that the protocol for the test will never mix a simulated sound and a real sound to constitute a pair (these two categories of sounds being very different): the pairs are either simulated or real. The results expected from this study concern the abilities of simulations to introduce changes

in the timbre or pitch of an instrument, and an assessment of how these changes are perceptible. An additional result concerns the indirect comparison of the results of the discrimination tests for the simulated sounds, and the real sounds. Section 2 provides details on the physical model of trumpets, the simulation method and its parameters and the description of the instruments used in the experiments. Section 3 describes the material and methods of the experiments, through the presentation of the same-different test, the design of the stimuli used for instrument and dynamics discrimination and the data analysis method. Section 4 describes the results and a discussion about the abilities of simulation to elicit perceptual differences between trumpet sounds. The main conclusions of this study are drawn in Section 5.

2. Background on the physical modeling of a trumpet

2.1. Physical model and harmonic balance technique

The trumpet is a wind instrument that produces sounds by a coupling of an excitator, the lips of the musician, to a resonator, the body of the instrument. The characteristics of the played note (frequency, timbre) depend on the inner shape of the resonator (called the “bore”) and of course on the musician’s ability. From an overpressure P_m in the mouth of the musician (from 1 kPa to about 10 kPa), the lips act as a vibrating valve that modulates the air flow into the instrument [1]. The column of air in the instrument vibrates, according to the resonance frequencies of the resonator. A regime of oscillations is created, result of a coupling between the resonator and the lips. It is important to mention that this coupling is the result of the reaction of the resonator on the lips: vibrations of the lips are facilitated at frequencies which correspond to the resonance frequencies of the bore. With the same bore, several notes can be obtained, corresponding to the different resonance frequencies.

To obtain a chromatic scale, the musician uses valves changing the bore length, thus the resonance frequencies. The main measurement characterizing the resonance frequencies of the resonator alone is the input impedance Z (Figure 3). It represents the response of the instrument in forced oscillations for a given frequency range and shows several peaks (corresponding to resonances) that are used to play the different notes. The basic physical model of the trumpet is described by three equations (Equation 1, 2, 3) that represent the behavior of the different parts of the system. See for example, [1].

$$p(j\omega) = Z(j\omega)v(j\omega), \quad (1)$$

$$\frac{d^2 H(t)}{dt^2} + \frac{2\pi f_l}{Q_l} \frac{dH(t)}{dt} + (2\pi f_l)^2 (H(t) - H_0) = \frac{P_m - p(t)}{\mu_l}, \quad (2)$$

$$v(t) = bH(t)\sqrt{\frac{2(P_m - p(t))}{\rho}}. \quad (3)$$

These three equations associate three time-varying variables: the opening height $H(t)$ of the two lips, the volume flow $v(t)$ and the pressure in the mouthpiece $p(t)$. Equation (1), with the frequency domain notation, describes the acoustical behavior of the resonator. It represents the impedance Z of the instrument, defined as the ratio of acoustic pressure in the mouthpiece $p(j\omega)$ to the acoustic volume flow $v(j\omega)$ entering the instrument (frequency domain notation). Equation (2) describes the mechanical behavior of the lips of the musician, considered as a one degree of freedom mechanical oscillator that controls the opening height $H(t)$ between the two lips ($H(t) = 0$ is numerically imposed if $H(t) < 0$). Equation (3) describes the coupling between the lips and the trumpet. Obtained by expressing the Bernoulli theorem, it represents a non-linear coupling between the pressure in the mouthpiece $p(t)$, the opening height of the lips $H(t)$ and the volume flow $v(t)$. Of course more complex models can be developed by refining the description of the system but this basic model is sufficient to represent the underlying physics of brass instruments. Several parameters are introduced in this model: (1) parameters concerning the “virtual musician”: P_m (pressure in the mouth), f_l (resonance frequency of the lips), μ_l (mass per area of the lips), b (width of the lips), H_0 (rest value of the opening height of the lips), Q_l (quality factor of the resonance of the lips); (2) the input impedance Z of the trumpet and (3) the air density ρ . Numerical solutions $p(t)$ of this system of equations can be computed to simulate the sound created by a trumpet, given by its input impedance Z , for a given virtual musician defined by its control parameters.

The harmonic balance technique is utilized to simulate sounds. It simulates sounds in permanent regime (steady state) in the frequency domain. This technique is a numerical method computing converging periodic solutions of the pressure $p(t)$ of the system taking into account a given finite number N of harmonics in a truncated Fourier series,

$$p(t) = \sum_{n=1}^N A_n \cos(2\pi n F_0 t + \varphi_n). \quad (4)$$

Assuming that the solution $p(t)$ of the system of equations is harmonic (Equation (4)), the unknowns of the simulations are the amplitudes of the harmonics A_n , the phases φ_n and the playing frequency F_0 . A numerical solution $p(t)$ satisfying equations (1), (2), (3) can be defined if the system converges towards a stable solution (more details can be found in [8]). From the values of the amplitudes A_n , the phases φ_n and the playing frequency F_0 , a steady state sound can then be synthesized by additive synthesis.

Let us not forget that the approached solution $p(t)$ computed from the elementary model corresponds to an acoustic pressure inside the instrument mouthpiece. In order to obtain the acoustic pressure outside the resonator, at the level of the bell, a spectrum transformation function is used, according to the works of Benade [12]. The resonator is considered in this case as a high pass filter, the filter’s envelope corresponding to the transfer function of

Table I. Values $vm1$ and $vm2$ of the control parameters for the simulations (virtual musician).

Definition	Notation	$vm1$	$vm2$
Resonance frequency of the lips	f_l (Hz)	400	412
Mass per area of the lips	μ_l (kg/m ²)	0.5	0.33
Pressure in the mouth	P_m (Pa)	16 000	
Width of the lips	b (mm)	10	
Rest value of the opening height	H_0 (mm)	0.5	
Quality factor of the resonance	Q_l	5	

the instrument. In practice, given that the instruments considered in the study are very similar, the same filter (cutoff frequency = 2000 Hz) is used for every instrument (computations show that differences between the transfer function of the different instruments are negligible). The amplitude of the harmonics of the external sound are computed by multiplying directly the magnitude of each harmonic A_n by the filter’s magnitude response. This generates the amplitudes a_n of the harmonics ($n = 1$ to N) characterizing the external sound of the trumpet.

2.2. Control parameters of the simulations

To simulate sounds, it is necessary to define relevant values (values that lead to a convergence towards a steady state sound) for the parameters of the virtual musician. These values are obtained from [13] and by a systematic exploration of the parameters’ space [6] that ensures convergence of the numerical model. One particular note, Bb4 (concert-pitch, $F_0 = 466.16$ Hz), located in the middle of the range of the trumpet, was chosen as a target sound for the experiments. This note corresponds to the fourth oscillation regime of the trumpet. Two virtual musicians $vm1$ and $vm2$ are considered in this study to simulate this note. The values of the control parameters of the simulations, that lead to a convergence toward a periodic solution, are given in Table I. We are aware that the values of the pressure in the mouth P_m are not realistic for a real musician (they are too high). But these values are necessary to create the oscillations with the chosen (minimal) physical model.

Given that the impedance measurement is limited in frequency (Figure 3), it is unnecessary to take into account many harmonics for the computation of the simulated sounds. Above 2500 Hz, the magnitude of the impedance is flat and no difference between trumpets is noticeable. For this reason, the number N of harmonics considered for the simulations was limited to not exceed 2500 Hz. For the note Bb4 (466.16 Hz), this corresponds roughly to $N = 6$ harmonics. All the sounds were then simulated with $N=6$ harmonics (examples of spectra of raw simulated sounds are given in Figure 5). In conclusion, for a given trumpet, characterized by its input impedance Z , the simulations give the playing frequency of the sound, the amplitudes and phases of its first six harmonics, according to the set of parameters describing the virtual musician (P_m , μ_l , f_l , b , H_0 , Q_l). **Raw simulated sounds**, corresponding to the

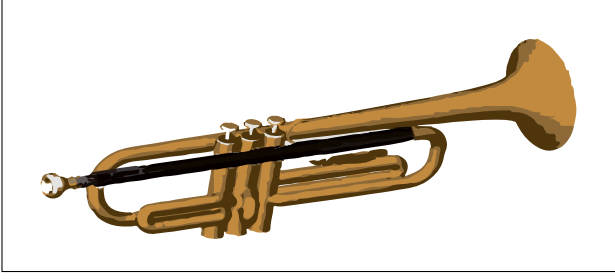


Figure 1. Location of the trumpet leadpipe (in black) in the whole trumpet.

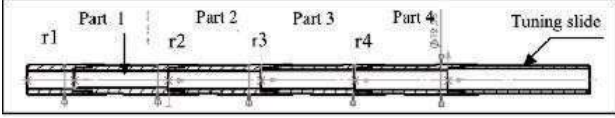


Figure 2. Dimensions of the parameterized leadpipe and definition of the four radii r_1 , r_2 , r_3 , r_4 .

Table II. Description of the dimensions (in mm) of the three leadpipes A, B, C of the study. N. B.: The radius of the last part r_5 is not a parameter, it is constant for all the leadpipes and corresponds to the cylindrical bore radius of the trumpet (*Bach* trumpet Medium Large (ML) bore, .459").

	Part 1	Part 2	Part 3	Part 4	
	r_1	r_2	r_3	r_4	r_5
A	4.64	4.64	4.64	4.64	5.825
B	4.64	5.45	5.50	6.00	5.825
C	4.64	5.00	5.50	5.70	5.825

pressure at the bell of the instrument, can then be synthesized by additive synthesis for a given duration.

2.3. Parameterized leadpipe

The different instruments considered in this study are obtained by small variations in the inner shape of the trumpet, called the bore. These variations concern a tube called the leadpipe, an important part of the bore that connects the mouthpiece to the tuning slide. This part, roughly conical, has a great influence on the intonation and the timbre of the instrument (Figure 1) [6]. A parameterized leadpipe was designed, divided into 4 parts of equal length ($l=55\text{mm}$), corresponding to a quarter of the length of a standard one (Figure 2). The parts were manufactured with various internal input and output radii. Given that we only consider continuous leadpipe profiles (the output radius of the $n-1$ part is equal to the input radius of the n part), the leadpipe is parameterized by 4 radii r_1 , r_2 , r_3 , r_4 .

Three leadpipes, named A, B, C, were considered for the study. Table II shows the radii values of these leadpipes.

Leadpipes B and C are very similar, with variations in the order of some tenths of mm. Leadpipe C corresponds roughly to the shape of the standard one provided with the trumpet. Leadpipe A presents significant variations in the

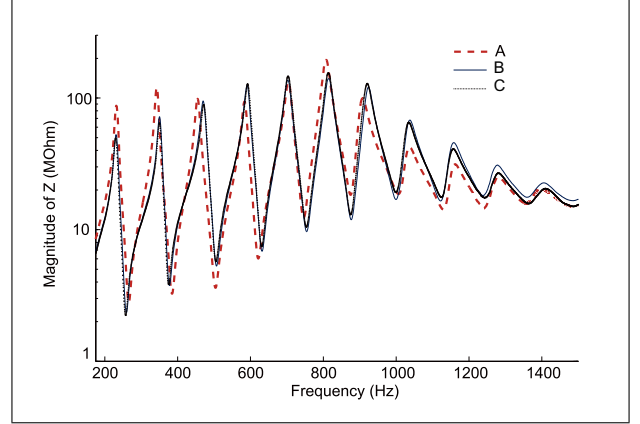


Figure 3. Input Impedance Z (magnitude) of the three trumpets A, B, C. The lowest resonance of Z (first regime), not played in the current trumpet tessiture, is not shown on this figure. The fourth peak (corresponding thus to the third maximum), is located around 460 Hz.

inner shape (in the order of a millimeter compared with the two others). Next, these leadpipes are assembled with the rest of the trumpet (a Bb trumpet, *Bach* model Vernon, bell 43). The trumpets are named A, B, C in the rest of the paper for simplicity. It is interesting to mention that trumpets B and C are very similar when played (trumpeters are generally unable to distinguish them in a blind test), while trumpet A is very different to the two others (players always recognize this leadpipe), due to intonation and timbral characteristics. Using the same mouthpiece (*Yamaha* 15B4), the input impedances Z of these three trumpets were measured with an impedance sensor developed by Dalmont and Le Roux [14] in an anechoic chamber. This measurement is the input of the sound simulations described in the previous section. To visualize the importance of the differences between the three trumpets, the magnitude of Z is presented in Figure 3. One notices that the differences between the trumpets concerning impedance are in agreement with the difference concerning the geometry: B and C are very close, while A is distinct from B and C.

3. Material and methods

3.1. Description of the perceptual tests

A same-different test [11] was implemented to study instrument discrimination and dynamics discrimination. The principle of this paired test is straightforward: in a single trial, a single pair of same or different sounds is presented; the participant is invited to indicate whether the two sounds are the same or different. The participant could play a pair several times and had to listen to it at least once. Headphones (identical for every participant, Beyerdynamic DT 990) and a computer interface were used to play the pairs. The participant was free to adjust the audio volume of the stimuli. The graphical interface for the test is given in Figure 4 for information.

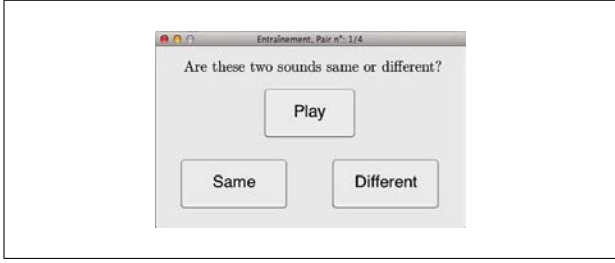


Figure 4. Description of the interface for the same-different test.

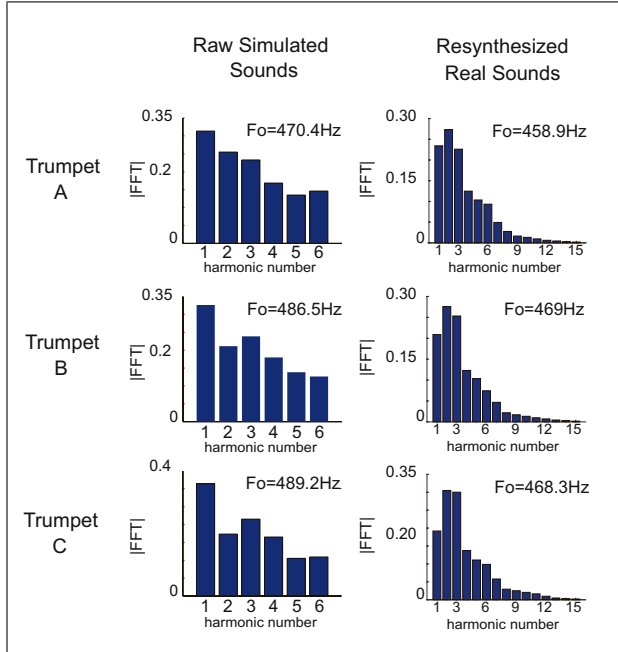


Figure 5. Spectra of the sounds listened to by the participants with 6 (resp. 15) harmonics for the raw simulated sounds (resp. resynthesized real sounds). The fundamental frequency of the sound is written in top right corner of each plot.

The pairs of sounds were constituted of two either simulated or real sounds. Simulated and real sounds are never associated in the same pair, given that they are very different in timbre (simulated sounds have only 6 harmonics – played sounds have 15 harmonics, see Figure 5). A detailed description of the pairs presented to the participants is given in the next section. 24 participants (9 women and 15 men – ages spanned from 20 to 23 years old) took part in the hearing experiment. All participants, engineering students, had normal hearing and a regular practice of a musical instrument. For instrument discrimination, two playing instances of the same instrument were included in the test in order to increase the diversity of sounds. Concerning simulated sounds, these two instances correspond to the two virtual musicians *vm1* and *vm2*. Regarding real sounds, they correspond to two distinct recordings of the instrument. The panel of subjects were divided in 2 groups of 12 participants. For simulated sounds, the first group (respectively the second group) listened to sounds generated by the virtual musician *vm1* (respectively *vm2*) and for processed real sounds, the first group (resp. the sec-

ond group) listened to real sounds played by a musician in a first recording (resp. in a second recording). Given that the results of the tests for the two groups were similar, in the end, all the data for the two groups were gathered.

3.2. Definition of the sound stimuli

For every experiment, the note generated to constitute the stimuli was a Bb4 (concert pitch) corresponding to the fourth regime of the trumpet, with no valve pressed. The tuning slide of the trumpet was adjusted to the same position (pulled 1cm) for all the measurements (impedance measurements – recording of the notes). The three trumpets A, B, C were considered for instrument discrimination, and only trumpet C was considered for dynamics discrimination. Every sound of the study were generated by additive synthesis (described thoroughly in this section) and their duration was the same (0.75 s). A 50ms linear envelope slope was added at the beginning and end of each sound. Every sound was normalized in amplitude so that the peak amplitude of the pressure remains in the interval $(-1, +1)$. Eventually, pairs of sounds were constituted of two sounds, separated by a 250 ms space.

3.2.1. Instrument discrimination

The objective of this first experiment is to study whether the differences between the sounds of the three trumpets A, B, C are perceptible, the sounds being either played by a musician or simulated with the harmonic balance technique. Since there are 3 instruments, 3 comparisons are considered for the same-different test (A versus B, B versus C and A versus C). For each comparison, pairs of sounds coming from different categories are considered in the test. These categories are described thereafter.

Simulated sounds: This category corresponds to sounds obtained by simulations of the physical model. For each trumpet, and each virtual musician *vm1* and *vm2* (see Table I), the playing frequency F_0 , the amplituden a_n , and phases φ_n of the six first harmonics representing the pressure at the bell of the instrument were computed. The sounds were next generated by additive synthesis (duration = 0.75s). This set is called **raw simulated sounds**. The spectra of the trumpets for the virtual musician *vm1* are plotted in Figure 5 (left column) for information.

Resynthesized Real sounds: This category corresponds to sounds played by a musician. For the 3 trumpets, the same note Bb4 was recorded with the same musician. All the recordings (sampling frequency 44100Hz, 16 bits) were made in the same room with a Shure SM 58 microphone. The microphone was placed in the axis of the bell (distance = 10 cm) and connected to a pre-amplifier Canford 20-266 and a Digigram Vx pocket V2 soundcard. The musician was asked to play a mezzo forte (*mf*) dynamics. Two repetitions of the note were recorded. In order to limit the variability inherent to the musician as much as possible, he was asked to play the note in the easiest and most natural way (without trying

to adjust the pitch or the timbre of the notes). However, these real sounds were too easily distinguishable in a same-different test because of the inherent impossibility of the trumpeter to produce sounds with a constant envelope, pitch, intensity and spectral content. For this reason, these sounds were resynthesized in permanent regime. The sounds were windowed to suppress the transient part of the signal. With the remaining part (considered as permanent regime), a Fourier analysis (FFT algorithm) was made to extract the playing frequency F_0 and the magnitude of the harmonics. In practice, only the first 15 harmonics were considered, the amplitudes of the harmonics above the 15th being negligible. In the end, the sounds were resynthesized by additive synthesis, using the playing frequency F_0 and the amplitude of the 15 harmonics (duration = 0.75 s). Two different recordings were made in order to constitute two different instances of the real trumpet sound providing different sounds to each group of participants. The spectra of the 3 trumpets corresponding to the first repetition are presented in Figure 5 (right column) for information. This set is called **resynthesized real sounds**.

Frequency normalization: The simulated and recorded sounds of the different trumpets present important differences in terms of playing frequencies. It is however interesting to study how the discrimination between the instruments changes when these playing frequency differences are removed. These frequency differences are even meaningless for instrument discrimination because one can tune the instrument to get the same pitch (by adjusting the tuning slide). An objective frequency normalization of the sounds is done to overcome this tuning problem and to assess how instruments are discriminated according to timbral characteristics only. Of course this tuning could be done by the musician, but this process is subjective, and furthermore solves only the problem for the “real sounds” category, not for the simulations. A target frequency was preferred to normalize the sounds in an objective way. 480 Hz was the target frequency (corresponding roughly to the average playing frequency of every Bb4 sound). All the sounds (simulated – resynthesized real sounds) were re-generated by additive synthesis (sum of the contribution of the harmonics) at a fundamental frequency of 480 Hz (duration = 0.75 s). Consequently, two more sets of sounds with normalized playing frequencies (480 Hz) were added to the two initial sound categories (simulated and resynthesized real sounds). These two new sets are called respectively **normalized simulated sounds** and **normalized resynthesized real sounds**. To sum up, the 2 initial sets of sounds are doubled producing 4 sets which are used to assess instrument discrimination).

3.2.2. Dynamics discrimination

The objective of this second experiment is to study whether the differences between the dynamics of the same trumpet are perceptible, the sounds either being played by

a musician or simulated. It is important to mention that the effect of the dynamics concerns only the timbre and playing frequency of the notes (sound amplitude is normalized in order to avoid the discrimination linked to sound loudness). According to the dynamics indication used in music, seven levels were considered to define the stimuli. Only one instrument is considered in this study: trumpet C (standard leadpipe).

Simulated sounds: Previous studies showed that the main control parameter that governs the dynamics of a simulated sound is the mouth pressure P_m (the physical model behaves actually in agreement with a real player) [6]. Seven increasing levels of the mouth pressure P_m were chosen to represent the evolution of simulated sounds dynamics (increasing from d1, $P_m = 8000$ Pa, to d7, $P_m = 20000$ Pa by a step of 2000 Pa). The lowest level corresponds to the minimum pressure necessary to obtain convergence of the simulation (regime of oscillation), the highest level corresponds to the value of the pressure from which there is no spectral enrichment anymore. The other values of the parameters of the virtual musician are chosen in order to obtain convergence of the simulations ($f_l = 400$ Hz, $\mu_l = +0,66$ kg/m²). These sounds are normalized in amplitude to prevent sound intensity difference (see below the normalization process). This set is named *Simulated sound dynamics*.

Resynthesized Real sounds: The same resynthesis described above over 7 gradual dynamics (spanning from pianissimo *pp* to fortississimo *fff*) played by the same musician formed a last set called **Resynthesized Real sound dynamics**.

It is important to mention that no frequency normalization was undertaken for dynamics discrimination because these sounds relate to the same instrument (C). Indeed, frequency differences are meaningful to represent dynamics differences (in other words, it would be totally irrelevant to change the tuning of an instrument according to the dynamic played).

3.3. Experimental design

This section describes the design of the pairs for the same-different test. Regarding instrument discrimination, for each considered set, every pair of instruments is presented as well as the pairs made of the same instrument, which produces 6 pairs (A/B, A/C, B/C, A/A, B/B, C/C). Concerning dynamics discrimination, it is useless to compare two dynamics that are too far from each other as the differences are easily audible (pilot tests showed a 100% success rate in the discrimination). Consequently, the dynamics discrimination is studied within three ranges $R1=\{d1, d2, d3\}$ (*low*), $R2=\{d3, d4, d5\}$ (*middle*), $R3=\{d5, d6, d7\}$ (*high*), producing pairs as follows: d1/d2, d2/d2, d2/d3. This process produces 9 pairs for each concerned set (Figure 6). For the instrument discrimination, sets of pairs were heard twice by the participant in order to have an insight into participant answer consistency. Dynamics discrimination pairs

		Instruments						dynamics								
		A		B		C		instr.	C	d1	d2	d3	d4	d5	d6	d7
		vm1	vm2	vm1	vm2	vm1	vm2									
		vm1	vm2	vm1	vm2	vm1	vm2									
A	vm1	XX		XX		XX		d1	X							
	vm2		XX		XX		XX	d2	X	X						
B	vm1			XX		XX		d3			X					
	vm2				XX		XX	d4			X	X				
C	vm1					XX		d5					X			
	vm2						XX	d6						X	X	
	vm1							d7								
	vm2															

Figure 6. Definition of the pairs presented in the experimental design for the listening test (simulated sounds). On the left, pairs presented to each participant regarding instrument discrimination involving two virtual musicians, *vm1* and *vm2*. Pairs were repeated for this test (xx stands for a pair presented twice). On the right, pairs of sounds regarding dynamics discrimination (x stands for a pair presented). 7 dynamics levels were proposed, from d1 (low dynamics) to d7 (high dynamics).

were not repeated so as to shorten the test duration as much as possible in order to keep participant concentration in this difficult test. Figure 6 summarizes the pairs proposed to the participants regarding instrument discrimination (left) and dynamics discrimination (right) in the case of simulated sounds. Identical tables can be drawn for resynthesized real sounds. Since the dynamics discrimination does not concern normalized sounds, only the table on the left can be reproduced for normalized synthesized sounds and normalized resynthesized real sounds. Each participant listened to a total of 66 pairs (2(repetitions) \times (6(simu) + 6(real) + 6(normalized simu) + 6(normalized real)) + 1(repetition) \times (9(simu) + 9(real))). Before the test, a short training introduced the task. The pairs were presented in a random order to each participant to control a possible order effect in the assessment. Furthermore, the sound order in the pair was also randomized to prevent an order effect. The participant listened to a pair and provided his/her answer (“same” or “different”) interacting with the graphical interface (see Figure 4). The average duration of the test was about 20 minutes.

3.4. Data analysis

The answers gathered by the same-different tests are analyzed using the confusion matrices [9]. These 2×2 matrices summarize the choices of the participants according to the nature of the pairs presented (see Figure 7). The columns represent the condition (“are the sounds of the presented pair different or the same?”), and the lines represent the participants’ answers. The numbers reported in the different cells comprise individual judgments from participants, including repeated trials for the instrument discrimination conditions. There are as many matrices as experimental contexts, namely 3 instrument discriminations (A/B, A/C, B/C) \times 4 sets (simulated sounds, normalized simulated sounds, resynthesized real sounds, normalized resynthesized sounds) plus 3 dynamics ranges (R1, R2, R3) \times 2 sets (simulated sounds, resynthesized real sounds), giving a total of 18 confusion matrices. Figure 7

		Condition	
		different (B/C)	same (B/B-C/C)
participants answers	Different	true positive (TP) 43	false alarm (FP) 19
		omission (FN) 3	true negative (TN) 73
	Same	Sensitivity $\frac{TP}{FN+TP} = 0,93$	Specificity $\frac{TN}{FP+TN} = 0,79$

Figure 7. Example of a confusion matrix of the same-different test regarding instrument B/C comparison using normalized simulated sounds. Furthermore, sensitivity and specificity are defined.

presents an example of the confusion matrix for the instrument B/C comparison in the case of normalized simulated sounds. From these matrices two main indices assess the quality of the discrimination: sensitivity and specificity [15]. The first one provides a fraction describing the subjects’ ability to differentiate sounds and the second one a fraction describing the subjects’ ability to detect similar pairs.

4. Results and Discussion

4.1. Discrimination between instruments

The results of the confusion matrices for each experimental condition are represented with a common plot in signal detection theory, the ROC plot (Receiver Operating Characteristic)[16]. This graphical plot represents the sensitivity (true positive rate) according to (1-specificity), i.e. the false positive rate. Each confusion matrix is represented by a point in the ROC space. A perfect detection is represented by the (0, 1) point, and a random detection is located along the diagonal line. The relevance of the discrimination can be assessed on this plot by the distance of the point above the diagonal: the greater this distance, the more relevant the discrimination. To qualify the performance in the detection relatively to random choices in the same-different test, an independence test using the Binomial Distribution is proposed [9]. From the Binomial Distribution, with a specific number of trials, and an observed number of correct identifications, a p-value can be calculated. This value represents the false rejection probability of the null hypothesis H_0 : subjects answered randomly to the tests. A threshold of 5% is considered for the significance test, allowing the computation of threshold values for true positive (TP) and true negative (TN). These thresholds allow the plotting of a 95% confidence limits on the ROC space, for sensitivity (lower bound) and specificity (upper bound). These limits are a first indicator helping us to estimate the subjects’ answers reliability.

Figure 8 and Figure 9 summarize the results of the same-different tests and show the positions of the pairs of sounds in the ROC space. The “raw” sounds (simulated sounds and resynthesized real sounds) are represented in Figure 8, while the normalized sounds (normalized simulated sounds and normalized resynthesized real sounds) are represented in Figure 9.

In Figure 8, for raw simulated sounds, the sensitivity of the detection is the weakest for the pair B/C (0.32). Furthermore, this pair is below the sensitivity confidence limit emphasizing how the participants miss the difference between the 2 instruments. We can conclude that the difference between the instrument B and C is almost inaudible, while differences for the two other pairs A/B and A/C are more audible (sensitivity > 0.8) and of the same magnitude. Concerning resynthesized real sounds, even if the pair B/C scores higher than its simulated counterpart, its discrimination is not as good as the one of the two other pairs. These observations are in agreement with the amplitude of the differences between the instruments. Indeed, the instruments B and C have a very similar geometry while A is more different to B and C. When the difference between instruments is important enough (A/B or A/C), the physical model is then able to generate perceptible differences between the instruments’ sounds.

In Figure 9, concerning normalized simulated sounds, the sensitivity is very low: each couple (A/B, A/C, B/C) is under the 95% confidence limit and has a sensitivity inferior to 0.5. When pitch differences are removed, the timbre differences between simulations are too small to be audible. Yet the detection for resynthesized real sounds is better (sensitivity > 0.8), every pair being above the two confidence limits and the pairs of instruments are not differentiated in the same way as raw sounds. The conclusion is that the physical model is not able to grasp timbre differences between instruments, while resynthesized real sounds are better discriminated. This could also be the result of the substantial difference between the spectra of the two sets of sounds. Indeed, simulated sounds contain six harmonics while real sounds contain fifteen harmonics leading to a greater timbre flexibility. This is the first limit of the physical model. The comparison of Figure 8 and Figure 9 indicates that the quality of the discrimination is always better when the sounds are not normalized in frequency. This result, in agreement with the expectations, is an indication on the validity of the test. Indeed, the difference between playing frequencies seems to be the main factor driving instrument discrimination, the timbre having a lesser importance in this experiment. To visualize the correlation between the sensitivity and the playing frequency differences for the different pairs of sounds, a scatter plot is shown in Figure 10: the larger the playing frequency difference, the more sensitive the instrument discrimination.

4.2. Discrimination between dynamics

Figure 11 shows the discrimination between dynamics in the ROC space. Simulated sounds are harder to discrimi-

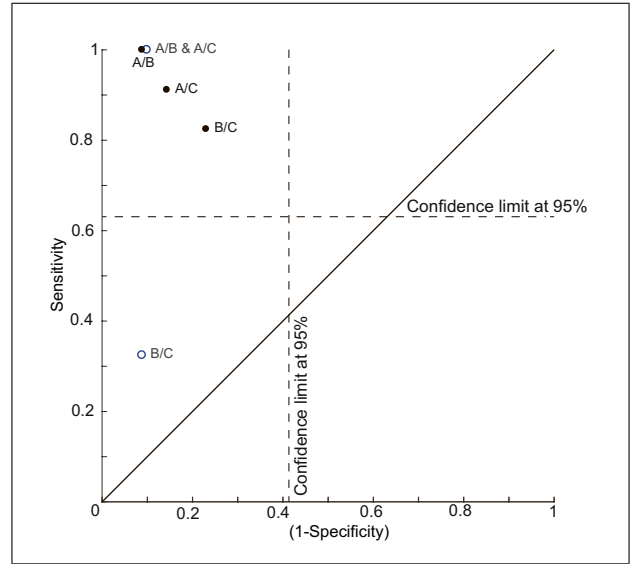


Figure 8. Instrument pair positions in the ROC space for simulated sounds (○) and resynthesized real sounds (●).

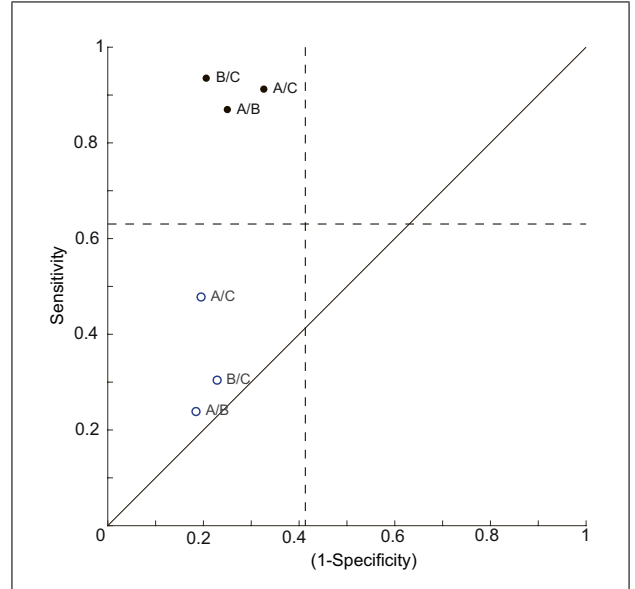


Figure 9. Instrument pair positions in the ROC space for normalized simulated sounds (○) and normalized resynthesized real sounds (●).

nate than resynthesized real sounds. Globally, only the R1 dynamics range of the resynthesized real sounds is above both confidence limits while all the other points are under the sensitivity confidence limit meaning the omission rates are too high. The position of the dynamics R1(d1d2d3) and R2(d3d4d5) on the diagonal line of the ROC space indicates that the detection is random for the simulation. Another general conclusion is the difficulty to perceive differences between mezzo piano, mezzo forte (R2={d3, d4, d5}) sounds for both sound categories. Regarding the low simulated sounds performance, a possible explanation is that the physical model used to simulate sounds does not enrich the sound spectrum enough while increasing dy-

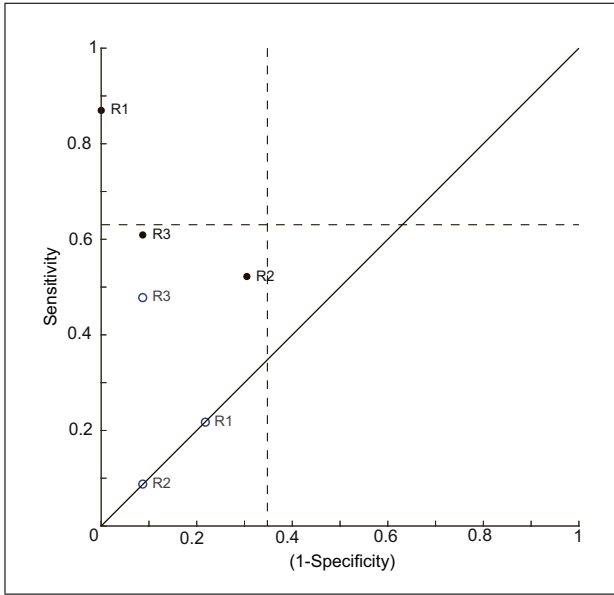


Figure 10. Scatter plot of sensitivity according to playing frequency differences for the different pairs of sounds (simulated sounds (○) and resynthesized real sounds (●)).

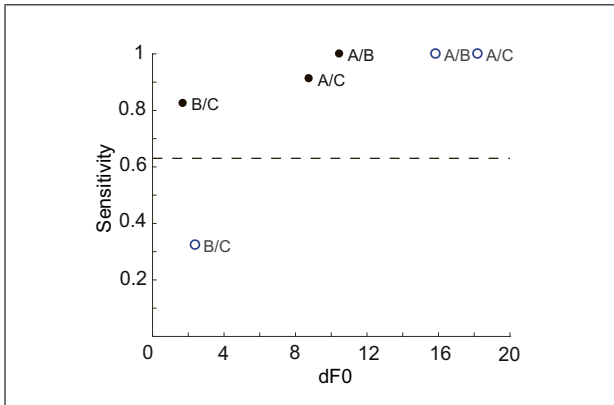


Figure 11. Position of triplets of dynamics (R1 low dynamics, R3 high dynamics) in the ROC space for the simulated sounds (○) and the resynthesized real sounds (●).

namics because it does not take into account the preponderant non-linear effect of high dynamics [17]. According to the confidence limits, the range that is the most discriminated is the pianissimo to mezzo piano resynthesized real sounds R1 with a sensitivity better than any other pair (sensitivity = 0.88). One explanation could be that the spectral enrichment of the trumpet is very important in the low dynamics for the 7 to 15 first harmonics since these harmonics are not shared between simulated and resynthesized real sounds.

5. Conclusions

The perceptual same-different test undertaken over simulated and resynthesized real sounds in permanent regime confirms the capability of the harmonic balance technique to represent coherent differences between instru-

ments. Regarding raw sounds, the results show that the simulations are able to generate reliable differences between trumpets when the differences between the instruments are sufficient (A/B–A/C). This is in accordance with the results with the “real” musician. When the differences are small (pair B/C), the differences between simulated sounds are not audible whereas they are with a “real” musician. This can be due to the musician, but additional tests should be carried out to confirm it (or an artificial mouth should be used [18]). Yet, the normalized sounds analysis shows that the physical model synthesizes too small timbre differences given the leadpipe modifications while the resynthesized real normalized sounds are more discriminable. This shows one limit of the model regarding the very small leadpipe differences. Besides, since the spectra of the resynthesized real sounds contain fifteen harmonics against six for the simulated sounds this may lead to greater timbre differences for the resynthesized real sounds. Concerning the dynamics, the mezzo piano to mezzo forte range is hard to discriminate, particularly for simulated sounds. The protocol selecting recorded sounds for dynamics discrimination should be modified classifying a priori sounds according to some descriptors related to loudness. The discrimination tasks are difficult for the subjects and the leadpipe modifications were too small to be perceptible in general. Consequently, one primary approach would be to refine our geometrical variables in order to objectively produce sharper sound differences. Furthermore, truncating the signal taking only the permanent regime is simplistic and temporal simulation will be studied in the future to provide more perceptually representative simulations. It is worth mentioning that deeper longer tests with many repetitions, assessing carefully the musician effect should be undertaken in order to seize more information in a near future.

Acknowledgments

Acknowledgments to the 24 subjects who took part in the perceptual tests with great care and patience.

References

- [1] M. Campbell: Brass instruments as we know them today. *Acta Acustica united with Acustica* **90** (2004) 600–610.
- [2] S. Carral, V. Chatziioannou: Single vs double reed conical woodwind sounds: Where does the difference lie? *Proceedings Forum Acusticum, Aalborg, Denmark, 2011*, 551–556.
- [3] J. Woodhouse, E. K. Y. Manuel, L. A. Smith, A. J. C. Wheble, C. Fritz: Perceptual thresholds for acoustical guitar models. *Acta Acustica united with Acustica* **98** (2012) 475–486.
- [4] S. Adachi, M. Sato: Trumpet sound simulation using a two-dimensional lip vibration model. *Journal of the Acoustical Society of America* **99** (1996) 1200–1209.
- [5] V. Välimäki, J. Pakarinen, C. Erku, M. Karjalainen: Discrete-time modelling of musical instruments. *Reports on progress in physics* **69** (2006) 1.

- [6] J. F. Petiot, J. Gilbert: Comparison of trumpets' sounds played by a musician or simulated by physical modelling. *Acta Acustica united with Acustica* **99** (2013) 629–641.
- [7] R. Tournemenne, J. F. Petiot, J. Gilbert: Study of perceived differences between simulated and real trumpets sounds. *Proceedings ISMA, Le Mans, France, 2014*.
- [8] J. Gilbert, J. Kergomard, E. Ngoya: Calculation of the steady-state oscillations of a clarinet using the harmonic balance technique. *Journal of the Acoustical Society of America* **86** (1989) 35–41.
- [9] J. Boley, M. Lester: Statistical analysis of abx results using signal detection theory. *Audio Engineering Society Convention* **127** (2009).
- [10] J. Bi: *Discrimination tests and measurements*. Blackwell Publishing, Ames, IA, 2006.
- [11] H. T. Lawless: *Quantitative sensory analysis. Psychophysics, models and intelligent design*. John Wiley & Sons, 2013.
- [12] A. H. Benade: Relation of air-column resonances to sound spectra produced by wind instruments. *Journal of the Acoustical Society of America* **40** (1966) 247–249.
- [13] J. S. Cullen, J. Gilbert, M. Campbell: Brass instruments: Linear stability analysis and experiments with an artificial mouth. *Acta Acustica united with Acustica* **86** (2000) 704–724.
- [14] J. P. Dalmont, J. C. Le Roux: A new impedance sensor for wind instruments. *Journal of the Acoustical Society of America* **125** (2008) 3014–3014.
- [15] D. G. Altman, J. M. Bland: Statistics notes: Diagnostic tests 1: sensitivity and specificity. *BMJ* **308** (1994) 1552.
- [16] M. Mahony, M. J. Hautus: The signal detection theory ROC curve: Some applications in food sensory science. *Journal of Sensory Studies* **23** (2008) 186–204.
- [17] L. Norman, J. P. Chick, D. M. Campbell, A. Myers, J. Gilbert: Player control of “brassiness” at intermediate dynamic levels in brass instruments. *Acta Acustica united with Acustica* **96** (2010) 614–621.
- [18] J. F. Petiot, F. Teissier, J. Gilbert, M. Campbell: Comparative analysis of brass wind instruments with an artificial mouth: first results. *Acta Acustica united with Acustica* **89** (2003) 974–979.