Study of perceived differences between simulated and real trumpets sounds

R. Tournemenne[a], J.F. Petiot[b] and J. Gilbert[c]
[a]IRCCYN Nantes, IRCCYN, 1 rue de la Noë, 44321 Nantes, France
[b]Ecole Centrale de Nantes, IRCCyN, 1 rue de la noe, 44321 Nantes, France
[c]LAUM - UMR CNRS 6613, Avenue Olivier Messiaen, 72085 Le Mans Cedex 9, France
robin.tournemenne@irccyn.ec-nantes.fr

This paper addresses the perception of differences between sounds of trumpets played by a musician or simulated by physical modeling. The harmonic balance technique is used to simulate trumpet sounds in permanent regime. The input parameter of the simulations is the input impedance of the trumpet (resonator), the control parameters are the characteristics of the virtual musician (excitator), and the outputs of the simulations are the playing frequency and the magnitude of the 6 first harmonics of the notes. Three different trumpets, obtained by small geodescriptoral variations of the leadpipe, are first simulated using several virtual musicians, and second played by a "real" musician. The objective of this paper is to define to which extent differences between sounds are noticeable by a panel of listeners. The factors of the experiment are the type of instrument used (type of trumpet), the playing dynamics of the sounds (pressure in the mouth for the simulated sounds), and the virtual musician (characterized by the control parameters of the simulations). For the two populations of sounds (simulated or real), a two alternatives forced choice hearing test was designed, with a panel of 26 participants. The analysis of the results of the tests with the signal detection theory allows the determination of the influence of the different factors on the noticeable differences between the sounds. The agreement between the results concerning the simulated sounds and the real sounds is assessed, to open the door to sound simulations for instrument design.

1 Introduction

Investigating physical models of musical instruments is an interesting mean to raise the knowledge about their functioning, which may enable us to provide a better assistance in their design. Yet, assistance in instrument design using physical modeling is still at its genesis [1] and further work is still necessary to enhance its reliability. The first phase of physical modeling consists of choosing a relevant physical model to represent the functioning of the instrument [2], [3]. Then, once a physical model simulating sounds is chosen, a second phase establishes in what extent the simulated sounds are in agreement with real sounds, as played by musicians [4]. In [5], the authors focused their work on how simulations by physical modeling could be used to predict certain characteristics of brass instruments sound. The present paper is in the continuation of [5] and focuses on perceptive aspects. In [5], the authors created parameterized trumpet leadpipes as shown in Figure 1, forming small differences between trumpets [6]. It is then possible to create simulated sounds using the model described in [5] and real natural sounds representing the same instrument. Audio descriptors based on the signal are defined to characterize the sounds, then the level of accordance between the simulated sounds and the "real" sounds is assessed. This concludes on the reliability of the physical model to represent differences between instruments (Cf. Figure 2).

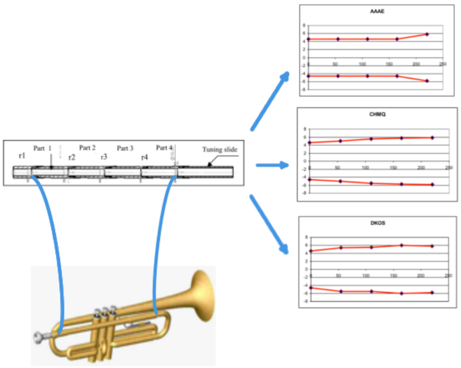


Figure 1: The trumpet leadpipe is replaced by four encastrable parts each of them being represented by a letter. The three trumpets considered in this study are labelled AAAE, CHMQ, DKOS.

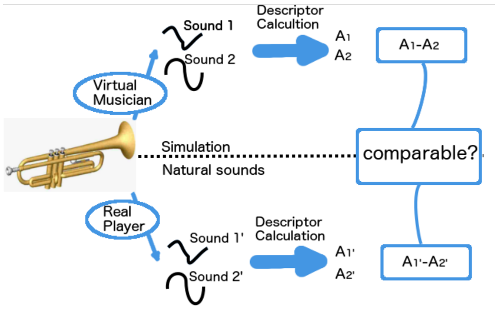


Figure 2: Model reliability estimation. Are the descriptor differences of natural and simulated sounds comparable?

The study in [5] showed that through two descriptors, which are the playing frequency and the spectral irregularity (eq. 1), the simulation and the natural sounds are in agreement:

- The trumpet AAAE is different from DKOS and CHMQ
- The trumpet CHMQ and DKOS are similar.

$$IRR = \sum_{k=2}^{N-1} \left| a_k - \frac{a_{k-1} + a_k + a_{k+1}}{3} \right|, \quad (1)$$

a_k being the amplitude of the k^{th} harmonic .

This fact was conform to the intuition given that the geometry of the instruments CHMQ and DKOS were very similar whereas AAAE was different. The study in [5] concluded that the harmonic balance technique was representative of the physics of brass instrument, and may be able to mimic some instrument behaviors. Yet, the differences observed over descriptors like the spectral irregularity and playing frequency were objectively measured, it is thus interesting to know if the human ear detects these differences. Therefore, the general aim of this paper is to know whether the differences between simulated and natural sounds are perceptible in order to strengthen the results of [5]. Perceptive tests were carried out to meet this global goal.

Section 2 provides more details about the simulation and its parameters. Section 3 introduces how the perceptive tests have been tackled and what kind of statistics evaluated the results. Section 4 describes the results of the perceptual tests.

Section 5 analyzes the link between the differences according to the objective descriptors and the perceived differences.

2 Background on the physical model of Brasses

The method, which produces simulation in the frequency domain, consists of the three equations 2, 3, 4 (more details can be found in [7]). These three equations are coupling the opening height $H(t)$ between the two lips, the volume flow $v(t)$ and the pressure in the mouthpiece $p(t)$. For one defined trumpet of impedance Z, the air density ρ and the width of the lips b can be considered as fixed, which reduces the number of parameters to three: P_m, the pressure in the mouth, f_l the resonance frequency of the lips, μ_l the mass per area of the lips (the formula of Q_l is fixed and depend on f_l and μ_l). Even though this basic model is rough, its advantages are the low number of parameters and their clear meaning.

$$P(j\omega) = Z(j\omega)V(j\omega), \tag{2}$$

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

$$\frac{d^2H(t)}{dt^2} + \frac{\omega_l}{Q_l}\frac{dH(t)}{dt} + \omega_l^2 H(t) = \frac{P_m - p(t)}{\mu_l}. \tag{4}$$

In order to find solutions in permanent regime, the harmonic balance technique has been implemented. It gives the playing frequency of the sound and the amplitudes of its 6 first harmonics, according to the set of parameters describing the virtual musician and the instrument. The compared sounds consider the note Bb4 played and simulated on three trumpets termed AAAE, DKOS, CHMQ (Cf. Figure 1). For natural sounds, the same musician played 20 times the note whereas simulations ran the harmonic balance algorithm over all combinations of the three parameters, described in the table 1.

Table 1: Ranges of the virtual musicians parameters

Definition	Range
$P_m(P_a)$	8000 to 20000 (step of 2000)
$f_l = \omega_l/2\pi(Hz)$	400 to 439 (step of 1Hz)
$\nu_l = 1/\mu_l(m^2/kg)$	-0.5 to -3 (step of -0.5)

P_m represents the sound dynamic and the 7 available levels are called d1(P_m = 8000) to d7 (P_m = 20000). This 7 levels are also used for the natural sound dynamic which spans from pianissimo to fortissimo.

3 Material and methods

3.1 Description of the perceptual tests

Given the objective of this study, 2 main interrogations emerge:

- are the differences between trumpets, and between different dynamics, noticeable?
- are the perceived differences in agreement with the objective differences determined in the previous study?

Consequently, a set of more precise questions emerges, firstly about the instrument differences, for simulated and natural sounds:

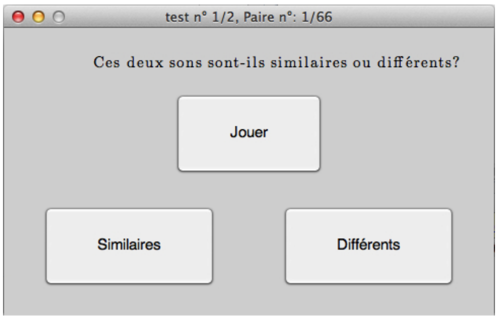
- are the differences between AAAE and DKOS noticeable?
- are the differences between AAAE and CHMQ noticeable?
- are the differences between CHMQ and DKOS noticeable?

and then about the dynamic perception:

- are the differences between d1 d2 d3 noticeable?
- are the differences between d3 d4 d5 noticeable?
- are the differences between d5 d6 d7 noticeable?

The sounds with the dynamic d1 (respectively d3) is not compared with the sounds having dynamics higher than d3 (respectively d5) because the difference was important and easily audible.

A two-alternative forced choice test (2 AFC test) was implemented in order to answer these questions. This perceptual test plays a pair of sounds that can be different or not and the listener has to indicate whether the two played sounds are different or not (Cf. Figure 3). Participants can play the pairs again and they have to listen to it at least once. 26 participants took part in a hearing test divided in 2 groups (13 participants in each of them). Only the effect of the instrument has to be considered, not random effects due to the musician or the session. That is why, in order to curb this undesirable effects, each group listened to a different set of sounds: as instance for simulated sounds, group 1 (respectively group 2) listened to sounds generated by a virtual musician 1 (respectively virtual musician 2). Finally, all the data are gathered over the two groups.


test n° 1/2, Paire n°: 1/66

Ces deux sons sont-ils similaires ou différents?

Figure 3: View of the interface displayed to the participants.

3.2 Experimental design

Every sound plays the Bb4 note lasting 1.5 second and an envelop slope has been implemented at the beginning and end of each sound in order to prevent too sharp movements of the headphone membrane. In order to answer the 6 questions, each pair is composed of two sounds taken in one of the six populations described below:

Raw Simulated Sounds played by two different virtual musicians following these characteristics:

- musician 1: $P_m = 16000, \quad f_l = 400, \quad v_l = -2$
- musician 2: $P_m = 16000, \quad f_l = 412, \quad v_l = -3$

Each group listened to the sounds made by one musician in order to only discern the possible instrument differences.

Normalized Simulated Sounds which have been normalized to the playing frequency of the well-tempered Bb4 (A4=440Hz).

Raw Natural Sounds which are resynthesized creating a signal with the prominent harmonics of the temporal signal played by the musician. Indeed, the natural sounds were too different for a same instrument taking different trial of the same musician certainly because of the difficulty of the player to always make exactly the same sound. 2 sounds from each instrument was chosen randomly in a set of 20 sounds played by the same musician. The 6 chosen sounds were separated in 2 sets and group 1 (respectively group 2) listened to the set 1 (respectively set 2). Indeed, similarly to simulated sounds, the aim of the set separation is to limit the effect that is induced by a specific trial (session effect).

Normalized Natural Sounds created from the same natural resynthesized sounds taking the well-tempered Bb4.

Simulated Sounds with different Dynamics taken from the instrument DKOS simulated with the 7 gradual dynamics P_m ($f_l = 400, \quad v_l = -2,5$).

Natural Sounds with different Dynamics taken from DKOS with the same musician played at 7 gradual dynamics.

For the 4 first populations the pair was listened twice by the participant in order to better assess the repeatability of the test. The tables 2 and 3 summarize every pair of simulated sounds the participants listened to (the table 3 has to be repeated once for the normalized sounds). Almost identical tables can be drawn for the natural sounds.

A part of the test was implemented with a normalized frequency in order to remove the effect of the playing frequency over the differentiation task. Every pair was mixed up randomly for each participant and the question asked to the participant was: these two sounds are similar or different? (Cf Figure 3) Finally, each participant listened to 66 pairs (2x12+1x9+2x12+1x9) and the test took about 15 minutes.

3.3 Data analysis

Once all the answers of every participant have been gathered over all pairs, the confusion matrix can be build. These 2X2 matrices represent the decision of the participants in comparison to the experimental condition (the truth). The column are the realness (are the sounds of the presented pair the same, or two different sounds), and the lines show the answer of the participant (Cf. Table 4).

Table 2: Pairs of simulated sounds that have been presented to each participant regarding the tests about dynamic variations. Each dynamic is represented by a number 1 for piano and 7 for fortissimo. Each x represents the pair that was listened.

DKOS	d1	d2	d3	d4	d5	d6	d7
d1		x					
d2	■	x	x				
d3	■	■		x			
d4	■	■	■	x	x		
d5	■	■	■	■		x	
d6	■	■	■	■	■	x	x
d7	■	■	■	■	■	■	

The confusion matrices were analyzed firstly thanks to the value of the Chi-square test of independence in order to first know if the participants answered randomly or not and in what extent [8]. If the assumption ”the participants answered randomly” can be rejected with a low risk, it still has to be estimated to which extent the sounds are different. The Specificity of the test gives this final conclusion. The Specificity is the ratio between the true negative occurrences (Cf. Table 4) and the sum of the true negative and false alarm occurrences. The higher the Specificity, the more the participants thought the two instruments were different (or three dynamics), and vice versa.

4 Results

Since the test, its proceeding and analyze methods have been introduced, the results of the Chi-square test of independence, conjugated with the Specificity value can be introduced. Technically, this Chi-square test compares the value of a statistic T (describing the independency between the participant answers and the truth) with the Chi-square law at 1 degree of freedom taken at a certain risk level (limit of 5% in this study).

4.1 Differences between instruments

The table 5 summarizes the results of the test evaluating the degree of similarity between the three instruments according to the ”raw” sounds while the table 6 summarizes the results over the sounds with a renormalized playing frequency.

Firstly the results are almost unequivocal for the raw sounds while ambiguous for the normalized sounds. This explanation may come from the fact that the playing frequency plays an important role in sound differentiation. Indeed, at 1kHz the human ear can approximately ear a difference of 2Hz. A tiny frequency difference will then be easily recognized by the participants. Besides, another interesting result is the fact that overall, the natural sounds

Table 3: Pairs of simulated sounds that have been presented to each participant regarding instrument differentiation tests. vm stands for virtual musician and each x represents the pair that was listened.

		AAAE		DKOS		CHMQ	
		vm1	vm2	vm1	vm2	vm1	vm2
AAAE	vm1	xx		xx		xx	
	vm2		xx		xx		xx
DKOS	vm1			xx		xx	
	vm2				xx		xx
CHMQ	vm1					xx	
	vm2						xx

Table 4: This confusion matrix describes the nature of the answers given by the participants while trying to discern if the Instrument CHMQ was similar to DKOS listening to the normalized natural sounds.

	is CHMQ similar to DKOS ?	Condition similar	 different
participants answers	yes, they are!	true positive 73	false alarm (FA) 3
	no, they are different!	omission 19	true negative (TN) 43
			Specifity $\frac{TN}{FA+TN}$ = 0,93

are better differentiated than the simulated ones. This comes from the fact that even if the sounds are resynthesized and come from the same player, a session effect creates timbre differences. For the raw simulated sounds, it was easier to discern AAAE from the two other instruments, while DKOS and CHMQ were hardly differentiable. For the normalized sound, every instrument seems similar, only AAAE and DKOS may have been discerned but the value of the Specificity shows that participants thought the two instruments may be similar. The same findings are drawn for the raw natural sounds, but for normalized sounds all instruments are well distinguished.

Table 5: Final answers to the statistical questions. The superior triangle of the table presents the results gathered by the comparisons on the natural sounds, while the lower triangle provide the results regarding the simulated sounds. The first number is the p-value, if it is inferior to 5%, the Specificity is then given.

is similar to	AAAE	DKOS	CHMQ
AAAE		p<0.0001 0.91	p<0.0001 1
DKOS	p<0.0001 1		p<0.0001 0.83
CHMQ	p<0.0001 1	0.0003 0.32	

Table 6: The answers regarding the normalized sounds are hereafter drawn following the same rules than the table 5.

is similar to	AAAE	DKOS	CHMQ
AAAE		p<0.0001 0.91	p<0.0001 0.87
DKOS	0.0005 0.48		p<0.0001 0.93
CHMQ	NS	NS	

4.2 Differences between dynamics

The table 7 shows the ability of participants to differentiate sounds with different dynamics. This model doesn't implement brassy sounds and non-linearity behaviors, it is thus legitimate that the participants couldn't differentiate these different dynamics. Even if they didn't answer randomly at high dynamics, Specificity is low meaning they thought the signals may be similar. For the natural sounds, participants recognize quite well the differences in low and high dynamic whereas it is hard to feel a timbre difference between the mezzopiano, mezzoforte dynamic range (d3, d4, d5).

4.3 Agreement Simulations/Real sounds

The only sounds population having the same interpretations for simulation and natural sounds is the population of raw sounds. The normalized simulated sounds are almost impossible to differentiate even if DKOS and AAAE may be recognized, while the normalized natural sounds are recognized. Finally, regarding the differences between dynamics, the results confirm the findings from the previous study: it is very hard to distinguish dynamics differences with simulated sounds. This is the reason why there are significant differences between the natural and the simulated sounds.

Table 7: dynamic influence on recognition. If results are significative, then the Specificity is given.

are similar?	d1, d2, d3	d3, d4, d5	d5, d6, d7
for simulated sounds	NS	NS	p<0.0001 0.47
for natural sounds	p=0.0002 0.87	NS	p<0.0001 0.61

5 Interpretation of perceptual differences with descriptors

Since the similarity issue between instruments/dynamics has been perceptually tackled, the correlation with the objective descriptors of the previous study [5] has to be estimated. The more the correlation, the more the descriptors may predict some instruments behaviors.

5.1 The playing frequency effect

The raw signals provide a more important differentiation between the instruments than the normalized one. It is then legitimate to think that the playing frequency plays an important role in the differentiation process. One can agree that the results of the study regarding the raw signals should follow the difference between the playing frequencies of these signals. If two signals have a close playing frequency, they should be then difficult to discern and vice-versa.

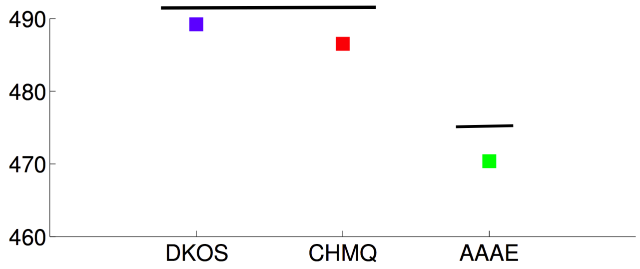


Figure 4: Average playing frequency of the simulated sounds made with the three trumpets played by the musician 1. The bars represent the group of instruments that weren't significantly different.

The Figure 4 shows that the simulated instrument AAAE played by the first virtual musician has a much lower playing frequency than the two other, which are comparable (the same results are observed with the second virtual musician). Consequently, the instrument AAAE should be distinguished from the two others while CHMQ and DKOS may be told as similar. The perceptual tests are completely in agreement with this observation strengthening the fact that the playing frequency plays a substantial role in the instrument differentiation task.

Almost the same graphic is found concerning natural sounds accentuating again the role of the playing frequency in a real context.

5.2 IRR and the normalized sounds

The IRR descriptor doesn't take into account the playing frequency, it measures local variation of the harmonics. Consequently, it may contains information about the timbre of the instrument and its value can then be compared with the results provided by the normalized pairs.

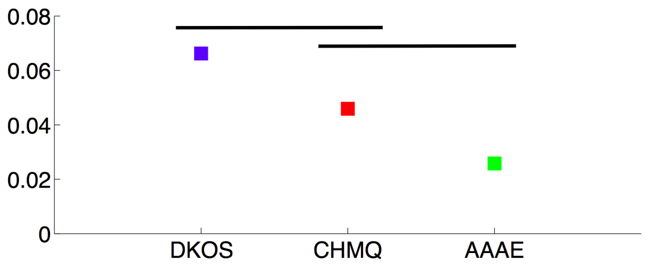


Figure 5: Average IRR of the simulated sounds made with the three trumpets played by the musician 1.

According to the objective values CHMQ and AAAE would have been distinguished in the same manner while DKOS and AAAE should be more easily differentiated. This result seems true for the Simulated sounds even if the Specificity value regarding the difference of AAAE and DKOS is low (0.48). For the natural sounds, every instrument is still well recognized, but the difference between AAAE and DKOS is not the highest which may signify either that the session effect over natural sounds is prominent or that the instrument effect doesn't follow the IRR value. Again, there are similar Figures than Figure 5 regarding the virtual musician 2 and the natural sounds.

6 Conclusions

The perceptual 2 AFC test undertaken over simulated and natural sounds composed of 6 populations enlightened the reliability of the harmonic balance technique and what are its limits. The interesting comparison between the objective descriptors and the participants feelings provides elements urging further studies to dive deeper in these descriptors to extract even more meaningful interpretations. Different simulated sounds are recognized thanks to the playing frequency and were difficult to recognize when normalized. Different natural sounds are recognized in both context: raw or normalized. For the normalized sounds it is impossible to say if the differentiation task was successful thanks to the instruments differences or a session effect. Consequently, further studies should focus on inter as well as intra instruments differentiation. In order to improve the meaning of the results slightly larger trumpet modifications should be taken into account. Besides, a look into time simulation [9] would be interesting to improve these steady-state simulations. Once trumpet modifications, model enhancements and new descriptors analyses are implemented, more extensive perceptual studies could be undertaken in order to validate improvements.

References

[1] Murray Campbell. Brass instruments as we know them today. *Acta Acustica united with Acustica*, 90(4):600–610, 2004.

[2] RL Pratt and JM Bowsher. The objective assessment of trombone quality. *Journal of Sound and Vibration*, 65(4):521–547, 1979.

[3] Vesa Välimäki, Jyri Pakarinen, Cumhur Erkut, and Matti Karjalainen. Discrete-time modelling of musical instruments. *Reports on progress in physics*, 69(1):1, 2006.

[4] Vasileios Chatziioannou and Sandra Carral. Single vs double reed conical woodwind sounds: Where does the difference lie? In *Forum Acusticum 2011*, pages 551–556. Aalborg University Denmark, 2011.

[5] Jean-Francois Petiot and Joel Gilbert. Comparison of trumpets' sounds played by a musician or simulated by physical modelling. *Acta Acustica united with Acustica*, 99(4):629–641, 2013.

[6] Emilie Poirson, Joël Gilbert, and Jean-François Petiot. Integration of user perceptions in the design process: application to musical instrument optimization. *Journal of Mechanical Design*, 129(12):1206–1214, 2007.

[7] Joël Gilbert, Jean Kergomard, and Edouard Ngoya. Calculation of the steady-state oscillations of a clarinet using the harmonic balance technique. *The Journal of the Acoustical Society of America*, 86(1):35–41, 1989.

[8] Sheldon Ross. A first course in probability, (2002).

[9] F. Silva, Ch. Vergez, Ph. Guillemain, J. Kergomard, and V. Debut. MoReeSC: A Framework for the Simulation and Analysis of Sound Production in Reed and Brass Instruments. *Acta Acustica united with Acustica*, 100(1):126–138, January 2014.