Giving Instruments a Voice: Are there vowel-like qualities in the timbres of musical instruments?

Christoph Reuter†[1], Charalampos Saitis[2], Isabella Czedik-Eysenberg[1], and Kai Siedenburg[3,4]

[1] *SInES, Musicological Department, University of Vienna, Austria, E-Mail: christoph.reuter@univie.ac.at*
[2] *Centre for Digital Music, Queen Mary University of London, UK, E-Mail: c.saitis@qmul.ac.uk*
[3] *Signal Processing and Speech Communication Laboratory, Graz University of Technology, Austria, E-Mail: kai.siedenburg@tugraz.at*
[4] *Dept. of Med. Physics and Acoustics, Carl von Ossietzky Universität Oldenburg, Germany*

Background

Based on the theory that the spectra of musical instrument sounds have similar formant ranges as vowels, vowels play a very prominent role in German-language timbre descriptions [1-10]. This is not the case in English-language literature, where formant ranges or vowels are rarely used to describe instrument timbres.

The typical vowel assignments to the individual orchestral instruments' timbres found in literature can be summarized in one table:

Table 1: Vowel similarities of musical instruments in the low, middle and high register according to [11][5][8][10].

	Low	Middle	High
Flute		A	
Oboe		A, Ä	
Clarinet			I
Bassoon		O	
Trumpet		A, Ä	
Trombone		Å	
French Horn		O	
Tuba		U	
Violin		O, A	
Viola	U	O	Å
Violoncello	U, O	Å	A
Double Bass	U, O	Å	A

Research Question

Although there seems to be some agreement in the literature about the vowel similarity of musical instrument sounds, this has never been tested empirically in listening experiments. Thus, we aim to provide empirical evidence whether vowel-like qualities can be significantly observed in the timbre of musical instruments as previously assumed in German-language timbre descriptions.

Methods

In a listening experiment 64 participants (20 ♂, 37 ♀, 7 ☿, age: 18-45, ø 22 years) were presented recordings of the most common Western orchestral instruments (flute, oboe, clarinet, bassoon, trumpet, trombone, French horn, tuba, violin, viola, cello, and double bass in three registers and two dynamic levels each). Their task was to assign German vowels and umlauts to the instruments' timbres and to rate the strength of the perceived vowel similarity on a scale from 1 to 100. The following method was used to measure the vowel similarity according to the aggregated participant data: e.g. the frequency of selections of the vowel "A" for the oboe sound (playing E4 in *ff*) was multiplied by the sum of all strength values set for this vowel. The same applies to the other vowels selected by the participants for this sound. Correlations between the vowel choices show, that the vowels "Ä" and "E" were relatively often assigned to the same sounds ($r = 0.779$). In all other cases, the correlations were much weaker. This indicates that most of the included vowels represent distinct qualities following different patterns of assignment.

Results

Fisher's exact tests were used to find out which vowels were chosen significantly more frequently for a specific stimulus than by chance by the listeners (considering Bonferroni correction). This test was conducted 1) for individual sounds (see Figure 1) as well as 2) for each group of stimuli representing an instrument as a whole (see Figure 2).

Instrument		Vowel	Instrument		Vowel	Instrument		Vowel
Flute	E4 *ff*	A	**Trumpet**	A3 *ff*	A	**Violine**	A3 *ff*	**E** (7,02E-06) Ä (1,62E-08)
	E4 *pp*	**U** (1,62E-12)		A3 *pp*	**A** (7,14E-07)		A3 *pp*	A
	E5 *ff*	U		A4 *ff*	Ä		A4 *ff*	E
	E5 *pp*	**Ü** (3,43E-05)		A4 *pp*	A		A4 *pp*	**E** (2,19E-05)
	E6 *ff*	**I** (2,80E-08) Ü (4,69E-09)		A5 *ff*	**I** (6,93E-05)		A5 *ff*	**I** (4,43E-27)
	E6 *pp*	**Ü** (9,60E-06), I (2,22E-12)		A5 *pp*	I		A5 *pp*	**I** (7,73E-22)
Oboe	E4 *ff*	**A** (5,08E-08)	**Trombone**	A2 *ff*	A	**Viola**	E3 *ff*	E
	E4 *pp*	**A** (4,07E-06)		A2 *pp*	**O** (3,24E-05)		E3 *pp*	E
	E5 *ff*	A		A3 *ff*	A		E4 *ff*	E
	E5 *pp*	A		A3 *pp*	**U** (5,30E-07)		E4 *pp*	E
	E6 *ff*	**I** (2,67E-19)		A4 *ff*	Ä		E5 *ff*	**I** (4,63E-07)
	E6 *pp*	**I** (8,15E-26)		A4 *pp*	U		E5 *pp*	**I** (5,62E-12)
Clarinet	E3 *ff*	Ö	**Fr.Horn**	E2 *ff*	Ö	**Cello**	E2 *ff*	Ö
	E3 *pp*	U		E2 *pp*	**O** (4,27E-07)		E2 *pp*	A
	E4 *ff*	U		E3 *ff*	Ä		E3 *ff*	Ä
	E4 *pp*	**U** (1,92E-16)		E3 *pp*	**O** (6,50E-08)		E3 *pp*	A
	E5 *ff*	**I** (5,43E-05)		E4 *ff*	Ä		E4 *ff*	E
	E5 pp	**U** (3,77E-07)		E4 *pp*	**O** (1,76E-06)		E4 *pp*	**Ä** (2,54E-05)
Bassoon	E2 *ff*	O	**Tuba**	A1 *ff*	O	**DbBass**	E1 *ff*	Ö
	E2 *pp*	**O** (4,27E-07)		A1 *pp*	**O** (9,05E-13)		E1 *pp*	O
	E3 *ff*	**O** (2,46E-06)		A2 *ff*	O		E2 *ff*	O
	E3 *pp*	**O** (4,91E-16)		A2 *pp*	**O** (6,06E-12)		E2 *pp*	Ö
	E4 *ff*	O		A3 *ff*	**O** (4,27E-07)		E3 *ff*	Ä
	E4 *pp*	**O** (2,07E-08)		A3 *pp*	**U** (3,11E-28)		E3 *pp*	Å

Figure 1: Vowels (marked in color) chosen significantly more often than random for individual instrumental sounds (p values in brackets, Fisher's exact test for the individual stimuli). Find an interactive map with sound examples at https://muwiserver.univie.ac.at/vowellike/v.htm

Instrument	U	O	Å	A	Ä	Ö	Ü	E	I
Flute	0,0000	1,0000	0,9996	0,8137	1,0000	0,9136	0,0000	1,0000	0,0000
Oboe	0,9950	1,0000	0,9567	0,0000	0,9752	0,9981	0,3397	0,9817	0,0000
Clarinet	0,0000	0,9997	0,8875	1,0000	0,9977	0,1759	0,0002	0,9846	0,8727
Bassoon	0,9983	0,0000	0,0215	0,9820	0,9898	0,0812	1,0000	0,9989	1,0000
Trumpet	1,0000	1,0000	0,9409	0,0000	0,0021	0,9766	0,5800	0,1695	0,0863
Trombone	0,2683	0,0737	0,0000	0,0243	0,6580	0,6045	0,9903	0,9865	1,0000
Fr. Horn	0,6009	0,0000	0,1176	0,9972	0,5346	0,0109	0,9990	0,7093	1,0000
Tuba	0,0000	0,0000	0,0042	0,9992	1,0000	0,9379	0,9917	1,0000	1,0000
Violine	1,0000	1,0000	0,9972	0,9845	0,0004	0,9993	0,9973	0,0000	0,0000
Viola	1,0000	1,0000	0,9999	0,9915	0,0015	0,3071	0,3805	0,0000	0,0000
Violoncello	1,0000	1,0000	0,2956	0,0243	0,0000	0,2443	0,9903	0,0000	0,9988
Db.Bass	0,6835	0,1734	0,0352	0,7971	0,2901	0,0003	0,2840	0,7337	1,0000

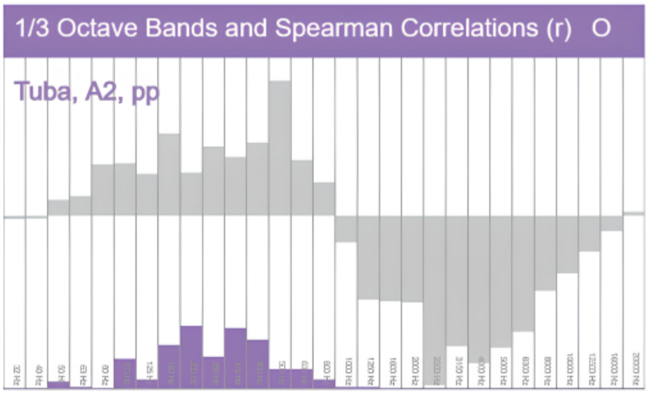
Figure 2: Vowels (marked in color) chosen significantly more often than random for instrumental timbres. The displayed numbers are the p values resulting from a Fisher's exact test for each instrument timbre as a whole.

A correlation analysis was carried out for the perceived vowel similarity with regard to the spectral energy in the 1/3rd octave bands of the corresponding musical instrument timbre (see Figure 3). It turned out that the strongest positive correlations were found at the respective vowel formant positions. This indicates that the more an instrument sound is associated with a certain vocal quality, the more it tends to have higher levels in the corresponding formant ranges.

A comparison of the measured vowel qualities with the vowel attributions found in literature (cf., Figure 1) reveals pitch- and dynamic-dependent similarities mostly for the woodwind and brass instruments, while there are almost no matching vowel attributions for the string instruments. This can be seen in individual pitches and dynamic levels as well as in instruments as a whole (see figure 4).

Register	measurement low		measurement middle		measurement high		literature low	literature middle	literature high
Dynamics	*pp*	*ff*	*pp*	*ff*	*pp*	*ff*			
Flute	U	A	Ü	U	I	I		A	
Oboe	A	A	A	A	I	I		A / Ä	
Clarinet	U	Ö	U	U	U	I			I
Bassoon	O	O	O	O	O	O		O	
Trumpet	A	A	A	Ä	I	I		A / Ä	
Trombone	O	A	U	A	U	Ä		Å	
French Horn	O	Ö	O	Ä	O	Ä		O	
Tuba	O	O	O	O	U	O		U	
Violin	A	E	E	E	I	I		O / A	
Viola	E	E	E	E	I	I	U	O	Å
Cello	A	Ö	A	Ä	Ä	E	U / O	Å	A
Double Bass	O	Ö	Ö	O	A	Ä	U / O	Å	A

Figure 4: Comparison of the perceived vowel similarity (left) to the vowel similarity of musical instrument sounds described in the literature (right).



1/3rd oct.filter	U	O	Å	A	Ä	Ö	Ü	E	I
32 Hz	0,405	-0,016	-0,226	-0,224	-0,483	-0,233	0,468	-0,404	0,169
40 Hz	0,378	-0,014	-0,223	-0,223	-0,459	-0,198	0,474	-0,406	0,144
50 Hz	0,465	0,075	-0,165	-0,245	-0,566	-0,233	0,388	-0,513	0,066
63 Hz	0,464	0,093	-0,139	-0,195	-0,435	-0,115	0,449	-0,461	0,025
80 Hz	0,29	0,243	0,044	-0,173	-0,359	0,067	0,26	-0,378	-0,168
100 Hz	0,368	0,248	0,156	-0,084	-0,396	0,03	0,285	-0,474	-0,226
125 Hz	0,427	0,198	0,032	-0,185	-0,446	-0,078	0,416	-0,538	-0,086
160 Hz	0,169	0,391	0,215	-0,109	-0,208	0,36	0,011	-0,211	-0,272
200 Hz	0,106	0,204	0,222	0,149	-0,224	-0,054	0,007	-0,297	-0,255
250 Hz	0,082	0,33	0,235	-0,004	-0,241	0,057	-0,062	-0,357	-0,359
315 Hz	0,086	0,278	0,106	-0,161	0,018	0,301	-0,067	0,055	-0,241
400 Hz	-0,01	0,348	0,397	0,098	-0,039	0,082	-0,323	-0,092	-0,413
500 Hz	-0,061	0,64	0,565	-0,065	-0,06	0,405	-0,499	-0,112	-0,604
630 Hz	0,131	0,263	0,112	0,101	-0,018	0,044	-0,2	0,039	-0,13
800 Hz	-0,194	0,158	0,323	0,177	0,031	0,064	-0,287	0,026	-0,179
1000 Hz	-0,296	-0,128	0,17	0,354	0,31	0,058	-0,243	0,252	-0,065
1250 Hz	-0,213	-0,401	-0,173	0,354	0,262	-0,03	0,098	0,254	0,376
1600 Hz	-0,399	-0,409	-0,049	0,38	0,571	0,109	-0,087	0,506	0,154
2000 Hz	-0,277	-0,415	-0,3	0,201	0,628	0,196	-0,056	0,707	0,334
2500 Hz	-0,546	-0,807	-0,613	0,068	0,56	-0,223	0,162	0,596	0,662
3150 Hz	-0,429	-0,621	-0,426	0,138	0,623	-0,095	0,031	0,702	0,482
4000 Hz	-0,369	-0,701	-0,631	-0,126	0,447	-0,272	0,167	0,531	0,628
5000 Hz	-0,219	-0,628	-0,639	-0,181	0,337	-0,232	0,281	0,409	0,583
6300 Hz	0,012	-0,553	-0,586	-0,297	0,134	-0,321	0,379	0,212	0,558
8000 Hz	0,157	-0,357	-0,513	-0,31	-0,127	-0,35	0,455	-0,052	0,484
10000 Hz	0,209	-0,275	-0,434	-0,302	-0,249	-0,332	0,503	-0,202	0,419
12500 Hz	0,334	-0,172	-0,368	-0,373	-0,352	-0,237	0,541	-0,275	0,323
16000 Hz	0,387	-0,075	-0,286	-0,326	-0,44	-0,217	0,528	-0,352	0,238
20000 Hz	0,424	0,019	-0,203	-0,226	-0,49	-0,192	0,473	-0,423	0,135

Figure 3: Correlations (Spearman r) of the perceived vowel similarity of musical instrument sounds to their 1/3rd octave band spectra. The example of the tuba, whose sound (playing A2 in ***pp***) was significantly assigned to the vowel "O", shows that its energy maxima in the 1/3rd octave bands spectrum match with the third octave ranges whose energy content corresponds positively with the frequency of the assigned vowel "O". Similar patterns can be seen for the other vowels and umlauts. Find an interactive map with sound examples at https://muwiserver.univie.ac.at/vowellike/t.htm

Thus, it is visible in Figure 4 that — when it comes to the musical instruments as a whole — vowel qualities are particularly assigned to the double reed and brass instruments in accordance with the observations in the literature ("A" to oboe and trumpet, "Å" to trombone, "O" to bassoon and French horn and "U" to tuba). A closer look at the individual sounds of the musical instruments reveals clearer vocal qualities in ***pp***. The flute and clarinet occasionally have "U" sounds in their lower and middle register, while the high register is often associated with the vowel "I" (this holds also for violin and viola). While the sounds of the violin in the low and middle registers are perceived as being similar to the vowel "E" (in accordance with the results in [12]: "front vowels"), the cello and double bass display almost no significant vowel similarities.

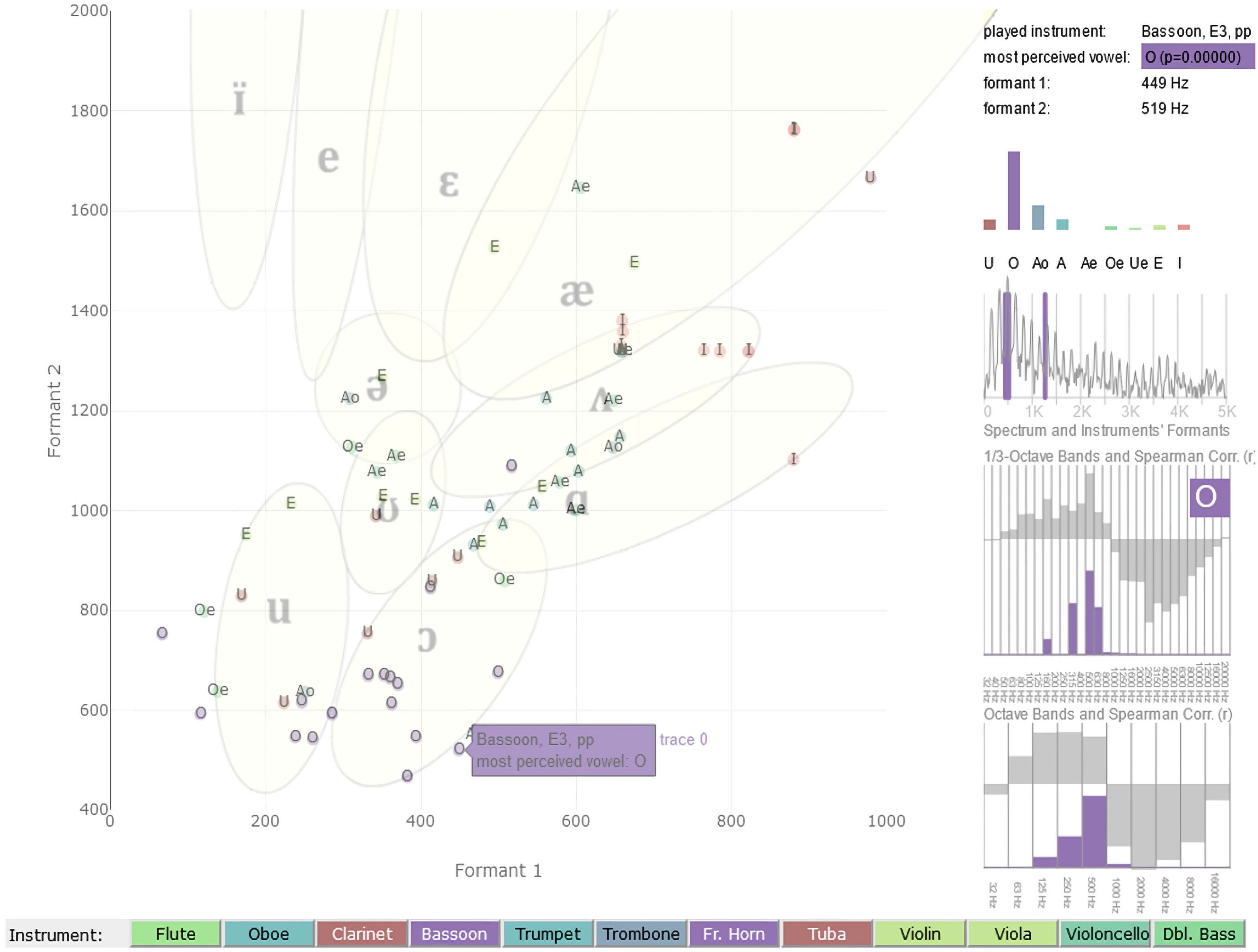


Figure 5: Interactive visualization of the vowel similarities of the instrument sounds in a vowel trapezoid (Formant 1 and 2 on the X and Y axes) combined with — on the right side — a display of the strength of the selected vowels, the formant ranges detected in the spectrum as well as their third octave and octave band spectra.

Scan the QR code, to explore this figure interactively at https://muwiserver.univie.ac.at/vowellike/ and move the mouse over the individual points to hear the sound examples. Individual instruments can be shown/hidden using the buttons at the bottom.

Since the instrument timbres with their respective vowel assignments show patterns in their third-octave band spectra matching the corresponding vowel formants in many cases, they can also be projected directly onto a vowel trapezoid (see Figure 5). In this representation, it becomes clear that the timbres with predominantly assigned to the vowel "O" and "A" can be mapped particularly well onto the vowel trapezoid, while sounds with an "I" vowel quality are characterized above all by an extraordinary high first formant, with the result that they only have a relationship to the vowel "I" via their second formant. This is due to the fact that the pitches of these sounds are almost exclusively in the high register, which means that the lower first formant typical for the vowel "I" cannot be produced in these cases.

Conclusions

All in all, it can be stated that musical instrument sounds do indeed have timbral vowel similarities. However, these are more or less pronounced depending on the instrument, register and dynamic level. Double-reed and brass instruments in particular have clearly recognizable vocal qualities, which also correspond to the information found in the existing literature. This also applies to a more limited extent to the clarinet. Vowel qualities are particularly evident in the ***pp*** of the wind instruments, while string instruments (especially violoncello and double bass) show almost no timbral associations to vowels. Since the energy distribution in the 1/3rd octave bands of musical instrument sounds often corresponds to the formant ranges of the vowels assigned to them, these timbres can also be visualized on a vowel trapezoid. Here, particularly the "A"- and "O"-like timbres have positions corresponding to their respective formants.

Acknowledgements

Our special thanks go to the Vienna Symphonic Library, which provided us with instrument sounds as stimuli for this study.

Literature

[1] Schumann, K. E. (1929). *Physik der Klangfarben*, Band II [habilitation thesis, University Berlin].

[2] Meyer, J. (1972). *Akustik und musikalische Aufführungspraxis: Leitfaden für Akustiker, Tonmeister, Musiker, Instrumentenbauer und Architekten*. Verlag Das Musikinstrunent.

[3] Fricke, J. P. (1975). Formantbildende Impulsfolgen bei Blasinstrumenten. *Fortschritte der Akustik – DAGA 1975*, 407-410.

[4] Voigt, W. (1975). *Untersuchungen zur Formantbildung in Klängen von Fagott und Dulziane*. G. Bosse.

[5] Gieseler, W., Lombardi, L., & Weyer, R.-D. (1985). *Instrumentation in der Musik des 20. Jahrhunderts: Akustik, Instrumente, Zusammenwirken*. Moeck.

[6] Auhagen, W. (1987). Dreieckimpulsfolgen als Modell der Anregungsfunktion von Blasinstrumenten. *Fortschritte der Akustik – DAGA 1987*, 709-712.

[7] Dickreiter, M. (1990). *Musikinstrumente: Moderne Instrumente, historische Instrumente, Klangakustik.* Deutscher Taschenbuch-Verlag.

[8] Reuter, C. (1996). *Die auditive Diskrimination von Orchesterinstrumenten: Verschmelzung und Heraushörbarkeit von Instrumentalklangfarben im Ensemblespiel*. Peter Lang.

[9] Oehler, M. (2008). *Die digitale Impulsformung als Werkzeug für die Analyse und Synthese von Blasinstrumentenklängen*. Peter Lang.

[10] Meyer, J. (2015). *Akustik und musikalische Aufführungspraxis: Leitfaden für Akustiker, Tonmeister, Musiker, Instrumentenbauer, und Architekten* (6., erweiterte Auflage). PPVMedien GbmH, Edition Bochinsky.

[11] Stumpf, C. (1926). *Die Sprachlaute. Experimentell-Phonetische Untersuchungen nebest einem Anhang über Instrumentenklänge*. Springer.

[12] Mores, R. (2017). Vowel Quality in Violin Sounds – A Timbre Analysis of Italian Masterpieces. In Schneider, A. (Ed.). Studies in Musical Acoustics and Psychoacoustics (p. 223-245). Cham: Springer.