

POSLECHOVÉ HODNOCENÍ VLASTNOSTÍ ZPĚVU – PŘEDBĚŽNÁ STUDIE

Marek Frič, Zdeněk Otčenášek, Klára Kadlecová

*Musical Acoustics Research Centre (MARC) Prague
Sound Studio of Music Faculty, Academy of Performing Arts in Prague,
Malostranské. nám. 13, 11800 Praha 1, Czech Republic
marekfric@centrum.cz, otcenasek@hamu.cz, klara.kadlecova@email.cz
<http://zvuk.hamu.cz/zvukovestudio.php>*

Abstract: The goal of presented experiment was to judge the singing voice in listening tests applied to a group of subjects in different level of voice education. Group consists of 18 subjects (2 operatic singers, 6 professional actors and 10 beginning students of acting). Recordings were performed on singing of a generally known Czech folk song in two tonalities (distance 5 semitones). Recordings were evaluated separately for each tonality in psychoacoustic listening tests. Firstly, by mean of the subjective rating procedure, in which were rated following properties: overall vocal performance, appropriate vibrato, resonance balance, ring, pitch accuracy, breath management, evenness trough the range, and strain. Secondly, evaluation was based on the dissimilarities between all stimuli pairs from all used recordings. Pair comparison test also consists of preferences in overall quality, and comparisons on the semantic scales bright-dull, wide-narrow and clear-rough. Results of listening tests were statistically analyzed by mean of multidimensional scaling (MDS). MDS analysis, based on dissimilarities, places the used stimuli into multi-dimensional model of a subject's perception space. MDS creates models with 4 dimensions for all 18 stimuli and 3D models for reduced 14 stimuli. These models were interpreted by emersion of different types of external parameters. In contribution there are discussed the results of experiment, which indicate a need for a next research.

1 Úvod

Studie zabývající se percepčním hodnocením hlasu se zaměřují jak na hlas patologický, tak umělecký. Tato hodnocení vycházejí ze základních postupů psychoakustického výzkumu, tedy z vyhodnocení výsledků různých druhů poslechových testů. Mezi nejčastěji používané možno zařadit hodnocení na subjektivních percepčních škálách, často v podobě vizuálně-analogových stupnic (VAS). I když tento typ hodnocení vyžaduje větší trénovanost hodnotitelů, je vhodným prostředkem pro vyhodnocování relativně velkého množství stimulů. Uvedený typ testování je zároveň i podkladem pro tvorbu percepčních protokolů, využívaných v klinické praxi [1], ale i v umění [2, 3]. Dalšími typy poslechových testů jsou seřazování podnětů (ranking) [4, 5] a párové (nebo i triadické) porovnávání [6]. Uvedené způsoby jsou však ve vokologii doposud méně používané a to zejména pro jejich časovou náročnost, z níž vychází nutnost hodnotit výrazně menší počet stimulů v jednom poslechovém testu.

Zpěvní hlas představuje velmi široce zkoumanou problematiku. V oblasti poslechového hodnocení se však omezuje hlavně na hodnocení kvality a barvy jednotlivých vokálů. Celkovému zpěvnímu projevu se obvykle věnují studie zaměřené na určitý specifický způsob použití hlasu. Tímto způsobem byl např. vytvořen prostředek pro hodnocení vlastností kvality operního zpěvu „**Singing**“ protokol [2].

U výzkumu barvy zvuku (témbru) hlasu, či zvuků obecně, se jako nejvýznamnější ukazuje metoda párového porovnání stimulů s hodnocení jejich podobnosti, resp. nepodobnosti, která následně využívá statistické zpracování výsledků pomocí multidimenzionálního škálování (MDS). Metoda MDS umožňuje geometrickou reprezentaci rozložení stimulů v odhadovaném

vícerozměrném percepčním prostoru tak, aby v něm vzdálenosti jednotlivých dvojic stimulů vypočtené MDS algoritmem co nejlépe odpovídaly mírám nepodobnosti těchto dvojic určeným v poslechovém testu. Pomocí dalších matematicko-statistických metod (například mnohonásobné lineární regrese) je pak možné řešit interpretaci popisovaných subjektivních vlastností stimulů.

Z MDS analýz nepodobnosti při čtení textu se např. ukázalo, že jako základní dimenzi obecného rozdílu hlasů je pohlaví [7]. Pokud se porovnávají separátně mužské a ženské hlasy (vokály), které ale nemají stejnou výšku, tak základní dimenzí rozdílů zvuků zas právě výška [6, 8, 9]. Studie [9, 10] pak dále ukazují, že druhá percepční dimenze souvisí s hlasovou klasifikací (Bar, Tenor, Alt, Soprán) a zároveň s těžištěm spektra u hodnocení jak mužských, tak ženských hlasů.

V oblasti kvazi stacionárních stimulů, jak vokálů, tak obecných zvuků, se dále ukazuje, že geometrická reprezentace nepodobností zvuků je dobře predikovatelná pomocí určení vzdáleností jejich spektrálních interpretací [6] (například pomocí 1/3 oktávových spekter). Samsonova studie percepce tónů různých typů stimulů [11] ukazuje rozlišení stimulů na základě počtu harmonických v první a třetí dimenzi a na základě časového průběhu náběhu a doznění v druhé dimenzi. Studie kombinující MDS s objektivními akustickými parametry poskytují různé interpretace dimenzí jak pro normální hlasy [12], tak pro patologické [13, 14].

Pomocí sémantického porovnání vzdáleností atributů popisu kvality a barvy zvuku s využitím MDS je možné hledat základní poslechové dimenze. Tímto způsobem byly ve studii [15] nalezeny 3 nezávislé (na sobě kolmé) dimenze popisu barvy hudebních zvuků: 1) tmavý – světlý, 2) úzký – široký a 3) drsný – jemný.

Cílem zde prezentované předběžné studie byl praktický vzhled do problematiky hodnocení vlastností zpěvního hlasu; dále ověření, jestli „Singing“ protokol je použitelný i pro porovnávání kvality zpěvu u různě trénovaných subjektů. Zároveň bylo snahou zjistit, jestli i u zpěvního hlasu jsou uvedené 3 percepční dimenze („Tmavost“, „Šířka“ a „Drsnost“) nezávislé a lze pomocí nich zpěvní hlasové projevy hodnotit.

2 Materiál a metody

2.1 Nahrávky

Dle metodiky [16] v Nahrávacím studiu DAMU Praha byly pořízeny nahrávky zpěvu lidové písně u 18 subjektů - mužů (2 operní zpěváci (subj.: 2, 5), 5 profesinálních herců (subj.: 8, 10, 11, 14, 15), 11 studentů herectví (subj.: 13, 22, 23, 26, 27, 28, 32, 33, 35, 39, 40)) ve dvou tóninách (A-dur, a d0-dur). Nahrávky byly pořízeny kondenzátorovými mikrofony Sennheiser s kulovou směrovou charakteristikou, typ (ME 62), umístěny ve standardních vzdálenostech od rtů subjektů 30 a 100 cm. Každý subjekt byl instruován, aby zpíval volně v komfortní hlasitosti, ale předepsané tónině, co nejkvalitněji dle svého uvážení. Pro poslechové testy byly použity nahrávky z bližšího mikrofону a vystřižen úsek 2 prvních veršů (cca 7 vteřin). Všechny nahrávky byly následně subjektivně vyrovnány na přibližně stejnou průměrnou hlasitost.

2.2 Poslechové testy

Autoři článku prováděli v předběžné studii 3 poslechové testy zaměřené na hodnocení kvality hlasu a vlastností hlasového projevu při zpěvu, každý pro 2 sady stimulů (obě tóniny zvlášť) po 18 nahrávkách. Všechny testy byly sestaveny v programu LiTEd [17], který umožňuje návrh sestavení obrazovek při provádění testů na PC a systematické řazení stimulů pro

každého posluchače. Použito bylo náhodné pořadí při škálovacích testech a řazení dle Rossova algoritmu pro párové porovnávací testy [18, 19].

První test (obrázek 2A) byl sestaven na základě „**Singing**“ protokolu [2], ve kterém pomocí subjektivního škálování na 20 bodové vizuálně-analogové stupnici byly hodnocené vlastnosti: „**Celkové provedení**“, „**Vhodnost vibrata**“, „**Vyvážení rezonance**“, „**Zvonivost barvy hlasu**“, „**Intonační přesnost**“, „**Vedení a použití dechu**“, „**Volnost zpěvního hlasu v celém rozsahu hlasu**“, „**Hlasové napětí**“.

Stimuly byly předváděny v náhodném pořadí (různém pro každého hodnotícího), hodnotící si mohl přehrát zvuk v libovolném počtu opakování, po ohodnocení všech vlastností stimulu se posouval na další stimul. První test byl s odstupem cca 20 dnů zopakován jako retest (v novém náhodném pořadí).

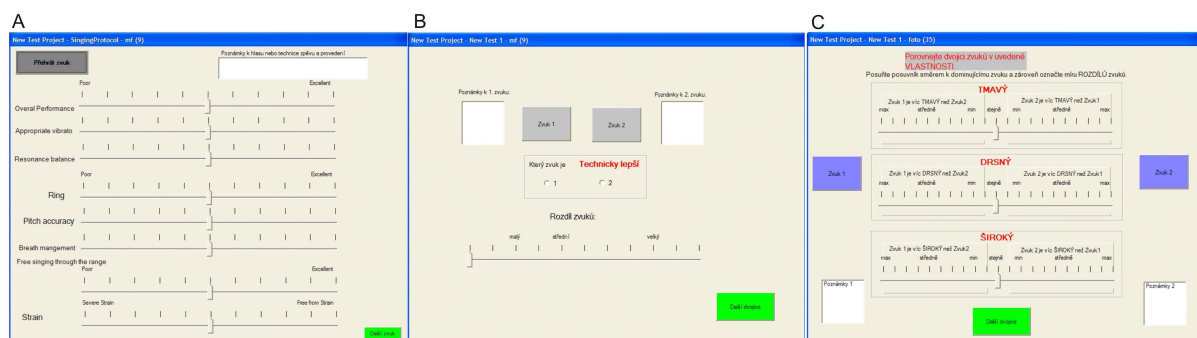
V druhém testu byly při párovém porovnání nucenou volbou – rozhodováním, který zvuk z páru má vyšší míru vlastnosti **technická kvalita hlasu** a jaký je **celkový rozdíl zvuků** (dissimilarity test) na 20 stupňové škále - obrázek 1B.

Třetí test v párovém porovnání na 20 stupňové stupnici hodnotil jak rozhodování, který zvuk z dvojice má větší zastoupení hodnocené vlastnosti, a také diferenci zvuků - jak se hodnocené zvuky od sebe odlišují v hodnocené vlastnosti. Hodnocené byly „**Tmavost**“, „**Drsnost**“ a „**Šířkost**“ zvuků (obrázek 1C). Tyto tři vlastnosti byly vybrány z předcházejícího výzkumu [15] jako tři základní dimenze popisu barvy hudebního zvuku.

Párové porovnávací testy byly hodnoceny pro každou možnou dvojici stimulů vždy pro konkrétní tóninu bez důležitosti pořadí stimulů (posluchač si mohl libovolně-krát přehrát kterýkoliv zvuk z dvojice). Tímto způsobem pro 18 stimulů v jednom setu vzniklo $18 \times 17 / 2 = 153$ dvojic.

2.3 Statistické zpracování výsledků

Pro ohodnocení reliability hodnotitelů (intra a inter-rater) byly vypočteny korelace pro jednotlivá hodnocení. Výsledky diferenčních testů byly zpracovány pomocí nemetrického multidimenzionálního škálování (MDS), z kterých byly následně vyhodnoceny dimenze percepčního hodnocení stimulů. Do nich byly vnořeny pomocí metody mnohonásobné lineární regrese (MLR) průměrné hodnoty parametrů získané subjektivním škálováním podle „**Singing**“ protokolu, a také průměrné pořadové koeficienty jednotlivých stimulů pro vlastnosti Tmavost, Šířka a Drsnost.



Obrázek 1. Vzhled obrazovek poslechových testů. A) Percepční „Singing“ protokol pro hodnocení zpěvu, B) Preferenční a diferenční párový test technické kvality, C) Párový diferenční a rozhodovací test pro vlastnosti Tmavost, Šířka a Drsnost.

3 Výsledky

Vzhledem k tomu, že pro úspěšné vyhodnocení všech vlastností obou setů stimulů bylo nutno vyhodnotit 378 dvojic resp. jednotlivých stimulů, jeden z autorů tyto testy nedokončil a jeho výsledky nebyly následně zpracovány.

Pro vyhodnocení shody hodnocení pomocí škálovacího testu „Singing“ protokolu obou hodnotitelů byly hodnoty z 1. testu a jeho retestu porovnány pomocí korelace. Výsledky korelační analýzy jsou uvedeny v Tab. 1. Výsledky párových porovnání nucenou volbou vedly k vytvoření tzv. preferenčních matic a následně k vytvoření pořadových koeficientů pro jednotlivé vlastnosti a stimuly obou hodnotitelů. Jejich výsledky byly porovnány pomocí Spearmanovy pořadové korelace a jsou zapsány v Tab. 1. v sloupcích Kvalita, Tmavost, Šířka, Drsnost.

Tabulka 1. Korelační porovnání výsledků „Singing“ protokolu (1. testu a retestu) a pořadových koeficientů párového porovnání stimulů (označeny šedě). Na řádcích jsou iniciály porovnávaných hodnotitelů, v případě stejných iniciál jde o test-retest korelaci, zobrazené jsou jenom statisticky významné korelace ($p < 0,05$).

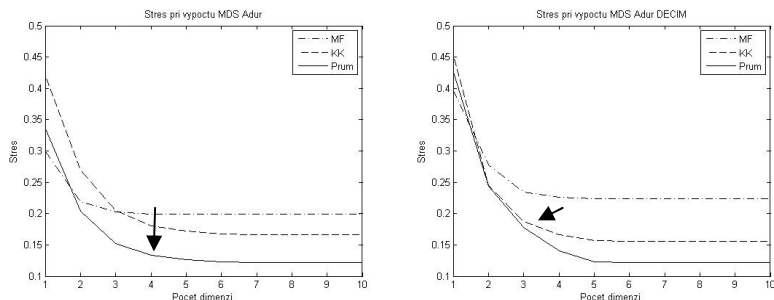
A-dur	Provedení	Vibrato	Rez. vyvážení	Zvonicost	Intonace	Dech	Volnost	Napětí	Kvalita	Tmavost	Šířka	Drsnost
MF_KK	0,54		0,61	0,70		0,73			0,80	0,78	0,56	0,52
KK_KK2		0,62			0,48	0,81	0,49					
MF2_KK2			0,60	0,64	0,69	0,51						
MF_MF2	0,73	0,58	0,62	0,57	0,72	0,76		0,78				
MF2_KK			0,55	0,59	0,50	0,62						
MF_KK2			0,52		0,52	0,55						

Korelace ukázala konzistenci výpovědí hodnotitelů jak v rámci testu a retestu (intrajudge) i v rámci jejich porovnání (interjudge). V Tab. 1 jsou zobrazeny jenom statisticky významné korelace, a tedy je možno uzavřít, že obecně výpovědi hodnotitelů nebyly stejně konzistentní pro všechny hodnocené parametry „Singing“ protokolu. Za konzistentní je možno považovat hodnocení parametrů Dech, Intonace, Resonanční vyvážení, naopak za nekonzistentní Vibrato, Napětí hlasu a Volnost hlasu. Parametr Provedení je jenom částečně konzistentní. Do další analýzy byly zahrnuty jenom konzistentní parametry s tím, že výsledné hodnoty parametrů byly vypočteny jako průměrné hodnocení konzistentních hodnocení obou hodnotitelů. Pořadové koeficienty, získané z preferenčních matic párových porovnání, ukazují statisticky významnou pořadovou korelaci pro všechny hodnocené parametry Kvalita, Tmavost, Šířka i Drsnost.

Výsledky diferenčního párového testu byly uspořádány do matic nepodobností a následně vytvořena matice průměrných nepodobností obou hodnotitelů. Z průměrné matice nepodobnosti byla metodou multidimenzionálního škálování (nemetrická metoda, kritérium úspěšnosti modelu „metricstress“, výpočet v prostředí MATLAB, funkce mdscale) vypočtena geometrická reprezentace pro 1 až 10 dimenzí. Hodnoty „stress“, které charakterizovaly úspěšnost MDS procesu u jednotlivých počtů dimenzí jsou uvedeny na Obr. 2A (hodnocení tóniny A-dur a všech 18 stimulů). Dle doporučení [18] bylo vybráno jako optimální řešení MDS 4dimenzionální řešení, protože vyhovuje podmínce *počet dimenzí* < (*počet stimulů* / 4), a také podmínce, že „stress“ MDS řešení se s přidáváním dalších dimenzí již dále výrazně nesnižuje.

Výsledky 4dimenzionální konfigurace MDS-analýzy reprezentují rozložení stimulů v jejich 4D modelu percepční reprezentace. Pomocí MLR analýzy je možno do 4D percepčního

prostoru „vnořit“ hodnoty subjektivních percepčních parametrů, konzistentních parametrů „Singing“ protokolu a také pořadové koeficienty parametrů z párového porovnání. Výsledky MLR pro 4D percepční prostor 18 stimulů A-dur jsou uvedeny v Tab. 2.

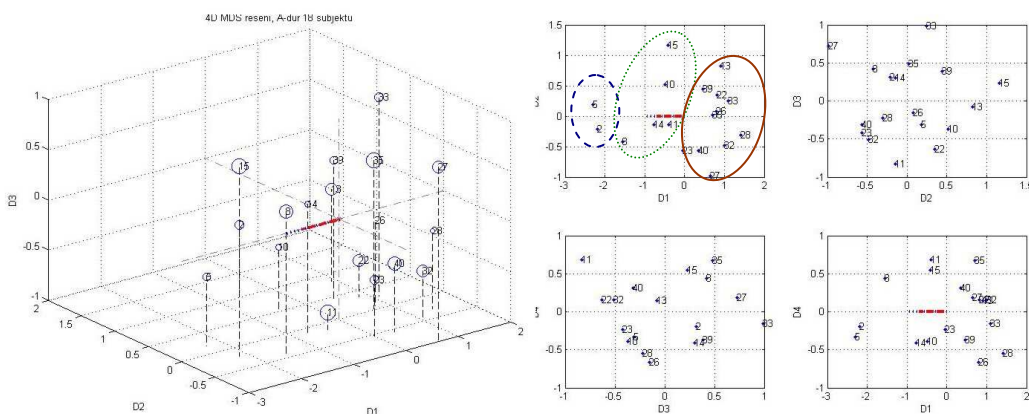


Obrázek 2. Hodnoty „stress“ geometrické reprezentace stimulů pro konkrétní počet dimenzí při multidimenzionálním škálování (pro konkrétního hodnotícího a z průměrné matice nepodobnosti). A) pro všech 18 stimulů, B) pro snížený počet 14 stimulů.

Výsledky MLR ukazují, že většina hodnocených vlastností, zejména Kvalita a Tmavost z párového porovnávání souvisí s první dimenzí MDS řešení. Vlastnosti Dech a Intonace souvisí i s 3 dimenzí. Zobrazení rozložení stimulů v 4D percepčním prostoru a „vnoření“ vlastností Kvalita a Tmavost jsou znázorněny na Obr. 3.

Tabulka 2. Koeficienty mnohonásobné lineární regrese hodnocených parametrů vzhledem k 4D řešení MDS analýzy percepčního prostoru 18 stimulů tóniny A-dur a 3D řešení MDS analýzy percepčního prostoru pro 14 (blízkých) stimulů. Zobrazeny jsou jenom statisticky významné koeficienty na hladině $p < 0,05$ pro konkrétní dimenze.

	A-dur 18 subj. – 4D řešení							A-dur 14 subj. – 3D řešení								
	Kvalita	Tmavost	Dech	Intonace	Provedení	Rez.vyv.	Zvon.	Kvalita	Tmavost	Šířka	Drsnost	Dech	Intonace	Provedení	Rez.vyv.	Zvon.
D1	-3,77	-2,66	-1,21	-1,22	-0,83	-1,44	-1,23	-3,66	-2,73			-1,23	-1,17	-0,98	-1,20	-1,60
D2								1,93		2,66		0,91				
D3			-0,81	-0,76							-3,37					
D4																

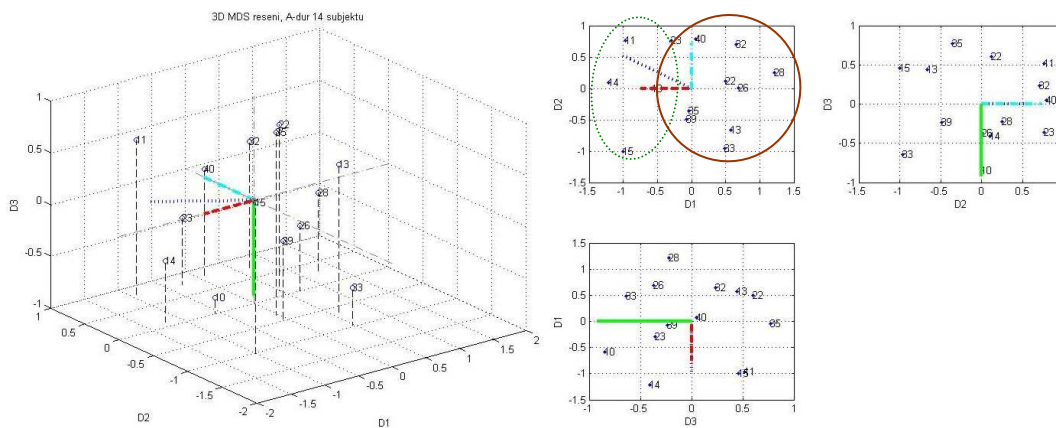


Obrázek 3. Rozložení hodnocených stimulů v 4dimenzionální řešení MDS tóniny A-dur pro všech 18 stimulů (4. dimenze je reprezentována velikostí kroužků jednotlivých stimulů). Vnořené vlastnosti Kvalita a Tmavost na základě MLR statisticky významně korelují jenom s první dimenzí. Na obrázcích vpravo je znázornění rozložení stimulů a „vnoření“ Kvality a Tmavosti vždy pro danou dvojici dimenzí.

Dle MLR se ukazuje, že kvalita pro všech 18 stimulů probíhá podél 1. dimenze. Na Obr. 3. je elipsami znázorněno seskupení jednotlivých kategorií subjektů dle jejich hlasové trénovanosti. Čárkovaná (modrá) elipsa označuje profesionální operní zpěváky, tečkovaná (zelená) – profesionální herce a plná (hnědá) elipsa označuje studenty (začátečníky). Dle uvedeného, takové seskupení odpovídá i korelaci 1. dimenze s vlastností Kvalita. Proto je možné ji nazvat Technickou kvalitou zpěvu, v níž nejvyšší hodnocení jednoznačně získali profesionální pěvci, následně školení herci a na nejhorším místě se umístila skupina studentů. Dále se ukázalo, že na daném vzorku byla Tmavost spojena s technickou kvalitou zpěvu (Tmavost významně korelovala s 1. dimenzí i Kvalitou, čím tmavší byl hlas, tím byl i kvalitnější).

Jelikož uvedený percepční prostor je výrazně protažen ve směru 1. dimenze (interpretovatelné jako technická kvalita zpěvu, nebo i tmavost), tak abychom omezili vliv dominantní kvalitativní vlastnosti geometrii prostoru vypustili jsme extrémní subjekty na obou stranách 1. dimenze (oba odlehlí profesionální zpěváci, subj. 2 a 5, jednoho profesionálního herce, subj. 8, a jednoho studenta, subj. 27) a postup jsme zopakovali. Hodnoty „stress“ pro MDS analýzu zbývajících 14 stimulů jsou znázorněny na Obr. 2B. Jako optimální bylo vybráno 3D řešení. Výsledky „vnoření“ konzistentních vlastností pomocí MLR pro 14 (blízkých) stimulů jsou uvedeny v Tab. 2. Tyto výsledky ukazují, že pro hodnocení percepčně bližších stimulů, je vnoření hodnocených vlastností méně ovlivněné dominující technickou kvalitou zpěvu. Vlastnosti Tmavost, Šířka a Drsnost se ukazují jako nezávislé (korelují postupně s dimenzemi 1, 2 a 3, které jsou na sebe kolmé). Kvalita už není interpretovatelná jen jednou dimenzí, ale vzniká kombinací Tmavosti a Šířky a je nezávislá na Drsnosti. Většina vlastností ze „Singing“ protokolu, ale stále koreluje jenom s první dimenzí a tedy i Tmavostí. Proto je možné předpokládat, že jednotlivé pro hodnocení použité vlastnosti ze „Singing“ protokolu mají jako složku Tmavost.

Grafické znázornění 3D řešení MDS pro 14 stimulů tóniny A-dur a vnoření vlastností Kvalita, Tmavost, Šířka a Drsnost jsou znázorněny na Obr. 4. Zelená – tečkovaná elipsa a hnědá – plná kružnice označují skupiny profesionálních herců respektive studentů. Vyšší kvalita zpěvu se stále ukazuje pro profesionální herce.

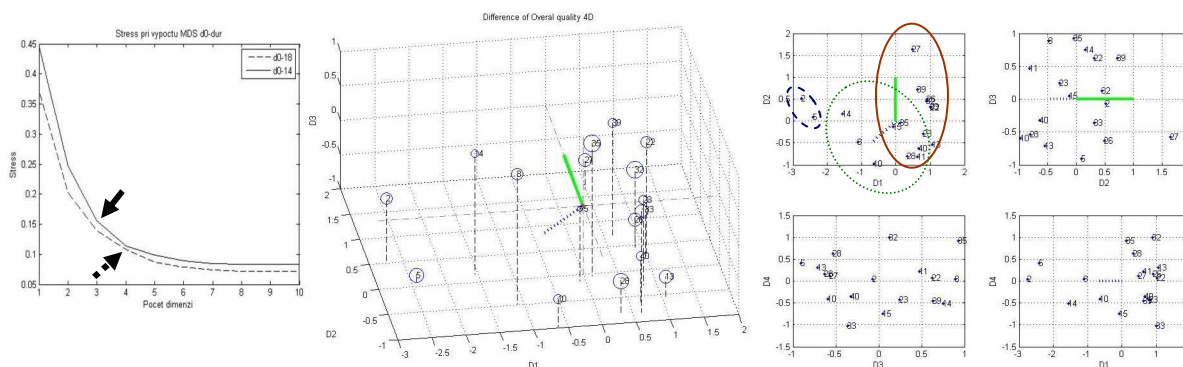


Obrázek 4. Rozložení hodnocených stimulů v 3dimenzionální řešení MDS tóniny Adur pro 14 (blízkých) stimulů. Vnořené vlastnosti Kvalita, Tmavost, Šířka a Drsnost jsou znázorněny tečkovanou, čárkovanou, čerchovanou respektive plnou tlustou čarou (modře, červeně, tyrkysově a zeleně v barevném provedení). Na obrázcích vpravo je znázornění rozložení stimulů a vnořených vlastností vždy pro danou dvojici dimenzí.

Podobným způsobem byly vyhodnoceny výsledky poslechových testů ve vyšší tónině - d0-dur. Nejdříve byly pomocí korelace zjištěny intra- a inte- individuální konzistence výpovědí. Výsledky, uvedené v Tab. 3, ukazují, že jako konzistentní je možno považovat parametry Dech, Volnost, Intonace, Vibrato, Kvalita a Drsnost; jako méně konzistentní Rezonanční vyvážení a Zvonivost. Jako nekonzistentní Provedení, Tmavost a Šířka.

Tabulka 3. Korelační porovnání výsledků „Singing“ protokolu (1. testu a retestu) a pořadových koeficientů párového porovnání stimulů pro tóninu d0-dur. Zobrazené jsou jenom statisticky významné korelace ($p < 0,05$).

d0-dur	Dech	Volnost	Intonace	Provedení	Rez. vyvážení	Zvonivost	Vibrato	Kvalita	Tmavost	Šířka	Drsnost
MF1_KK1	0,82	0,58	0,56		0,50		0,48	0,78	ns	ns	0,61
MF2_KK2											
MF1_MF2	0,88	0,63	0,82	0,88	0,86	0,67					
KK1_KK2	0,61						0,54				
MF1_KK2	0,63						-0,55				
MF2_KK1	0,81	0,49	0,56								



Obrázek 5. A) Stres MDS analýzy párového porovnání rozdílů stimulů v d0-dur tónině pro všech 18 stimulů. B) Zobrazení 4D řešení MDS analýzy s vnořenými vlastnostmi Kvalita a Drsnost (4. dimenze zobrazena velikostmi kroužků). C) Planární zobrazení rozložení 4D MDS řešení pro jednotlivé dvojice dimenzí, a průměty vlastností Kvalita a Drsnost.

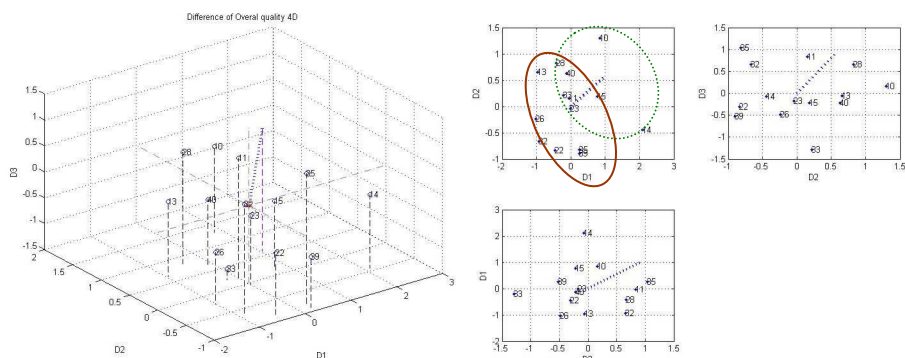
Jak je naznačeno na Obr. 5, na základě hodnot „stress“ bylo vybráno 4D řešení, které je zobrazeno na Obr. 5. Výsledky mnohonásobné lineární regrese jsou v Tab. 4.

Tabulka 4. Koeficienty mnohonásobné lineární regrese hodnocených parametrů vzhledem k 4D řešení MDS analýzy percepčního prostoru 18 a 14 stimulů tóniny d0-dur. Zobrazeny jsou jenom statisticky významné koeficienty na hladině $p < 0,05$ pro konkrétní dimenze D1-D4.

	d0-dur 18 subj. – 4D řešení							d0-dur 14 subj – 3D řešení				
	Kvalita	Drsnost	Dech	Volnost	Intonace	Res.vyváž.	Zvonivost	Dech	Intonace	Res.Vyv.	Zvonivost	Kvalita
D1	-2,35		-1,13	-0,89	-1,53	-1,14	-1	0,98	1,14	1,11	1,14	2,77
D2	-1,59	3,24										1,55
D3												2,49
D4												

Na Obr. 5., na grafu rozložení 4D MDS řešení pro D1 a D2 dimenzi, jsou v elipsách znázorněny skupiny operních pěvců, profesionálních herců a studentů, kde možno podobně jako při A-dur tóniny pozorovat odlehlost subjektů 2 a 5 – operních pěvců, a studenta 27. Proto byly, podobně jako pro tóninu A-dur, z další analýzy eliminovány (stejně i subj. 8), aby

byly výsledky porovnatelné mezi tóninami. Nová skupina 14 stimulů v tónině d0-dur ukazuje jako optimální řešení MDS ve 3 dimenzích, vnoření percepčních vlastností je znázorněno v Tab. 4 a Obr. 6.



Obrázek 6. Zobrazení 3D řešení MDS analýzy 14 stimulů v tónině d0-dur s vnořenou vlastností Kvalita. Planární zobrazení rozložení 3D MDS řešení pro jednotlivé dvojice dimenzí, a průměty vlastnosti Kvalita.

U skupiny 14 stimulů v tónině d0-dur probíhá Kvalita šikmo percepčním prostorem, koreluje se všemi třemi dimenzemi. Hodnocené vlastnosti „Singing“ protokolu, však stále korelují jen s první dimenzí. Drsnost nevykazovala při MLR statisticky významnou korelaci s některou z dimenzí.

4 Diskuze

V tomto příspěvku jsou uvedeny a porovnány výsledky třech typů poslechových testů vlastností zpěvu ve dvou tóninách na nahrávkách zpěvu lidové písně. Tato studie byla navržena jako předběžná studie s cílem praktického náhledu na problematiku hodnocení zpěvního hlasu, proto se jí zúčastnili jenom 2 autoři. Jeden z hodnotitelů je amatérský sborový, ale i sólový zpěvák klasického zpěvu (v této studii zastupuje netrénovanou skupinu v oblasti hodnocení zpěvu), druhý hodnotitel je hlasový pedagog a terapeut (svým zaměřením reprezentuje specificky trénovanou skupinu pro hodnocení vlastností hlasu při výuce zpěvu). Oba hodnotitelé již minimálně 5 let prakticky řeší problematiku hodnocení vlastností patologického hlasu, proto je možno je považovat za informovanou skupinu hodnotitelů.

První test používá metodu subjektivního škálování, kde bylo hodnoceno 9 vlastností tzv. „Singing“ protokolu. Výsledky tohoto testu ukázaly relativně velkou nekonzistenci soudů mezi hodnotiteli, která se postupně zhoršuje pro vyšší tóninu. Toto zjištění není překvapivé a je doloženo vícero studiemi o subjektivním hodnocení trénovaných a netrénovaných hodnotitelů nebo i různém způsobu hodnocení trénovaných hodnotitelů [10, 15, 20]. Neshody v hodnocení je možné obecně vysvětlit:

- různým chápáním hodnocených vlastností – tedy různými vnitřními modely;
- různou mírou a vahou hodnocených vlastností;
- nestabilitou (netrénovaností) hodnocení konkrétních vlastností;
- nevhodnou skupinou hodnocených stimulů, která neumožňuje konzistentní hodnocení;
- nevhodnou vybranou skupinou hodnocených vlastností, které nejsou jednoznačně definované pro vybranou skupinu stimulů.

Za nejkonzistentnější vyhodnocované, v rámci obou tónin, je možno považovat vlastnosti Dech a Intonace, naopak za nekonzistentní Napětí, Volnost, Vibrato a pro vyšší tóninu i Provedení. Dle výsledků korelační analýzy se ukazuje, že statisticky byl konzistentnější

v hodnocení méně trénovaný hodnotitel zpěvu. Dle další analýzy vzájemné korelace hodnocených vlastností se ukazuje, že u tohoto hodnotitele dominovala jedna subjektivní složka, která pak ovlivnila hodnocení všech vlastností zpěvu v „Singing“ protokolu. Obdobné zjištění uvedl i [6]. Podobným způsobem, dle vzájemné diskuze obou hodnotitelů, je možno vysvětlit, paradoxně nejlépe konzistentní, hodnocení vlastnosti Dech pro obě tóniny. Ve skutečnosti, sledování použití dechových svalů, ani ekonomičnost a správnost nadechování v rámci skladby, nebylo na ukázkách možné hodnotit. Proto se zdá, že vlastnost „Vedení dechu“ oba posuzovatelé nehodnotili na základě skutečných informací o použití dechu, ale o jeho předpokládaném způsobu použití, které následně ovlivnilo celkovou kvalitu provedení zpěvu. Jinými slovy, hodnotitelé ve skutečnosti hodnotili, jestli kvalita hlasu a projevu podporuje jejich subjektivní představu o správném použití dechu.

Hodnocení Celkového provedení v „Singing“ protokolu obecně nepatřilo mezi dobře konzistentní, což je možno vysvětlit také nevhodně definovaným účelem provedení. Profesionální pěvci získávali nízké hodnocení Celkového provedení při hodnocení hlasového pedagoga, i když jejich hodnocení technické kvality zpěvu bylo nejlepší, což vysvětlil nevhodností zvolené interpretace pro zpěv lidové písně.

Druhý poslechový test pomocí párového srovnávání hodnotil vzdálenosti (dissimilarity) jednotlivých stimulů. Třetí skupina testů použila nucenou volbu, také v párovém porovnávání, pro určení větší míry vlastností Kvalita, Tmavost, Šířka a Drsnost. Na základě preferenčních matic byly vypočteny pořadové koeficienty stimulů v hodnocení jednotlivých vlastností. Tyto koeficienty vykazovaly dobrou korelaci mezi hodnotiteli u hlubší tóniny ve všech vlastnostech. Ve vyšší tónině byla konzistence jenom pro vlastnosti Kvalita a Drsnost. Párový porovnávací test hodnotitelé považovali subjektivně za jednodušší a přesnější, protože měli odpovídat jenom na otázku, který z předkládaných stimulů má vyšší míru hodnocené vlastnosti). Pokud však bylo porovnání míry komplikovanější (například vícedimenzionální), nebo naopak – hodnocené stimuly nevykazovaly výrazné rozdíly, výsledky porovnávacího testu v preferenčních maticích obsahovaly velké množství tzv. kruhových triád. Výsledky tak nebyly dobře interpretovatelné [6], v tomto případě se tak dělo pro vyšší tóninu.

Do optimálních řešení percepčních prostorů na základě diferenčních testů byly „vnořeny“ pomocí mnohonásobné lineární regrese (MLR) jenom konzistentně vyhodnocené vlastnosti. MLR interpretuje jednotlivé percepční modely a umožňuje hledání vztahů mezi hodnocenými percepčními vlastnostmi v použitých poslechových testech. Pro obě tóniny, se při hodnocení všech 18 stimulů ukazuje, že první dimenze modelu MDS (největší rozptyl v hodnocení rozdílů stimulů) nejvíce koreluje s vlastností Kvalita a také s většinou vlastností „Singing“ protokolu. V obou tóninách vyhovovala 4dimenzionální řešení MDS. Podél první dimenze se rozložily jednotlivé skupiny subjektů na základě jejich hlasové trénovanosti. Naopak nekonzistence v hodnocení Provedení „Singing“ protokolu naznačuje, že hodnocení této vlastnosti je složitější. Na základě vyjádření hlasového pedagoga, bylo hodnocení Provedení bylo výrazně poznamenáno vkusem a účelem provedení (interpretace), který bohužel v této studii nebyl při nahrávání dostatečně předdefinován.

Protože v obou tóninách, byly nepodobnosti způsobeny výrazně odlišnou technickou kvalitou zpěvu operních pěvců, byly separátně vyhodnoceny prostory i pro skupinu 14 bližších stimulů (po odstranění odlehlých stimulů pěvců a velmi nekvalitních zpěváků). V případě hlubší tóniny se tak umožnila lepší interpretace získaného 3D MDS percepčního prostoru (vlastnosti Tmavost, Šířka a Drsnost korelovaly vždy s jinou dimenzí a byly percepčně nezávislé). Tento výsledek potvrzuje předpoklad předchozího výzkumu [15]. Výsledky pro hlubší tóninu

ukazují, že vyšší Tmavost je průvodní jev technicky kvalitního zpěvu. Pro skupinu všech 18 stimulů leží Tmavost ve stejném směru jako Kvalita. U skupiny 14 (blízkých) stimulů hlubší tóniny Kvalita souvisela s Tmavostí a zároveň s Šířkou, naopak byla nezávislá na Drsnosti.

Ve vyšší tónině se hodnocení Kvality v souvislosti s jinými vlastnostmi značně lišilo od hlubší tóniny. Vlastnosti Tmavost, Šířka a ani Provedení nevykazovaly konzistenci v hodnocení hodnotitelů. Podobně jako studie [9, 10], které dokládají menší schopnost rozlišení tmavosti hlasu s nárůstem výšky hlasu, jak pro muže, tak pro ženy, je možné podobným způsobem interpretovat i zhoršenou konzistenci pro vyšší tóninu soudů vlastnosti Tmavost v naší studii. Nekonzistence hodnocení Šířky může mít podobnou příčinu. Naopak, při hodnocení technické kvality zpěvu ve vyšší poloze se ukazuje nárůst významu vlastnosti Drsnost, která pro skupinu všech 18 stimulů leží v 4D MDS percepčním modelu téměř v opačném směru než vlastnost Kvalita. Drsnost však neukazovala ve vyšší tónině žádnou interpretaci pro 3D MDS model zmenšené skupiny 14 stimulů.

Ze vzájemné diskuze autorů o příčinách nekonzistence jednotlivých subjektivních hodnocení vyplynulo, že použití „Singing“ protokolu nebude vhodné na hodnocení kvality zpěvu obecně. Tento protokol byl vytvořen zejména pro hodnocení operního zpěvu a je též částečně zaměřen na interpretaci. Z výsledků také vyplynula nutnost změnit uspořádání a definici stimulů pro příští percepční hodnocení zpěvu. Je nutno napřed pevně stanovit účel studie a odlišit interpretační, technické a jiné hodnocení kvality. „Singing“ protokol pravděpodobně dobře pokrývá hodnocení techniky zpěvu v propojení s klasickou operní interpretací, naopak vlastnosti Tmavost, Drsnost a Šířka hodnotí kvalitativní vlastnosti zvuku jako takového. Naše studie pro hlubší tóninu dokonce potvrzuje, že tyto vlastnosti jsou nezávislé a dimenzionální. Pokud bude snahou studie zjistit základní barvotvorné vlastnosti, měla by být již při tvorbě stimulů, kromě stejné tóniny a hlasitosti, zachována co nejmenší variabilita v interpretačních vlastnostech. Tedy například omezení ve variabilitě druhu vokálů, způsob nasazení hlasu a zejména striktní dodržení tempo-rytmu (například pomocí metronomu). Hodnocení interpretace by zase naopak mělo mít dobře definovaný interpretační účel (například vyvolání konkrétní emoce, nebo jiného subjektivního pocitu). Další doporučení tímto směrem jsou ale nad rámec této studie.

5 Závěr

Poslechové hodnocení vlastností zpěvu je komplikovaný problém. Dle výsledků studie se ukazuje, že vlastnosti zpěvu je možno rozdělit na interpretační (svou podstatou více umělecké), technické a barvotvorné. Párový diferenční test ukazuje jeho použitelnost v hodnocení a možnosti interpretace barvotvorných vlastností Tmavost, Šířka a Drsnost. Naopak, použitý „Singing“ protokol se neukazuje jako vhodný na obecné hodnocení vlastností zpěvu, výrazně v sobě zahrnuje interpretační složku a dostatečně neumožňuje rozlišení vlastností zpěvu netrénovaných osob.

Poděkování

Práce vznikla za podpory projektu Ministerstva školství a mládeže, ČR Pr.No. 1M0531.

Literatura

- [1] Kempster G. B., Gerratt B. R., Verdolini A. K., Barkmeier-Kraemer J., Hillman R. E. (2009): Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol, *Am J Speech Lang Pathol*, 18, 2, 124-32.

- [2] Oates J. M., Bain B., Davis P., Chapman J., Kenny D. (2006): Development of an auditory-perceptual rating instrument for the operatic singing voice, *J Voice*, 20, 1, 71-81.
- [3] Wapnick J., Ekholm E. (1997): Expert consensus in solo voice performance evaluation, *J Voice*, 11, 4, 429-36.
- [4] Granqvist S., Lindestad P. A. (2001): A method of applying Fourier analysis to high-speed laryngoscopy, *J Acoust Soc Am*, 110, 6, 3193-7.
- [5] Bele I. V. (2005): Reliability in perceptual analysis of voice quality, *J Voice*, 19, 4, 555-73.
- [6] Bloothoof G., Plomp R. (1988): The timbre of sung vowels, *J Acoust Soc Am*, 84, 3, 847-60.
- [7] Singh S., Murry T. (1978): Multidimensional classification of normal voice qualities, *J Acoust Soc Am*, 64, 1, 81-7.
- [8] Baumann O., Belin P. (2010): Perceptual scaling of voice identity: common dimensions for different vowels and speakers, *Psychol Res*, 74, 1, 110-20.
- [9] Erickson M. L. (2003): Dissimilarity and the classification of female singing voices: a preliminary study, *J Voice*, 17, 2, 195-206.
- [10] Erickson M. L. (2008): Dissimilarity and the classification of male singing voices, *J Voice*, 22, 3, 290-9.
- [11] Samson S., Zatorre R. J., Ramsay J. O. (1997): Multidimensional scaling of synthetic musical timbre: perception of spectral and temporal characteristics, *Can J Exp Psychol*, 51, 4, 307-15.
- [12] Murry T., Singh S. (1980): Multidimensional analysis of male and female voices, *J Acoust Soc Am*, 68, 5, 1294-300.
- [13] Murry T., Singh S., Sargent M. (1977): Multidimensional classification of abnormal voice qualities, *J Acoust Soc Am*, 61, 6, 1630-5.
- [14] Kreiman J., Gerratt B. R., Berke G. S. (1994): The multidimensional nature of pathologic vocal quality, *J Acoust Soc Am*, 96, 3, 1291-302.
- [15] Štěpánek J., Moravec O. (2004): Percepční prostory barvy hudebního zvuku a jejich slovní popis, GA ČR 202/02/1370.
- [16] Frič M., Kadlecová K., Szymiková R. (2007): Záznam mluvního a zpěvního hlasu pro psychoakustické analýzy, MARC-Technologický list čís. 13, *Zvukové studio HAMU*.
- [17] Otčenášek Z., Štěpánek J., Prokop J., Frič M., Hrb J. (2010): LiTEd -Software pro editaci, provádění a vyhodnocování poslechových testů, Technologický list č. 36: *Zvukové studio HAMU*.
- [18] Melka A. (2005): *Základy experimentální psychoakustiky*, Akademie múzických umění v Praze, ISBN 80-7331-043-0.
- [19] Otčenášek Z. (2008): *O subjektivním hodnocení zvuku*, Akademie múzických umění v Praze, ISBN 978-80-7331-113-1.
- [20] Shrivastav R. (2006): Multidimensional scaling of breathy voice quality: individual differences in perception, *J Voice*, 20, 2, 211-22.