

Vliv temporálních manipulací na vnímání kompetence mluvčího

Zuzana Berkovcová — Štěpánka Černíková — Radek Skarnitzl

ABSTRACT:

Effect of Temporal Manipulations on the Perception of Speaker Competence. Speech communication research based on psychological methods currently stands at the forefront of scientific interest. Speech is an integral part of the social identity of a person and has a significant impact on the perception of the speaker by his surroundings. The present study aims to chart the effect of the temporal organization of utterances on the perception of a speaker's competence. Recordings of four Spanish native speakers were manipulated in a way which destabilized the regular temporal structure of their utterances. A perception test was administered to forty Czech listeners differing in level of proficiency with the Spanish language. The aim of the test was to reveal the listeners' positive or negative judgments of the original (regular) and manipulated (dysfluent) items. The basis for the perception test was the Big Five personality traits model, with the factor evaluated being the speakers' competence that is the ability and readiness to effectively deal with tasks. The results confirmed our main hypothesis, which assumed that both groups of listeners will, in terms of competence, evaluate the manipulated items negatively. Students of Spanish studies were more perceptive of the temporal manipulations, most likely due to their familiarity with the prosodic structure of Spanish and their understanding of the meaning of the tested utterances.

KLÍČOVÁ SLOVA / KEY WORDS:

kompetence, pětifaktorový model osobnosti, psychofonetika, španělština, temporální členění řeči Big Five personality traits, competence, psychophonetics, Spanish, temporal patterning of speech

ÚVOD

Řečí přenášíme kromě primárních lingvistických obsahů naší mysli i další informace, které můžeme souhrnně nazvat indexové. Ty se týkají různých vlastností mluvčího, poukazují (angl. *index*) na různé jeho charakteristiky, ať už se jedná o charakteristiky fyziologické povahy (např. vzrůst či věk), nebo povahy sociální (tzn. příslušnost k různým sociálním skupinám), viz např. Nygaardová (2005) nebo Belin et al. (2011). Řečový projev tak představuje významnou součást sociální identity jedince a zároveň výrazně ovlivňuje způsob, jakým posluchači mluvčího vnímají (Edwards, 1999). Je dobře známé, že na základě řečového projevu si lidé automaticky a velmi rychle vytvářejí (ne vždy přesné) úsudky o povahových vlastnostech mluvčího (viz např. Carney — Colvin — Hall, 2007; McAleer — Todorov — Belin, 2014), a že proto — z obráceného pohledu — řečový projev jedince může mít závažný a kvantifikovatelný vliv na povahu jeho sociálních interakcí a na jeho život obecně. Nesčetné množství studií se zabývalo rozdíly v hodnocení mluvčích, kteří hovoří různými jazyky (např. Lambert et al., 1960), různými akcenty v rámci jednoho jazyka (Purnell — Idsardi — Baugh, 1999; Dixon — Mahoney — Cocks, 2002) či cizineckým



přízvukem (Cargile — Giles, 1997; Lindemann, 2003; Lev-Ari — Keysar, 2010). Tyto a další studie téměř bez výjimky ukazují, že nestandardní akcent a cizinecký přízvuk jsou spojovány s negativními stereotypy, předsudky, či dokonce diskriminací (shrnutí tohoto výzkumu přináší Gluszek — Dovidio, 2010; Collins — Clément, 2012; Lippi-Green, 2012).

Analogicky můžeme o odlišném hodnocení mluvčího uvažovat nejen v případě nestandardních či nerodilých variet, tedy „globálních“ vlastností řečového projevu, ale i v případě „lokálně“ se projevujících jevů, které řečový projev daného jedince nějakým způsobem odlišují od projevu (ze statistického hlediska) normálního. Mezi takovéto lokální jevy můžeme řadit například řečové vady, neobvykle rychlé či pomalé mluvní tempo nebo různé druhy dysfluencí. Jsou to právě temporální aspekty, na něž se zaměříme v této studii.

Rytmická a temporální pravidelnost řečového projevu se ukazuje jako důležitý indikátor „kvality“ mluvčího. Jak shrnuje Volín (2010), cit pro rytmus a pravidelnost nejen v řeči je pro člověka vrozenou vlastností, a proto k pravidelnému uspořádání jevů kolem sebe přirozeně inklinuje (důkazem mohou být děti snadno se učící rytmické říkanky či dav sportovních fanoušků skandujících do rytmu). Při běžné řečové komunikaci se jako klíčový aspekt jeví předvídatelnost: výzkumy ukazují, že percepce je do značné míry založena na opakovaném ověřování a potvrzování predikcí. Pokud je přicházející řečový signál synchronizován s mozkovými vlnami, je tento proces bezproblémový a velmi efektivní: dochází k tzv. neurální rezonanci (viz Grossberg, 2003; 2013). Pokud je však temporální struktura výpovědi narušena, mozek se na ni nemůže „naladit“ (adaptovat) a je třeba signál rekonstruovat z krátkodobé paměti, což je pro mozek podstatně náročnější. Řečová percepce je tedy relativně snadná, když je temporální uspořádání vstupních řečových jednotek předvídatelné, zatímco v opačné situaci je zpracování řeči náročnější. To dokazují i studie založené na měření kognitivní zátěže spojené s percepcí řeči: pokud je řeč z temporálního hlediska pravidelná, jsou reakční doby posluchačů na cílové řečové jednotky výrazně kratší (Quené — Port, 2005) a práh detekce řečových jednotek je nižší (Quené — van Delft, 2010).

V této studii se zabýváme vlivem temporálního členění výpovědi na hodnocení osobnosti mluvčích posluchači. Zkoumání lidské osobnosti patří k tradičním úlohám psychologie. V posledních desetiletích se výrazně prosazuje pětifaktorový model osobnosti, zejména Normanův model Big Five z roku 1963 (McCrae — Costa, 1987; Barrick — Mount, 1991; Blatný et al., 2010, s. 43; Hřebíčková, 1999; 2011). Tento model se týká pěti základních rozměrů (dimenzí), které se v psychologii používají k popisu lidské osobnosti. Důležitou vlastností pětifaktorového modelu je skutečnost, že se jedná o faktory ortogonální, tedy na sobě nezávislé.

Jeden z metodologických přístupů pětifaktorového modelu se zakládá na lexikálním přístupu, který je odvozen z hypotézy, že všechny relevantní rysy a odlišnosti v osobnosti člověka máme potřebu pojmenovat, to znamená, že jsou zachycené v jazyce a máme pro ně existující slovo. Množina takovýchto slov byla na základě rozsáhlých výzkumů odborníky rozdělena do okruhů, které shrnují všechny rysy osobnosti člověka. Jde o těchto pět faktorů (ke každému zároveň uvádíme několik příkladů adjektiv; českou terminologii přejímáme od Hřebíčkové, 1999; 2011):

- živost / extraverte (povídavý, společenský),
- přívětivost (dobrosrdečný),
- svědomitost (důkladný, pečlivý),
- emocionální stabilita / neuroticismus (vyrovnaný, klidný),
- kultura / intelekt / otevřenost vůči zkušenosti / integrita (vzdělaný, originální).

Pro náš výzkum jsme zvolili faktor svědomitosti, jehož těžištěm z hlediska osobnosti je vztah k práci a vědomí povinnosti (Hřebíčková, 2011, s. 41), a konkrétně jednu z jeho podškál, rys *kompetence* (neboli *způsobilosti*). Kompetenci můžeme popsat jako schopnost a připravenost efektivně řešit úkoly (*ibid.*, s. 98), odkazuje k charakteristikám jako schopný, prozíravý a efektivní. Opačný pól tohoto rysu je formulován jako neschopnost postavit se k práci (*ibid.*, s. 105).

Ačkoli jsou snahy o propojení řečového výzkumu s psychologií osobnosti patrné již delší dobu — viz např. na svou dobu velmi pokrokové výzkumy Addingtona (1968) nebo Aronovitche (1976) — intenzivnější snahy můžeme pozorovat až v posledních několika letech. Například studie Imhofové (2010) je založena na hodnocení hlasů jednoho muže a jedné ženy pomocí všech pěti faktorů osobnosti. Autorka snížila a zvýšila původní základní frekvenci (Fo) mluvčích a zjistila, že mluvčím s vyšším hlasem je připisována vyšší extraverte, avšak také nižší svědomitost a emocionální stabilita.

Za zmínku stojí i studie, které se pokoušejí o automatickou klasifikaci osobnostních rysů z řeči. Polzehl, Möller a Metze (2010) porovnávali hodnocení posluchačů s automatickým klasifikátorem a dosáhli šedesátiprocentní úspěšnosti, přičemž nejlepších výsledků bylo dosaženo u neurotičnosti a extraverte. Z foneticky interpretovatelných parametrů se jako nejužitečnější koreláty těchto dimenzí ukázaly parametry jako intonační rozpětí či rychlost melodických změn. Mohammadi a Vinciarelli (2012) použili nahrávky ve francouzském jazyce, jimž respondenti nerozuměli, a nebyli proto ovlivněni obsahovou stránkou řeči. Dobré výsledky byly zjištěny — podobně jako u předchozí studie — při identifikaci extraverte a nejuspěšnějším prediktorem byla opět melodická variabilita a také nízká četnost pauz. Zajímavé také je, že osoby s vyššími hodnotami vokálních formantů byly hodnoceny jako méně přívětivé. Další příklady ve své knize uvádí Polzehl (2015, 2. kapitola).

Pro účely naší studie jsou nejzajímavější výzkumy, v nichž autoři provádějí manipulace vybraného akustického parametru a v percepčních testech pak zjišťují, jestli se změna odrazila v hodnocení mluvčího posluchači. Hlavní výhodou takového přístupu je samozřejmě to, že se jedná o kontrolované manipulace, a výsledné položky se tedy liší pouze v tomto jediném aspektu. Lze tak předpokládat, že případné zjištěné rozdíly v hodnocení spárovaných položek (téže věty v původní a manipulované podobě) jsou způsobeny právě provedenou manipulací. První takové pokusy sahají až do 70. let minulého století, kdy Bruce Brown a jeho kolegové pomocí speciálně upraveného magnetofonu manipulovali řečovým tempem a základní frekvencí (Fo) a pomocí řečového analyzátoru a syntezátoru byli schopni měnit dokonce i rozpětí Fo. Posluchači pak manipulované nahrávky posuzovali z hlediska laskavosti a kompetence. Výsledky těchto studií ukazují, že zvýšené intonační rozpětí se odráží v lepším skóre laskavosti (a naopak) nebo že zvýšená hladina Fo způsobuje (u mužských hlasů) nižší skóre u laskavosti i kompetence. Rychlejší mluvčí jsou hodnoceni jako kompe-





tentnější, avšak méně laskaví (Brown — Strong — Rencher, 1973; 1974; Smith et al., 1975). Podobným způsobem manipulovali řečovými parametry i Apple, Streeterová a Krauss (1979), kteří zjistili, že muži s vyššími hlasy jsou posluchači vnímáni jako méně pravdomluvní, jako nervóznější a slabší nebo že pomalejší tempo v posluchačích vzbuzuje dojem nižší plynulosti a mluvčí jsou vnímáni jako méně pravdomluvní a méně přesvědčiví.

V dnešní době je možné provádět kromě zmíněných globálních změn i změny lokální. Volín, Poesová a Skarnitzl (2014) prováděli lokální temporální změny na nahrávkách 14 profesionálních mluvčích angličtiny, a to tak, že trvání přízvukných samohlásek (které v angličtině bývají výrazně delší než samohlásky nepřívzvučné) snížili na polovinu, zatímco trvání samohlásek v nepřívzvučných slabikách zdvojnásobili. Posluchači posuzovali původní a manipulované položky z hlediska faktoru neuroticismu. Výsledky ukazují, že mírné narušení temporální struktury vede k statisticky významně horšímu hodnocení neuroticismu.

V této studii se rovněž zaměříme na vliv temporálních manipulací na hodnocení osobnosti mluvčích. Oproti studii Volína, Poesové a Skarnitzla (2014) manipulace v této studii nesimulují narušený řečový rytmus, ale kolísání tempa. Sledovaným osobnostním rysem je vnímaná kompetence rodilých mluvčích španělštiny u dvou skupin posluchačů — u hispanistů a u posluchačů, kteří španělštinu neznají. Volba dvou skupin posluchačů lišících se obeznámeností se španělštinou vycházela z hypotézy, že právě znalost cílového jazyka může být zajímavým faktorem, který hodnocení respondentů ovlivní. Jinými slovy bylo naším cílem zjistit, zda posluchači, kteří se španělštinou zkušenosti nemají, budou na změny tempa reagovat podobným způsobem jako posluchači, kteří s mluvenou španělštinou běžně přicházejí do styku.

Na základě výsledků dřívějších studií budeme pracovat s následujícími dvěma hypotézami:

- H₁: Posluchači budou z hlediska kompetence hodnotit manipulované položky, tedy ty, u kterých je patrné kolísavé tempo, negativněji.
- H₂: Rozdílné hodnocení se projeví u obou zkoumaných skupin respondentů, avšak u hispanistů očekáváme větší citlivost na provedené manipulace, a to z důvodu znalosti prozodické stavby španělštiny i díky porozumění významu testovaných výpovědí.

METODA

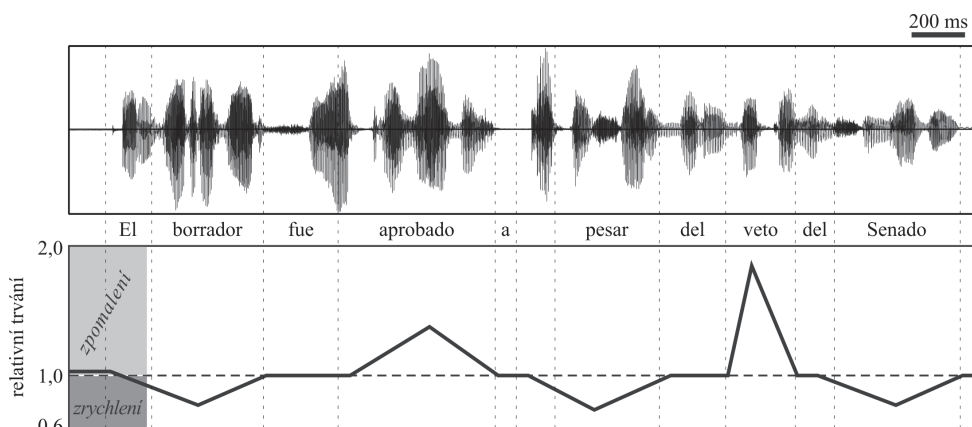
MATERIÁL

Pro náš výzkum byly použity nahrávky čtyř rodilých mluvčích kastilské španělštiny, kteří přečetli tutéž relaci španělských rozhlasových zpráv ze stanice Radio Praga. Konkrétně se jednalo o dva muže a o dvě ženy pocházející z různých oblastí Španělska, žijící delší dobu v České republice a pracující jako lektori španělského jazyka. Veškeré nahrávky byly pořízeny v nahrávacím studiu Fonetického ústavu Filozofické fakulty Univerzity Karlovy za použití kondenzátorového mikrofonu AKG C 4500 B-BC.

Z nahrávek rozhlasových zpráv byly vybrány 3–5 sekund dlouhé pasáže, přičemž bylo vždy dbáno na to, aby všechny tyto úseky byly syntakticky a sémanticky celistvé. Po obsahové stránce byly věty neutrální, povětšinou z oblasti politiky, kultury a regionálního dění v České republice. Vyvarovali jsme se jakýchkoliv pasáží, které by mohly svým obsahem v respondentech vyvolat emoce či je nasměrovat k přisouzení nějaké charakteristiky mluvčímu. Zároveň byly vybrány jen takové úseky, které byly mluvčími proneseny bez chyb, dysfluencí či jiných rušivých elementů.

Pro větší variabilitu hlasů jsme v programu Praat (Boersma — Weenink, 2015) využili možnost manipulace nastavení základní frekvence a formantů tak, že výsledný hlas vyvolával dojem opačného pohlaví (jedná se o funkci *Change gender*, která kombinuje změny vzorkovací frekvence s metodou PSOLA; viz Moulines — Charpentier, 1990). Protože u takovýchto globálních zásahů do originálního zvukového signálu mohou vznikat nežádoucí artefakty, byl tímto způsobem úspěšně vytvořen pouze jeden další, „mužský“ hlas. Celkově jsme tedy pracovali s pěti hlasy, třemi mužskými a dvěma ženskými, a s 35 položkami.

Následné manipulace simulující kolísání mluvního tempa byly prováděny rovněž v programu Praat, a to za pomoci algoritmu PSOLA, který umožňuje v průběhu promluvy měnit tempo řeči. Tyto manipulace je nutné dělat pečlivě, jelikož je třeba najít taková místa v nahrávce, kde je zachována modální fonace, PS v názvu algoritmu odpovídá slovům *pitch-synchronous*, což znamená, že dochází k manipulaci celých hlasivkových period. Nezbytným požadavkem tedy je, aby byla detekovatelná základní frekvence (F_0). Manipulace jsme proto prováděli manuálně, vždy v závislosti na povaze konkrétní nahrávky. Cílem manipulací bylo simulovat reálně znějící, plynulé temporální výkyvy, tedy zrychlení či zpomalení ovlivňující vždy celé slovo, popřípadě dvojice slov (např. předložku a podstatné jméno); příklad jedné položky ukazuje obrázek 1. Míra manipulací byla poměrově různá v závislosti na charakteru nahrávky a jednotlivých slov v ní, v průměru se však jednalo o lokální zpomalení o 70 % (= koeficient 1,7 na obrázku 1) a lokální zrychlení o 25 % (= koeficient 0,75) původního tempa.



OBŘÁZEK 1: Ukázka tvoření kolísavého tempa v programu Praat: zrychlování (tedy snížení relativního trvání) a zpomalování (zvýšení relativního trvání) slov ve větě.



Každá položka byla opakovaně a detailně poslouchána a konzultována; položky byly vybrány všemi členy výzkumného týmu tak, aby výsledný zvuk zněl zcela přirozeně a aby v něm nebyl patrný jakýkoli umělý zásah. Důležité je, že vytvořené varianty by hypoteticky mohly odpovídat reálnému řečovému projevu zdravého člověka.

Takto vytvořené nahrávky — 35 původních a 35 manipulovaných položek — pak sloužily jako stimuly v percepčním testu, jehož cílem bylo ověřit vliv popsaných temporálních modifikací na hodnocení osobnosti mluvčích. Kontrolované manipulace nahrávek nám umožnily využít tzv. techniku spojených masek (*matched-guise technique*; viz Lambert et al., 1960; Chromý, 2009), kdy posluchači posuzují dvojice položek, které se liší pouze provedenými manipulacemi.

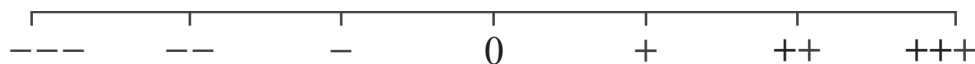
PERCEPČNÍ TEST A POSLUCHAČI

Cílem percepčního testu bylo zjistit, jak výše popsané temporální manipulace ovlivňují vnímání mluvčího posluchači, konkrétně nakolik se bude lišit hodnocení kompetence mluvčího. Obsahová stránka percepčního testu vycházela především z publikací zabývajících se pětifaktorovým modelem osobnosti (viz výše). Vytvořili jsme seznam tvrzení vztahujících se ke kompetenci mluvčího, tato tvrzení byla formulována pomocí klíčových adjektiv pro tyto osobnostní charakteristiky (např. *svědomitý*, *ctížádostivý* atd.) či pomocí činností, ze kterých tyto charakteristiky vyplývají (např. *často řeší povinnosti na poslední chvíli*).

Percepční test obsahoval dvojice původních ($n = 35$) a manipulovaných ($n = 35$) položek, pseudonáhodné pořadí přitom bylo stanoveno tak, aby původní a manipulovaná verze téže nahrávky byly umístěny co nejdále od sebe. Pro ověření konzistence hodnocení respondentů bylo 10 položek (5 původních a 5 manipulovaných) zopakováno. Percepční test tedy celkově zahrnoval 80 položek.

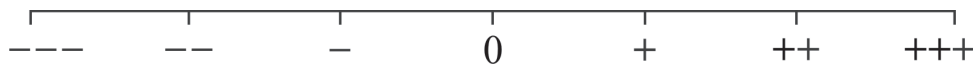
Ke každé dvojici položek (původní a odpovídající upravené) bylo přiřazeno stejné tvrzení. To znamená, že případné pozorované rozdíly v hodnocení při párovém srovnání skutečně vypovídají o vlivu provedené manipulace. Samotná hodnocení těchto výroků byla respondenty vyznačována na bipolární sedmistupňové Likertově škále, která se v podobných výzkumech v minulosti poměrně dobře osvědčila (viz např. Norman, 2010; Volín — Poesová — Skarnitzl, 2014). Jak ukazuje obrázek 2 a 3, středobodem této škály je nula a škála sahá od „---“ (= vůbec nesouhlasím) po „+++“ (= naprosto souhlasím). Tvrzení byla formulována pozitivním (obrázek 2) i negativním (obrázek 3) způsobem.

Tento člověk efektivně řeší úkoly.



OBRÁZEK 2: Příklad formulace pozitivního tvrzení a škály k jeho hodnocení.

Tato osoba má potíže s plněním svých povinností.



OBRÁZEK 3: Příklad formulace negativního tvrzení a škály k jeho hodnocení.

Takto strukturovaný percepční test byl zadán dvěma skupinám respondentů: dvaceti studentům hispanistiky na FF UK (16 ženám s průměrným věkem 25,4 let a 4 mužům s průměrným věkem 24 let). Všichni posluchači z této skupiny mají velmi pokročilou znalost španělského jazyka, dle evropského referenčního rámce na úrovni B2–C2. Respondenti studují jazyk minimálně od začátku středoškolského studia (tj. 5–10 let v závislosti na typu školy a věku posluchače). Drtivá většina z nich jsou absolventy bilingvních gymnázií a mají za sebou též studijní pobyty ve španělsky mluvících zemích, díky svému studiu jsou vysoce motivováni a na jazyku průběžně pracují. Dále byl test zadán stejnému počtu osob (12 ženám s průměrným věkem 23 let a 8 mužům s průměrným věkem 25,5 let), které španělský jazyk neovládají, a při hodnocení se tedy řídily pouze zvukovou, nikoliv obsahovou stránkou jazyka. Tato skupina respondentů byla tvořena bez výjimky studenty vysokých škol hovořícími minimálně jedním cizím jazykem (angličtina) na různě pokročilé úrovni. Pro zjednodušení budeme první skupinu v následujícím popisu nazývat „hispanisté“ a druhou skupinu „laičtí“ posluchači. Percepční test byl zadáván pomocí skriptu v programu Praat, a to individuálně v tiché místnosti do uzavřených sluchátek Sennheiser HD-201.

Test byl sestaven do čtyř bloků po 20 položkách, mezi jednotlivé bloky byla zařazena pauza ke krátkému odpočinku. Respondenti slyšeli každou položku dvakrát s přibližně půlvteřinovým odstupem, poté měli dostatek času (ISI = 2 sekundy) na zanesení hodnocení. Celý test včetně instruktáže a pauz trval přibližně 40 minut; každý z respondentů obdržel za svou účast finanční odměnu.

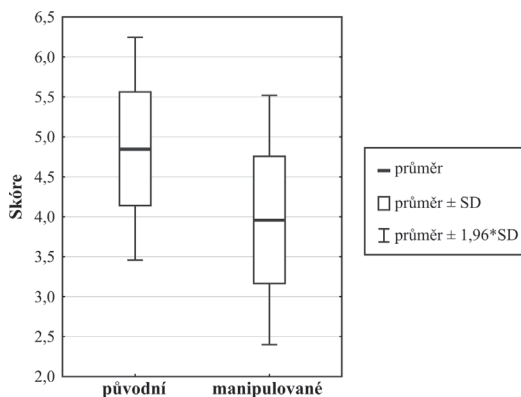
ZPRACOVÁNÍ VÝSLEDKŮ

Zanesené odpovědi byly převedeny na číselnou škálu 1–7. Okrajový bod odpovídající slovnímu vyjádření „vůbec nesouhlasím“ byl označen nejnižším číslem, opačný pól (slovně „naprosto souhlasím“) byl označen číslem 7. U negativně formulovaných tvrzení (např. „Mluví je v zaměstnání nesvědomitý a lenivý“) bylo samozřejmě třeba změnit polaritu hodnocení. Data byla dále zpracována v programu Statistica (StatSoft, verze 10). Statistická významnost dat byla ověřena pomocí analýzy rozptylu (ANOVA) pro opakovaná měření a pomocí post-hoc testů (t-testů pro opakovaná měření). Vzájemně porovnávána byla hodnocení původních a manipulovaných položek mezi oběma skupinami respondentů a dále dílčí hodnocení jednotlivých mluvčích.



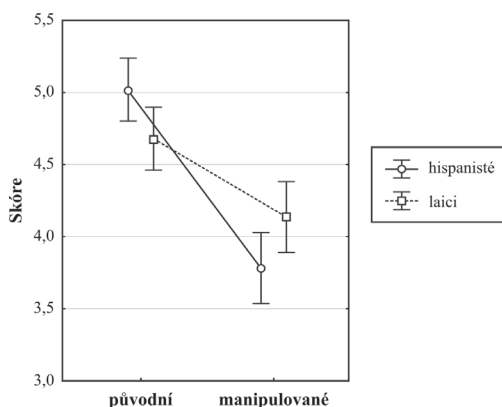
VÝSLEDKY A DISKUSE

Hodnocení původních a manipulovaných položek ukazuje krabicový graf na obrázku 4; rozdíl v hodnocení je vysoce významný, jak ukazuje výsledek t-testu pro opakovaná měření: $t(79) = 10,71, p < 0,0001$. Je patrné, že vnímaná kompetence mluvčího byla temporálními manipulacemi negativně ovlivněna.



OBRÁZEK 4: Hodnocení původních a manipulovaných nahrávek. Krabice zobrazuje jednu směrodatnou odchylku (SD) od průměru, vousy zahrnují 95 % všech dat. Hodnocení na stupnici od 1 (nejnižší míra kompetence) do 7 (nejvyšší míra kompetence).

Dále nás zajímá, nakolik se v hodnocení španělských mluvčích liší posluchači, kteří se španělštině věnují ve svém studiu („hispanisté“), a posluchači, kteří se španělštinou zkušenosti nemají („laičtí“ posluchači). Jak je zřejmé z obrázku 5, způsob hodnocení se skutečně liší a celkový vztah (interakce mezi typem položky a skupinou respondentů) je vysoce významný: $F(1, 78) = 22,09, p < 0,0001$.



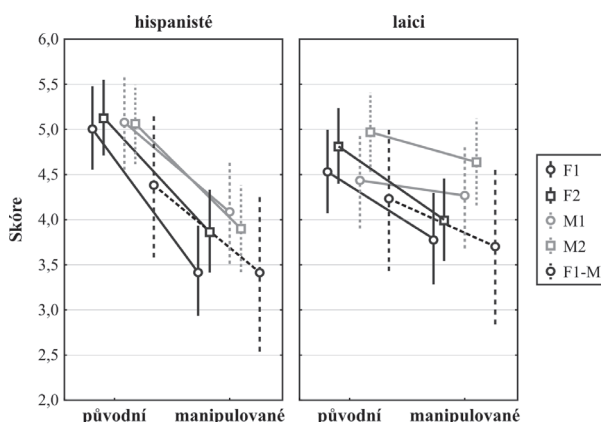
OBRÁZEK 5: Hodnocení původních a manipulovaných nahrávek hispanisty a laickými posluchači. Vousy zobrazují 95% interval spolehlivosti odhadu průměrné hodnoty. Hodnocení na stupnici od 1 (nejnižší míra kompetence) do 7 (nejvyšší míra kompetence).



Skupina respondentů rozumějících španělštině ve srovnání se skupinou laiků hodnotila původní nahrávky pozitivněji a manipulované nahrávky negativněji. Celkový rozdíl ve vnímání plynulých a kolísavých položek je tedy pro hispanisty markantnější (t -test pro opakovaná měření: $t(39) = 10,56, p < 0,0001$) než pro laiky, i pro ty je však stále vysoce významný ($t(39) = 6,04, p < 0,0001$). Hodnocení laických posluchačů blíže střední hodnotě oproti hispanistům je významné v případech původních ($t(78) = 2,20, p < 0,05$) i manipulovaných položek ($t(78) = 2,03, p < 0,05$). Je tedy zřejmé, že znalost prozodické stavby daného jazyka a cílového jazyka obecně ovlivňuje způsob, jakým posluchači posuzují osobnostní vlastnosti mluvčích.

Z obrázku 5 vyplývá, že hodnocení manipulovaných položek se pohybuje u obou skupin posluchačů okolo čísla 4. Jak je uvedeno výše, tato hodnota na škále odpovídá neutrálnímu bodu (v dotaznících byla značena jako 0), z čehož můžeme vyvodit několik závěrů. Položky s manipulací tempa jsou sice v relativním srovnání s původními nahrávkami hodnoceny hůře, avšak z absolutního umístění hodnocení kolem středového bodu škály se kolísání tempa pro posluchače nezdá úplně nepřijatelné: posluchači jsou v hodnocení poměrně tolerantní a vykazují určitou rezistentnost v ohledu na nepřijatelnost projevu. Zároveň to svědčí i o tom, že provedené temporální manipulace na posluchače nepůsobily nijak nepřírozeně, že manipulované položky stále zněly jako reálná řeč.

Dále jsme se zaměřili na hodnocení jednotlivých mluvčích. Z obrázku 6 je opět jasně patrné horší hodnocení položek s temporálním kolísáním a zároveň i kategoričtější hodnocení hispanistů oproti laickým posluchačům (úsečky spojující původní a manipulované položky konkrétního mluvčího jsou v levé části obrázku strmější než napravo). Při bližším porovnání grafů na obrázku 6 si dále můžeme všimnout, že skupina respondentů znala španělštinu hodnotila především nemanipulované nahrávky všech mluvčích podobněji. To může být způsobeno více faktory, jako pravděpodobný se jeví předpoklad, že skupina laických posluchačů se právě kvůli neznalosti obsahu či prozodie španělštiny orientuje více na celkový dojem z hlasu mluvčího, tedy například na barvu či výšku hlasu, fonační nastavení apod. Tyto faktory v naší studii kontrolovány nebyly.



OBRAZEK 6: Hodnocení původních a manipulovaných nahrávek hispanisty a laickými posluchači, rozdělené podle jednotlivých mluvčích. Vousy zobrazují 95% interval spolehlivosti odhadu průměrné hodnoty. Hodnocení na stupnici od 1 (nejnižší míra kompetence) do 7 (nejvyšší míra kompetence).



Výsledky nenaznačují žádný systematický vliv pohlaví mluvčích, jen u laických posluchačů je viditelná tendence k příznivějšímu hodnocení manipulovaných položek u mužských mluvčích. Poněkud odlišné je také hodnocení mluvčího F1-M, tedy původně ženského hlasu, který byl změněn na mužský. Ačkoli respondenti na základě osobních rozhovorů po skončení testu tento umělý zásah většinou nezaznamenali, hodnocení tohoto hlasu z hlediska kompetence bylo horší, především u položek, u nichž nebyly prováděny temporální manipulace.

Celkově se potvrdily obě dvě námi formulované hypotézy: posluchači hodnotili jako méně kompetentní a svědomité ty mluvčí, u jejichž projevu bylo patrné kolísání tempa. Tento jev se potvrdil u španělsky hovořících i laických hodnotitelů, s tím rozdílem, že první skupina vykazuje větší homogenitu a markantnější rozdíl mezi pozitivními hodnoceními u původních položek a negativními hodnoceními u kolísavých položek. To je pravděpodobně způsobeno znalostí prozodické a obsahové stránky jazyka, která jim umožnila sledovat tyto jevy lépe, na rozdíl od respondentů, kteří svůj dojem neměli možnost oprít o lexikální stránku a obeznámenost s prozodickou podobou španělského jazyka.

OBECNÁ DISKUSE A ZÁVĚR

Tato studie se zabývá dopadem změn v temporální struktuře řeči na vnímání osobnostních vlastností člověka. Jádrem výzkumu je experiment, který srovnává soudy posluchačů vůči mluvčím španělského jazyka s plynulým nebo kolísavým řečovým tempem. Jako základ ke zmapování tohoto jevu byl uplatněn v psychologii uznávaný pětifaktorový model osobnosti, z něhož jsme se zaměřili na vnímanou kompetenci mluvčích.

Výsledky prokázaly statisticky signifikantní odlišnosti ve vnímání původních (a z hlediska tempa pravidelných) a uměle manipulovaných (kolísavých) položek. To se potvrdilo u skupiny respondentů hovořících španělsky i u laické skupiny, ačkoliv hispanisté byli na změny mluvního tempa citlivější. Odlišnou citlivost na fonetické detaily v cizím jazyce nepřímo dokládají například i studie zaměřené na identifikaci mluvčího: schopnost rozpoznat mluvčího hovořícího cizím jazykem se ukazuje jako nižší než v případě mluvčích, kteří hovoří stejným jazykem jako posluchači (viz např. Wester, 2012). Obě pracovní hypotézy tedy byly potvrzeny.

Tato studie poskytuje další důkazy o tom, že temporální stránka řeči — tedy její pravidelnost a rytmičnost, případně očekávatelně realizované zpomalování a zrychlování — představuje významný faktor ovlivňující kvalitu mezilidské komunikace (Volín, 2010). Výchylinky oproti pravidelnému temporálnímu uspořádání na posluchače působí negativně, a mohou tak mluvčího stigmatizovat. V budoucím výzkumu se budeme snažit odhalovat další řečové faktory, které mohou stigmatizovat mluvčího v uších posluchačů nebo které naopak přispívají k úspěchu komunikačního procesu. Nové možnosti v oblasti počítačové manipulace řečového signálu, například tzv. morfovací nástroje jako STRAIGHT (Kawahara et al., 2008), představují slibný směr, který umožní kvalitní manipulaci většího množství akustických parametrů.

PODĚKOVÁNÍ

Tento výstup vznikl v rámci projektu Podpora vědeckého výzkumu studentů doktorandů v oblasti obecné lingvistiky, srovnávací jazykovědy a fonetiky na FF UK v roce 2015, podprojektu Temporální variace výpovědi ve španělštině a jejich vliv na vnímání kompetence mluvčího řešeného na Filozofické fakultě Univerzity Karlovy z prostředků Specifického vysokoškolského výzkumu na rok 2015. Radek Skarnitzl byl podpořen projektem GA ČR 14-08084S.

LITERATURA:

- ADDINGTON, David W. (1968): The relationship of selected vocal characteristics to personality perception. *Speech Monographs*, 35(4), s. 492–503.
- APPLE, William — STREETER, Lynn A. — KRAUSS, Robert M. (1979): Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology*, 37(5), s. 715–727.
- ARONOVITCH, Charles D. (1976): The voice of personality: stereotyped judgments and their relation to voice quality and sex of speaker. *Journal of Social Psychology*, 99(2), s. 207–220.
- BARRICK, Murray R. — MOUNT, Michael K. (1991): The Big Five personality dimensions and job performance: a meta-analysis. *Personnel Psychology*, 44(1), s. 1–26.
- BELIN, Pascal — BESTELMEYER, Patricia E. G. — LATINUS, Marianne — WATSON, Rebecca (2011): Understanding voice perception. *British Journal of Psychology*, 102(4), s. 711–725.
- BLATNÝ, Marek et al. (2010): *Psychologie osobnosti: Hlavní témata, současné přístupy*. Praha: Grada.
- BOERSMA, Paul — WEENINK, David (2015): *Praat: Doing Phonetics by Computer [verze 6.0]*. Cit. 28. 10. 2015. Dostupné z WWW: <<http://www.praat.org/>>.
- BROWN, Bruce L. — STRONG, William J. — RENCHER, Alvin C. (1973): Perceptions of personality from speech: effects of manipulations of acoustical parameters. *Journal of the Acoustical Society of America*, 54(1), s. 29–35.
- BROWN, Bruce L. — STRONG, William J. — RENCHER, Alvin C. (1974): Fifty-four voices from two: the effects of simultaneous manipulations of rate, mean fundamental frequency and variance of fundamental frequency on ratings of personality from speech. *Journal of the Acoustical Society of America*, 55(2), s. 313–318.
- CARGILE, Aaron Castelan — GILES, Howard (1997): Understanding language attitudes: exploring listener affect and identity. *Language and Communication*, 17(3), s. 195–217.
- CARNEY, Dana R. — COLVIN, C. Randall — HALL, Judith A. (2007): A thin slice perspective on the accuracy of first impressions. *Journal of Research in Personality*, 41(5), s. 1054–1072.
- COLLINS, Katherine A. — CLÉMENT, Richard (2012): Language and prejudice: direct and moderated effects. *Journal of Language and Social Psychology*, 31(4), s. 376–396.
- DIXON, John A. — MAHONEY, Berenice — COCKS, Roger (2002): Accents of guilt? Effects of regional accent, 'race' and crime type on attributions of guilt. *Journal of Language and Social Psychology*, 21(2), s. 162–168.
- EDWARDS, John (1999): Refining our understanding of language attitudes. *Journal of Language and Social Psychology*, 18(1), s. 101–110.
- GLUSZEK, Agata — DOVIDIO, John F. (2010): The way they speak: a social psychological perspective on the stigma of nonnative accents in communication. *Personality and Social Psychology Review*, 14(2), s. 214–237.



- GROSSBERG, Stephen (2003): Resonant neural dynamics of speech perception. *Journal of Phonetics*, 31, s. 423–445.
- GROSSBERG, Stephen (2013): Adaptive resonance theory: how a brain learns to consciously attend, learn, and recognize a changing world. *Neural Networks*, 37, s. 1–47.
- HŘEBÍČKOVÁ, Martina (1999): Obecné dimenze popisu osobnosti: Big Five v češtině. *Československá psychologie*, 43(1), s. 1–12.
- HŘEBÍČKOVÁ, Martina (2011): *Pětifaktorový model osobnosti: Přístupy, diagnostika, uplatnění*. Praha: Grada.
- CHROMÝ, Jan (2009): Postoje k jazyku, technika spojitých masek a čeština. *Naše řeč*, 92(5), s. 252–262.
- IMHOF, Margarete (2010): Listening to voices and judging people. *International Journal of Listening*, 24(1), s. 19–33.
- KAWAHARA, Hideki — MORISE, Masanori — TAKAHASHI, Toru — NISIMURA, Ryuichi — IRINO, Toshio — BANNO, Hideki (2008): Tandem-STRAIGHT: a temporally stable power spectral representation for periodic signals and applications to interference-free spectrum, FO, and aperiodicity estimation. In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2008, March 30 - April 4, 2008, Caesars Palace, Las Vegas, Nevada, USA*. Las Vegas, NV: IEEE, s. 3933–3936.
- LAMBERT, Wallace E. — HODGSON, Richard C. — GARDNER, Robert C. — FILLENBAUM, Samuel (1960): Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology*, 60(1), s. 44–51.
- LEV-ARI, Shiri — KEYSAR, Boaz (2010): Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, 46, s. 1093–1096.
- LINDEMANN, Stephanie (2003): Koreans, Chinese or Indians? Attitudes and ideologies about non-native English speakers in the United States. *Journal of Sociolinguistics*, 7(3), s. 348–364.
- LIPPI-GREEN, Rosina (2012²): *English with an Accent: Language, Ideology, and Discrimination in The United States*. London — New York, NY: Routledge.
- MCALÉER, Phil — TODOROV, Alexander — BELIN, Pascal (2014): How do you say 'hello'? Personality impressions from brief novel voices. *PLoS ONE*, 9(3), s. 1–9.
- MCCRAE, Robert R. — COSTA, Paul T. (1987): Validation of the Five-Factor Model of Personality across the instruments and observers. *Journal of Personality and Social Psychology*, 52(1), s. 81–90.
- MOHAMMADI, Gelareh — VINCIARELLI, Alessandro (2012): Automatic personality perception: prediction of trait attribution based on prosodic features. *IEEE Transactions on Affective Computing*, 3(3), s. 273–284.
- MOULINES, Eric — CHARPENTIER, Francis (1990): Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication*, 9(5–6), s. 453–467.
- NORMAN, Geoff (2010): Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*, 15(5), s. 625–632.
- NYGAARD, Lynne C. (2005): Perceptual integration of linguistic and nonlinguistic properties of speech. In: David B. Pisoni — Robert E. Remez (eds.), *The Handbook of Speech Perception*. Oxford: Blackwell, s. 390–413.
- POLZEHL, Tim (2015): *Personality in Speech: Assessment and Automatic Classification*. Heidelberg: Springer.
- POLZEHL, Tim — MÖLLER, Sebastian — METZE, Florian (2010): Automatically assessing acoustic manifestations of personality in speech. In: *Proceedings of IEEE Spoken Language Technology Workshop (SLT)*. Berkeley, CA: IEEE, s. 7–12.
- PURNELL, Thomas — IDSARDI, William — BAUGH, John (1999): Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology*, 18(1), s. 10–30.
- QUENÉ, Hugo — PORT, Robert (2005): Effects of timing regularity and metrical expectancy on spoken-word perception. *Phonetica*, 62(1), s. 1–13.

- QUENÉ, Hugo — VAN DELFT, L. E. (2010): Non-native durational patterns decrease speech intelligibility. *Speech Communication*, 52(11-12), s. 911-918.
- SMITH, Bruce L. — BROWN, Bruce L. — STRONG, William J. — RENCHER, Alvin C. (1975): Effects of speech rate on personality perception. *Language and Speech*, 18(2), s. 145-152.
- VOLÍN, Jan (2010): On the significance of the temporal structuring of speech. In: Markéta Malá — Pavlína Šaldová (eds.), ... *for thy speech bewrayeth thee (A Festschrift for Libuše Dušková)*. Praha: Filozofická fakulta Univerzity Karlovy v Praze, s. 289-305.
- VOLÍN, Jan — POESOVÁ, Kristýna — SKARNITZL, Radek (2014): The impact of rhythmic distortions in speech on personality assessment. *Research in Language*, 12(3), s. 209-216.
- WESTER, Mirjam (2012): Talker discrimination across languages. *Speech Communication*, 54(6), s. 781-790.



Zuzana Berkovcová | Fonetický ústav FF UK
<berkovcova.z@gmail.com>

Štěpánka Černikovská | Ústav románských studií FF UK
<stephanie@seznam.cz>

Radek Skarnitzl | Fonetický ústav FF UK
<radek.skarnitzl@ff.cuni.cz>