

Josef Vachek:

Poznámky k fonologii českého lexika.

Za deset let od svého zrodu vytrřibila fonologie značnou měrou své pojmy a zjemnila své metody. Jaký pokrok tu byl učiněn, vyplývá nejlépe ze srovnání fonologických příspěvků obsažených v prvním svazku *Travaux du Cercle linguistique de Prague* (1929) a kodifikujících výkladů zvěčnělého N. S. Trubeckého v sedmém svazku téže sbírky, vyšlém loni (viz o nich můj ref. v tomto ročníku LF.). Avšak pozornějšímu pohledu neujde, že praktické fonologické bádání, t. j. fonologický průzkum existujících jazyků, nezachovávalo ve všech úsecích svého pracovního pole stejné pracovní tempo.

Zvláště zřetelně se projevuje nepoměr mezi průzkumem fonémů a fonémových vztahů v jazykových systémech na jedné straně a průzkumem funkčního zatížení těchto fonémů v jazykových systémech na straně druhé. Tento nepoměr zvláště jasně ukáže malé srovnání. V prvním svazku pražských *Travaux* najdeme vedle studií B. Trnky, Trubeckého a B. Havránka¹, které se zabývají jednak otázkami obecně fonologickými, jednak vzájemnými vztahy fonémů ve fonologickém systému, také pojednání V. Mathesiuse² o fonologické struktuře české slovní zásoby, v kterém se srovnává funkční zatížení českých fonémů ve slovech o rozsahu jednoho až čtyř fonémů se zatížením fonémů německých ve slovech téhož rozsahu. Je pozoruhodné, že příspěvky prvního druhu zahajují dlouhou řadu prací o obdobných problémech, takže jejich bibliografie jde dnes již do set, kdežto k uvedenému pojednání Mathesiusovu se řadí — nehledě k jeho vlastním vývodům ve čtvrtém svazku *Travaux*³ — jen Trubeckého nečetné poznámky v ruské morfonologii⁴ a příslušné partie Trnkova fonologického rozboru moderní angličtiny⁵. Avšak Trubeckého zajímá jen fonologická stavba ruských morfémů, nikoli funkční zatížení ruských fonémů, a Trnka, který na rozdíl od Trubeckého látku důkladně vyčerpává, volí si za východisko nikoli anglická slova, nýbrž morfémy.

¹ B. Trnka, *Méthode de comparaison analytique et grammaire comparée historique*, TCLP I, 1929, 33—38; N. Trubetzkoy, *Zur allgemeinen Theorie der phonologischen Vokalsysteme*, ibid. 39—66, B. Havráněk, *Influence de la fonction de la langue littéraire sur la structure phonologique et grammaticale du tchèque littéraire*, ibid. 106—120.

² V. Mathesius, *La structure phonologique du lexique du tchèque moderne*, TCLP I, 1929, 67—84.

³ V. Mathesius, *Zum Problem der Belastungs- und Kombinationsfähigkeit der Phoneme*. TCLP IV, 1931, 148—152.

⁴ N. Trubetzkoy: *Das morphonologische System der russischen Sprache*. TCLP V—2/1934.

⁵ B. Trnka, *A Phonological Analysis of Present-Day Standard English*, *Práce z vědeckých ústavů filoz. fakulty univ. Karlovy XXXVII*, Praha 1935.

Nepoměr, na který tu bylo poukázáno, lze ovšem do značné míry vysvětliti tím, že průzkum fonologické struktury lexika předpokládá, že už byl v jazyce spolehlivě určen jeho fonologický systém; dokud se tak nestalo, jsou výsledky zkoumání v oboru struktury jazyka nutně provisorní, poněvadž každá pozdější korektura ve výkladu fonologického systému daného jazyka má nutně za následek i korektury ve výsledcích zkoumání na poli fonologie lexikálně strukturální. Ale na druhé straně je třeba uvážit, že nehotovost a nedefinitivnost výsledků je nutným znakem vši průkopnické vědecké práce, a proto může býti obava před ní jen částečnou omluvou dosavadnímu fonologickému zkoumání. Příští fonologické studium se bude musit více zabývat průzkumem fonologické struktury lexika, než tomu bylo dosud, už proto, že teprve po získání dat, která se jí týkají, bude možno charakterisovat jazyk po stránce fonologické zcela výstižně. Přispěti několika poznámkami k průzkumu fonologické struktury českého lexika je skromným cílem těchto řádků.

Zkoumal jsem fonologickou strukturu lexika moderní spisovné češtiny, pokud je složeno ze slov pětifonémových. Navázal jsem tak na výše uvedenou práci Mathesiusovu, která probrala slova o jednom až čtyřech fonémech. Moje pracovní východisko se liší od Mathesiusova hlavně tím, že upravuje výčet českých fonémů, jenž v práci Mathesiusově, psané v samých počátcích fonologického bádání, dnes přirozeně v některých bodech zastaral. Výčet českých fonémů, z kterého vycházím sám, je tento:

Samohláskové fonémy: *a, á, e, é, i, í, o, u, ú; ó* jen ve slovech rázu emocionálního a ve slovech synchronicky cizích ⁶.

Souhláskové fonémy: *p, b, t, d, l', d', k, f, v, s, z, š, ž, ř, c, č, h, ch, m, n, ň, r, l, j; g* jen ve slovech rázu emocionálního a synchronicky cizích, naproti tomu *f* i ve výrazech intelektuálních domácích (srv. *doufatí, houf, rozšafný*).

T. zv. dvojhlásky v češtině (*aj, ej, ij, oj, uj, áj, íj, új; ou*) jsou vesměs skupinami dvou fonémů ⁷; slabičné hlásky *r, l, m* jsou kombinatorními variantami neslabičných *r, l, m* ⁸.

Jinak jsem technicky, pokud to bylo možno, zachovával při práci obdobný postup jako Mathesius. I u mne je brán slovní materiál jen z jazyka spisovného (arci s přihlédnutím k těm hovorovým prvkům, jež v spisovném jazyce zdomácněly a necítí se v něm jako rušivé). Shodně s Mathesieusem sčítám i výrazy rázu emocionálního a slova zvukodobná, ale mimo to i běžná slova syn-

⁶ O cizích slovech se stanoviska synchronického psal V. Mathesius v *Časopise pro moder. filologii* roč. 18, 1932, 231 nn.

⁷ Srv. J. Vachek, *Über die phonologische Interpretation der Diphthonge*, *Práce z věd. ústavů fil. fakulty univ. Karlovy XXXIII*, 1933, str. 87 nn.

⁸ Srv. J. Vachek, *Fonologický poměr hlásek i a j v češtině a slovenštině*, *Slavia* 11, 1932, 265 nn.

chronicky cizí a výrazy interjekční. Těchto posledních je však tak málo, že z nich nelze vyvozovat obecnějších závěrů. Shodně s Mathesieusem počítám dále za jediné slovo všechny flektivní tvary téhož substantiva nebo slovesa, pokud v nich nedochází k alternaci ve kmeni nebo ke změně v souhláskové skupině. Tak byly na př. výrazy *louka, louku, louky* počítány za slovo jediné, ale ovšem výraz *louce* za slovo samostatné. Homonyma byla, jako u Mathesiuse, počítána dvojmo po př. trojmo atd. Důsledně jsem přihlížel k morfematické složenosti slov a také při zpracování výsledkových dat jsem morfémových hranic všude dbal. Dále jsem jako Mathesius soustavně rozlišoval slova domácí rázu intelektuálního od slov rázu emocionálního na jedné straně a od slov synchronicky cizích na straně druhé. V datech, která dále uvádím, je počet slov resp. fonémových skupin synchronicky cizích označován hvězdičkou u příslušné číslice, u slov emocionálních volím označení \times , u slov zvukodobných $^{\circ}$, u interjekcí i .

Slova jsem řadil do typů, jež označuji schematicky tak, že *a* značí kterýkoli samohláskový, *b* kterýkoli souhláskový foném (tak na př. slovo *unese* patří k typu *ababa*, slovo *nenos* do typu *babab*).

Pro nedostatek místa není zde možno uváděti jednotlivé doklady (je jich celkem na 11.000). Podávám proto alespoň přehled všech slovních typů v češtině zastoupených s počtem zjištěných dokladů:

Typ:	příklad:	počet dokladů:					celkem
		intel. dom.	ciz.	emoc.	zvukod.	interj.	
<i>aabaa</i>	uořou	1	2	0	0	0	3
<i>aabab</i>	uoráš	2	7	1	0	0	10
<i>aabba</i>	uondá, ouška	1	2	1	0	0	4
<i>abaaa</i>	ideou	0	1	0	0	0	1
<i>abaab</i>	obout	2	4	0	0	0	6
<i>ababa</i>	unese	452	44	9	3	0	508
<i>ababb</i>	účast	65	17	2	0	0	84
<i>abbaa</i>	orvou	28	5	0	1	0	34
<i>abbab</i>	oprav	475	45	6	7	0	533
<i>abbba</i>	ostří	67	9	0	3	0	79
<i>abbbb</i>	ustrň	14	0	0	1	0	15
<i>baaba</i>	moucha, vyučí	89	16	27	9	0	141
<i>baabb</i>	couvl	7	0	2	4	0	13
<i>babaa</i>	radou	239	29	7	6	0	281
<i>babab</i>	nenos	3319	280	189	140	1	3929
<i>babba</i>	návsi	2045	136	190	122	6	2459
<i>babbb</i>	návrh	83	4	6	22	0	115
<i>bbaab.</i>	proud	28	2	4	1	0	35
<i>bbaba</i>	prodá	1409	54	43	50	1	1557
<i>bbabb</i>	prost	196	5	13	24	6	244

<i>bbbaa</i>	mstou	30	0	1	1	0	32
<i>bbbab</i>	strop	397	0	19	19	0	465
<i>bbbbb a</i>	stkví, vrhli	160	0	8	34	0	202
<i>bbbbb b</i>	smršť, svrhl	20	0	1	6	0	30
Celkem		9121	662*	532 ×	483 ⁹	14 ⁱ	10812

Nejsou zastoupeny typy: *aaaaa*, *aaaab*, *aaaba*, *aaabb*, *aabbb*, *baaaa*, *baaab*, *baaaa*.

Úhrnný počet našich dokladů převyšuje tedy 3.625krát český materiál Mathesiusův (2800 slov, ovšem bez slov cizích, která M. nesčítal); mimo to je třeba uvážiti, že by se Mathesiusův materiál ještě zmenšil, kdyby z něho byla vypuštěna slova obsahující české dvojhhlásky, která M. počítal jako fonémy jednoduché, jež se však dnes vykládají jako skupiny fonémů dvou. I počet uskutečněných slovních typů je u slov pětifonémových větší než u slov jednofonémových až čtyřfonémových dohromady (24 proti 22; zde ovšem opět upravuji data Mathesiusova materiálu ve smyslu nového fonologického hodnocení českých dvojhhlásek).

Číselný vzestup materiálu i slovních typů se přirozeně vysvětluje větším počtem kombinačních možností při pěti fonémech, než byl počet možností u fonémů jednoho až čtyř. Jakou měrou tyto možnosti vzrůstají, ukáže malé srovnání. Mathesius na str. 83 cit. čl. dovedl, že čeština používá svých kombinačních možností ve skupině nejpočetnějších typů z jeho materiálu (jde o typy *bab*, *abab*, *baba*, *babb*, *bbab*) tak, že se z těchto možností realizují 3.1%. Ptáme-li se, jak je využito kombinačních možností češtiny u slov pětifonémových a zkoumáme-li za tím účelem poměry u nejpočetnějšího z pětifonémových slovních typů (*babab*), docházíme k závěru, že celkový počet možností je $24 \cdot 9 \cdot 24 \cdot 9 \cdot 17^9$, t. j. 793.152, takže možností uskutečněných (3649)¹⁰ je jen 0.44%! — Ještě mizivější zlomek procenta je uskutečněn z kombinačních možností těch typů, které se začínají fonémem samohláskovým. U nejpočetnějšího z nich (*abbab*) je z celkového počtu možností $9 \cdot (17 \cdot 24 + 7 \cdot 17) \cdot 9 \cdot 17$, t. j. 725.679, využito (bez cizích slov) jen 488, t. j. 0.06%. U druhého nejpočetnějšího typu se samohláskovým počátkem (*ababa*) je počet možností $9 \cdot 24 \cdot 9 \cdot 24 \cdot 9$, t. j. 419.904, kdežto počet realizací (zase bez cizích slov) jen 464, t. j. pouhých 0.11%. Naše čísla tedy ukazují, jak u slov pětifonémových proti slovům o menším rozsahu jedná vysoko stoupá počet teoreticky možných kombinací, jednak procentuálně hluboko klesá počet možností uskutečněných. Tím se velmi výrazně potvrzuje typický rys české fonologické soustavy, na který upozornil Mathesius (str. 84), že se v ní vy-

⁹ Ze samohláskových fonémů není bráno v počet *é*, ze souhláskových *g*, neboť nám jde jen o slova domáci.

¹⁰ Bez cizích slov (aby srovnání s daty Mathesiusovými bylo provedeno za stejných podmínek).

užívá bohatých fonologických možností s poměrně nepatrnou intenzitou.

Hořejší naše propočty však také znovu ukazují a potvrzují nechuť češtiny k vokalickému počátku slova, kterou poznal na svém materiálu i Mathesius. Stojí za povšimnutí, že jen dva nejpočetnější typy s počátkem souhláskovým (*babab*, *babba*) samy o sobě představují 59.45%, tedy téměř dvě třetiny veškerého našeho materiálu. Přes to, že počet typů se samohláskovým počátkem se téměř vyrovnává počtu typů s počátkem souhláskovým (11 : 13), v počtu slovních dokladů je mezi oběma typovými skupinami nepoměrně mnohem pronikavější (1277 : 9535 v neprospěch slov se samohláskovým počátkem, v procentech 11.8% : 88.2%). Nepřihlížíme-li při svém srovnávání k slovům cizím, u nichž je tento poměr vyjádřen čísly 20.5% : 79.5%, ukáže se samohláskový slovní počátek v nevýhodě ještě o něco větší (11.24% ku 88.76%). Za zmínku stojí, že ještě v menší oblibě než u slov intelektuálních je samohláskový počátek u slov emocionálních (3.6% : 96.4%) a zvukodobných (3.1% : 96.9%).

Nechuť češtiny k samohláskovému počátku slovnímu můžeme však sledovat do dalších podrobností. Již Mathesius upozornil, že velká většina českých domácích slov s počáteční samohláskou se počíná předponovými morfémy *j. o-, u-, ú-* (str. 79). Jak vypadá po té stránce náš materiál? Přihlédneme-li zase k typu *abbab* jako nejpočetnější skupině se samohláskovým počátkem, shledáváme, že slova počínající se na *o-, u-, ú-* jsou v něm v drtivé převaze proti slovům s počátečním *a-, e-, i-* (92.31 : 7.69%), bereme-li v počet i slova cizí; bez nich je poměr 98.56% : 1.44%! Uvnitř typů s počátečním *o-, u-, ú-* pak konstatujeme velkou převahu slov, jež se začínají předponovými morfémy *o-* (po případech *ob-, od-*), *u-, ú-*: číselný poměr těchto slov k slovům, kde *o-* resp. *u-, ú-* patří ke kmeni, je výmluvný. U slov s počátečním *o-* je to 84.17% : 15.83% (počítáme-li i slova cizí, tedy 81.69% : 18.31%), u počátečního *u-* 84.11% : 15.89% (s cizími slovy 83.01% ku 16.99%), u počátečního *ú-* pak 78.12% : 21.88% (cizích slov v této skupině není). K tomu je třeba ještě poznamenati, že značný podíl na skupině slov s kmenovým *o-* mají slova, kde morfémová hranice kdysi bývala, ale synchronicky není patrná (j. *otruby*, *otrok*, *oblek*) a pak různá slova rázu terminologického, jež i jinak ve své fonologické stavbě jeví různé zvláštnosti (na př. *orlem*, *octem*, *oměj* — fonol. *omňej* atp.). Jeví se tedy vskutku samohláskový slovní počátek jako něco češtině velmi nezvyklého.

Jiná zajímavá drobnost, týkající se výskytu českých samohláskových fonémů, vyplývá ze zkoumání, jak se spojuje v našem materiálu na počátku slova foném souhláskový s fonémem samohláskovým. Vyjděme z dat zjištěných pro naše dva nejpočetnější typy, jež se počínají spojením souhlásky plus samohláskou! Jsou to typy *babab* a *babba*, o nichž tu již jednou byla řeč. Podrobný

propoččet ukazuje, že ve slovech intelektuálních domácích se s počátečním souhláskovým fonémem nejčastěji spojuje samohláskový foném *e* (1036 případů), pak *o* (978), *a* (929), *i* (809), dále *í* (509), *á* (505), *u* (468), *ú* (80) a konečně *é* (50). Sestavíme-li podobnou stupnici pro slova cizí, najdeme v jejím čele *a* (113 dokladů), po němž následuje *i* (77), dále *o* (74), *e* (63), *u* (38), *ó* (23), *á* (15), *í* (7), konečně pak *ú* a *é* (po 3 dokladech). Naproti tomu u slov emocionálních je v čele tabulky *u* (138) sledované *a* (118), dále *e*, *o* (po 38), *i* (26), *á*, *í*, *ú* (po 7); pro *é*, *ó* není dokladů. Podobná je situace i u výrazů zvukodobných: vede *u* (76 doklady) před *e* (50), dále následují *a* (47), *i* (41), *í* (26), *o* (12), *á* (10), *é*, *ó*, *ú* bez dokladů.

Srovnání těchto čtyř stupnic zajímavě ukazuje důležitou úlohu samohlásky *u* ve struktuře slov povahy neintelektuální; kdežto ve stupnicích pro slovní zásobu intelektuální zaujímá místo sedmé (u slov domácích) resp. páté (u slov cizích), ve stupnicích pro slova rázu emocionálního a zvukodobného bezpečně vede¹¹.

Pokud se týká souhláskových fonémů, jež stojí na počátku slov v těchto nejpočetnějších typech, lze u domácích slov intelektuálních (celkový počet jejich je 5364) konstatovati největší funkční zatížení u *p* (567), za nímž následují *v* (548), *n* (489), *l* (342), *k* (338) a *z* (324). Nutno ovšem míti na paměti, že počet dokladů u *p*, *v*, *n* a *z* značně vzrůstá tím, že mnohá ze slov jimi se začínajících obsahují předpony j. *po-*, *na-*, *ne-*, *za-* atp. Nejméně domácích intelektuálních slov se počíná fonémy *f* (28), *ň* (55), *l'* (57), *d'* (80), *ch* (90) a *ř* (98). — Ze slovní zásoby synchronicky cizích slov intelektuálních (celkem 416 slov) nejvíce se počíná fonémem *p* (43), pak následují *r* (35), *m* (33), *k* (32), *s* (30) a *f*, *t* (po 29). Je však třeba zdůrazniti, že relativně nejvíce cizích výrazů je — nehledíme-li ovšem ke 20 výrazům počínajícím se na *g-*, proti nimž není vůbec slov intelektuálních domácích — u slov s počátečním *f-* (29 z 81, t. j. 35.80%). Nejméně cizích výrazů lze konstatovati při počátečním *d'*, *l'*, *ň*, *ř* (0), *č* (1), *ž* (3).

Slovní zásoba emocionální (celkem 379 slov) má vedoucí souhláskový foném *v*, jímž se počíná 44 slov; za ním následují *k* (40), *h* (38), *p* (30), *l* (26) a *š* (25). Relativně nejpočetnější je však v této skupině *d'* (22 výrazů z 80), t. j. 27.5 %, sledováno absolutně vedoucím *č* (44 ze 180, t. j. 24.44%). Nejméně emocionálních výrazů se počíná *ř* (0), *v* (1), *n* (2), *s*, *z* (po 3). — V zásobě slov zvukodobných (celkem 262 slov) vede *c* (28 výrazů) před *l'* (27), *b* (25), *h*, *r* (po 21), *ř* (16). Relativně nejvíce slov zvukodobných

¹¹ V jiné souvislosti jsem (v těchto LF. 40, 1933, 312 n.) statisticky potvrdil pozorování Mathesiusovo (viz Naše řeč 15, 38 n.), že se v slovech této kategorie spojují samohl. *o*, *u*, *ú* zvláště s předcházejícími souhláskami „měkkého“ typu (t. j. *č*, *š*, *ž*, *ř*, *c*, *j*, *d'*, *l'*, *ň*).

se počíná rovněž *l* (27 z 94, t. j. 28.72%). Na dolních místech této stupnice najdeme *d'* a *l* (0), *č* (1), *ch* (2), *ň* (3) a *v* (4).

Zajímavá jsou i fakta, týkající se souhláskových skupin počátečních i koncových. Pro nedostatek místa se zde musíme omezit na pouhá číselná data. Jak je přirozeno, mají počáteční skupiny souhláskových fonémů velkou početní převahu nad skupinami koncovými. Počátečních skupin dvou souhláskových fonémů shledal jsem v příslušných typech (*bbaab*, *bbaba*, *bbabb*) celkem 161, t. j. $150 + 5^* + 3 \times + 3^0$; ze 150 v domácí intelektuální zásobě slovní je 20 skupin vždy rozděleno morfémovou hranicí, 28 se vyskytuje jak ve skupinách kompaktních, tak v rozdělených, zbytek 102 jen ve skupinách kompaktních. Skupiny vyskytující se jen ve slovech cizích, emocionálních a zvukodobných jsou vesměs kompaktní. — V počátečních skupinách třífonémových si musíme především povšimnout těch, které obsahují (na druhém místě) slabičnou variantu fonému *r* nebo *l*. Je jich celkem 113, z toho 83 ve slovech intelektuálních, 10 ve slovech emocionálních a 20 ve výrazech zvukodobných; cizích výrazů v příslušných typech vůbec není, tedy není ani skupin pro ně typických; všechny skupiny zde zjištěné jsou kompaktní. Mimo tuto skupinu jsou tu však také třífonémová spojení, jež slabičného vrcholu neobsahují. Těch je celkem 81 (t. j. $76 + 1 \times + 4^0$; ze 76 skupin v domácích intelektuálních slovech se 12 vyskytuje jen rozdělených morfémovou hranicí, 31 jak ve skupinách rozdělených, tak kompaktních, 33 pak jen v skupinách kompaktních; skupiny emocionální a zvukodobné jsou vesměs kompaktní). — Počáteční skupiny čtyřfonémové neobsahující slabičný vrchol jsou v mém materiálu jen 2 (obě intel. dom.). Skupiny obsahující slabičnou variantu *r* nebo *l* jsou tu ještě početnější než u skupin dvojfonémových. Celkem je jich 184, t. j. 57 kompaktních (t. j. $42 + 3 \times + 12^0$) a 127 rozdělených morfémovou hranicí (t. j. $103 + 4 \times + 20^0$). — Konečně u slov pětifonémových (typ *bbbbb*) je z jejich struktury patrné, že tu je kolik slov, tolik souhláskových skupin ($20 + 4 \times + 6^0$); ve všech arci najdeme slabičné *r* nebo *l* a jen dvě z nich jsou morfémově kompaktní; zbývajících 28 se skládá z více než jednoho morfému.

Koncové skupiny fonémů jsou mnohem méně početné než počáteční a jsou také jednotvárnější. Koncových skupin dvojfonémových je celkem 37, z toho slabičné *l* nebo *r* obsahují $15 + 2 \times$ skupiny (z tohoto počtu jen 3 skupiny jsou kompaktní, kdežto $12 + 2 \times$ jsou morfematicky rozděleny). Bez slabičného *l* po př. *r* vyskytuje se tu celkem 20 skupin, z nichž $13 + 5^*$ je kompaktních, 1 morfematicky rozdělená a 1 přípustná jak v podobě kompaktní, tak rozdělené. — Ze tří fonémů je složeno celkem 66 koncových skupin; z toho jen 3 (t. j. $2 + 1 \times$) neobsahují slabičné *r* nebo *l* (2 z nich jsou skupiny kompaktní, jedna morfematicky rozdělená). Ze 63 skupin třífonémových vyznačujících se slabič-

ným *r* nebo *l* je kompaktních $27 + 4^* + 11^0$, morfematically rozdělených $12 + 5 \times + 4^0$. Konečně pak skupin čtyřfonémových, jež se vyskytují na konci slov, je v našem materiálu tolik, jako slov typu *abbbb*, totiž 15; z nich ani jedna není bez sonantního *r* resp. *l*; 12 z nich, t. j. $11 + 1^0$, je morfematically rozděleno, 3 jsou kompaktní.

Výčet a rozbor těchto fonémových skupin by ukázal četné další zajímavosti. Není možno provésti jej v rámci tohoto článku; vrátíme se k němu, stejně jako k jiným otázkám fonologie českého lexika, při jiné příležitosti.