

Poznámky ke str. 11 (Jak probíhá syntéza řeči)

1) Nejčastěji se syntéza řeči provádí převodem psaného textu do mluveného (tzv. **text-to-speech**, TTS).

2) Převod ale nelze provést přímo, protože běžný pravopis nezaznamenává přesně mluvenou podobu jazyka. Či přesněji řečeno: běžný pravopis je **nejednoznačný**, protože grafémy mohou odpovídat různým fonetickým prvkům podle významu nebo kontextu užití slova. Tuto nejednoznačnost známe z angličtiny, ale týká se i češtiny (byť se nám to nemusí zdát) a pravděpodobně všech rozvinutých jazyků.

Grafém = grafická značka užitá v písmu. Může to být písmeno, ale také diakritika (čárka, háček) nebo interpunkce (tečka, vykřičník), popř. jiné značky podle povahy písma (srov. čínské písmo, egyptské hieroglyfy nebo značky jako %, +, číslice 1, 2, 3 atd.).

Fonetickým prvkem se myslí jakýkoliv zvukový prvek jazyka nebo jeho část. Na nejmenší úrovni sem patří hlásky (též nazývané *fony*, jednotné číslo *fon*). Dále sem patří přízvuk nebo intonace. Nejsou to konkrétní zvuky, ale jejich zobecnění/abstrakce.

3) Kvůli zmíněné nejednoznačnosti běžného písma se jako východisko pro TTS užívá **fonetická transkripce**. Je to soubor grafémů, které mají **jednoznačnou** hodnotu, tj. odpovídají jednomu a právě jednomu fonetickému prvku. Zaznamenává se do hranatých [] (někdy též lomených //) závorek.

Na stránce 11 jsou uvedeny příklady z přepisu pomocí mezinárodní fonetické transkripce IPA (International Phonetic Alphabet), což je v současnosti nejužívanější způsob transkripce. Schémata ukazují příklady přepisu souhlásek a samohlásek. Podrobnosti na internetu.

Poznámky ke str. 12–13 (Psaný jazyk ≠ mluvený jazyk)

4) Důvody nejednoznačnosti pravopisu (tj. neshody psané a mluvené podoby jazyka) vyplývají především z rozdílné rychlosti změn pravopisu oproti výslovnosti. Přinejmenším od vynálezu knihtisku se psaný/tištěný jazyk mění pomaleji než mluvený jazyk, protože „co je psáno, to je dáno“.

V případě češtiny nebo angličtiny je pravopis více méně už několik století neměnný. Avšak především se výslovnost angličtiny silně změnila.

5) Navíc zvukové změny jsou vlastně neviditelné, a to mimo jiné i proto, že máme svůj jazyk silně spjatý s pravopisem, takže když se pravopis nemění, myslíme si, že se nemění ani výslovnost.

Příklady výslovnostních změn, které probíhají v češtině: a) hláska [l] se zabarvuje do [u] (tzv. velarizované [l]); b) hláska [r] je vyslovována s jedním kmitem jazyka oproti více kmitům v dřívější výslovnosti. Oba příklady jsou potvrzeny srovnáním současné výslovnosti se staršími nahrávkami (či popisy) těchto hlásek.

6) Příklady ukazují nejednoznačnost běžného pravopisu češtiny a angličtiny.

7) V následujících příkladech se podtržené grafémy (včetně mezery) běžně nevyslovují, ačkoliv v jiných slovech a za jiných podmínek se vyslovují. V anglických příkladech se zde nevyslovují nikdy, v českých mohou být vysloveny ve velmi pečlivé (a mnohdy i nepřírozeně znějící) mluvě.

čeština: *jsí*, *leccos*, *srđce*

angličtina: *late*, *knight*, *write*

mezery (!): *dá_mi* × *dámy*

8) Pravopis je nejednoznačný, protože nezaznamenává úplně všechny fonetické prvky, které vyslovujeme. Příkladem je přízvuk. Uvedená anglická slova se liší tím, že podstatné jméno má přízvuk na první slabice, sloveso na druhé. České příklady se liší tím, že v izolovaném slově *stole* je přízvuk na slabice /sto/, zatímco v předložkovém spojení *na stole* je přízvuk na slabice /na/.

přízvuk: angl. *import*_{podst. jm.} × *import*_{sloveso}, č. *na stole* × *stole*

9) Dalším příkladem jsou hlásky, které pravopis češtiny nezaznamenává, ačkoliv je běžně vyslovujeme. Je jím tzv. ráz v předložkovém spojení *v oku* nebo *s uchem* (je to něco jiného než *suchem* < *sucho*) nebo hláska [j] ve slovech jako *Marie*.

některé hlásky: [fʔoku] *v oku*, [marije] *Marie*

10) Některé grafémy mohou odpovídat více různým fonetickým prvkům (fonům, tj. hláskám).

Posloupnost písmen *ou* může odpovídat dvojhlásce (*pouček*, tj. malý *puk*) nebo dvěma samohláskám (*pouček*, 2. pád, mn. č. < *poučka*, tj. < *učit*).

pouček [pouček] × [poʔuček]

Posloupnost písmen *io* může odpovídat hláskám [jo] (*ion*) nebo [ijo:] (*špion*; dvojtečka značí dlouhou samohlásku).

ion [jo:n] × *špion* [špijo:n]

Písmeno *a* může odpovídat hlásce [a] (v domácích slovech jako *baby* < *baba*) nebo skupině hlásek [ej] (v přejatých slovech jako *baby*, z angličtiny).

baby [babi] × [bejbi]

11) V běžném pravopisu někdy zapisujeme celé pojmy spíše než zvuky. Proto značka % má jinou výslovnost v češtině než v angličtině. A dále tato značka v češtině může odpovídat podstatnému jménu *procent* (100 % studentů) nebo přídavnému jménu *procentní* (100% úspěšnost).

100 % studentů × 100% úspěšnost

12) Posloupnost číslic se může vyslovovat různým způsobem podle kontextu užití. Srovnej:

777 123 567 = sedm set sedmdesát sedm milionů sto dvacet tři tisíc pět set šedesát sedm (jako hodnota v matematickém příkladu)

×

777 123 567 = tři sedmičky sto dvacet tři pět set šedesát sedm (jako telefonní číslo; existují i jiné způsoby výslovnosti – takto bych to vyslovil já)

13) V běžném pravopise používáme zkratky, které mohou mít jiný význam a hlavně jinou výslovnost podle kontextu užití.

sv. Tomáš [tj. svatý Tomáš] × sv. první [tj. svazek první]

IPA [ipa] × [i:pe:a:] – první výslovnost bývá častější pro označení druh piva (Indian Pale Ale), ta druhá pro mezinárodní fonetickou transkripci (International Phonetic Alphabet)

14) Má-li TTS proběhnout zcela automaticky, tedy je nutné, aby stroj (počítač) věděl, jaký je význam nebo kontext užití konkrétních slov nebo grafémů. Mnoho takových případů (vlastně všechny výše uvedené) lze samozřejmě stroj naučit rozpoznávat, ale jako mezikrok je jednodušší použít fonetickou transkripci. Užití fonetické transkripce také dovoluje lepší zpětnou kontrolu správnosti umělé/syntetizované řeči.