



# Manuál pro práci s nástrojem Prokus

Ing. Markéta Jůzová

27.6.2016

# Obsah

<b>1</b>	<b>Úvod</b>	<b>1</b>
1.1	Požadavky na práci s nástrojem <i>Prokus</i> . . . . .	1
<b>2</b>	<b>Stručný úvod do syntézy řeči konkatenací metodou</b>	<b>2</b>
2.1	Zpracování textu pro účely syntézy řeči . . . . .	2
<b>3</b>	<b>Práce s nástrojem Prokus</b>	<b>3</b>
3.1	Přihlášení . . . . .	3
3.2	Základní nabídka . . . . .	3
3.3	Základní použití nástroje <i>Prokus</i> . . . . .	3
3.3.1	Výběr hlasu . . . . .	3
3.3.2	Zadání požadovaného textu . . . . .	4
3.3.3	Pravidla pro psaní textu . . . . .	6
3.3.4	Syntéza věty . . . . .	7
3.4	Možnosti práce se syntetizovanou promluvou . . . . .	7
3.4.1	Možnosti přehrávání . . . . .	8
3.4.2	Prohlížení signálu – posun, přiblížení . . . . .	8
3.4.3	Uložení vygenerované promluvy . . . . .	8
3.5	Úpravy syntetizované promluvy . . . . .	8
3.5.1	Výskyt řečového artefaktu . . . . .	9
3.5.2	Nahlašování artefaktů . . . . .	10
3.5.3	Vygenerování alternativ . . . . .	11
3.5.4	Zkoušení různých alternativ . . . . .	12
3.5.5	Jak vybrat „správnou“ jednotku . . . . .	12
<b>4</b>	<b>Závěr</b>	<b>13</b>

# 1 Úvod

Tento dokument slouží jako návod pro práci s nástrojem *Prokus*, který je součástí nástroje *SpeechLab* umístěném na webové adrese <https://speechlab.zcu.cz/> (vytvořeno Ing. Jakubem Vítem v roce 2016). Nástroj *Prokus* slouží k používání systému syntézy řeči ARTIC a spravování jeho řečových databází, vyvíjeném na KKY/NTIS Fakulty aplikovaných vět Západočeské univerzity v Plzni ve spolupráci s firmou SpeechTech. Umožňuje:

- syntetizování promluvy zvoleným řečníkem
- grafické znázornění průběhu signálu a spektra
- přehrání syntetizované promluvy nebo její části
- znázornění hranic jednotek a míst napojení
- vygenerování seznamu alternativních jednotek
- nahrazení vybrané jednotky jinou jednotkou
- uložení syntetizované promluvy

## 1.1 Požadavky na práci s nástrojem *Prokus*

Pro práci s nástrojem *Prokus* je nutné používat jeden z následujících webových prohlížečů (v ostatních prohlížečích nezaručujeme plnou funkčnost):

- Google Chrome
- Chromium

Dále je samozřejmě zapotřebí připojení počítače k internetu.

Těm, kteří se syntézou řeči nezabývají, doporučuji pro pochopení práce s nástrojem *Prokus* přečtení následující sekce. Ostatní mohou přejít na sekci 3, která popisuje možnosti a způsob práce s tímto nástrojem.

## 2 Stručný úvod do syntézy řeči konkatenáční meto- dou

Pojmem *syntéza řeči* rozumíme převod psaného textu do zvukové podoby, který probíhá automaticky pomocí počítače, mobilního telefonu či tabletu. Metod syntézy řeči existuje více, protože ale tento popis slouží k pochopení práce s nástrojem *Prokus*, popíšeme zde pouze jednu z nich – metodu konkatenáční.

Základem této metody je namluvení velkého počtu vět (tisíce) zkušeným řečníkem. Tyto věty jsou poté anotovány (je jim přiřazen text, který řečník přečetl) a segmentovány (rozděleny na krátké řečové úseky – jednotky), a to tak, že každá jednotka začíná v polovině jedné hlásky a končí v polovině hlásky následující (nazýváme je *difóny*). Všechny tyto jednotky jsou uloženy do databáze a slouží k tvorbě syntetizované řeči. Pro jednotlivé nahrané promluvy jsou dále spočteny různé charakteristiky, které se využívají během syntézy.

Samotná syntéza probíhá následovně: Zadaná věta se převede do výslovnostní podoby a vytvoří se z ní posloupnost jednotek. Pro každou jednotku jsou v databázi nalezeny všichni kandidáti, kteří jsou k dispozici, a pomocí algoritmu se z těchto kandidátů vybere pro každou jednotku jeden tak, aby kandidáti jednotlivých jednotek na sebe „dobře navazovali“ a aby jejich spojením vznikla odpovídající syntetická řeč (např. s klesavou intonací na konci oznamovací věty).

Výstupy této metody se v současné době považují za nejpřirozenější řečovou syntézu. Přesto ale občas nastane situace, kdy je napojení dvou po sobě jdoucích jednotek rušivé pro posluchače, což může způsobit třeba větší „skok“ v intonaci či energii, někdy určitá jednotka zní v daném kontextu nepřirozeně. A právě pro opravy těchto *řečových artefaktů* vznikl nástroj *Prokus*.

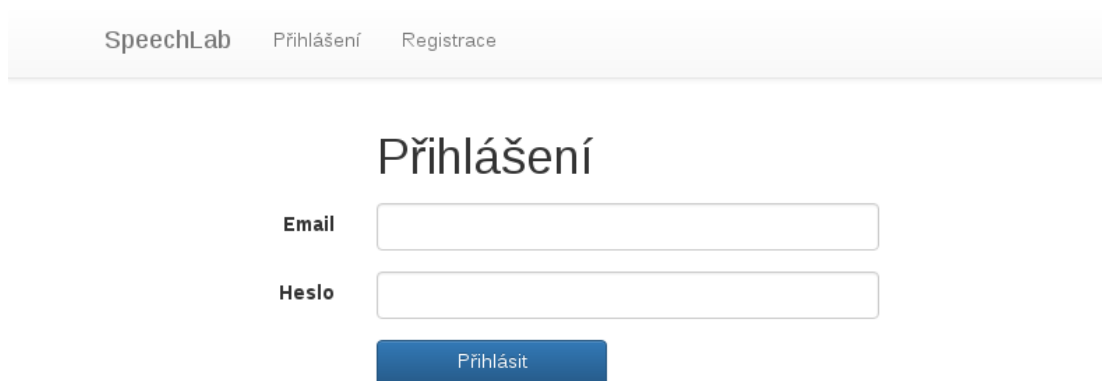
### 2.1 Zpracování textu pro účely syntézy řeči

Přepisu syntetizované věty do výslovnostní podoby předchází ještě tzv. *normalizace textu*, tedy přepis číslic a dalších nepísmenných znaků do slovní podoby, rozepisování zkratk apod. Český jazyk je pro tuto úlohu ale poměrně obtížný, neboť vyjadřuje jednotlivé gramatické funkce pomocí časování a skloňování. Přestože jednoduchou normalizaci textu systém syntézy řeči ARTIC provádí, stále má problémy se správným skloňováním jednotlivých výrazů.

## 3 Práce s nástrojem Prokus

### 3.1 Přihlášení

Pro práci s nástrojem Prokus je nutné mít založený uživatelský účet a přihlásit se. Na <https://speechlab.zcu.cz/> se po zvolení tlačítka *Přihlášení* objeví přihlašovací formulář, viz obr. 1.



The image shows a web interface for logging in. At the top, there is a navigation bar with the text "SpeechLab" on the left and "Přihlášení" and "Registrace" on the right. Below this, the main heading "Přihlášení" is centered. Underneath the heading, there are two input fields: "Email" and "Heslo". Below these fields is a blue button labeled "Přihlásit".

Obr. 1: Přihlašovací formulář.

### 3.2 Základní nabídka

Po přihlášení se zobrazí základní nabídka nástroje *SpeechLab*. Nás z ní bude zajímat pouze nástroj *Prokus*, který spustíte vybráním položky **Prokus** v horní nabídce (obr. 2).

### 3.3 Základní použití nástroje *Prokus*

Základní funkcí nástroje *Prokus* je syntéza zadané věty, nebo-li její převod do zvukové podoby.

#### 3.3.1 Výběr hlasu

Prvním krokem je výběr hlasu. *Prokus* nabízí několik mužských a ženských českých hlasu, jeden ženský slovenský a jeden ženský ruský hlas a dva hlasy anglické. Hlas vyberete z nabídky kliknutím myši a jeho popis se zobrazí v okýnku **Hlas**.

SpeechLab Prokus Artefakty Řečový inventář Nahrávání Markéta Jůzová ▾

## Prokus

**Hlas**

/artic\_images/spkr\_AJ.rev640.img ▾

**Text věty**

**Fonetická transkripce**

Provést syntézu

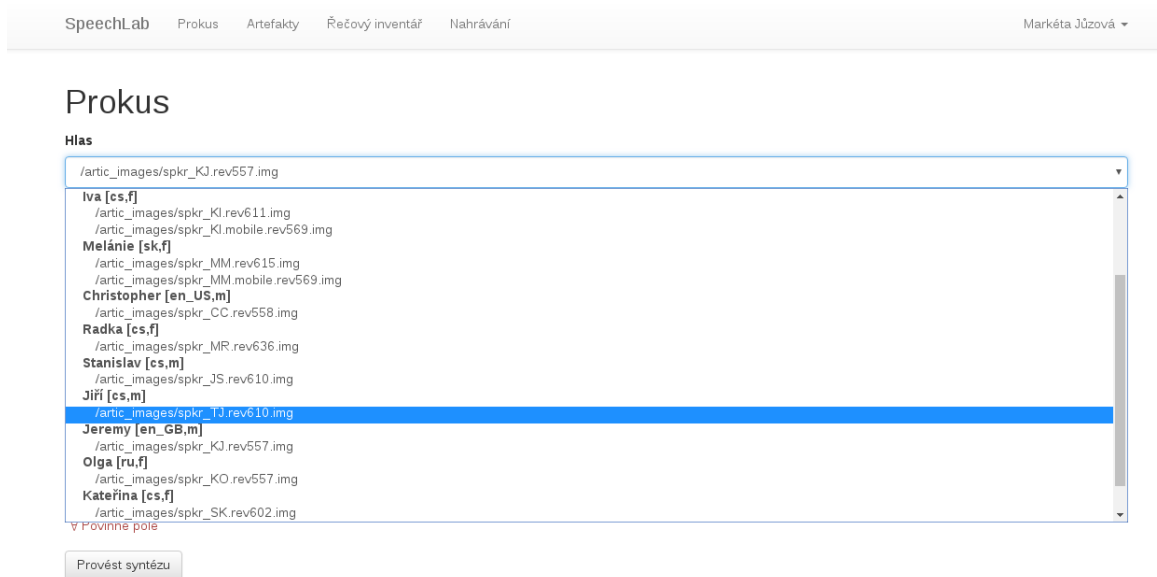
Obr. 2: Spuštění nástroje *Prokus*.

U některých z nich je navíc k dispozici více verzí (revizí), ze kterých je možné vybírat – např. pro český ženský hlas **Iva** je v čase vytváření tohoto dokumentu k dispozici verze `rev611` a `mobile.rev569` (viz obr. 3). Obecně platí, že vyšší číslo revize označuje novější verzi daného hlasu. Pro získání co nejkvalitnější syntetizované promluvy je rovněž vhodné **nevybírat** redukované verze pro mobilní zařízení (v popisku slovo *mobile*), které obsahují menší počet jednotek, a proto výsledná promluva většinou nezní tak dobře, jako při použití „normální“ verze (`/artic_images/spkr_XX.revYYY.img`).

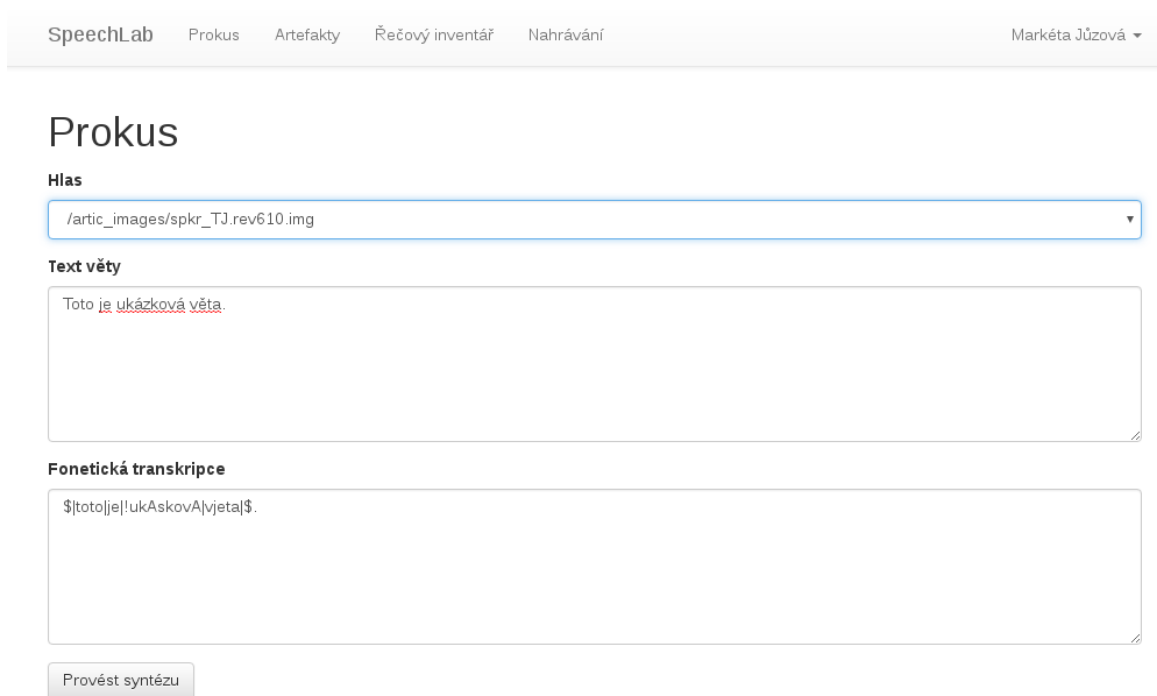
### 3.3.2 Zadání požadovaného textu

Po výběru hlasu zadejte do pole **Text věty** text, který chcete vysyntetizovat, a to včetně čárek a koncové interpunkce. Můžete najdenou syntetizovat i více vět. V poli **Fonetická transkripce** následně uvidíte Vámi zadaný text převedený do fonetické (výslovnostní) podoby (obr. 4). Používané znaky české fonetické abecedy jsou uvedeny v příloze.

V tomto formátu jsou jednotlivá slova od sebe oddělena znakem |, na začátku a konci každé věty je znak \$ (pauza) a před čárkami se vyskytuje znak # (značící krátkou pauzu). Fonetická transkripce textu dále odráží fonetické jevy, jako je spodoba či ztráta znělosti (např. *s* v přepisu slova *ukázková*) či měkčení hlásek před písmenem *i/í*, přidání vkladného



Obr. 3: Výběr hlasu v nástroji *Prokus*.



Obr. 4: Psaní věty v nástroji *Prokus*.

j před ě (viz *vjeta*) apod. Doufáme, že zobrazení fonetické transkripce Vám pomůže při kontrole správné výslovnosti cizích slov, zkratek a čísel.

### 3.3.3 Pravidla pro psaní textu

Před provedením samotné syntézy doporučujeme alespoň rychlou kontrolu výslovnosti cizích slov a zkratk v poli **Fonetická transkripce**. Přestože je ale možné tuto transkripce upravovat, obecně to nedoporučuji – úpravy by měla provádět pouze osoba seznámená s používanou fonetickou abecedou a jednotlivými fonetickými jevy. Pokud výslovnost některého slova ze zadané věty podle Vás neodpovídá skutečnosti, je vhodnější upravit text zadaný v poli **Text věty** „tak, jak se čte“. To platí i pro „rozepisování“ různých zkratk, zvláště nejsou-li v prvním pádě. Jako příklad uveďme výraz

s Ing. Davidem Jonesem,

jehož fonetická transkripce bude vypadat pravděpodobně takto:

$\$|s|!inZenIr|davidem|jonesem|\$.^1.$

Pokud přepíšeme zadaný text následujícím způsobem:

s inženýrem Dejvidem džounsem,

nedojde ve fonetickém přepisu už k žádné chybě a bude vypadat následovně:

$\$|s|!inZenIrem|dejvidem|Wynsem|\$.^2.$

Pro zjednodušení práce s nástrojem *Prokus* tedy doporučujeme při zadávání textu následující:

- rozepsat zkratky
- přepsat cizí jména a názvy „tak jak se vyslovují“
- rozepsat číslovky
- doplnit čárky do míst, kde má být pauza (i přes to, že je to z gramatického hlediska špatně)

Na závěr této sekce doplníme, že text jde samozřejmě po provedení syntézy upravit (pokud zjistíte nějaké nedostatky ve výslovnosti).

---

<sup>1</sup>Znak ! se nazývá *ráz* a zjednodušeně řečeno se vyskytuje na začátku slovo začínajících samohláskou.

<sup>2</sup>Znak W odpovídá [dž] a znak y dvojhlásce [ou].

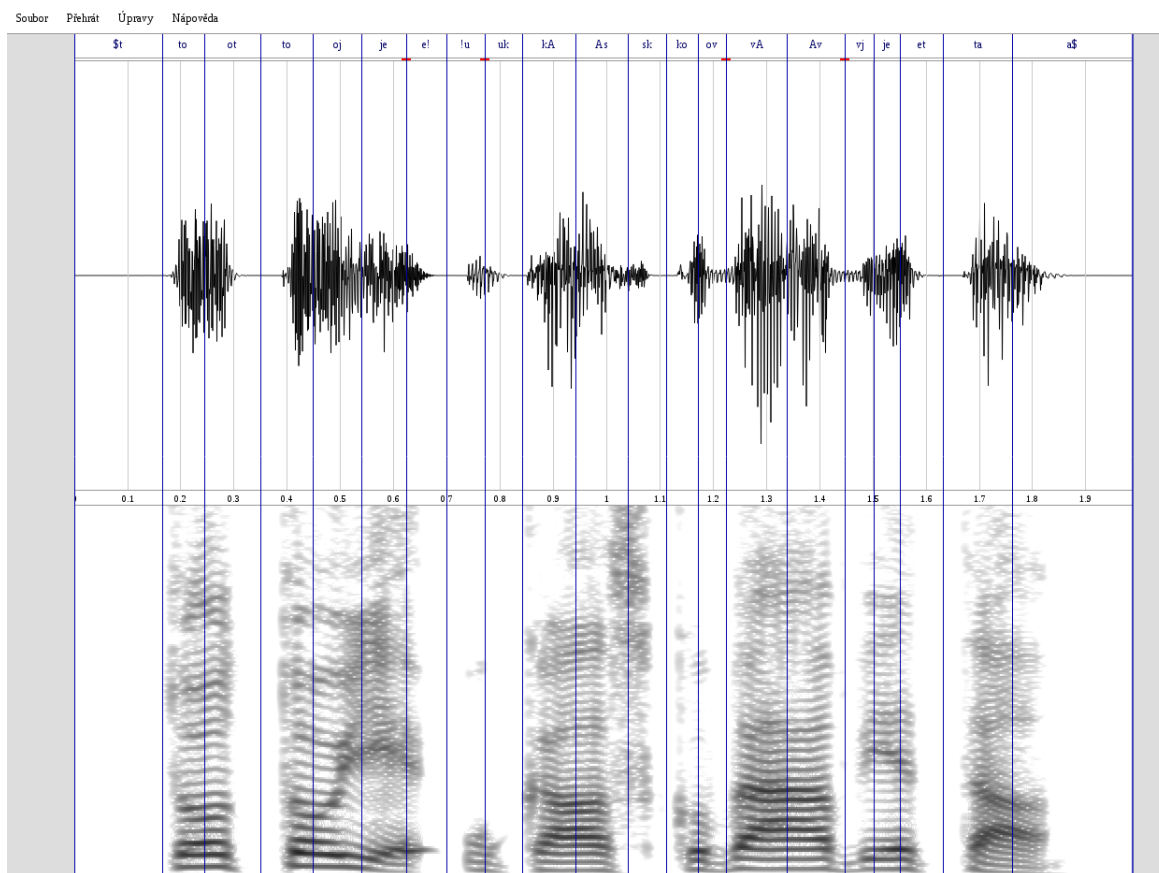


### 3.3.4 Syntéza věty

Po výběru hlasu, zadání požadovaného textu a případné kontrole fonetické transkripce provedete syntézu stiskem tlačítka **Provést syntézu**.

## 3.4 Možnosti práce se syntetizovanou promluvou

Po vygenerování promluvy (pro delší souvětí či více vět může trvat několik vteřin) se v horní části zobrazí průběh zvukového signálu (viz obr. 5). Svislé čáry označují hranice jednotek, ty jsou navíc vypsané v horním řádku zobrazení (jak bylo uvedeno v sekci 2, jedná se o difóny). Hranice jednotek, které jsou v 2. řádku (pod vypsanými jednotkami) zvýrazněné červeně, označují místa napojení signálu; neoznačená místa znamenají, že jednotky pochází ze stejné nahrané věty, kde spolu sousedily. Zhruba v polovině zobrazení je časová osa a v dolní části pak spektrum signálu.



Obr. 5: Ukázka vygenerované promluvy.

V následujících sekcích jsou popsány jednotlivé možnosti základní práce se syntetizova-

nou promluvou – její přehrávání, zoom, označení části, ukládání apod. Náповědu k těmto operacím naleznete rovněž v horní nabídkové liště.

### 3.4.1 Možnosti přehrávání

Celou promluvu lze přehrát stiskem klávesy **o** či výběrem z nabídkové lišty **Přehrát** → **Vše**. V případě, že chcete přehrát pouze její část, je třeba nejprve stisknutím levého tlačítka myši a táhnutím označit požadované jednotky a poté zmáčknout klávesu **p**, která vyvolá přehrávání pouze označené části. Pro přehrávání libovolné části signálu (ne zarovnané na hranice jednotek) označte požadovaný úsek stiskem klávesy **Shift**, levým tlačítkem myši a táhnutím, po označení opět zmáčkněte klávesu **p**.

### 3.4.2 Prohlížení signálu – posun, přiblížení

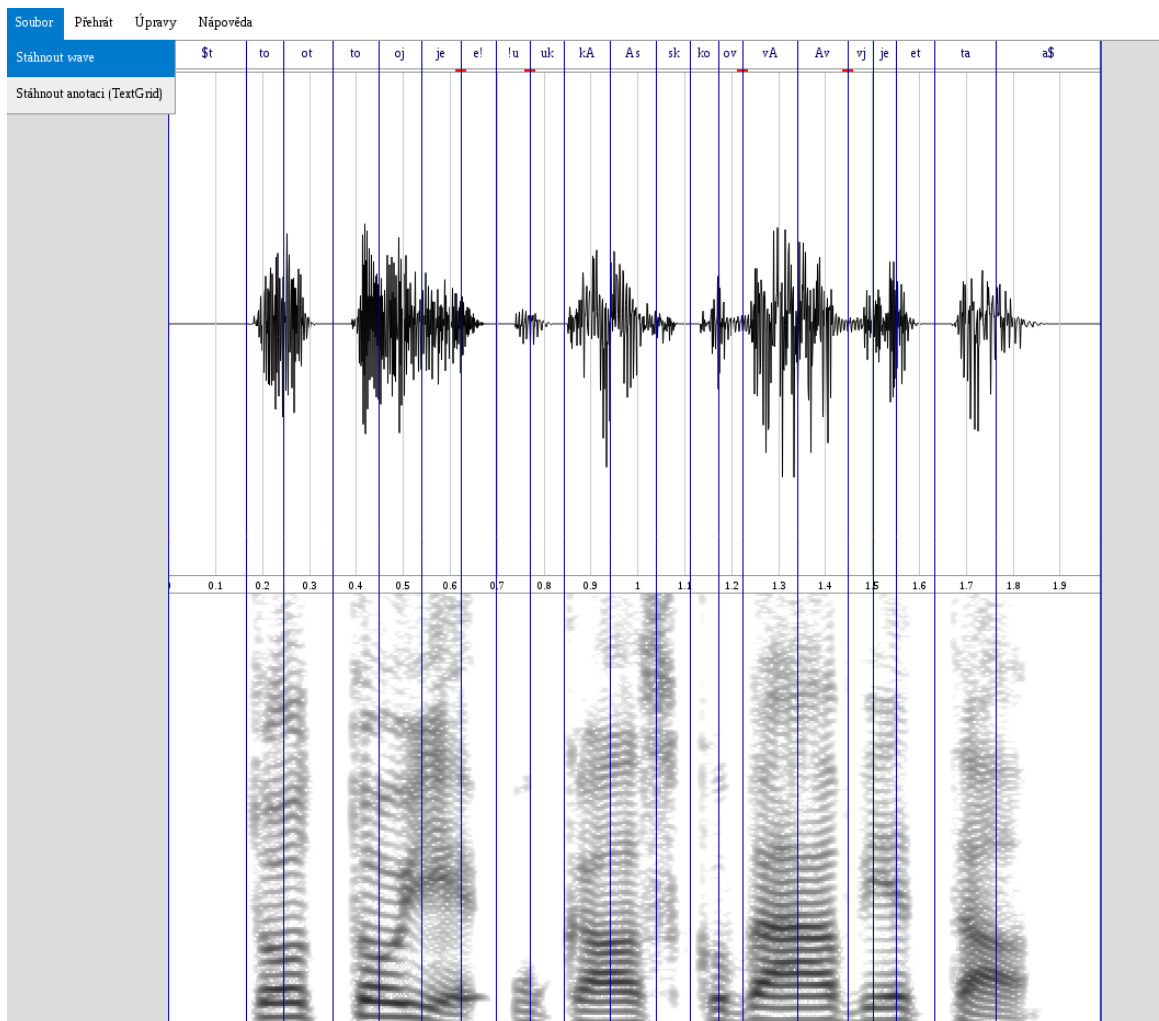
V případě delších (nebo několika) vět nejsou po vygenerování promluvy viditelné hranice jednotek a celá promluva je zmenšená tak, aby se vešla na šířku Vašeho prohlížeče. Pro přiblížení určité části signálu posuňte kurzor myši na toto místo a stiskem klávesy **Ctrl** a kolečkem myši můžete signál přibližovat/oddalovat. Pro posun v čase (posun doleva/doprava) použijte pouze kolečko myši.

### 3.4.3 Uložení vygenerované promluvy

Pro uložení vygenerované promluvy zvolte v horní nabídkové liště **Soubor** → **Stáhnout wave** (viz obr. 6). Promluva se uloží ve formátu **wav** do složky pro umístění stahovaných souborů nastavené ve Vašem prohlížeči – většinou se jedná o složku **Downloads** či **Stažené**. Název **wav** souboru bude odpovídat unikátnímu řetězci znaků v adresním řádku; např. promluva na obr. 7 bude uložena jakou soubor s názvem **wcegai0w.wav**.

## 3.5 Úpravy syntetizované promluvy

Pokud neodpovídá vygenerovaná promluva Vaším představám, je několik možností, jak ji upravit. První, nejjednodušší způsob lze použít, pokud syntetizovaná promluva obsahuje chybnou výslovnost některého slova nebo číslovku či zkratku rozepsanou ve špatném pádě/čísele. V tomto případě použijte tlačítko **Zpět** ve Vašem prohlížeči (návrat na předchozí stránku), čímž se opět dostanete na úvodní stránku nástroje *Prokus* (obr. 2), kde najdete Vámi zadanou větu i vybraný Vámi zvolený hlas. V zadaném textu věty upravte problematická slova/výrazy (doporučujeme dodržovat pokyny uvedené v 3.3.3) a opět klikněte na volbu **Provést syntézu**. Pokud si myslíte, že by na některém místě ve větě měla



Obr. 6: Uložení vygenerované promluvy.

<https://speechlab.zcu.cz/prokus/wcegai0w>

Obr. 7: Adresní řádek vygenerované promluvy.

být pauza, přidejte do textu čárku (přestože tam z hlediska pravopisu být nemají) – čárka v zadaném textu způsobí vložení pauzy, mírné zpomalení a pravděpodobně i mírně stoupavou intonaci v posledních slabikách před čárkou.

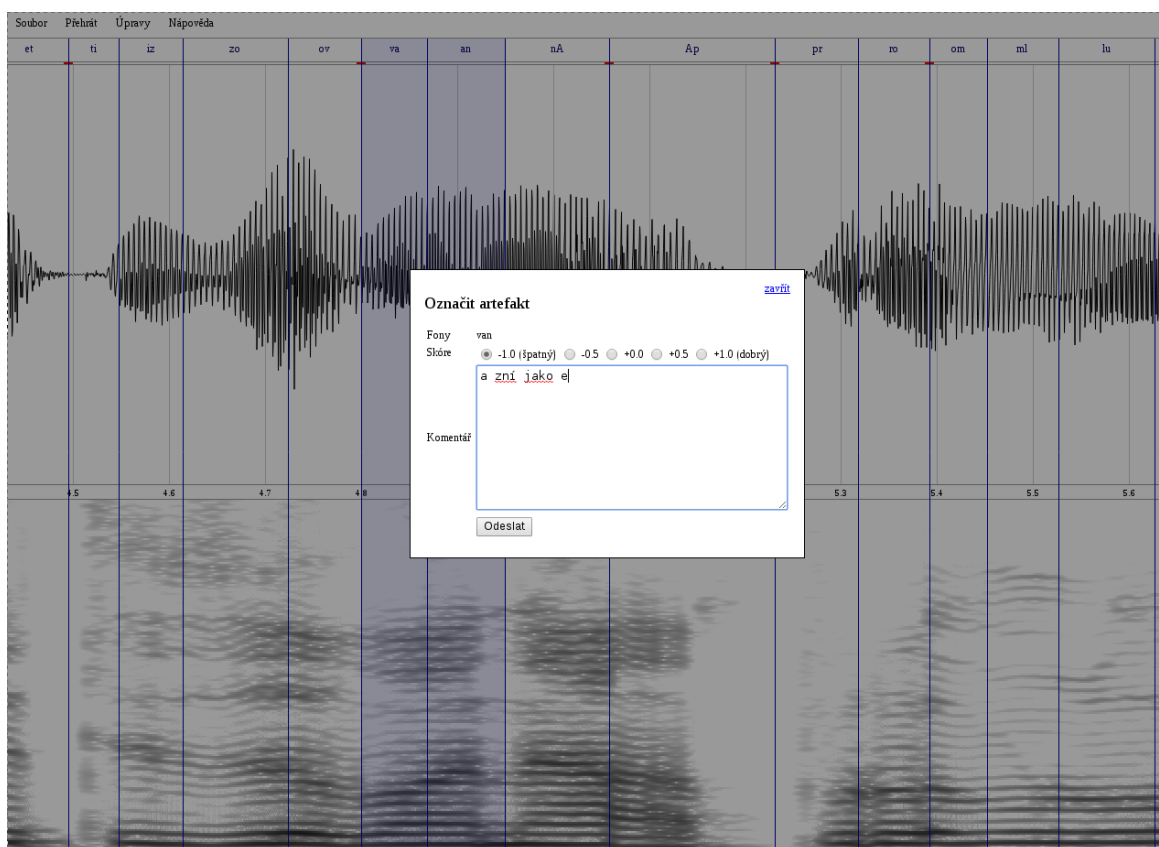
### 3.5.1 Výskyt řečového artefaktu

V případě, že některá jednotka zní nepřírozně (např. slyšíte krátké a místo dlouhého á či e místo a, nebo slyšíte nějakou hlásku navíc) nebo v některém z napojení jednotek

je slyšitelný přechod (přeskočení hlasu, „zazpívání“), je třeba využít další funkci nástroje *Prokus*, a to možnost vygenerování alternativních jednotek a náhrada jednotek v syntetizované promluvě. Jedná-li se o větší chybu (jako např. krátká/dlouhá samohláska, záměna hlásek apod.), označte nám prosím tuto jednotku/jednotky jako *artefakt*, abychom mohli zjistit, proč tato chyba nastala a snažit se ji opravit (takové chyby bývají způsobeny chybnou segmentací nahraných promluv či anotací). Po „nahlášení chyby“ teprve provádějte další úpravy.

### 3.5.2 Nahlašování artefaktů

Označte kliknutím levým tlačítkem myši jednotku, kterou chcete nahlásit (v případě více jednotek použijte klávesu **Shift**), a nabídky v horní liště vyberte **Úpravy** → **Označit artefakt** či stisknete klávesu **m**. Objeví se okno s názvem **Označit artefakt**, ve kterém vyberte **Skóre**  $-1.0$  a napište krátký komentář (viz obr. 8).



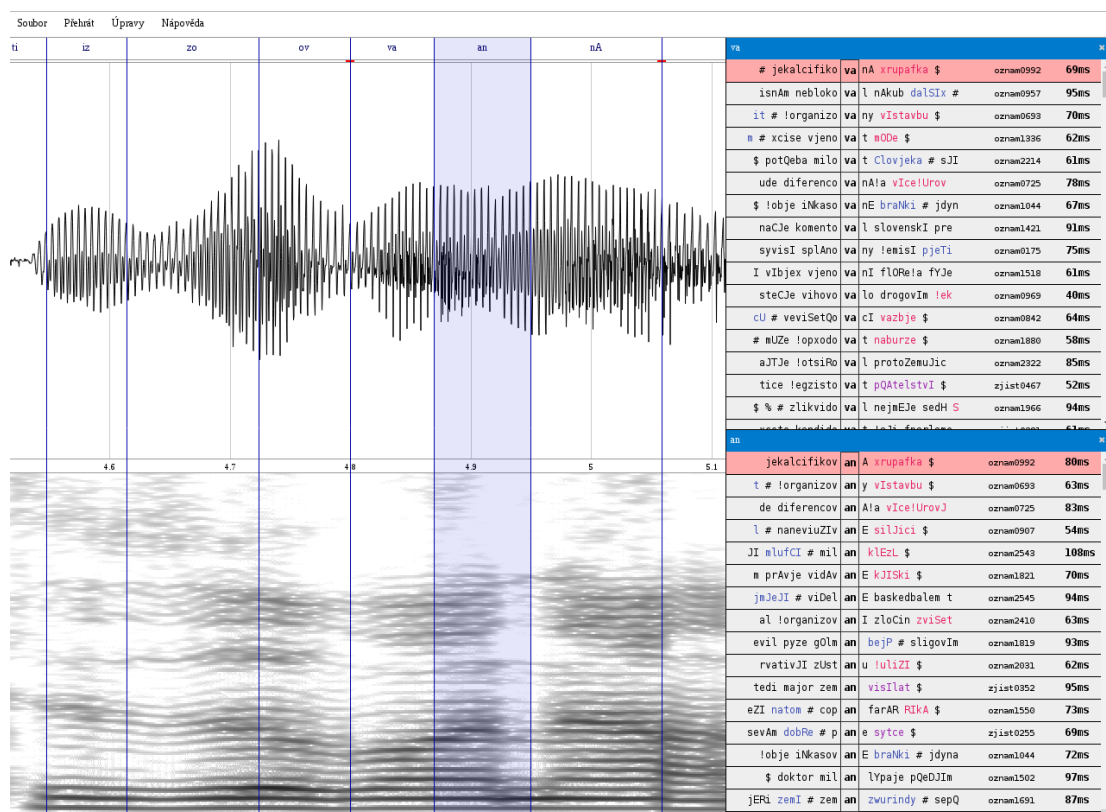
Obr. 8: Nahlášení artefaktu.

Poznámka: Nahlašujte nám prosím opravdu jen závažné chyby – drobnější nedostatky v

napojování jednotek zatím nehlašte!

### 3.5.3 Vygenerování alternativ

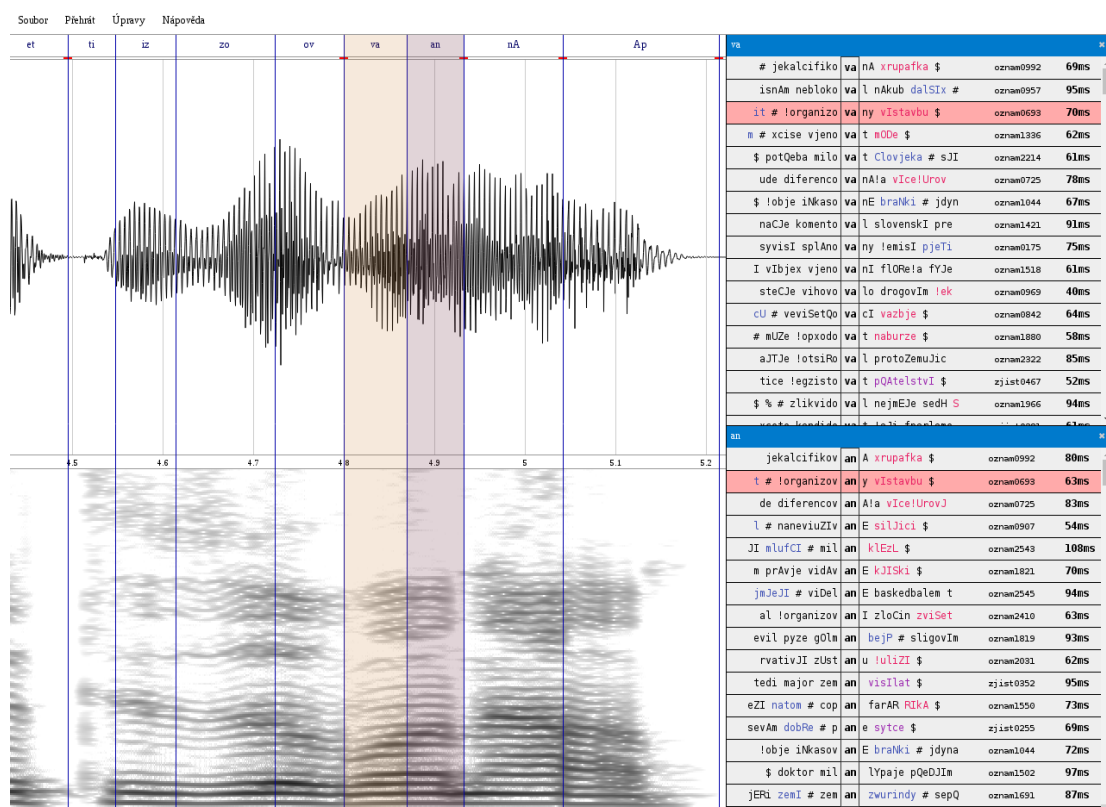
Pro zobrazení seznamu alternativních jednotek vyberte konkrétní jednotku a stiskněte klávesu **a** či vyberte z horní nabídkové lišty **Úpravy** → **Alternativy**. Pokud se problém týká více jednotek (např. slyšíme **e** místo **a**), je třeba vygenerovat si alternativy postupně pro všechny tyto jednotky, kterých se problém týká. Obecně doporučuji řešit najednou max. 3 sousedící jednotky, zkoušení alternativ pro více jednotek najednou by bylo podle našeho názoru velmi chaotické. Na obr. 9 je zobrazen výchozí stav při vygenerování alternativ pro sousední jednotky **va** a **an**. V seznamu alternativních jednotek je vlevo vidět textový kontext dané jednotky, označení nahrávky a vpravo délka dané alternativní jednotky (tento údaj se může hodit v případě, že potřebujete určitou jednotku zkrátit/prodloužit). Jako první v seznamu je vždy původní jednotka, která byla při syntéze vybrána (měla nejlepší – minimální – skóre). Poté následují další jednotky seřazené vzestupně podle vypočteného skóre.



Obr. 9: Vygenerované alternativy pro 2 zvolené jednotky.

### 3.5.4 Zkoušení různých alternativ

V seznamech se jednotky vybírají kliknutím myši. Po vybrání určité alternativy (obr. 10) je okamžitě původní jednotka touto jednotkou nahrazena v signále, což si můžete poslechnout stisknutím tlačítka o či výběrem z nabídkové lišty.



Obr. 10: Vybrání alternativ pro zvolené jednotky.

### 3.5.5 Jak vybrat „správnou“ jednotku

Neexistuje jednoduchý způsob, jak vybrat správnou jednotku. Obecně platí, že je třeba zkoušet měnit alternativy a poslouchat, dokud věta nebude znít tak, jak si přejete. Vodítkem Vám může být průběh signálu či jeho spektrum, pomoci může i údaj o délce jednotky v seznamu alternativ a kontext jednotky.

## 4 Závěr

Věříme, že nástroj *Prokus* Vám dobře poslouží pro generování syntetizovaných promluv a jejich úpravu. V případě nejasností nás můžete kontaktovat prostřednictvím e-mailu [artic@kky.zcu.cz](mailto:artic@kky.zcu.cz).

## Příloha: Česká fonetická abeceda

EPA	SAMPA	Příklad	EPA	SAMPA	Příklad
i	i	pivo	W	d_Z	léčba
e	e	pes	f	f	facka
a	a	máma	v	v	vlak
o	o	bok	s	s	osel
u	u	rum	z	z	koza
I	i:	víno	R	P\<	moře
E	e:	lépe	Q	Q\<	tři
A	a:	táta	S	S	pošta
O	o:	jód	Z	Z	žena
U	u:	růže	j	j	voják
y	o_u	pouto	x	x	chyba
Y	a_u	auto	h	h\<	had
F	e_u	eunuch	G	G	abych byl
@	@	hláskování	r	r	rak
p	p	prak	l	l	loď
b	b	bod	m	m	mír
t	t	otec	n	n	nos
d	d	dům	N	N	banka
T	c	kutil	J	J	laň
D	J\<	děti	M	F	nymfa
k	k	oko	L	l=	vlk
g	g	guma	H	m=	osm
!	?	ráz	P	r=	bratr
c	t_s	ocel	\$	-	dlouhá pauza
w	d_z	leckdo	#	-	krátká pauza
C	t_S	oči	%	-	nádech