

Webový nástroj pro opravy anotací řečového inventáře

Jakub Vít¹

1 Úvod

Syntéza řeči převádí vstupní text na lidskou řeč. Dobře znějící syntéza vyžaduje kvalitní inventář řečových nahrávek. Inventář obsahuje věty nahrané ve zvukové komoře namluvené profesionálním řečníkem, jehož hlas syntéza reprodukuje. Pro namluvené věty jsou vytvořeny anotace. To znamená, že jsou do nich vloženy časové značky, které označují začátky jednotlivých písmen. Při nahrávání hlasu jsou anotace prováděny automaticky pomocí rozpoznávače řeči. Tento proces není bezchybný. Občas se mohou vyskytnout chyby, které je potřeba ručně opravit. Chyby například vznikají, když řečník přečte některé slovo jinak, než bylo napsáno.

Pro opravy anotací na našem pracovišti byl vytvořen speciální nástroj, který tento proces zjednodušuje. V dnešní době se stále více programů přesouvá do prostředí webového prohlížeče. Tento nástroj není výjimkou. Původní verze tohoto nástroje byla okenní aplikace. Nyní byla nahrazena webovou verzí. Tento přechod umožňuje větší pohodlí uživatele, jelikož konfigurace hlasů a výpočet syntézy jsou prováděny na serveru. Zároveň je zajištěna kompatibilita a stejná funkčnost ve všech operačních systémech, což u nativní aplikace vyžadovalo nemalé úsilí.

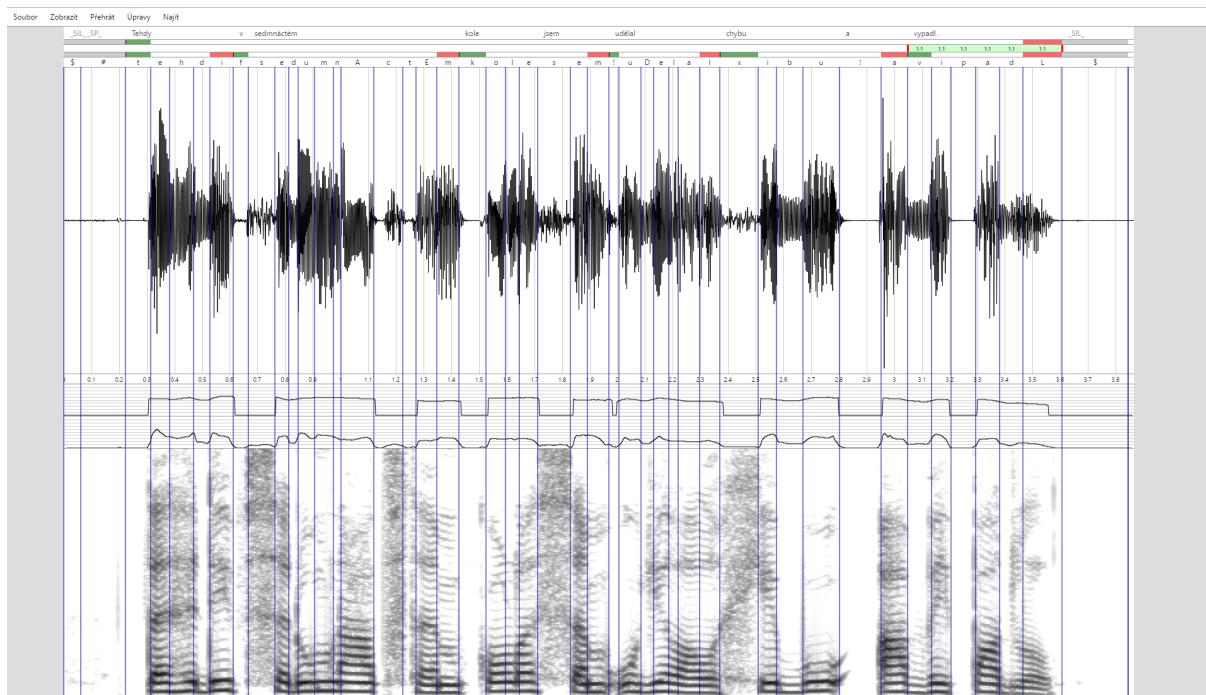
2 Úprava anotací

Základní funkcí je oprava anotačních chyb. Uživatel může procházet řečový inventář a otevřít libovolnou větu. Po otevření se mu zobrazí pohled na větu (obrázek 1). Věta lze přibližovat a oddalovat. Je rovněž možné přehrávat libovolnou část věty. Nejvíce důležitý je panel časového průběhu řečového signálu (oscilogram) a frekvenčního průběhu (spektrogram). Dále jsou tu panely zobrazující hranice slov a frází. V signálu jsou zaneseny časové hranice jednotlivých hlásek. Ty je možné pomocí myši posouvat. Dále je možné přidávat nové, mazat, upravovat, přidávat slova, měnit výslovnosti a mnoho dalšího. Tento nástroj umožňuje kompletně spravovat řečový inventář.

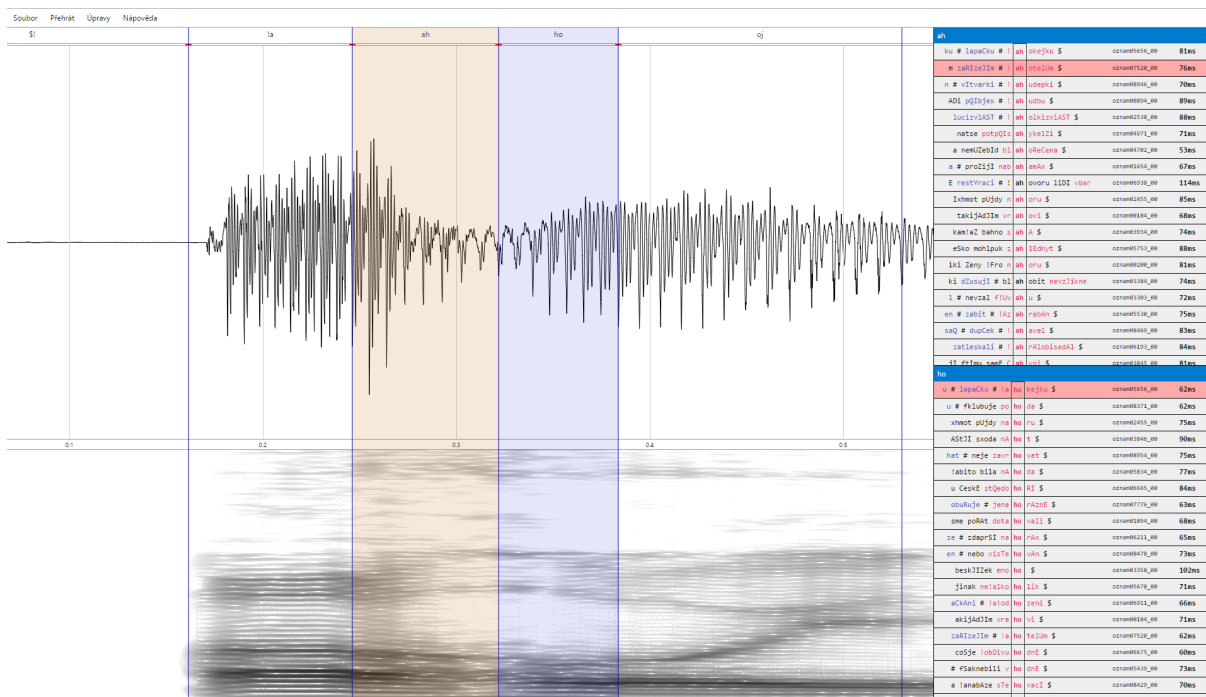
3 Analýza výstupu syntézy řeči

Oprava anotací je užitečná věc. Její přínos je však umocněn ve spojení s druhou funkcí, kterou nástroj nabízí. Tou je syntéza řeči a analýza jejího výstupu. Nástroj umožňuje syntetizovat libovolný text a poté zobrazit výstupní větu včetně informací o použitých řečových jednotkách, které byly použity pro sestavení výsledné věty. Uživatel má možnost prokliknout konkrétní jednotku a zobrazit si originální větu v editoru anotací. Tímto způsobem je možné velmi rychle opravit nahlášenou chybu od zákazníka. Dále je možné volit jiné alternativní jednotky, než které vybral algoritmus syntézy řeči (obrázek 2). Tímto způsobem lze ručně odstranit řečové artefakty, které vznikly napojením špatných jednotek.

¹ student doktorského studijního programu Aplikované vědy a informatika, obor Kybernetika, specializace Umělá inteligence, e-mail: vit89@ntis.zcu.cz



Obrázek 1: Úprava anotací řečového inventáře



Obrázek 2: Volba alternativ v syntetizované větě