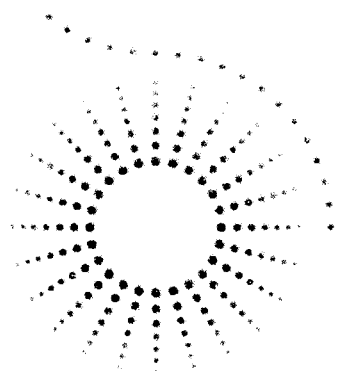


Univerzita Karlova v Praze

Filozofická fakulta



Diplomová práce

2010

Jana Heranová

Univerzita Karlova v Praze

Filozofická fakulta

Fonetický ústav

Diplomová práce

Jana Heranová

Harmonicita jako možný indikátor hranic mezi segmenty v češtině
Harmonicity as a possible indicator of segment boundaries in Czech

Praha 2010

vedoucí práce: Mgr. Radek Skarnitzl, Ph.D.

Poděkování

Ráda bych vyjádřila své poděkování všem, kteří mě podporovali během vzniku této práce. V první řadě patří uznání vedoucímu diplomové práce Mgr. Radkovi Skarnitzlovi, Ph.D. za podnětné připomínky a rady, které tuto práci v mnoha ohledech obohatili. Ráda bych také poděkovala Ing. Marku Dohnalovi za technickou podporu při zpracování dat.

Prohlášení

„Prohlašuji, že jsem tuto diplomovou práci vypracovala samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.“

V Praze dne 20.8. 2010

podpis

Jana Horová

Abstrakt

Cílem této diplomové práce je prozkoumat harmonicitu jako indikátor hranic mezi segmenty v češtině. V naší analýze jsme se soustředili na spojení znělých a neznělých segmentů, konkrétně na spojení neznělých obstruentů a vokálů. Nejdříve jsme se pokusili zmapovat průběh křivky harmonicity v oblasti přechodu mezi segmenty. Na základě výsledků jsme sestavili klasifikaci průběhů harmonicity, která sloužila jako východisko pro posouzení vztahu mezi průběhem křivky harmonicity a umístěním manuálně určené hranice mezi segmenty. Výsledky, tj. časové intervaly, v kterých je umístěna většina hranic, naznačují, že se harmonicitu projevila jako slibný indikátor hranic mezi segmenty. Dalším krokem výzkumu, jehož cílem je implementace harmonicity do softwaru pro automatickou segmentaci řeči, bude prozkoumání přesnosti segmentace za použití různých nastavení pro měření harmonicity.

Klíčová slova: harmonicitu, HNR, průběh harmonicity, křivka harmonicity, hranice mezi segmenty, automatická segmentace řeči, přesnost segmentace řeči

Abstract

The aim of this diploma thesis is to examine harmonicity as an indicator of segment boundaries in Czech. In our analysis we have focused on sequences of voiced and unvoiced segments, more precisely the combination of voiceless obstruents and vowels. First we examined the behaviour of harmonicity in the area of segment transitions. Based on the results we established a classification of harmonicity behaviours, which was further used as a basis for examining the relationship between the placement of segment boundaries and the behaviour of harmonicity. The results, i.e. established time intervals, in which most boundaries were situated, imply harmonicity to be a promising indicator of segment boundaries in Czech. The further course of research, which attempts to use harmonicity for purposes of automatic segmentation, is to examine the accuracy of segmentation while using different settings for HNR estimation.

Keywords: harmonicity, HNR, behaviour of harmonicity, harmonicity curve, segment boundaries, automatic segmentation of speech, accuracy of speech segmentation

Obsah

Seznam zkratk	2
1. Úvod	3
2. Teoretická část	5
2.1 Znělost jako kategorie	5
2.2 Harmonicita	6
2.2.1 Co je to harmonicita?	6
2.2.2 Využití HNR	7
2.2.3 Metody měření HNR	8
2.2.4 Měření HNR v programu PRAAT	17
2.3 „Prague Labeller“ – fonetická segmentace na bázi HMM	18
2.4 Segmentace hlásek – problémy a pravidla	20
3. Praktická část	27
3.1 Materiál a metoda	27
3.2 Výsledky	31
3.2.1 Klasifikace průběhů HNR	31
3.2.2 Průběh křivky HNR s ohledem na řečový signál	37
3.2.3 Vztah umístění hranic a průběhu HNR	51
4. Závěr a diskuze	58
5. Použitá literatura	63
A. Příloha – grafy skupiny 1	66
B. Příloha – grafy skupiny 2	70
C. Příloha – grafy skupiny 3	74
D. Příloha – obsah přiloženého CD	78

Seznam zkratek

A – afrikáty

CV – spojení konsonant – vokál

ČVUT – České vysoké učení technické v Praze

E – explozivy

F – frikativy

FFT – rychlá Fourierova transformace

FF UK – Filozofická fakulta Univerzity Karlovy v Praze

HMM – skryté Markovovy modely

HNR – harmonicita

VC – spojení vokál – konsonant

1. Úvod

Tato práce se věnuje problematice automatické segmentace řeči, přičemž si klade za cíl ověřit, zda lze ke zpřesnění segmentace využít harmonicitu jakožto indikátor hranic mezi segmenty v češtině.

Nejdřív představíme teoretická východiska, z nichž tato práce vychází, a rozebereme je v kapitolách týkajících se znělosti, harmonicity a automatické segmentace řeči. V druhé části práce provedeme analýzu materiálu, ve kterém se zaměříme na spojení znělých a neznělých hlásek, přičemž budeme sledovat průběh křivky harmonicity v oblasti jejich hranic.

Stručně shrnuto: Znělé hlásky představují většinu ze všech řečových zvuků (Vaissière, 1977). Jejich základ tvoří znělost, jazyková kategorie, která se v jazycích hojně využívá k odlišení hlásek (Vaissière, 1977; Palková, 1997). Korelátém znělosti ve frekvenční doméně je harmonicity, která zachycuje vlastnosti harmonického spektra (Möbius, 2004), respektive vyjadřuje podíl harmonických a šumových složek v signálu. Znělé segmenty tedy dosahují vyšších hodnot harmonicity než segmenty neznělé, které postrádají harmonickou složku (Heranová, 2009). Z toho vyplývá, že by se harmonicity mohla využít i v oblasti automatické segmentace, protože postihuje rozdíl mezi znělými a neznělými segmenty a podle výsledků práce Heranová (2009) i rozdíly mezi jednotlivými třídami hlásek jako jsou vokály, sonory, znělé obstruenty a neznělé obstruenty. Tato práce si tedy klade za cíl ověřit, zda je tato logická úvaha správná, přičemž se soustředí na konkrétní spojení neznělých obstruentů s vokály. V rámci těchto spojení budeme mapovat průběh křivky harmonicity, který se pokusíme dát do souvislosti s umístěním manuálně a automaticky určených hranic mezi segmenty. Na základě kvalitativní analýzy dat v praktické části práce se pokusíme sestavit klasifikaci průběhů harmonicity na hranicích segmentů, z které budeme dále vycházet při zkoumání vztahu mezi průběhem harmonicity a umístěním hranic mezi segmenty. Nulová hypotéza, jež se nám v tomto směru nabízí, říká, že se na hranicích segmentů neobjevuje žádný klasický průběh harmonicity, respektive že je průběh křivky harmonicity v těchto kontextech náhodný. Vyvrátí-li se tento vztah, budeme moci na základě naší analýzy sestavit klasifikaci jednotlivých průběhů, přičemž očekáváme, že klasický průběh bude nabývat podoby logistické křivky neboli sigmoidy, kdy se záporné hodnoty harmonicity neznělého obstruentu budou měnit s přechodem na následující vokalický segment na kladné hodnoty a naopak. Co se týče vztahu mezi průběhem harmonicity a umístěním segmentálních hranic, nulová hypotéza opět předpokládá, že mezi průběhem harmonicity a umístěním hranic mezi segmenty žádný vztah neexistuje. Další hypotéza, se kterou budeme pracovat, vyvrátíme-li tuto nulovou hypotézu, předpo-

kládá, že umístění hranice mezi segmenty bude korelovat s kladnými hodnotami harmonicity, které indikují znělost signálu respektive přítomnost formantové struktury u vokálů.

Tato práce je zaměřená na prozkoumání možného přínosu harmonicity v oblasti hláskové segmentace, a proto mohou být její výsledky využity hlavně v oblasti automatické segmentace řeči. V současné době se segmentace řeči vyžaduje v mnoha odvětvích. Manuální segmentace ale představuje časově náročný úkol. Pokud je navíc materiál zpracováván skupinou anotátorů, můžou se mezi jednotlivými anotátory objevit odchylky v segmentaci (Machač & Skarnitzl, 2009, str. 13 – 15), což znamená, že výsledný materiál není zpracován jednotným způsobem. Žádaná jednotnost a přesnost segmentace, jako i požadavek na objem segmentovaného materiálu, který je zapotřebí k vytvoření korpusu nebo pro fonetický výzkum, daly podnět ke vzniku softwaru pro automatickou segmentaci řeči. Ve Fonetickém ústavu je za tímto účelem využíván software *Prague Labeller*, který sice značně ulehčuje práci, ale představuje pouze jeden krok v procesu zpracování materiálu pro fonetický výzkum, a to kvůli chybovosti automatické segmentace. Data, jež se pokusí poskytnout tato práce, mohou sloužit jako motivace nebo východisko pro další výzkum, jenž by napomohl ke zpřesnění automatické segmentace.

Celá práce sestává ze dvou hlavních částí, teoretické a praktické. Ty mají následující strukturu: teoretická část představí teoretická východiska práce, která popíše v kapitole 2.1 týkající se kategorie znělosti, v kapitole 2.2 představující harmonicitu, její využití a metody měření a v kapitolách 2.3 a 2.4, které pojednávají o problémech manuální a automatické segmentace hlásek. V praktické části práce provedeme analýzu, přičemž v kapitole 3.1 představíme použitý materiál a metody měření a výsledky analýzy budeme prezentovat v kapitole 3.2. Kapitola týkající se výsledků bude obsahovat tři dílčí části: klasifikaci průběhů křivky harmonicity (3.2.1), popis jednotlivých typů průběhů s ohledem na vlastnosti řečového signálu (3.2.2) a zhodnocení vztahu mezi umístěním hranic a průběhem harmonicity mezi segmenty (3.2.3). Ve 4. kapitole nakonec shrneme závěry, ke kterým jsme se dopracovali, a jako náměty k diskusi načrtneme případné problémy a doporučení pro další výzkum.

2. Teoretická část

2.1 Znělost jako kategorie

Většina řečových zvuků je znělá, tj. je tvořena za pomoci kmitání hlasivek. V důsledku proudění vzduchu přes hlasivky, které jsou během produkce řeči přiblíženy k sobě, dochází k opakovanému přibližování a oddalování hlasivkových vazů, neboli k otevírání a zavírání hlasivkové štěrbině, a tím ke vzniku hlasu. Předpokladem pro kmitání hlasivek je dostatečně velký subglotální tlak, který způsobí rozražení hlasivkových vazů, které se vzápětí v důsledku své elasticity a fungování Bernoulliho efektu opět vrátí k sobě. Tato kvaziperiodická modulace výdechového proudu vzduchu, nazývaná fonace, představuje primární zdroj energie pro tvoření znělých zvukových segmentů (Vaissière, 1997).

Na základě přítomnosti nebo nepřítomnosti základního tónu, který vzniká jako důsledek fonace, rozlišujeme hlásky znělé a neznělé (Palková, 1997, str. 113). Hlásky můžeme dále klasifikovat podle míry účasti základního tónu na jejich charakteru, tedy podle stupně sonority hlásek, na vokály, sonorní konsonanty, znělé obstruenty a neznělé obstruenty, přičemž vokály jsou na stupnici sonority nejvýše a neznělé obstruenty nejnižší (Palková, 1997, str. 101 – 110). Zde je nutné poznamenat, že v případě znělých obstruentů představuje znělost až sekundární zdroj energie pro jejich tvoření. Znělost jako kategorie nám tedy slouží k vytvoření kontrastu mezi skupinami hlásek, její využití v jednotlivých jazycích se nicméně liší (Vaissière, 1997). Například podíl znělých segmentů v řeči je ve francouzštině 78 %, kdežto v kantonské čínštině pouze 41 % (Catford, 1977, citováno ve Vaissière, 1997). Toto je způsobeno různou mírou uplatnění znělosti k vytvoření kontrastu v jazykových systémech. Například francouzština rozlišuje znělé a neznělé obstruenty, kdežto v čínštině mezi sebou kontrastují neznělé aspirované a neznělé neaspirované obstruenty. V čínštině se tedy znělost k vytvoření kontrastu mezi jednotlivými řadami konsonantů nevyužívá vůbec (Vaissière, 1997). Také fonetická realizace fonologické znělosti respektive neznělosti se liší v závislosti na tom kterém jazyku. Například Möbius (2004) srovnával znělostní profily německých a čínských exploziv a zjistil, že fonologicky neznělé neaspirované čínské explozivy jsou svojí fonetickou realizací blízké spíš německým znělým neaspirovaným explozivám.

Znělost představuje v jazykovém systému důležitou kategorii, která nám umožňuje vytvářet základní kontrast mezi znělými a neznělými hláskami. Akusticky se znělost v signálu projeví přítomností základního tónu, tedy přítomností nízkofrekvenční energie ve spektru, přičemž v časové doméně ji

pozorujeme jako periodicitu signálu a ve frekvenční oblasti je jejím korelátem harmonické spektrum respektive harmonicita (Möbius, 2004).

2.2 Harmonicita

2.2.1 Co je to harmonicita?

Znělost v řeči vzniká na základě kmitání hlasivek. Tomuto kmitání odpovídá v akustické rovině analýzy řečového signálu základní frekvence f_0 . Kmitání hlasivek je kvaziperiodické, lidský vokální trakt je nestacionární, a proto se v oblasti fonace do jisté míry projeví také určitá nepravidelnost, která se označuje jako aperiodicita v hlasovém signálu. Pro akustiku fonace z toho vyplývá, že se setkáváme s odchylkami mezi jednotlivými cykly základní periody, tedy s perturbacemi v oblasti trvání a amplitudy základní periody (česky třas a tremolo) (Klingholz & Martin, 1985; Gelfer & Fendel, 1995; Brockmann et al., 2008) a s přídavným šumem (Murphy, 2006). Přídavný šum definoval Hillenbrand (citováno v Awan & Frenkel, 1994) jako vedlejší produkt turbulentního proudění vzduchu, který se produkuje na glotis během fonace. Tato turbulentní frikce vzniká v důsledku neúplného dovření hlasivek během fonace (Ferrand, 2002), které vzniká zejména v zadní části hlasivek (Hanson et al., 2001). Všechny tyto aspekty ovlivňují vnímanou kvalitu fonace, a tedy i kvalitu hlasu. Hlasy, u kterých je ve spektru vokalického segmentu přítomné větší množství šumu, bývají percepčně hodnoceny jako drsné (Yumoto et al., 1982). U patologických hlasů se naopak setkáváme s pojmem chrapotu, kterým se označuje vnímaná abnormalita v hlase (Yumoto et al., 1982). Kvalita fonace se může měnit s věkem (Linville, 2000; Schötz & Müller, 2007), ale i s nástupem onemocnění hrtanu či poruch hlasu (Ferrand, 2002). Abychom byli schopni objektivně posuzovat změny nebo vyhodnocovat kvalitu fonace, je zapotřebí objektivního vyjádření její kvality. To znamená, že potřebujeme míru, která by byla schopna objektivně zachytit percepční vlastnosti hlasového signálu. Za tímto účelem byly v 80. letech 20. století navrženy první algoritmy pro měření HNR.

HNR představuje zkratku anglického výrazu „harmonics-to-noise ratio“ (česky poměr harmonických a šumu) a také se označuje jako harmonicita. Jedná se o míru, která indikuje celkovou periodicitu signálu (Murphy, 2006). V tradičním pojetí, kde harmonicitu využíváme ke zhodnocení kvality hlasu, se HNR vyjadřuje v jednotce [dB] jako logaritmus poměru energie periodických a neperiodických složek v signálu (Boersma, 1993) a kvantifikuje tedy i relativní množství přídavného šumu v hlasovém signálu (Awan & Frenkel, 1994; Qi & Hillman, 1997).

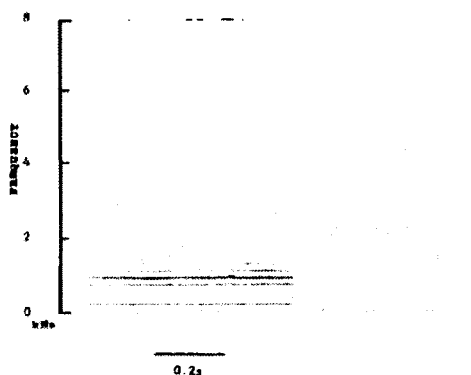
Budeme-li chtít definovat HNR, máme na výběr ze dvou možností: buď chceme zdůraznit míru periodicity, nebo míru aperiodicity hlasového signálu. Podle první z možností je HNR ukazatelem vokální funkce, který reflektuje dominantnost harmonické struktury signálu nad přítomným šumem (Ferrand, 2002), měří periodicitu signálu a úplnost jeho harmonické struktury (Yu & Wang, 2006). Podle druhého pohledu na tuto problematiku měří HNR relativní množství šumu ve frekvenčním spektru vokálů (Cox et al., 1989), využívá se ke stanovení úrovně šumu v hlasovém signálu (Qi & Hillman, 1997), a tedy i k evaluaci stupně drsnosti hlasu (Yumoto et al., 1982). Podle studie Murphy (2006) je HNR funkcí glotálního šumu a jiných faktorů jako je například třas (angl. jitter), tremolo (angl. shimmer), aspirační šum, variabilita glotálních cyklů nebo nestacionárnost vokálního traktu, které jsou zodpovědné za aperiodickou složku hlasu (srov. také Gelfer & Fendel, 1995).

2.2.2 Využití HNR

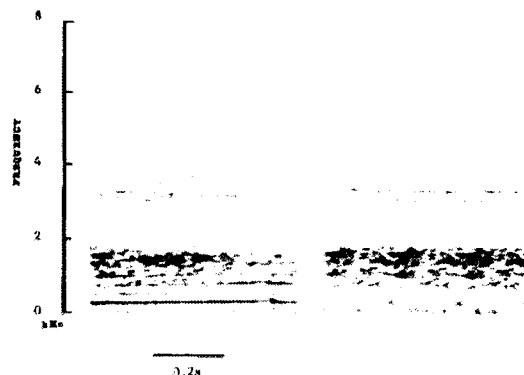
Fyziologické aspekty tvoření hlasu můžeme posuzovat percepčně, tj. subjektivně nebo objektivně za pomoci akustické analýzy řečového signálu. Dochází-li při fonaci k produkci přídavného šumu nebo k větší variabilitě hlasivkových cyklů, vnímáme daný hlas subjektivně jako drsný a ve spektru hlasové nahrávky se setkáme s vyšší úrovní šumu. Vztah mezi hladinou spektrálního šumu a vnímanou drsností hlasu potvrdil už Emanuel a spolupracovníci ve studiích z let 1969, 1970 a 1979 (citováno v Yumoto et al., 1982). Yumoto et al. (1982) dále referují o tom, že stupeň chrapotu, jakožto patologické drsnosti hlasu, jasně závisí na akustických vlastnostech signálu, a to tak, že se stoupajícím stupněm chrapotu přibývá šumu, který nahrazuje harmonickou strukturu vokálu. U drsných hlasů se ve spektru vokálu objevují šumové komponenty v oblasti hlavních formantů, dochází ke ztrátě harmonických komponent ve vysokých frekvencích a k výskytu vysokofrekvenčního šumu (Yanagihara, 1967; citováno v Yumoto et al., 1982). Na základě těchto poznatků byl pro kvantifikaci chrapotu u patologických mluvčích navržen algoritmus pro měření HNR (Cox et al., 1989). V oblasti hlasové patologie představuje HNR vhodný nástroj pro neinvazivní analýzu hlasu a hlasových poruch (Murphy, 2006), a to i díky její robustnosti, jelikož je závislá primárně na úrovni perturbací v signálu. Jiné akustické charakteristiky jako f_0 nebo formantová struktura by ji neměly výrazně ovlivňovat (Cox et al., 1989). V práci Ferrand (2002) se HNR dále uplatňuje i jako index stárnutí hlasu.

Pro stanovení HNR u patologických mluvčích se využívají nahrávky držených vokálů (Yumoto et al., 1982). V odborné literatuře se setkáváme s různými referenčními hodnotami HNR zdravého hlasu: Yumoto et al. (1982) uvádějí pro zdravé hlasu rozpětí HNR mezi 7 – 17 dB s průměrem 11,9 dB (pro 95 % mluvčích se předpokládá dosažení HNR vyšší než 7,4 dB), jiné práce (citováno ve Ferrand,

2002) uvádějí pro normální dospělé hlasy hodnoty HNR mezi 11 a 13 dB a Awan & Frenkel (1994), kteří vylepšili Yumotův algoritmus, uvádějí normativní hodnotu HNR 15,38 dB a předpokládají, že 95 % subjektů bude mít HNR vyšší než 13,6 dB. U patologických hlasů se u držených vokálů podle stupně postižení hlasu setkáváme se zápornými hodnotami. Na obrázcích č. 1 a 2 můžeme srovnat spektrogramy normálního a patologického hlasu.



Obrázek č. 1 (Yumoto et al., 1982, str. 1547).
Vlevo spektrogram fonace zdravého hlasu, vpravo spektrogram extrahovaného šumu. Harmonicita 14,6 dB.



Obrázek č. 2 (Yumoto et al., 1982, str. 1548).
Vlevo spektrogram fonace patologického hlasu s těžkým postižením, vpravo spektrogram extrahovaného šumu. Harmonicita -7,2 dB.

HNR jako výraz periodicity signálu se využívá i v oblasti zpracování řeči pro detekci znělosti (Murphy, 2006). Novou sadu harmonicit, které komplexněji popisují harmonickou strukturu vokálů, představili Yu & Wang (2006). Její aplikaci navrhují v oblasti detekce f_0 nebo pro účely detekce znělých úseků řeči z hlučného prostředí (angl. voicing detection) a následné zvýraznění harmonické struktury v zašuměném řečovém materiálu (angl. speech enhancement).

Tato práce se bude snažit prozkoumat možnosti využití HNR v oblasti automatické segmentace řeči jako indikátoru hranic mezi segmenty v češtině. Přesněji se budeme snažit prozkoumat, zda lze na základě průběhu HNR vylepšit úspěšnost segmentátoru, který pracuje na bázi HMM algoritmu. Ve zkoumaném materiálu se přitom zaměříme na hranice mezi neznělými obstruenty a vokály.

2.2.3 Metody měření HNR

HNR vyjadřuje míru periodicity signálu. Z tohoto důvodu se primárně předpokládá přítomnost periodicity v signálu, pro který má být HNR stanovena. Vhodným segmentálním materiálem pro tento účel jsou vokály, které představují skupinu nejsonornějších řečových segmentů. Kmitání hlasivek při produkci vokálů zajišťuje přítomnost f_0 jako periodické složky v signálu a způsob tvoření vokálů, kdy se výdechovému proudu vzduchu nekladou do cesty překážky, umožňuje tuto periodickou slož-

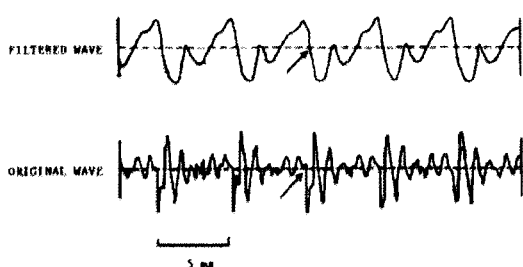
ku díky nastavení vokálního traktu dále modifikovat, takže dochází ke vzniku vokalických formantů. Způsob tvoření vokálů zároveň představuje nejméně náročný způsob artikulace vhodný i pro pořizování materiálu u patologických mluvčích, kterým může působit problémy už jen udržení fonace (Yumoto et al., 1982). Jako takové jsou vokály výsledkem působení periodické produkce zdroje, tj. kmitání hlasivek, a filtru vokálního traktu. Kmitání hlasivek se v signálu akusticky odráží jako f_0 a její harmonické, t.j. celočíselné násobky f_0 . Filtr vokálního traktu formuje formantovou strukturu vokálů. Obě tyto složky vytvářejí periodickou složku signálu. Abychom byli schopni poskytnout odhad HNR, je nutné zastoupit signál modelem, který nám umožní určitý stupeň abstrakce. Proto vycházejí všechny algoritmy měřící periodicitu signálu z předpokladu, že vokál lze modelovat jako sumu periodických a neperiodických složek (Yumoto et al., 1982; Cox et al., 1989; Qi & Hillman, 1997). Periodická respektive harmonická složka signálu se v každé periodě f_0 opakuje (Cox et al., 1989). Její energii zastupuje energie harmonických. Aperiodickou respektive šumovou složku signálu tvoří perturbace f_0 a přídavný šum, který vzniká na glotis během fonace.

Přítomnost vyšší úrovně šumu v hlasovém signálu lze identifikovat jak v oscilogramu podle tvaru vlny, tak i ve spektrogramu daného signálu (Awan & Frenkel, 1994). Podle toho, který způsob reprezentace signálu zvolíme (tj. oscilogram nebo spektrum), můžeme HNR měřit buď v časové, nebo ve frekvenční doméně. Oscilogram zobrazuje signál jako funkci času. Pomocí Fourierovy transformace získáme spektrum, které signál zobrazuje jako funkci frekvence. V časové doméně vychází počítání HNR z tvaru vlny akustického signálu, ve frekvenční doméně se k výpočtu HNR používá reprezentace této vlny (Qi & Hillman, 1997), a to pomocí spektra nebo kepra. Protože jsou obě domény, tedy časová i frekvenční, vzájemně propojeny skrze Fourierovu transformaci, odhady HNR v obou doménách by měly vzájemně korelovat (Qi & Hillman, 1997).

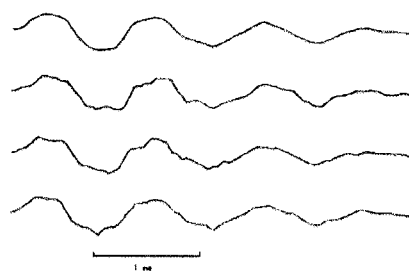
Časové metody: Jednu z prvních časových metod představila studie Yumoto et al. (1982). Východiskem časových metod pro stanovení HNR je získání jedné periody tzv. průměrné vlny, která by reprezentovala energii harmonické složky signálu (Murphy, 2006). Ve studii Yumoto et al. (1982) se vycházelo z předpokladu, že akustický signál je zřetěžením vln, které se mezi sebou liší tvarem a trváním periody. Tyto odchylky v trvání periody a tvaru vlny, které se projeví při porovnání jednotlivých period, do signálu vnášejí šumovou komponentu. Tu lze odstranit zprůměrováním dostatečně velkého počtu po sobě následujících period signálu, přičemž jako výsledek získáme jednu periodu průměrné vlny, jejíž energie zastupuje energii periodické složky signálu. Všechny jednotlivé periody se poté srovnávají s takto získanou průměrnou vlnou a jejich odchylky od tvaru průměrné vlny slouží

k získání energie šumu (Qi & Hillman, 1997). HNR následně získáme jako poměr energie průměrné vlny a průměru odchylek jednotlivých period (Murphy, 2006).

Časové metody jsou poměrně málo náročné na výpočet, ale jako vstup vyžadují zpracovaný signál, ve kterém jsou přesně určeny začátky a konce jednotlivých period (angl. pitch marks) (Murphy, 2006). Ty odpovídají okamžiku efektivního uzavření glotis neboli GCI. Yumoto et al. (1982) k výpočtu HNR použili 600 ms signálu v oblasti, kde byla fonace nejstálější. Za pomoci nízkopropustného filtrování (500 Hz) byla provedena extrakce f_0 (viz obrázek č. 3), ve které se manuálně vyznačily hranice jednotlivých period. Na základě takto zpracovaného signálu byl poté proveden výpočet HNR. Ukázku průměrné vlny a odchylek individuálních period znázorňuje obrázek č. 4.



Obrázek č. 3 (Yumoto et al., 1982, str. 1545).
Filtrovaný signál jako východisko pro měření HNR. Dole ukázka vstupního signálu, nahoře vlna po nízkopropustném filtrování. Šipky označují místo průchodu nulou, jako začátek periody.

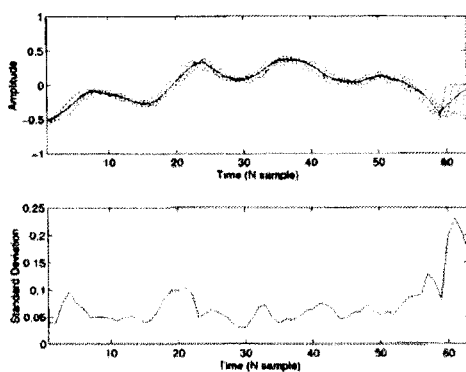


Obrázek č. 4 (Yumoto et al., 1982, str. 1546).
Ukázka průměrné vlny (nahore) a individuálních period (ostatní) jednoho mluvčího.

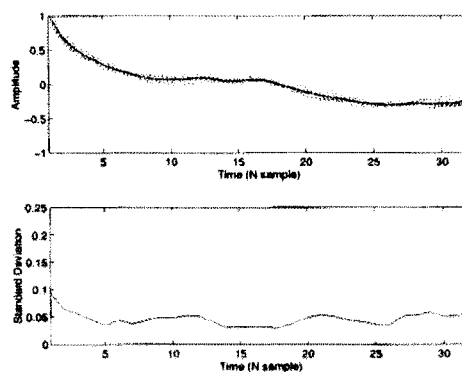
Yumotova časová metoda odhadu HNR sice vykazovala silnou korelaci s výsledky subjektivního hodnocení množství šumu ve spektrogramu, ale objevily se u ní rozdíly 5 – 7 dB mezi zjištěnou a předpokládanou HNR (Awan & Frenkel, 1994). Yumotův algoritmus a problematiku vyznačování jednotlivých hranic hlasivkových cyklů dál zkoumal Hillenbrand a zjistil, že právě manuálním určováním začátků a konců period se zvyšuje chybovost v odhadu HNR (Hillenbrand, 1987; citováno v Awan & Frenkel, 1994). Tato chybovost vzniká právě v důsledku závislosti algoritmu na správné identifikaci jednotlivých hlasivkových pulzů. O její snížení se snažili autoři Awan & Frenkel (1994), kteří navrhli algoritmus pro automatickou identifikaci hranic jednotlivých cyklů, čímž se celková chybovost snížila pod hladinu 2 dB.

Yumotova časová metoda je značně náchylná na přítomnost třasu (angl. jitter) v hlasovém signálu, který dokáže významně zvětšit odchylku od průměrné vlny. Už Yumoto et al. (1982) poukazují na to, že daný výpočet HNR není vhodný pro signály, ve kterých třas převládá nad jinými šumovými

komponentami. Vliv třasu, tj. perturbací f_0 , lze omezit za pomoci vhodné časové normalizace signálu aplikované ještě před výpočtem odchylky (Qi & Hillman, 1997). Yumoto et al. (1982) použili jako normalizační metodu doplnění nulami (angl. zero-padding). Další modifikace Yumotova algoritmu představil Qi. Ve své práci navrhol použít jako vhodnější normalizační metodu transformaci s nulovou fází (angl. zero-phase transformation) (citováno v Qi & Hillman, 1997). Úspěšnost jednotlivých normalizačních metod můžeme srovnat na následujících obrázcích: doplnění nulami na obrázku č. 5 a transformaci s nulovou fází na obrázku č. 6.



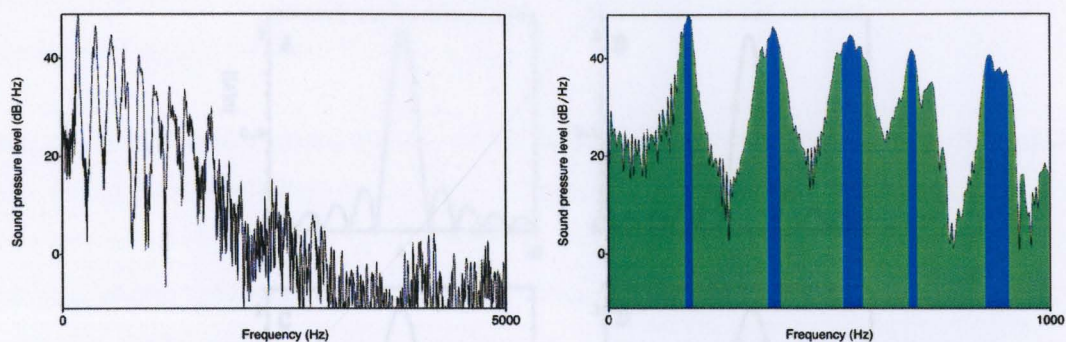
Obrázek č. 5 (Qi & Hillman, 1997, str. 539). Signál normalizovaný metodou doplnění nulami. Signál vykazuje kolem 60. vzorku, kde bylo použito doplnění nulami, značnou odchylku.



Obrázek č. 6 (Qi & Hillman, 1997, str. 539). Signál normalizovaný metodou transformace s nulovou fází. V porovnání s obrázkem č. 5 lze pozorovat více vyrovnanou odchylku.

Mezi časové metody měření HNR se řadí i autokorelační metoda, se kterou se setkáváme v práci Boersma (1993). Normalizovaná autokorelace představuje jednoduchou a přitom efektivní metodu zjišťování periodicity v signálu (Yu & Wang, 2006). Měření HNR v autokorelační doméně využívá software, se kterým se pracovalo při zpracování materiálu a měření HNR v této práci, a proto bude tato metoda detailněji popsána v samostatné kapitole (viz kapitola 2.2.4).

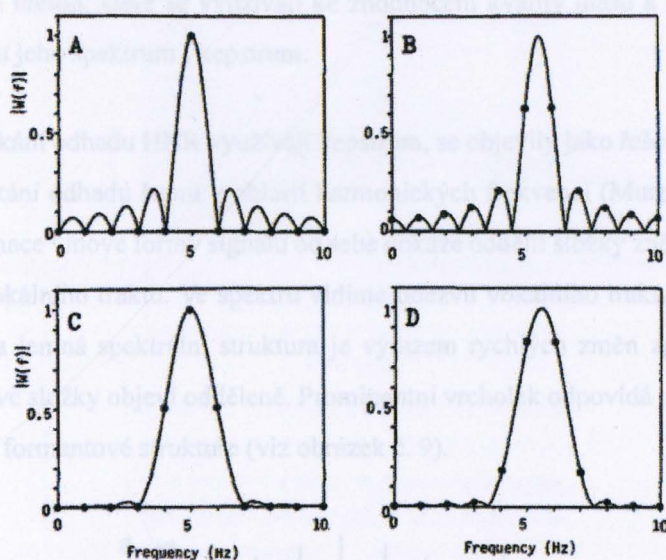
Frekvenční metody: Metody měřící HNR ve frekvenční doméně používají reprezentaci signálu pomocí spektra (Qi & Hillman, 1997). Ze spektra získáme odhad HNR opět jako podíl energie periodické a aperiodické složky. Periodickou složku reprezentuje energie v oblasti harmonických frekvencí a aperiodickou složku tvoří energie šumu v oblastech mezi harmonickými i v oblastech harmonických (viz obrázek č. 7) (Murphy, 2006).



Obrázek č. 7 (zpracovaný materiál – nahrávka MANA30_M12). Spektrum dlouhého vokálu [a:] získané pomocí FFT (angl. Fast Fourier Transform). Vlevo spektrum vokálu ve frekvenčním rozsahu 0 – 5 kHz. Vpravo část spektra vokálu (0 – 1 kHz) s modře vyznačenými oblastmi, které reprezentují energii harmonické složky signálu. Zeleně vyznačené části reprezentují frekvence přispívající k energii šumu.

Metody frekvenční domény sice nemusí řešit nutnost identifikace hranic jednotlivých period f_0 , ale při měření HNR se potýkají s problémem odhadu šumu v oblasti harmonických frekvencí (Murphy, 2006). Odhady HNR znesnadňuje spektrální rozšiřování a prosakování (angl. spectral broadening, spectral leakage), jejichž příčiny můžeme hledat v procesu transformace signálu jako výsledek aplikace okénkové funkce nebo ve vlastnostech signálu samotného – přítomnost třasu vede k rozšiřování harmonických a ke snižování jejich energie (Qi & Hillman, 1997). Cox et al. (1989) nám poskytují zjednodušený popis získávání spektra pomocí Fourierovy transformace a problémů s tím spojených: je-li signál dlouhý, tedy mění-li se spektrum signálu v čase, je pro analýzu zapotřebí zvolit si segment dat, to znamená použít okénkovou funkci, přičemž se předpokládá, že signál v rámci okna je stacionární. To v praxi odpovídá vynásobení signálu okénkovou funkcí (Cox et al., 1989), tedy konvoluci spektra signálu a spektra okna (Qi & Hillman, 1997). Výsledné spektrum signálu se následně získá jako suma výsledků Fourierových transformací všech oken (Cox et al., 1989), jinak řečeno, jako vážená suma všech spektrálních komponent, které určuje váhová funkce okna (Qi & Hillman, 1997). Cox et al. (1989) dále také vysvětlují, proč je i pro frekvenční metody měření HNR důležitá synchronizace spektrální analýzy se základní frekvencí. Důvodem je nutnost sřetězení harmonických komponent ve výsledném spektru vokálů. Rozčlenění signálu na vhodný počet datových segmentů, které jsou umístěny v celočíselných násobcích délky periody f_0 , toto sřetězení umožňuje a zabraňuje vzniku spektrálního prosakování. Energie harmonické složky se pak projeví v každém celočíselném koeficientu Fourierovy řady, ostatní koeficienty Fourierovy řady, kromě koeficientu nulové frekvence, budou konstituovat energii šumu (Cox et al., 1989).

Na výsledné složení spektra, a tím i na odhad HNR pro daný signál, mají značný vliv tvar a délka analyzačního okna. Vliv tvaru okna můžeme pozorovat na obrázku č. 8.



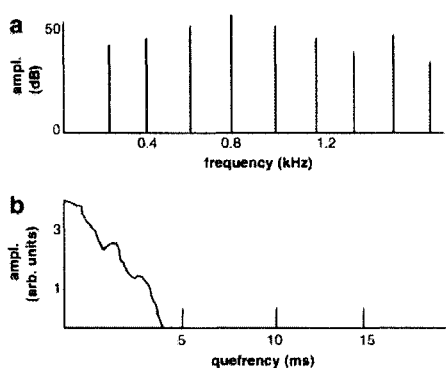
Obrázek č. 8 (Cox et al., 1989, str. 217). Frekvenční spektra sinusoid. Grafy A a B znázorňují spektra sinusoid o frekvenci 5 a 5,5 kHz za použití obdélníkového okna, ve kterých můžeme pozorovat velké postranní sukně. V grafech C a D byly sice postranní laloky minimalizovány po aplikaci Hanningova okna, zato zde můžeme pozorovat zhoršené spektrální rozlišení, tj. rozšíření harmonické.

Aplikace obdélníkového okna způsobuje vznik velkých postranních laloků (angl. side lobes) ve spektru, tedy rozšiřování spektrální energie (Cox et al., 1989). Tento jev, známý též jako spektrální prosakování, můžeme pozorovat v grafech A a B na obrázku č. 8. Spektrální prosakování lze omezit aplikací okna, které má vyhlazené okraje, tj. například Hanningova okna. Tím se významně omezí vliv spektrálního prosakování, ale zhorší se spektrální rozlišení, což má za následek větší rozptýlení energie harmonické kolem její nominální frekvence, které se označuje také jako rozšiřování harmonických (Cox et al., 1989). Tento jev můžeme pozorovat v grafech C a D na obrázku č. 8. Vliv délky analyzačního okna na odhad HNR shrnuli Qi & Hillman (1997) následovně: Použití krátkého okna způsobuje velké spektrální prosakování, což má za následek nižší hodnotu HNR. Dlouhé okno na druhou stranu znesnadňuje počítání harmonicity pro nestacionární signál.

Z předchozího popisu je zřejmé, že frekvenční metody jsou oproti těm časovým výpočetně náročné, přičemž je tato výpočetní náročnost obecně charakteristická pro všechny algoritmy analýzy řečového signálu (P lutka et al., 2006). I přes veškeré technické problémy při používání frekvenčních metod, tj. okénkový efekt a spektrální prosakování, se zdá, že vývoj metod měření HNR nebo periodicity jako takové jde spíše tímto směrem a odvrací se od časové domény. O tom svědčí i následující popis

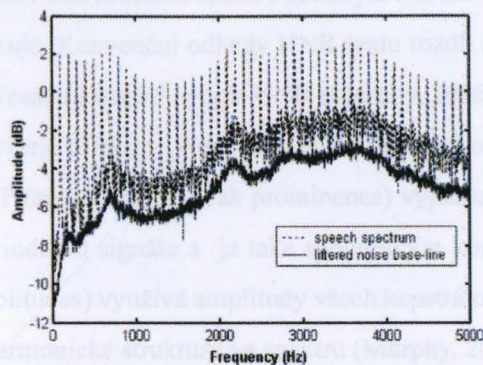
starších i novějších metod, které se využívají ke zhodnocení kvality hlasu a které kombinují jako reprezentaci signálu jeho spektrum i kepstrum.

Metody, které k získání odhadu HNR využívají kepstrum, se objevily jako řešení problému spektrálních metod při získání odhadu šumu v oblasti harmonických frekvencí (Murphy, 2006). Kepstrum jako další transformace vlnové formy signálu od sebe dokáže oddělit složky znělé řeči, tedy produkci zdroje a odezvu vokálního traktu. Ve spektru vidíme odezvu vokálního traktu jako pomalé změny spektrální obálky a jemná spektrální struktura je výrazem rychlých změn způsobených zdrojem. V kepstru se tyto dvě složky objeví odděleně. Prominentní vrcholek odpovídá produkci zdroje a širší vrcholek odpovídá formantové struktuře (viz obrázek č. 9).



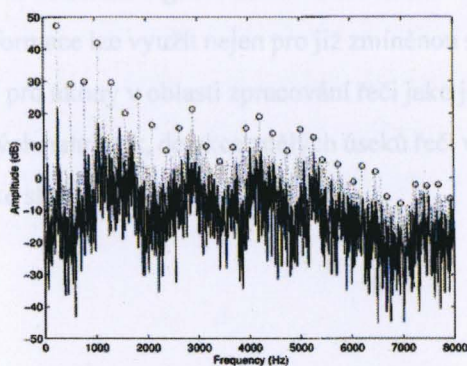
Obrázek č. 9 (Murphy, 2006, str. 1706). a) Schematická reprezentace spektra periodického signálu. b) Korespondující kepstrum. Harmonické frekvence představují prominentní vrcholky umístěny v celočíselných násobcích základní periody: 5, 10 a 15 ms.

První z těchto „kepstrálních“ metod představil de Krom (1993; citováno také v Qi & Hillman, 1997 a Murphy, 2006). Kepstrum se zde využívá k získání informace o energii šumu ve zkoumaném materiálu (Qi & Hillman, 1997). Hřebenovým filtrováním se v kepstru odstraní harmonické frekvence a opětovnou transformací filtrovaného kepstra na spektrum získáme spektrum šumu. V tomto spektru se ještě musí upravit umístění spektrální obálky šumu na amplitudové ose, protože původní odhad je příliš vysoký. Energie získaného spektra šumu se odečte od celkové energie původního harmonického spektra a poté lze provést výpočet HNR (Murphy, 2006). Výsledek de Kromova algoritmu můžeme vidět na obrázku č. 10.



Obrázek č. 10 (Murphy, 2006, str. 1707). Aplikace de Kromova algoritmu: tečkovaně je vyznačeno původní harmonické spektrum, plnou čarou spektrum šumu získané hřebenovým filtrováním z odpovídajícího kepra signálu.

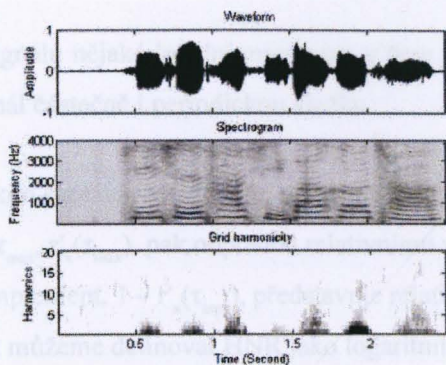
Modifikaci de Kromova algoritmu představili Qi & Hillman (1997). Autoři poukázali na problematické posouvání hladiny šumu ve filtrovaném spektru a také na to, že de Krom předpokládal, že všechny spektrální komponenty nad hladinou šumu budou přispívat k energii periodické složky. Navrhli tedy, aby energie periodické složky sestávala pouze z energie spektrálních vrcholů. Energie aperiodické složky pak byla reprezentována pouze šumem přítomným v těchto harmonických frekvencích. Tím se odstranila potřeba počítání energie celého spektra šumu a posouvání hladiny šumu ve filtrovaném spektru. Za pomoci nízkopropustného keprálního filtrování lze získat informaci o energii šumu v různých frekvencích. Průměr energie této množiny diskrétních frekvencí pak zastupuje energii aperiodické složky ve výpočtu HNR. Nízkopropustné keprální filtrování také zajistí vyhlazenější spektrální obálku šumu, jak je vidět na obrázku č. 11.



Obrázek č. 11 (Qi a Hillman, 1997, str. 539). Spektrální obálka šumu po nízkopropustném keprálním filtrování. Spektrální obálka je vyznačena plnou čarou a můžeme vidět, že je vyhlazenější než spektrální obálka šumu na obrázku č. 10. Původní spektrum je naznačeno tečkovaně.

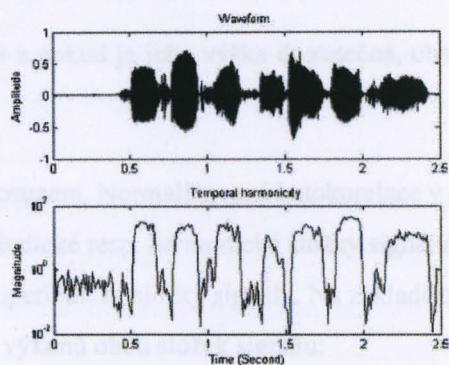
Teoreticky se u patologických hlasů můžeme setkat s odlišným tvarem vlny, která se ovšem pravidelně, tj. každou periodu, opakuje. Konvenční odhady HNR tento rozdíl nepostihnou, a proto je zapotřebí dalších metod, které přesahují tradiční popis HNR (Murphy, 2006). Příkladem takových metod můžou být keprálně založené přístupy, které nám poskytují údaj o periodicitě hlasu, míry CPP a SRA (Murphy, 2006). CPP (angl. cepstral peak prominence) vyjadřuje amplitudu první keprální „harmonické“, indikuje periodicitu signálu a je také spolehlivým korelátem hlasové kvality. SRA (angl. sum of harmonic amplitudes) využívá amplitudy všech keprálních „harmonických“ a přináší tak informaci o bohatosti harmonické struktury ve spektru (Murphy, 2006).

Zcela nový přístup, který konvenční HNR rozšiřuje na skupinu harmonicit, představili Yu & Wang (2006). Motivací autorů pro tuto novou míru harmonické struktury (angl. harmonic structure measure) byl fakt, že tradiční HNR nebo kontura f_0 neposkytují o harmonické struktuře signálu dost informací, které by se daly využít pro její další systematickou analýzu. Tato míra harmonické struktury sestává ze čtyř harmonicit: mřížkové harmonicity (angl. grid harmonicity), časové harmonicity (angl. temporal harmonicity), segmentálně-spektrální harmonicity (angl. segment-spectral harmonicity) a segmentální harmonicity (angl. segmental harmonicity). Mřížková harmonicita (viz obrázek č. 12) představuje harmonicitu na nejnižší úrovni signálu. Měří úplnost jednotlivých harmonických v každém analyzačním okně a slouží jako základ pro odvození ostatních harmonicit. Časová harmonicita (viz obrázek č. 13) reprezentuje sílu harmonicity a získáme ji jako váženou sumu mřížkových harmonicit v daném analyzačním rámci. Segmentálně-spektrální harmonicita vyhodnocuje integritu jednotlivých harmonických v rámci segmentu a segmentální harmonicita vyjadřuje celkovou sílu harmonické struktury segmentu. Celkově vzato nám tato sada harmonicit poskytuje údaje o tom, kde a do jaké míry se harmonická struktura signálu zachovává nebo ztrácí v přítomnosti nepříznivých podmínek. Takto získané informace lze využít nejen pro již zmíněnou systematickou analýzu harmonické struktury signálu, ale i pro úkony v oblasti zpracování řeči jako je přesnější detekce f_0 , analýza řeči z nekvalitních/zašuměných nahrávek, detekce znělých úseků řeči v hlučném prostředí a následné zvýraznění jejich harmonické struktury.



Obrázek č. 12 (Yu & Wang, 2006, str. 2940).

Mřížková harmonicitá: v pořadí ze shora následuje oscilogram, spektrogram a záznam mřížkové harmonicity, které byly získány na základě realizace čínské věty (mandarinská čínština) „qi2 shi2 ta1 xi3 huan1 he1 cha2.“ Český překlad: „Opravdu má ráda čaj.“



Obrázek č. 13 (Yu & Wang, 2006, str. 2941).

Časová harmonicitá: nahoře oscilogram a pod ním záznam časové harmonicity realizace čínské věty (mandarinská čínština) „qi2 shi2 ta1 xi3 huan1 he1 cha2.“ Český překlad: „Opravdu má ráda čaj.“

2.2.4 Měření HNR v programu PRAAT

Pro práci s materiálem v experimentální části této práce byl využit program *PRAAT* autorů Boersma & Weenink (2007). Metodu zjišťování HNR v tomto softwaru popisuje Boersma (1993). Zakládá se na detekci periodicity v řeči pomocí časové autokorelace signálu. Autokorelační doména se v *PRAATu* využívá jak k detekci f_0 , tak i k odhadům HNR.

Pozice maxima autokorelační funkce představuje v autokorelační doméně nejlepšího kandidáta základní frekvence a relativní výška tohoto maxima odráží stupeň periodicity signálu. Určení přesné pozice a výšky tohoto maxima ve frekvenční doméně je v důsledku vzorkování a používání oken problematické. Metoda autokorelace signálu tento problém řeší a dle autora se vyznačuje také větší přesností při detekci f_0 a měření HNR než metody frekvenční domény. Východiskem pro odhad HNR je zjištění autokorelace signálu $r_x(\tau)$ a toho dosáhneme vydělením autokorelace okénkového signálu $r_a(\tau)$ autokorelací okna $r_w(\tau)$:

$$r_x(\tau) \approx (r_a(\tau) / r_w(\tau))$$

Autokorelace je funkcí časového posunu τ . Funkce má globální maximum v čase $\tau = 0$ a platí, že pokud existuje nějaké globální maximum pro $\tau > 0$, mluví se o periodickém signálu. Pokud takové globální maximum neexistuje, mohou se ve funkci vyskytnout lokální maxima a pak platí, že pokud

v signálu nějaké lokální maximum v čase τ_{\max} existuje a pokud je jeho výška dostatečná, obsahuje signál částečně i periodickou složku.

Výkon signálu reprezentuje autokorelace s nulovým posunem. Normalizovaná autokorelace v posunu τ_{\max} , $r'_x(\tau_{\max})$, pak odpovídá relativnímu výkonu periodické resp. harmonické složky signálu a její komplement, $1 - r'_x(\tau_{\max})$, představuje relativní výkon aperiodické složky signálu. Na základě tohoto pak můžeme definovat HNR jako logaritmický poměr výkonů obou složek signálu:

$$HNR [v \text{ dB}] = 10 \log (r'_x(\tau_{\max}) / (1 - r'_x(\tau_{\max})))$$

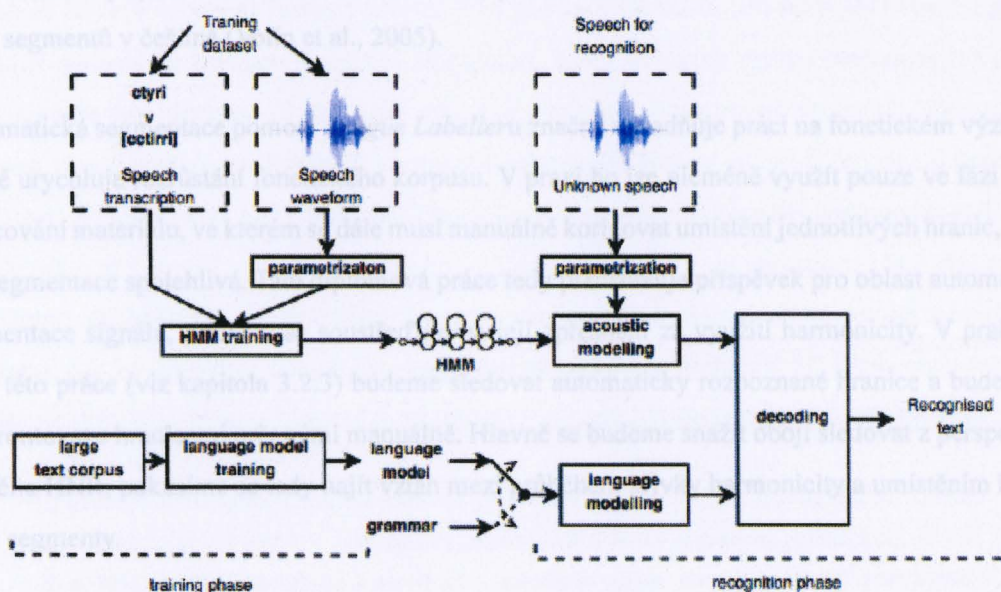
Jak zmiňuje předcházející kapitola (viz kapitola 2.2.3), je pro měření HNR důležitá i délka analyzačního okna. Ta závisí na minimální hodnotě f_0 . Pro detekci základní frekvence v autokorelační doméně je potřebné analyzační okno o délce tří základních period. Pro měření HNR je potřeba dvojnásobek, to znamená, že *PRAAT* při měření HNR používá okno dlouhé šest period. Při standardním nastavení *Minimum Pitch* na 75 Hz to odpovídá 80 ms.

2.3 „Prague Labeller“ – fonetická segmentace na bázi HMM

Potřeba fonetické segmentace v dnešní době roste s rozvojem řečových technologií a také s nároky na fonetický výzkum. V oblasti řečových technologií potřebujeme segmentovaný materiál například pro trénování neuronových sítí (Pollák et al., 2007) a pro fonetický výzkum mají zase zásadní význam databáze mluvené řeči, které obsahují segmentovaný materiál, protože čím větší je množství zpracovaných dat, tím spolehlivější výsledky může fonetický výzkum poskytnout (Volín et al., 2005). Manuální segmentace je ovšem časově náročná, a proto se začaly hledat metody automatického rozpoznávání řeči, které by tento úkol řešily rychle a efektivně (Volín et al., 2005). Jednou z nejčastěji používaných metod v řečových technologiích jsou skryté Markovovy modely HMM (angl. hidden Markov models). Tato metoda se využívá v mnoha systémech rozpoznávání řeči pracujících v reálném čase (např. systémy pro diktování nebo pro hlasové ovládání různých přístrojů) a své uplatnění našla i v oblasti fonetické segmentace (Pollák et al., 2007).

Pro potřeby fonetického výzkumu byl v rámci spolupráce Fonetického ústavu FF UK s ČVUT navržen nástroj pro segmentaci, tzv. *Prague Labeller*. Tento segmentátor pracuje na základě nuceného zarovnání trénovaných HMM modelů monofonů (angl. forced alignment of trained HMMs), byl konstruován na bázi HTK toolkitu a následně integrován do prostředí *PRAATu* (Pollák et al., 2007).

Pro přesnější představu se na základě přednášek docenta Polláka (2008) pokusím vysvětlit princip tohoto algoritmu: Obecné východisko skrytých Markovových modelů spočívá v nahrazení vzoru, tj. například vzoru konkrétního monofonu, se kterým se při segmentaci daný materiál srovnává, statistickým modelem parametrů, které získáme zprůměrováním velkého množství promluv ve fázi trénování. Rozpoznávání řeči na bázi HMM pak sestává z natrénovaných modelů, které jsou tvořeny více stavy, a z hledání modelu, pro který je, pravděpodobnost průchodu všemi stavy největší. Pro zefektivnění výpočtu lze využít Viterbiův algoritmus, který hledá pouze jednu pravděpodobnost jakožto optimální průchod modelem, přičemž sekvence stavů zůstává skryta. Pro identifikaci jednotlivých segmentů mluvené řeči se obvykle využívají levo-pravé modely monofonů se třemi emitujícími stavy bez přeskoků. Emitující stavy jako n-rozměrné Gaussovské funkce modelují parametry řečového signálu pomocí hustoty pravděpodobnosti. Neemitující stavy pak slouží zejména k zarovnání modelů. Parametrizace řeči, která představuje základ pro vytvoření modelu, se skládá z kepstrálních parametrů, dynamických a akceleračních parametrů a energetických parametrů, které popisují řečový signál. Nejčastěji se jako kepstrální parametry používají mel-frekvenční kepstrální koeficienty MFCC. Blokové schéma fungování rozpoznávače řeči na bázi HMM znázorňuje obrázek č. 14.



Obrázek č. 14 (Pollák, 2008). Blokové schéma fungování rozpoznávače řeči na bázi HMM.

Segmentátor *Prague Labeller* pracuje také s levo-pravými modely monofonů, které sestávají ze tří emitujících stavů a neumožňují přeskoky. Celkově bylo pro češtinu modelováno 44 monofonů a hlasivková exploziva, tzv. ráz, který byl modelován jako krátká pauza. V modelech se nepočítá se šva [ə] jako s výsledkem redukce vokálů a ani se znělým alofonem [ɣ] fonému /x/ (Pollák et. al, 2007),

2007), stejně tak se nerozlišuje mezi znělou a neznělou realizací fonému /i/ (Volín et al., 2005). Přesnost segmentace, tedy přesnost umístění hranic jednotlivých hlásek, závisí na mnoha faktorech, jako je použitá parametrizace, proces trénování modelů, struktura modelů atd. Pro *Prague Labeller* byl výběr těchto nastavení zvolen na základě předchozího výzkumu ve studii Volín et al. (2005). Trénování modelů pro segmentátor probíhalo na databázích Speechdat (1052 mluvčích) a SPEECON (550 mluvčích), což umožnilo odstranit vliv mluvčího, jelikož právě velké databáze obsahující mnoho různých mluvčích poskytují dost materiálu pro natrénování modelů nezávislých na mluvčím. Jako parametry pro popis signálu se používají mel-frekvenční keprální koeficienty MFCC a optimálním vstupem pro segmentaci v segmentátoru *Prague Labeller* je materiál se vzorkovací frekvencí 16 kHz. Pro analýzu tohoto signálu, tedy pro jeho segmentaci a následnou parametrizaci, se používá analyzační okno dlouhé 25 ms s časovým krokem posunu 10 ms nebo 16ms okno s krokem 8 ms (Pollák et al., 2007). Vliv délky okna a časového kroku detailněji prozkoumali Volín et al. (2005), kteří za tímto účelem provedli analýzu délky segmentů v češtině, a konstatují, že pro rozpoznávání českých segmentů je vhodnější kratší okno. Minimální délka segmentu, který lze rozpoznat pomocí modelu se třemi emitujícími stavy za použití okna dlouhého 32 ms s krokem 16 ms, totiž činí až 48 ms, zatímco v případě 16ms okna s krokem 8 ms je tato délka 24 ms, což lépe odpovídá trvání krátkých segmentů v češtině (Volín et al., 2005).

Automatická segmentace pomocí *Prague Labelleru* značně usnadňuje práci na fonetickém výzkumu a také urychluje rozrůstání fonetického korpusu. V praxi ho lze nicméně využít pouze ve fázi předzpracování materiálu, ve kterém se dále musí manuálně korigovat umístění jednotlivých hranic, má-li být segmentace spolehlivá. Tato diplomová práce tedy představuje příspěvek pro oblast automatické segmentace signálu, přičemž se soustřeďuje na její zpřesnění za využití harmonicity. V praktické části této práce (viz kapitola 3.2.3) budeme sledovat automaticky rozpoznané hranice a budeme je konfrontovat s hranicemi určenými manuálně. Hlavně se budeme snažit obojí sledovat z perspektivy průběhu HNR, pokusíme se tedy najít vztah mezi průběhem křivky harmonicity a umístěním hranic mezi segmenty.

2.4 Segmentace hlásek – problémy a pravidla

V současnosti není přesnost segmentace *Prague Labelleru* stoprocentní a automaticky detekované hranice segmentů se často liší od těch, které byly určeny manuálně. Závislost přesnosti segmentace na různém nastavení modelů a parametrizaci v rámci rozpoznávání na bázi HMM zkoumala studie Volín et al. (2005). Jako nejlepší z testovaných nastavení se ve studii prokázal algoritmus *mfcc16x*,

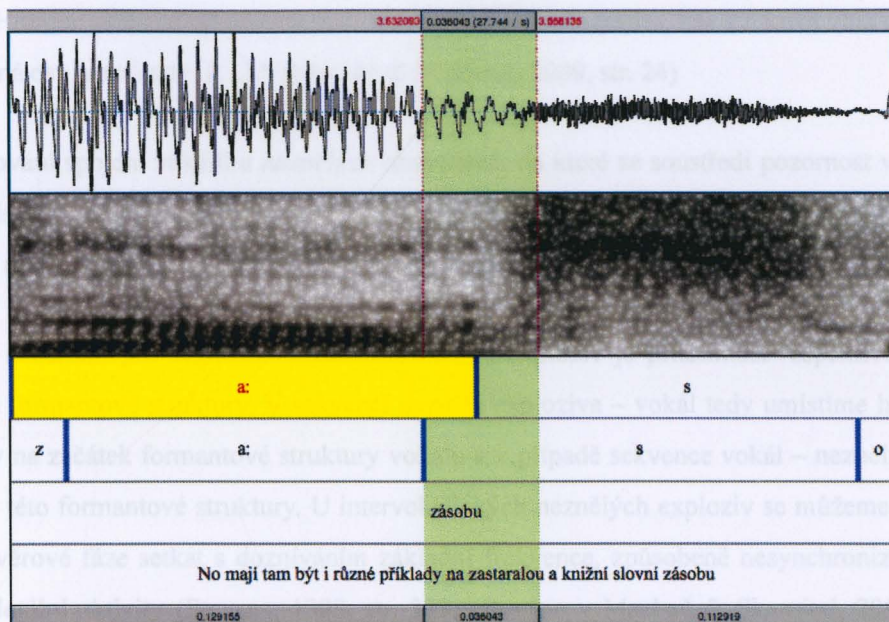
který pro modelování emitujících stavů využívá vážené směsi Gaussovských funkcí a při parametrizaci signálu pracuje s mel-frekvenčními kepstrálními koeficienty získanými na základě analýzy pomocí 16ms okna s 50% překryvem. Nicméně i tento algoritmus vykazoval průměrný rozdíl mezi automatickým a manuálním umístěním hranic 19,3 ms, přičemž u vokálů činila tato odchylka 18,1 ms a u neznělých obstruentů 14,7 ms (Volín et al., 2005).

Tato diplomová práce si klade za cíl ověřit možnosti využití harmonicity jako indikátoru hranic mezi segmenty v češtině, a proto se zaměří právě na hranice mezi vokály a neznělými obstruenty, kde v signálu dochází ke změně znělosti. Výběr těchto skupin hlásek byl motivován předešlou prací, která prokázala, že rozdíl mezi hodnotami harmonicity u těchto dvou skupin hlásek je největší, tj. že vokály dosahují nejvyšších hodnot HNR a neznělé obstruenty naopak těch nejnižších (Heranová, 2009).

Segmentace plynulé řeči s sebou přináší jistá úskalí, protože řeč jako taková nesestává ze sledu diskrétních jednotek, které by byly jasně oddělitelné, ale představuje kontinuum volně plynoucí v čase. Fonetická segmentace je tedy z hlediska komunikace a používání řeči nepřirozená a využíváme ji hlavně za účelem uchopení systému jazyka, jeho popisu a výzkumu. Při realizaci hlásek v proudu řeči totiž dochází ke změnám nastavení artikulátorů, které vedou ke vzniku přechodových oblastí mezi jednotlivými hláskami. V akustické rovině se tyto přechodové oblasti vyznačují změnou tvaru zvukové vlny a změnou spektrálních vlastností/spektrálního složení signálu. Zejména u nedbalejší artikulace, při pozvolném přechodu jedné hlásky na tu následující, mohou být tyto přechodové oblasti mezi hláskami dlouhé a přechod mezi nimi může být pozvolný (viz obrázek č. 15). Určit příslušnost těchto přechodných pásem k tomu či onomu segmentu, respektive umístit hranici mezi danými hláskami, pak bývá sporné, což představuje jeden ze zásadních problémů při segmentaci řečového signálu na jednotlivé hlásky. Machač & Skarnitzl (2009, str. 22) tento problém popisují následovně:

„Plynulost artikulačních pohybů a způsob synchronizace glotálních a supraglotálních úkonů má za následek, že některé rysy mohou přesahovat hranice segmentu, jemuž jsou vlastní. To se týká například tendence k otevřenosti vokálního traktu, nazality, znělosti a neznělosti. Zmíněné rysy pak spoluutvářejí artikulační a akustický charakter sousedních hlásek, kde působí jako rysy nevlastní. Tento jev může umístování hranic segmentů značně znesnadňovat.“

Na obrázku č. 15 můžeme vidět realizaci hlásek, při které mezi hláskami vzniká relativně dlouhá a pro segmentaci problematická přechodová oblast.



Obrázek č. 15 (zpracovaný materiál – nahrávka MANA30_M12). Realizace segmentů [a:s] s dlouhým přechodovým pásmem. Zelený pruh označuje přechodové pásmo dlouhé 36 ms. V oscilogramu je toto přechodové pásmo signalizováno zjednodušeným tvarem zvukové vlny a přibývajícím šumem. Ve spektrogramu (0 – 8 kHz) můžeme v oblasti přechodového pásma pozorovat slábnoucí formantovou strukturu vokálu a přibývající šum. V další vrstvě pod spektrogramem je naznačena manuálně korigovaná hranice mezi [a:] a [s], ve vrstvě následující je k porovnání načrtnuta automaticky rozpoznaná hranice.

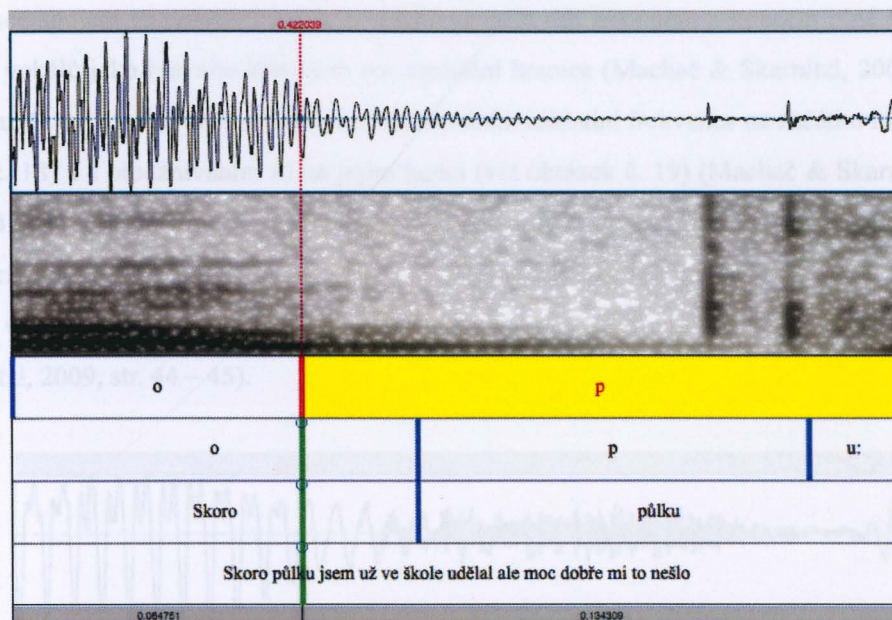
Má-li být fonetický výzkum validní, měl by se zakládat na materiálu, který je homogenní ve smyslu svého zpracování. Aby bylo toto kritérium splněno, je potřeba poskytnout jednotný rámec pro segmentaci, tedy stanovit určitá pravidla segmentace, která zaručí jednotnost segmentace i jednotný způsob řešení sporných případů (viz obrázky č. 15 – 19). Za tímto účelem vznikla ve Fonetickém ústavu příručka Machač & Skarnitzl (2009), která pro segmentaci stanovuje následující základní pravidla:

1. „Pro segmentaci mnoha sekvencí je velice důležitá přítomnost formantové struktury; hranice se pokoušíme umísťovat na kraj tzv. formantových sloupků (tj. tmavé vertikální oblasti ve spektrogramu, které představují maxima akustické energie při každém hlasivkovém pulzu.) Plná formantová struktura je charakterizována sledem zřetelných formantových sloupků.“ (Machač & Skarnitzl, 2009, str. 23)
2. „Je-li v signálu přítomna přechodová část [...], umístíme hranici do temporálního středu této oblasti.“ (Machač & Skarnitzl, 2009, str. 24)

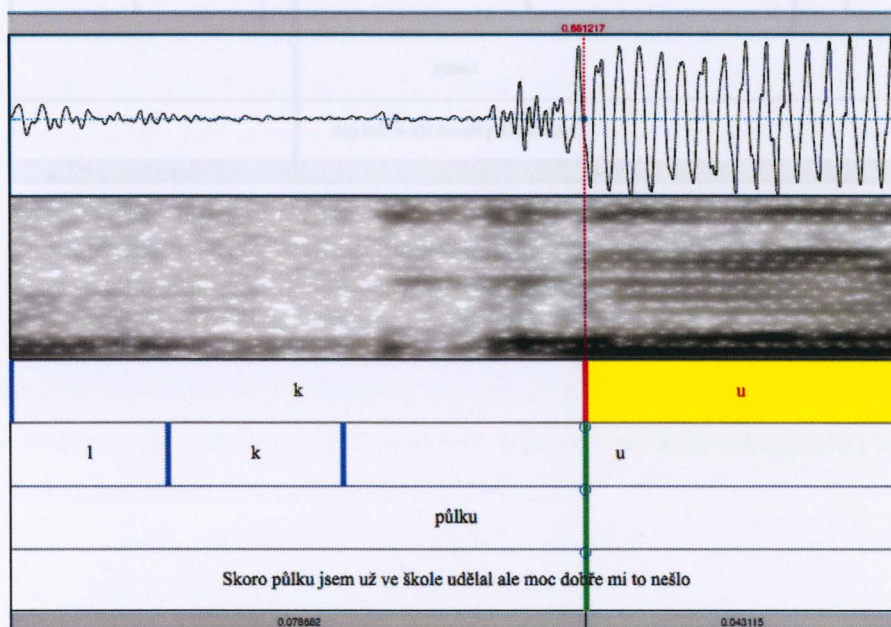
3. „Hranice se umísťujú do průchodu nulou (bod v oscilogramu, kde zvuková vlna protíná osu amplitudy) [...]“ (Machač & Skarnitzl, 2009, str. 24)

Segmentování spojení vokálů a neznělých obstruentů, na které se soustředí pozornost v této práci, dále popisují pravidla pro segmentaci intervokálních exploziv a frikativ z příručky Machač & Skarnitzl (2009, kapitola 2 a 3). Jedná se o následující pravidla:

Hlavním kritériem pro segmentaci intervokálních exploziv je přítomnost respektive nepřítomnost plné formantové struktury. V sekvenci neznělá exploziva – vokál tedy umístíme hranici mezi segmenty na začátek formantové struktury vokálu a v případě sekvence vokál – neznělá exploziva na konec této formantové struktury. U intervokálních neznělých exploziv se můžeme na začátku jejich závěrové fáze setkat s dozníváním základní frekvence, způsobené nesynchronizací glotální a supraglotální aktivity (Stevens, 1998, str. 333; citováno v Machač & Skarnitzl, 2009, str. 27). U sekvence vokál – neznělá exploziva se s tímto jevem setkáváme poměrně často (viz obrázek č. 16), vymezení hranic segmentů se nicméně pořád řídí kritériem plné formantové struktury. Stejný postup platí i pro předznívání základní frekvence ještě před nástupem plné formantové struktury u sekvencí neznělá exploziva – vokál (viz obrázek č. 17) (Machač & Skarnitzl, 2009, str. 39). V některých případech se přechodové oblasti mezi hláskami mohou vyznačovat nízkým kontrastem mezi akustickými vlastnostmi jednotlivých hlásek. Ve spektrogramech takovýchto realizací se můžeme setkat s pozvolným rozpadem formantové struktury, malou zřetelností formantových sloupků nebo s překryvem formantové struktury šumem. Tyto projevy jsou časté zejména pro začátky intervokálních exploziv, což je způsobeno delší dobou přibližování artikulátorů v porovnání s jejich oddálením v případě sekvencí exploziva – vokál (Machač & Skarnitzl, 2009, str. 35 – 36). Hranice se v těchto případech umísťuje do časového středu přechodové fáze.

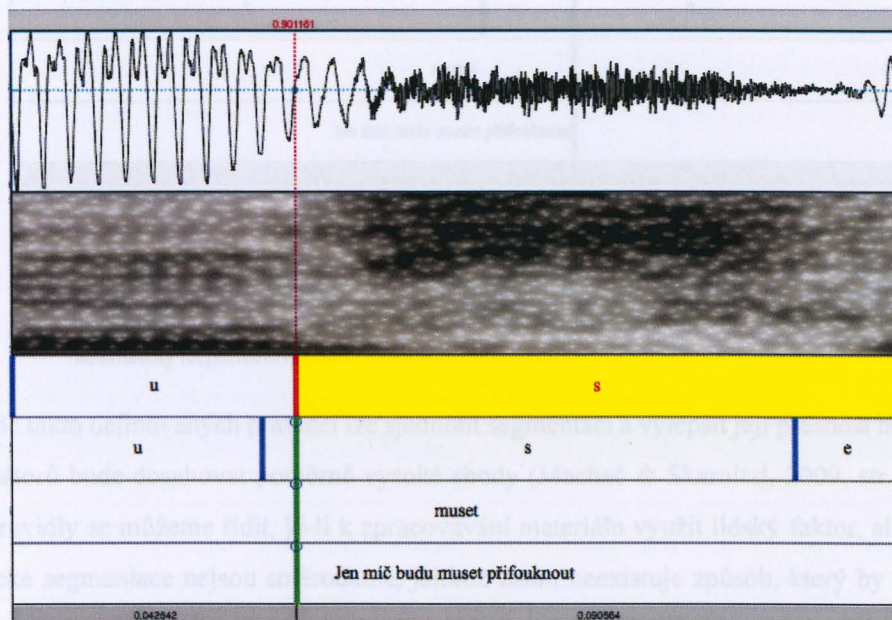


Obrázek č. 16 (zpracovaný materiál – nahrávka SLZA62_M13). Realizace segmentů [op] s dlouhým přezníváním znělosti do závěrové fáze explozivny. V další vrstvě pod spektrogramem, který znázorňuje frekvenční rozpětí 0 – 8 kHz, je naznačena manuálně korigovaná hranice mezi [o] a [p], ve vrstvě následující je k porovnání naznačena automaticky rozpoznávaná hranice.

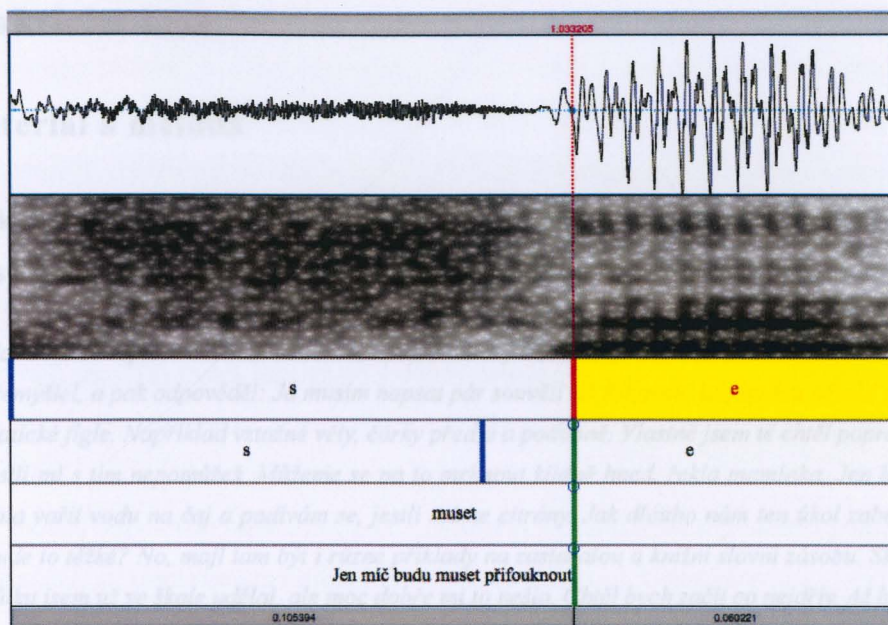


Obrázek č. 17 (zpracovaný materiál – nahrávka DVOA33_M13). Realizace segmentů [ku] s předzníváním základní frekvence ještě před nástupem formantové struktury vokálu. V další vrstvě pod spektrogramem, který znázorňuje frekvenční rozpětí 0 – 8 kHz, je naznačena manuálně korigovaná hranice mezi [k] a [u], ve vrstvě následující je k porovnání naznačena automaticky rozpoznávaná hranice.

Segmentace spojení neznělých frikativ a vokálů se opět řídí nástupem a koncem plné formantové struktury vokálů jako hlavním kritériem pro umístění hranice (Machač & Skarnitzl, 2009, str. 42), i když se u těchto spojení opět setkáváme s dozníváním základní frekvence na začátku frikativy (viz obrázek č. 18) a s předzníváním f_0 na jejím konci (viz obrázek č. 19) (Machač & Skarnitzl, 2009, str. 40 – 41). U sekvencí vokál – frikativa se dále poměrně často setkáváme s pozvolnějším oslabováním formantové struktury. V těchto případech, stejně tak jako u pozvolného nástupu formantové struktury ve spojeních frikativa – vokál, umísťujeme hranici do středu přechodové oblasti (Machač & Skarnitzl, 2009, str. 44 – 45).



Obrázek č. 18 (zpracovaný materiál – nahrávka DVOA33_M19). Realizace segmentů [us] s předzníváním znělosti do frikativy. V další vrstvě pod spektrogramem, který znázorňuje frekvenční rozpětí 0 – 8 kHz, je naznačena manuálně korigovaná hranice mezi [u] a [s], ve vrstvě následující je k porovnání naznačena automaticky rozpoznaná hranice.



Obrázek č. 19 (zpracovaný materiál – nahrávka DOHA58_19). Realizace segmentů [se] s předzníváním základní frekvence ještě před nástupem formantové struktury vokálu. V další vrstvě pod spektrogramem, který znázorňuje frekvenční rozpětí 0 – 8 kHz, je naznačena manuálně korigovaná hranice mezi [s] a [e], ve vrstvě následující je k porovnání naznačena automaticky rozpoznaná hranice.

Na základě takto definovaných pravidel lze sjednotit segmentaci a vylepšit její přesnost tak, že i skupina anotátorů bude dosahovat poměrně vysoké shody (Machač & Skarnitzl, 2009, str. 14 a 142). Těmito pravidly se můžeme řídit, je-li k zpracování materiálu využit lidský faktor, ale pro účely automatické segmentace nejsou směrodatná, jelikož zatím neexistuje způsob, který by umožňoval jejich implementaci do softwaru jako je *Prague Labeller*. I tento fakt motivoval vznik této práce, ve které se budeme snažit prozkoumat možnosti HNR jako indikátoru hranic segmentů, který by dosahoval větší korelace s hranicemi určenými na základě zmíněných segmentačních pravidel než s hranicemi určenými na základě natrénovaných modelů, které využívá *Prague Labeller*. Tyto modely jsou pro nás totiž foneticky neinterpretovatelné, a proto představuje tato práce přínos i v tom ohledu, že ověřuje možnost využití foneticky průhledné míry v procesu automatické segmentace řeči. Ta by tak dále již nemusela zůstat výhradní doménou techniků.

3. Praktická část

3.1 Materiál a metoda

Východisko naší analýzy tvořily nahrávky 12 mluvčích pořízené ve Fonetickém ústavu. Mluvčí byli nekuřáci starší 30 let (viz tabulka č. 1) a četli následující text:

Maminka se zeptala Milana: Milánku, už máš hotový úkol? Kdy ho budeš psát? Milan chvilku přemýšlel, a pak odpověděl: Já musím napsat pár souvětí na Řihovou, kde budou nějaké gramatické figle. Například vztažné věty, čárky před a a podobně. Vlastně jsem tě chtěl poprosit, jestli mi s tím nepomůžeš. Můžeme se na to mrknout klidně hned, řekla maminka. Jen bych dala vařit vodu na čaj a podívám se, jestli máme citróny. Jak dlouho nám ten úkol zabere? Bude to těžké? No, mají tam být i různé příklady na zastaralou a knižní slovní zásobu. Skoro půlku jsem už ve škole udělal, ale moc dobře mi to nešlo. Chtěl bych začít co nejdřív. Až budu hotov, došel bych ti do lékárny pro ten pneumocyt. A potom půjdu hrát fotbal. Včera jsem dal čtyři góly. Nebýt Láďových faulů, mohlo jich být víc. Jen míč budu muset přifouknout. Neboj se, dám pozor na auta.

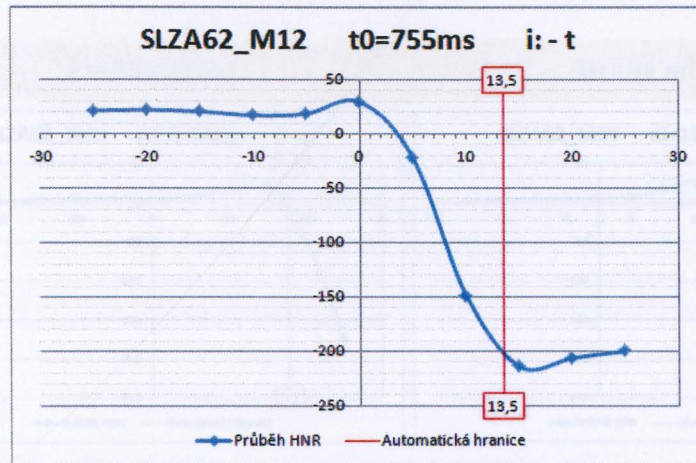
Tabulka č. 1. Seznam mluvčích.

Kód mluvčího	Věk	Pohlaví
CERV37	37	muž
DOHA58	58	žena
DVOA33	33	žena
MANA30	30	žena
MURK30	30	muž
MURA34	34	žena
SLZA62	62	žena
SLEA36	36	žena
SOLA30	30	žena
STES31	31	muž
STEP43	43	muž
VERA33	33	žena

Získaný materiál byl zpracováván v softwaru *PRAAT* (verze 4.6.36). Pomocí *Prague Labelleru* byla provedena automatická segmentace nahrávek. Ta v anotaci využívá transkripční abecedu SAMPA, která se také nachází v popiscích grafů a obrázků. Následně bylo umístění automaticky určených hranic manuálně korigováno podle zásad segmentační příručky Machač & Skarnitzl (2009). Z takto

zpracovaného materiálu se dále vycházelo při zjišťování průběhu harmonicity. Zaměřili jsme se na následující spojení hlásek: neznělý obstruent – vokál a vokál – neznělý obstruent. Skupinu vokálů reprezentovaly hlásky [ɪ i: e e: a a: o o: u u: \widehat{ou} \widehat{au} \widehat{eu}] a skupina neznělých obstruentů zahrnovala neznělé explozivy, frikativy a afrikáty [p t c k f s f x \widehat{ts} $\widehat{tʃ}$]. Z analýzy jsme vynechali kontexty obsahující neznělou frikativu [x], ve kterých docházelo k její znělostní asimilaci (například ve spojení slov „bych začít“ nebo „jich být“). Tento fakt zmiňujeme proto, že *Prague Labeller* modeluje foném /x/ pouze jako neznělou frikativu [x] bez ohledu na jeho znělostní charakteristiku závislou na kontextu. Palková (1997, str. 213) uvádí jako znělé varianty fonému /x/ hlásky [ɣ] a [ɦ], nicméně v *Prague Labelleru* není k dispozici model ani jedné z těchto hlásek jakožto možný výsledek znělostní asimilace hlásky [x]. Z toho vyplývá, že bylo nutné z analýzy vyřadit takové kontexty, které sice podle transkripce v anotaci obsahovali hlásku [x], podle znělostních charakteristik se ovšem jednalo o znělou variantu fonému /x/, která vznikla v důsledku regresivní asimilace. Celkově bylo tedy [x] zastoupeno pouze třemi výskyty, a to v kontextech „tě chtěl“, „bych ti“ a „Láďových faulů“.

Pomocí křížové korelace v programu *PRAAT* byly nahrávky za použití standardního nastavení převedeny na objekt *Harmonicity*, ve kterém byly s krokem 5 ms zjišťovány hodnoty HNR u zmíněných spojení hlásek, a to v intervalu ± 25 ms v okolí manuálně určené hranice. Výsledkem bylo naměřených 1955 položek, ze kterých jsme pro naši analýzu udělali užší výběr na základě kritéria umístění automatické hranice v porovnání s manuálně korigovanou hranicí. Tento výběr obsahuje 459 položek, u kterých se rozdíl v umístění obou hranic pohybuje v rozmezí 10 – 20 ms. Toto kritérium bylo zvoleno na základě úvahy, že časový úsek 10 ms lze u mužských hlasů považovat za délku jedné periody, kdežto 20 a více ms již představuje časový úsek, který může v některých případech odpovídat délce jednoho segmentu. Tato práce se tedy soustředí právě na ty rozdíly v umístění hranic, které se nacházejí v rámci daného segmentu a zároveň jsou větší než délka jedné periody u hlubokých hlasů, ale menší než možné trvání celého segmentu. Tím se z analýzy vyloučí případy, ve kterých došlo k chybnému umístění automatické hranice v důsledku přechů nebo hezitací, které *Prague Labeller* nereflektuje, protože provádí nucené zarovnání modelů na základě poskytnutého ortografického zápisu.

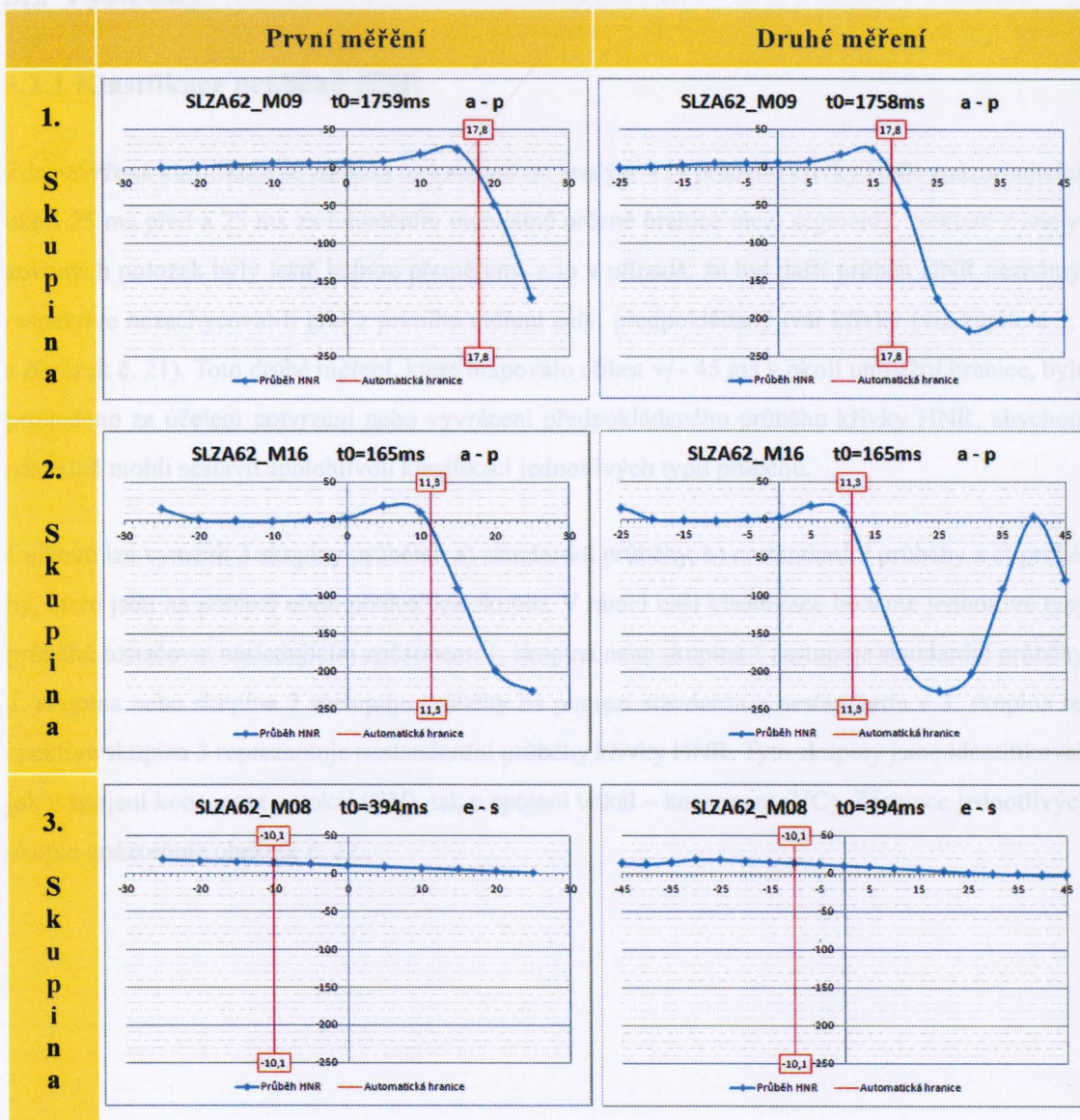


Obrázek č. 20 (zpracovaný materiál – nahrávka SLZA62_M12).

Příklad typického průběhu křivky HNR.

Údaje o hodnotách HNR na hranicích jednotlivých spojení hlásek jsme si nechali zobrazit v grafech, přičemž jsme zjistili, že můžeme často pozorovat průběh křivky HNR, v kapitole 3.2.1 označovaný jako standardní průběh, který je poměrně jednoznačný (viz obrázek č. 20). V jiných případech jsme se setkali s grafy, jejichž křivka naznačovala podobnost se standardním průběhem, ale dané zobrazení hodnot HNR v rozmezí pouze ± 25 ms kolem manuálně určené hranice neposkytovalo dostatek informací pro spolehlivé potvrzení hypotézy o shodnosti obou křivek. Proto jsme došli k závěru, že pro jednoznačnou a jasnou klasifikaci je třeba provést druhé měření, které by poskytovalo větší průzor do průběhu křivky HNR mezi segmenty. Toto měření jsme provedli pomocí upraveného skriptu, který HNR měřil v časovém rozmezí ± 45 ms od umístění manuálně korigované hranice. Porovnání grafů z prvního a druhého měření je vidět na obrázku č. 21. Pouze průběh HNR ve spojení [ap] v nahrávce SLZA62_M09 potvrdil předpoklad o standardním průběhu křivky. Ostatní křivky naznačily jiný průběh, který byl reflektován ve výsledné klasifikaci (viz kapitola 3.2.1).

Obrázek č. 21. Porovnání grafů podle prvního a druhého měření. Jednotlivé grafy v pořadí obsahují přílohu nahrávky, ve které se dané spojení nachází, čas umístění manuálně korigované hranice - 0 a konkrétní spojení hlásek. Každý graf reprezentuje...



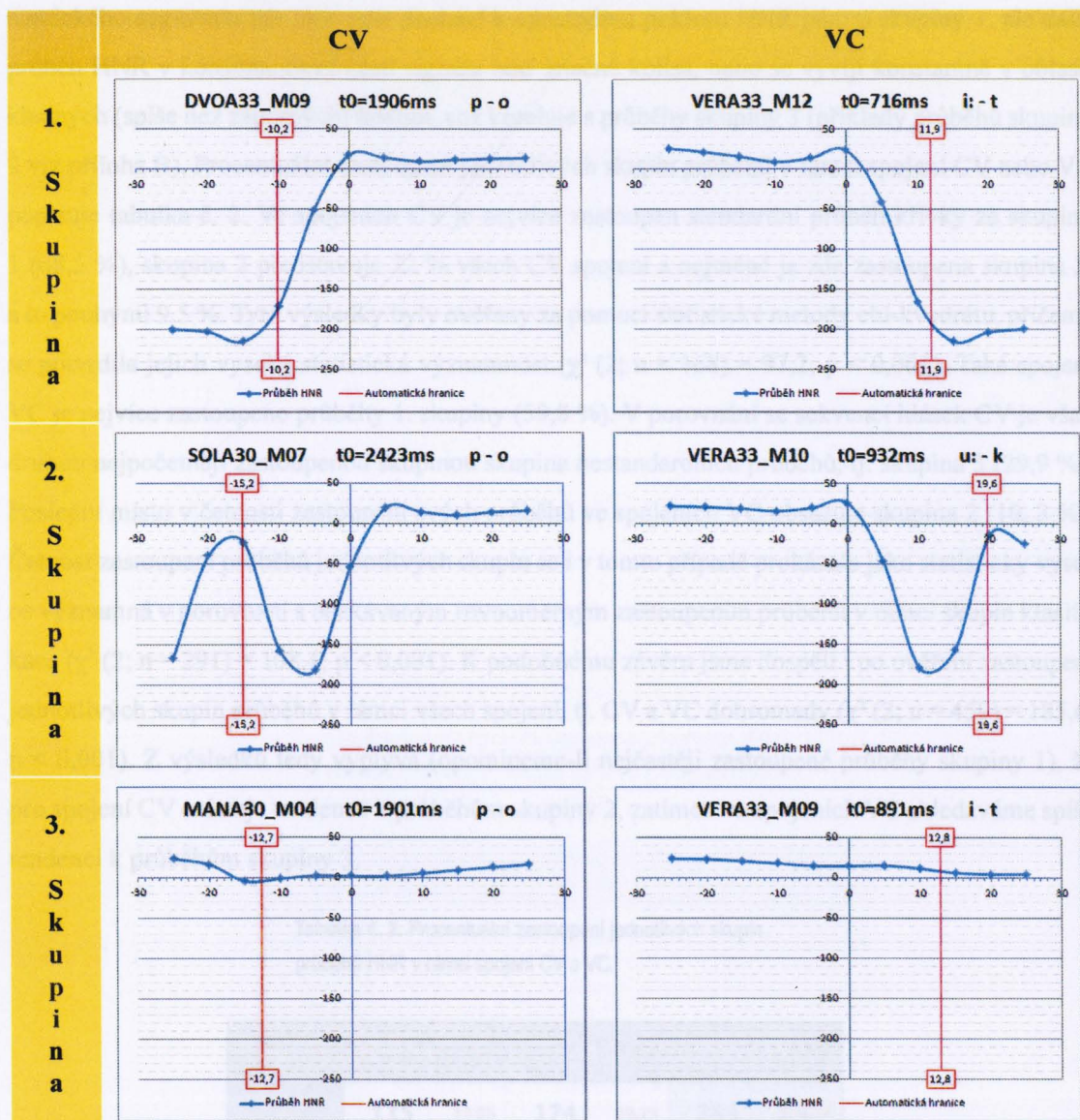
Obrázek č. 21. Porovnání grafů podle prvního a druhého měření. Jednotlivé grafy v popisku obsahují jméno nahrávky, ve které se dané spojení nacházelo, čas umístění manuálně korigované hranice – t_0 a konkrétní spojení hlásek, které graf reprezentuje.

3.2 Výsledky

3.2.1 Klasifikace průběhů HNR

Zde navržená klasifikace se zakládá na kvalitativní analýze 459 průběhů křivky HNR znázorňujících okolí 25 ms před a 25 ms za umístěním manuálně určené hranice mezi segmenty. Některé z analyzovaných položek byly ještě jednou přeměřeny, a to v případě, že byl další průběh HNR neznámý, respektive nezachycoval-li graf z prvního měření celý, předpokládaný tvar křivky (viz kapitola 3.1 a obrázek č. 21). Toto druhé měření, které mapovalo oblast ± 45 ms v okolí umístění hranice, bylo provedeno za účelem potvrzení nebo vyvrácení předpokládaného průběhu křivky HNR, abychom následně mohli sestavit spolehlivou klasifikaci jednotlivých typů průběhů.

Celkově lze vymezit 3 skupiny průběhů: a) standardní průběhy, b) nestandardní průběhy a c) průběhy, které jsou na pomezí obou předešlých skupin. V rámci naší klasifikace budeme jednotlivé typy průběhů označovat následujícím způsobem: 1. skupina nebo skupina 1 zastupuje standardní průběhy, 2. skupina nebo skupina 2 zastupuje průběhy na pomezí standardu a nestandardu a 3. skupina respektive skupina 3 reprezentuje nestandardní průběhy křivky HNR. Tyto skupiny jsme identifikovali jak u spojení konsonant – vokál (CV), tak u spojení vokál – konsonant (VC). Zástupce jednotlivých skupin znázorňuje obrázek č. 22.



Obrázek č. 22. Zástupci tří skupin z klasifikace průběhů křivky HNR. Jednotlivé grafy obsahují v popisku jméno nahrávky, ve které se dané spojení nacházelo, čas umístění manuálně korigované hranice – t0 a konkrétní spojení hlásek, které graf reprezentuje.

Standardní průběh neboli skupinu 1 charakterizuje výrazná změna v hodnotách HNR, která se objeví v důsledku přechodu neznělého konsonantického segmentu na segment vokalický a naopak (příklady průběhů skupiny 1 viz příloha A). V porovnání s touto skupinou představuje skupina 3 přesný opak a jsou pro ni charakteristické průběhy, kde k takovéto zásadní změně nedochází a průběh HNR lze popsat jako konstantní (příklady průběhů skupiny 3 viz příloha C). Skupina 2 pak byla vyčleněna jako kombinace skupin 1 a 3, tedy kombinace standardního a nestandardního průběhu. V rámci této skupiny můžeme pozorovat standardní průběh HNR ve vokalické části signálu. S nástupem konso-

nantického segmentu zde sice opět dochází k výraznému poklesu HNR jako u skupiny 1, ale další průběh HNR v konsonantické části signálu buď značně kolísá, nebo se vyvíjí konstantně v oblasti kladných (spíše než záporných) hodnot, což koreluje s průběhy skupiny 3 (příklady průběhů skupiny 2 viz příloha B). Procentuální zastoupení jednotlivých skupin průběhů v rámci spojení CV nebo VC popisuje tabulka č. 2. Ve spojení CV je nejvíce zastoupen standardní průběh křivky ze skupiny 1 (68,5 %), skupina 2 představuje 22 % všech CV spojení a nejméně je zde zastoupena skupina 3, a to pouhými 9,5 %. Tyto výsledky byly ověřeny za pomoci statistické metody chí-kvadrátu, přičemž se potvrdila jejich vysoká statistická významnost ($\chi^2(2; n = 168) = 97,2; p < 0,001$). Také spojení VC je nejvíce zastoupeno průběhy 1. skupiny (59,8 %). V porovnání se sekvencí hlásek CV je však druhou nejpočetněji zastoupenou skupinou skupina nestandardních průběhů, tj. skupina 3 (29,9 %). Poslední místo v četnosti zastoupení svých průběhů ve spojení VC obsazuje skupina 2 (10,3 %). Četnost zastoupení průběhů jednotlivých skupin se i v tomto případě prokázala jako statisticky vysoce významná v porovnání s očekávaným rovnoměrným zastoupením průběhů v rámci skupin klasifikace ($\chi^2(2; n = 291) = 108,4; p < 0,001$). K podobnému závěru jsme dospěli i po ověření zastoupení jednotlivých skupin průběhů v rámci všech spojení, tj. CV a VC dohromady ($\chi^2(2; n = 459) = 185,6; p < 0,001$). Z výsledků tedy vyplývá (opomineme-li nejčastěji zastoupené průběhy skupiny 1), že pro spojení CV existuje tendence k průběhům skupiny 2, zatímco ve spojení VC shledáváme spíše tendenci k průběhům skupiny 3.

Tabulka č. 2. Procentuální zastoupení jednotlivých skupin průběhů HNR v rámci spojení CV a VC.

	CV		VC		CV + VC	
1. Skupina	115	39.8%	174	60.2%	289	100.0%
	68.5%	25.1%	59.8%	37.9%	63.0%	63.0%
2. Skupina	37	55.2%	30	44.8%	67	100.0%
	22.0%	8.1%	10.3%	6.5%	14.6%	14.6%
3. Skupina	16	15.5%	87	84.5%	103	100.0%
	9.5%	3.5%	29.9%	19.0%	22.4%	22.4%
Celkem	168	36.6%	291	63.4%	459	100.0%
	100.0%	36.6%	100.0%	63.4%	100.0%	100.0%

Segmentace spojení CV a VC se podle pravidel segmentační příručky Machač & Skarnitzl (2009) řídí v první řadě kritériem plné formantové struktury vokálů. Znělost a formantová struktura vokálu by přitom měly korelovat s kladnými hodnotami HNR, zatímco průběh křivky HNR neznělého konsonantu, který postrádá znělost a jiné harmonické složky, by měly popisovat záporné hodnoty. Na hranici spojení VC by tedy mělo docházet k přechodu z kladných do záporných hodnot, tj. k výraznému poklesu hodnot v rámci průběhu křivky harmonicity. Budeme-li chtít využít průběh HNR pro segmentaci, můžeme pro tyto účely počítat pouze s průběhy ze skupin 1 a 2, u kterých dochází k popisované razantní změně hodnot v rámci křivky HNR při přechodu na konsonantický segment. Průběhy skupiny 3 jsou pro nás s ohledem na segmentaci nepoužitelné, jelikož reprezentují takové realizace hlásek, v rámci kterých mají hodnoty HNR spíše konstantní průběh. Podle kritéria využitelnosti/nevyužitelnosti křivky HNR pro segmentační účely tedy můžeme vyčlenit dvě skupiny průběhů: skupinu využitelných průběhů obsahující skupiny 1 a 2 a skupinu nevyužitelných průběhů obsahující skupinu 3. Tabulka č. 3 nám dle toho znázorňuje procentuální zastoupení jednotlivých skupin (tj. pro segmentaci využitelných nebo nevyužitelných průběhů) v našem vzorku dat.

Tabulka č. 3. Procentuální zastoupení jednotlivých skupin průběhů HNR v rámci spojení CV a VC z hlediska jejich využitelnosti pro segmentační účely.

	CV		VC		CV + VC	
Využitelný průběh	152	42.7%	204	57.3%	356	100.0%
	90.5%	33.1%	70.1%	44.4%	77.6%	77.6%
Nevyužitelný průběh	16	15.5%	87	84.5%	103	100.0%
	9.5%	3.5%	29.9%	19.0%	22.4%	22.4%
Celkem	168	36.6%	291	63.4%	459	100.0%
	100.0%	36.6%	100.0%	63.4%	100.0%	100.0%

Jak nám naznačují výsledky z tabulky č. 3, většina (77,6 %) průběhů křivky HNR vykazuje příslušnost k typu průběhu, který je pro segmentaci využitelný. Statistické ověření četnosti zastoupení průběhů v jednotlivých skupinách, které jsme provedli pro skupinu sdružující kontexty CV i VC (viz tabulka č. 3), potvrdilo vysokou významnost našich výsledků ($\chi^2(1; n = 459) = 139,5; p < 0,001$), a to v porovnání s rovnoměrným zastoupením průběhu v obou skupinách. Podle jiného úhlu pohledu na klasifikaci průběhů křivky HNR, který najdeme v tabulce č. 2, tvoří skupinu čistě standardních průběhů, tedy skupinu 1, až 63 % všech analyzovaných položek. Z výsledky obou klasifikací tedy vyplývá, že určitý standardní průběh křivky HNR představuje vždy většinu všech analyzovaných

položek. Na základě této skutečnosti tedy můžeme vyvrátit nulovou hypotézu, která říká, že průběh HNR v rámci spojení námi sledovaných segmentů je náhodný. Dále jsme předpokládali, že průběh HNR bude podobný logistické křivce. Tato hypotéza se ovšem zcela nepotvrdila, a to hned z několika důvodů. Za první je to samotná existence klasifikace, jež naznačuje, že průběh HNR nelze popsat pomocí jediné funkce jakou je logistická křivka. Za druhé tuto hypotézu vyvrací tvar křivky v rámci jednotlivých skupin klasifikace, zejména pak v rámci skupin 2 a 3. Jak křivky skupiny 2, tak i křivky ze skupiny 3 vykazují průběh HNR, který je zcela vzdálen tvaru logistické křivky. V jejich případě totiž buď nedochází k poklesu hodnot vůbec (a křivka probíhá konstantně na jedné úrovni bez větších změn), nebo se tento pokles následně promění v růst hodnot. Zmíněnou sigmoidu tedy nejvíc připomíná tvar křivky ve skupině 1. Oproti klasické logistické křivce se ovšem daná křivka liší, a to tak, že se u ní objevuje mírný nárůst hodnot bezprostředně před jejich razantním poklesem u sekvencí VC a také mírný pokles hodnot bezprostředně před jejich nárůstem v případě spojení CV.

Dále jsme provedli analýzu segmentálního zastoupení konsonantů, konkrétně exploziv, frikativ a afrikát, abychom zjistili, zda existuje souvislost mezi tvarem křivky HNR a charakterem konsonantického segmentu. Tabulka č. 4 uvádí zastoupení jednotlivých konsonantických skupin v rámci spojení CV a VC. Tyto výsledky jsme opět prověřili za použití chí-kvadrátu, přičemž se potvrdilo, že zastoupení jednotlivých konsonantů v rámci CV a VC kontextů není rozděleno rovnoměrně, jak bychom předpokládali. Pro CV kontexty vyšlo $p < 0,001$, přičemž $\chi^2 (2; n = 168) = 53,2$ a podobně pro VC kontexty bylo $p < 0,001$ a $\chi^2 (2; n = 291) = 108,5$.

Tabulka č. 4. Zastoupení konsonantických skupin exploziv, frikativ a afrikát v rámci analyzovaných spojení CV a VC.

	CV		VC		CV + VC	
Explozivy	53	25.4%	156	74.6%	209	100.0%
	31.5%	11.5%	53.6%	34.0%	45.5%	45.5%
Frikativy	96	44.7%	119	55.3%	215	100.0%
	57.1%	20.9%	40.9%	25.9%	46.8%	46.8%
Afrikáty	19	54.3%	16	45.7%	35	100.0%
	11.3%	4.1%	5.5%	3.5%	7.6%	7.6%
Celkem	168	36.6%	291	63.4%	459	100.0%
	100.0%	36.6%	100.0%	63.4%	100.0%	100.0%

Výsledky z tabulky č. 4 implikují poměrně vyrovnané zastoupení exploziv (45,5 %) a frikativ (46,8 %) v analyzovaných datech. Pouze výskyt afrikát je poněkud nižší (7,6 %), což ovšem ve složení dat nezpůsobuje nerovnováhu, jelikož lze způsob tvoření afrikát popsat jako kombinaci exploziv a frikativ. Některé koncepce (respektive některé popisy jazyků) dokonce o afrikátách neuvažují jako o amostatné třídě hlásek, ale popisují je jako kombinaci segmentů (Palková, 1997, str. 94). Zastoupení jednotlivých konsonantických skupin v naměřených datech také vyplývá z frekvence jejich výskytu v češtině. Jak uvádějí Těšilová et al. (1987, str. 94), v češtině jsou rovnoměrně využívány explozivy (48,47 %) a frikativy (47,59 %), zatímco hlásky polozávěrové, tedy afrikáty se vyskytují v míře mnohem menší (3,94 %).

Tabulka č. 5 dále dává do souvislosti zastoupení daných konsonantických skupin a klasifikaci průběhů křivky HNR, tj. vyčísluje procentuální zastoupení konsonantických skupin exploziv, frikativ a afrikát v rámci klasifikace průběhů křivky HNR.

Tabulka č. 5. Zastoupení konsonantických skupin exploziv, frikativ a afrikát v klasifikaci průběhů křivky HNR. Sloupce označeny jako „E“ reprezentují zastoupení exploziv, „F“ zastupuje frikativy a „A“ afrikáty. Podbarvení jednotlivých sloupců s daty naznačuje jejich příslušnost ke skupinám CV, VC nebo CV+VC, a tím také indikuje směr, jakým se mají data v tabulce interpretovat, tj. po sloupcích se shora dolů.

	CV			VC			CV + VC		
	E	F	A	E	F	A	E	F	A
1. Skupina	39	66	10	119	40	15	158	106	25
	73.6%	68.8%	52.6%	76.3%	33.6%	93.8%	75.6%	49.3%	71.4%
2. Skupina	9	22	6	17	13	0	26	35	6
	17.0%	22.9%	31.6%	10.9%	10.9%	0.0%	12.4%	16.3%	17.1%
3. Skupina	5	8	3	20	66	1	25	74	4
	9.4%	8.3%	15.8%	12.8%	55.5%	6.3%	12.0%	34.4%	11.4%
Celkem	53	96	19	156	119	16	209	215	35
	100%	100%	100%	100%	100%	100%	100%	100%	100%

Jak bylo stanoveno na základě výsledků z tabulky č. 2, klesá četnost zastoupení průběhů u spojení CV v pořadí: skupina 1, skupina 2, skupina 3. Podle tabulky č. 5 můžeme stejnou klesající tendenci sledovat i v počtu exploziv, frikativ a afrikát, které se objevují v jednotlivých skupinách průběhů

HNR v CV kontextech. Jinak řečeno, nejvíce frikativ z jejich celkového počtu v rámci CV spojení najdeme ve skupině 1 (66), méně pak ve skupině 2 (22) a nejméně ve skupině 3 (8). Stejně relace platí i pro zastoupení exploziv a afrikát ve spojeních CV, což naznačuje, že je zastoupení daných konsonantických skupin v rámci jednotlivých skupin průběhů HNR rovněž rovnoměrné, respektive že toto zastoupení nevykazuje významné rozdíly.

Pro skupinu VC klesala četnost zastoupení skupin průběhů HNR v tomto pořadí: skupina 1, skupina 3 a skupina 2 (viz tabulka č. 2). Výsledky z tabulky č. 5 obsahující zastoupení konsonantických skupin v jednotlivých typech průběhů HNR s touto tendencí korelují pouze v případě exploziv a afrikát. Podíváme-li se na zastoupení frikativ v jednotlivých typech průběhů, zjistíme, že nejvíce frikativ se objevuje ve skupině 3 (66, což představuje 55,5 % z celkového počtu 119 frikativ ve VC kontextech), tzn. že u frikativ ve spojeních VC můžeme sledovat výraznou tendenci k nestandardnímu průběhu křivky HNR, v rámci kterého nedochází k markantní změně hodnot HNR a který sleduje spíše konstantní vývoj hodnot na jedné úrovni (obvykle v oblasti kladných čísel). Zdá se tedy, že v rámci VC spojení budou frikativy s největší pravděpodobností sledovat nestandardní průběh HNR skupiny 3 a až v druhé řadě standardní průběh skupiny 1.

3.2.2 Průběh křivky HNR s ohledem na řečový signál

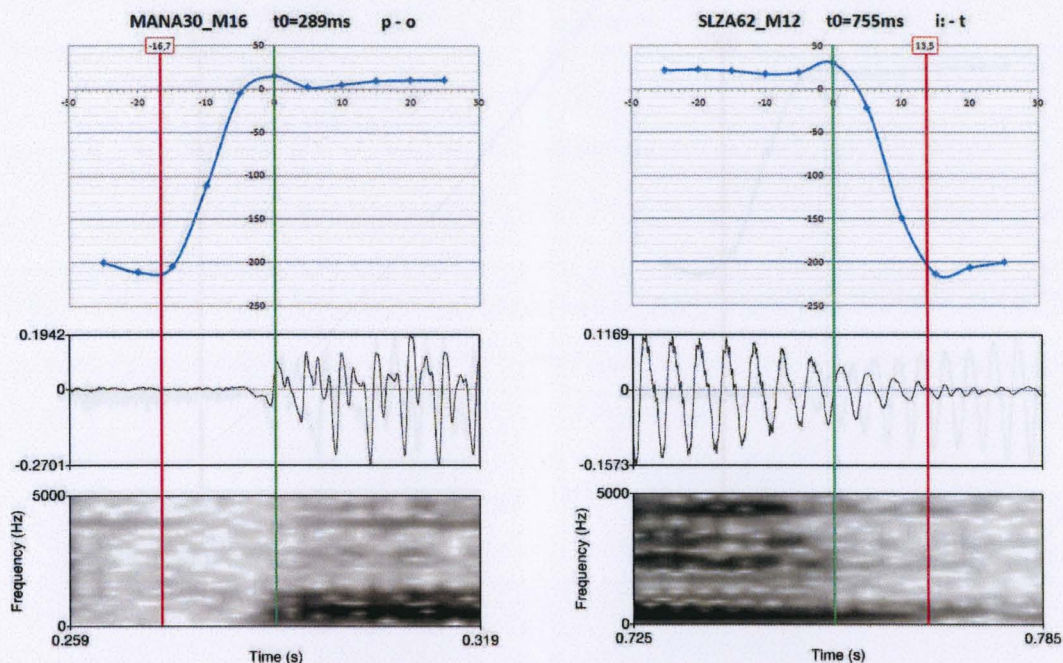
V této kapitole se pokusíme o detailnější popis jednotlivých skupin průběhů HNR podle klasifikace. Za tímto účelem budeme sledovat průběh HNR v závislosti na vlastnostech řečového signálu. Pro jednotlivé skupiny průběhů byly vybrány obrázky tak, aby byly rovnoměrně zastoupeny sekvence CV i VC a zároveň aby bylo reprezentováno i různé segmentální zastoupení konsonantů v daných kontextech.

Skupina 1 – standardní průběh: Dle tabulky č. 2 se standardní průběh křivky HNR objevuje v CV kontextech u 68,5 % realizací, v rámci sekvencí VC je jeho podíl menší a tvoří 59,8 % všech VC kontextů. Celkově představuje 1. skupina 63 % všech analyzovaných spojení (viz tabulka č. 2), což je poměrně vysoké číslo potvrzující převahu tohoto standardního průběhu nad všemi ostatními (příklady průběhů skupiny 1 viz příloha A).

Tvar křivky v této skupině se vyznačuje přítomností výrazné změny v hodnotách HNR, ke které dochází v důsledku přechodu vokálu na konsonant respektive obráceně. Ve výsledcích naší analýzy se u těchto křivek, jež jsou označeny jako standardní a představují tak měřítko pro ostatní průběhy HNR, zároveň objevuje mírný nárůst hodnot bezprostředně před jejich prudkým poklesem a obdobně

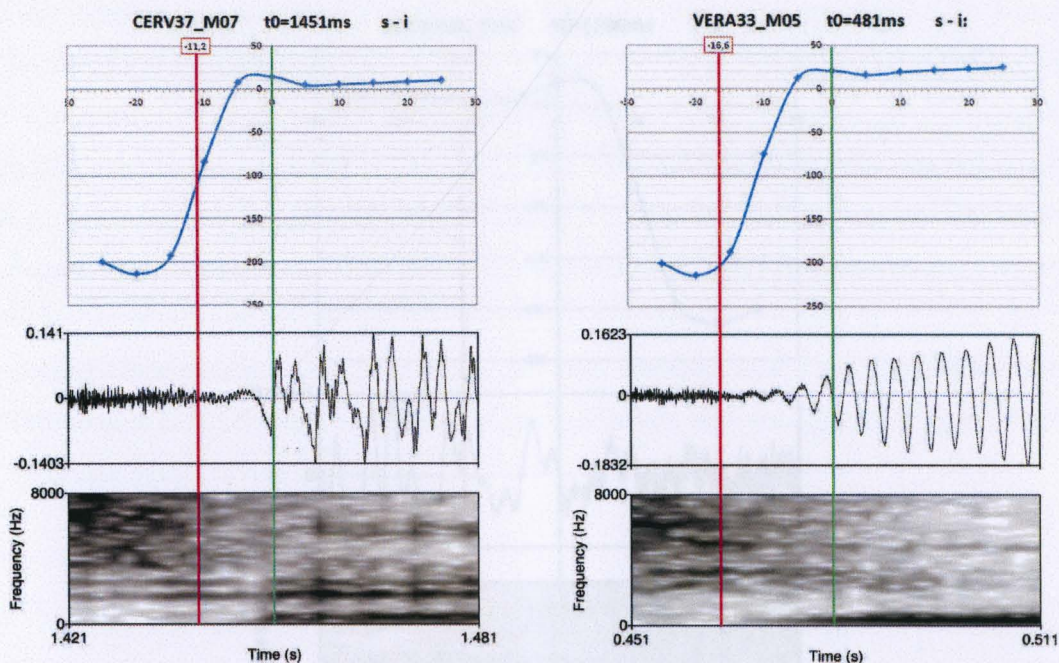
i írný pokles hodnot bezprostředně před jejich razantním nárůstem, což tyto křivky odlišuje od námi předpokládaného tvaru logistické křivky. Tento nárůst/pokles hodnot by tedy naznačoval, že začátky a konce vokálů jsou harmoničtější nežli samotné vokalické jádro, a obdobně, že na začátcích a koncích konsonantů, jež mohou být do jisté míry ovlivněny sousedícím vokalickým segmentem, převládají šumové komponenty nad těmi harmonickými ještě ve větší míře, než ve zbytku konsonantického segmentu. Tyto tendence se projeví v tak velké míře, že se důvody pro tuto zvlnění křivky hledaly v metodě měření, respektive v použitých nastaveních. Nakonec jsme dospěli k závěru, že zmíněná zvlnění křivky vznikají v důsledku použití kubické interpolace dat při zjišťování HNR v konkrétním čase (více viz kapitola 4).

Následují obrázky č. 23 – 25, které ilustrují standardní průběh křivky HNR pro sekvence CV a VC. Konsonanty v sekvencích jsou přitom zastoupeny jak explozivami, tak i frikativami. Na obrázku č. 23 můžeme porovnat standardní průběh křivky pro spojení CV a VC s různými projevy v přechodové části mezi vokálem a konsonantem, které znázorňuje oscilogram a spektrogram spojení hlásek. V sekvenci hlásek [po] můžeme v oscilogramu pozorovat půl periody, která ve spektrogramu odpovídá předznívání znělosti a rozvíjející se formantové struktuře. V grafu průběhu HNR to odpovídá stoupání hodnot HNR z oblasti záporných do oblasti kladných čísel, kde tyto hodnoty dále setrvávají po celou zbývající část vokalického segmentu. Naopak v sekvenci hlásek [i:t] můžeme pozorovat oslabování formantové struktury těsně před manuálně určenou hranicí (zelená linka v grafu) a hlavně setrvávající hlasivkovou činnost za touto hranicí, která v grafu odpovídá klesajícím hodnotám HNR. Zde bychom mohli namítat, že tato činnost hlasivek by měla korelovat spíše s kladnými hodnotami harmonicity. Klesající hodnoty HNR lze v tomto případě vysvětlit na základě algoritmu pro měření HNR v *PRAATu*. V rámci tohoto postupu se používá okno dlouhé 6 period, což u nízkých hlasů (tj. při standardním nastavení *Minimum Pitch* = 75 Hz) odpovídá 80 ms. Toto okno se posouvá s určitým časovým krokem, zatímco se v jednotlivých časových bodech tohoto posunu zaznamenává hodnota HNR získaná jako průměr hodnoty HNR celého okna. Pro náš konkrétní případ to tedy znamená, že za kmitání hlasivek korelující s klesajícími hodnotami HNR zodpovídá další vývoj konsonantu, který má šumovou povahu, respektive neobsahuje dále harmonické komponenty.



Obrázek č. 23. Skupina 1: porovnání dvou průběhů křivky HNR. Graf vlevo znázorňuje průběh křivky pro CV, graf vpravo pro VC. Konsonant je v obou grafech reprezentován explozivou. V popisku jednotlivých grafů najdeme jméno nahrávky, ve které se dané spojení nacházelo, čas umístění manuálně korigované hranice – t_0 a konkrétní spojení hlásek, které graf reprezentuje. Zelená linka označuje manuálně určenou hranici, červená linka reprezentuje automatickou hranici.

Na dalším obrázku (viz obrázek č. 24) můžeme srovnat spektrální průběh a průběh HNR ve spojení CV, kde je konsonant reprezentován frikativou. Podobně jako na předešlém obrázku můžeme pozorovat poměrně krátký a jasný přechod ve spojení hlásek [sɪ], kde slábnoucí šum frikativy přejde do tvaru složené vlny vokálu, přičemž manuální hranice mezi hláskami (zelená linka v grafu) označuje nástup formantové struktury. Spojení hlásek [si:] pak ilustruje pozvolnější přechod mezi konsonantem a vokálem. V oblasti mezi manuální a automatickou hranicí (mezi zelenou a červenou linkou v grafu) můžeme v oscilogramu pozorovat pozvolný nárůst amplitudy u vlny vokálu, ve spektrogramu pak ubývá šumových komponent a můžeme zde rozpoznat také přítomnost znělosti, za kterou zodpovídá činnost hlasivek, jež je přítomna ještě dřív, než je filtr vokálního traktu nastaven, tj. cca 15 ms před nástupem formantové struktury.

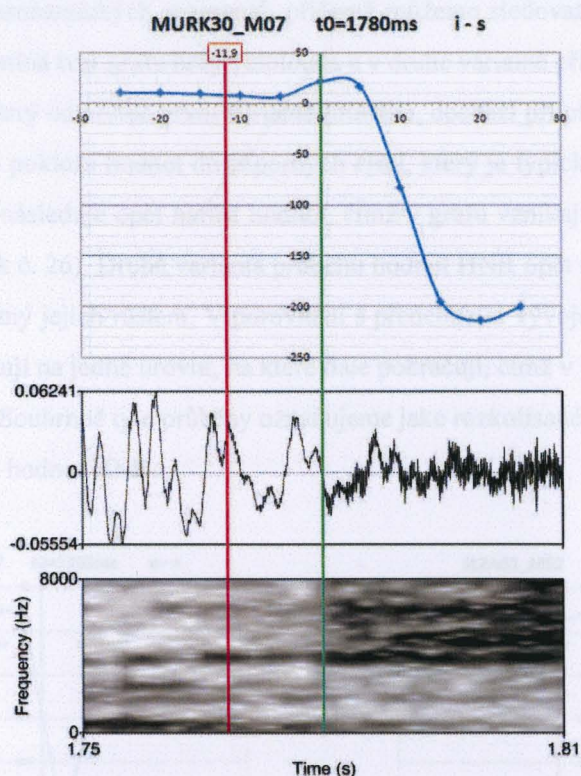


Obrázek č. 24. Skupina 1: porovnání dvou průběhů křivky HNR ve spojení CV, u kterých můžeme pozorovat jiné projevy ve spektrogramu. V obou kontextech se jako konsonant vyskytuje frikativa [s]. V popisku jednotlivých grafů najdeme jméno nahrávky, ve které se dané spojení nacházelo, čas umístění manuálně korigované hranice – t0 a konkrétní spojení hlásek, které graf reprezentuje. Zelená linka označuje manuálně určenou hranici, červená linka reprezentuje automatickou hranici.

Za pomoci obrázku č. 25 můžeme provést srovnání spojení CV se sekvencí VC, kde je konsonant reprezentován frikativou. V realizaci hlásek (viz obrázek č. 25) dochází s poklesem hodnot HNR ke zjednodušení tvaru vlny vokálu, která se dále opakuje, respektive dále doznívá, ale obsahuje navíc přídatný šum.

Celkově 1. skupina obsahuje 289 ze 459 analyzovaných položek, což představuje 63 % všech realizací. Konsonanty jsou přitom zastoupeny 153 explozivami (54,7 %), 106 frikativami (36,7 %) a 25 afrikátami (8,5 %). V rámci CV spojení převažují kontexty s frikativami (66 výskytů frikativ oproti 39 explozivám a 10 afrikátám), kdežto v sekvencích VC představují převážnou většinu explozivy (119 výskytů exploziv v porovnání se 46 frikativami a 15 afrikátami). Tento výrazný rozdíl v četnosti zastoupení exploziv a frikativ v rámci VC sekvencí lze vysvětlit již zmíněnou tendencí frikativ sledovat ve spojeních VC spíše nestandardní průběh HNR skupiny 3 (viz kapitola 3.2.1).

2. Skupina – rozkolísaný průběh: Do této skupiny byly zařazeny průběhy, jež se pohybují na pomezí standardního a nestandardního průběhu křivky HNR (příklady průběhů skupiny 2 viz příloha B). Standardní vývoj hodnot HNR sledujeme v rámci vokalizovaných segmentů, neobvyklý průběh HNR

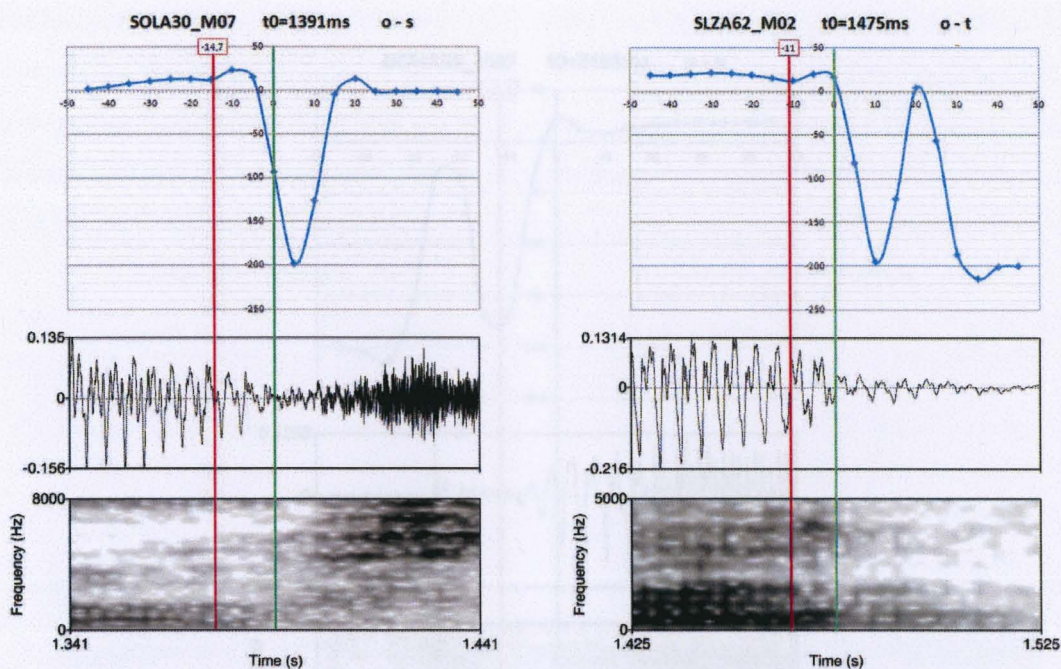


Obrázek č. 25. Skupina 1: průběh křivky HNR ve spojení VC, kde je konsonant reprezentován frikativou. V popisku grafu najdeme jméno nahrávky, ve které se dané spojení nacházelo, čas umístění manuálně korigované hranice – t0 a konkrétní spojení hlásek, které graf reprezentuje. Zelená linka označuje manuálně určenou hranici, červená linka reprezentuje automatickou hranici.

Celkově 1. skupina obsahuje 289 ze 459 analyzovaných položek, což představuje 63 % všech realizací. Konsonanty jsou přitom zastoupeny 158 explozivami (54,7 %), 106 frikativami (36,7 %) a 25 afrikátami (8,6 %). V rámci CV spojení převládají kontexty s frikativami (66 výskytů frikativ oproti 39 explozivám a 10 afrikátám), kdežto v sekvencích VC představují převážnou většinu explozivy (119 výskytů exploziv v porovnání se 40 frikativami a 15 afrikátami). Tento výrazný rozdíl v četnosti zastoupení exploziv a frikativ v rámci VC sekvencí lze vysvětlit již zmíněnou tendencí frikativ sledovat ve spojeních VC spíše nestandardní průběh HNR skupiny 3 (viz kapitola 3.2.1).

2. Skupina – rozkolísaný průběh: Do této skupiny byly zařazeny průběhy, jež se pohybují na mezích standardního a nestandardního průběhu křivky HNR (příklady průběhů skupiny 2 viz příloha B). Standardní vývoj hodnot HNR sledujeme v rámci vokálních segmentů, nečekaný průběh HNR

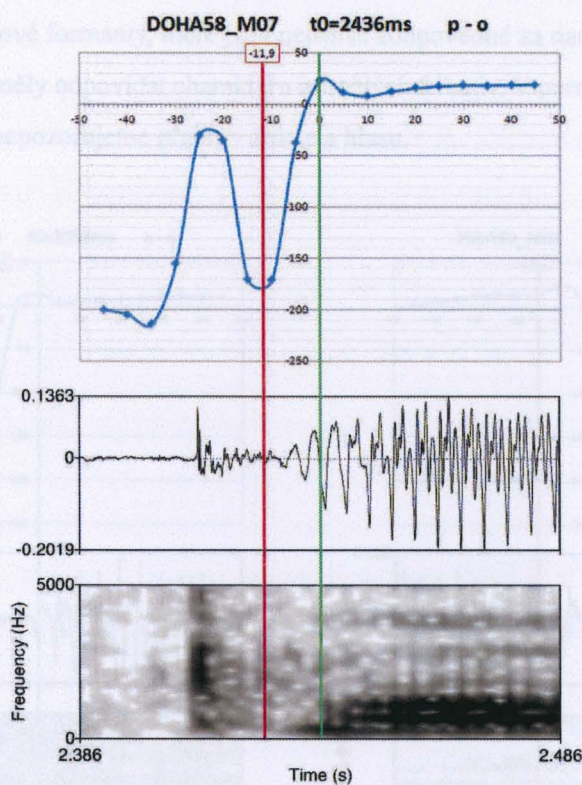
se pak objevuje u konsonantických segmentů, přičemž můžeme sledovat dvě linie vývoje hodnot: v první variantě připomíná tvar grafu hrby velblouda a v druhé variantě připomíná graf tvar písmene V. V průběhu HNR, který odpovídá první variantě průběhu, dochází při přechodu na konsonantický segment k výraznému poklesu hodnot do záporných čísel, který je typický pro průběhy skupiny 1, ale po tomto poklesu následuje opět nárůst hodnot, čímž v grafu vznikají tvary připomínající hrby velblouda (viz obrázek č. 26). Druhá varianta průběhu hodnot HNR opět zaznamenává výrazný pokles hodnot následovaný jejich růstem. V porovnání s předchozím vývojem se v těchto realizacích hodnoty HNR stabilizují na jedné úrovni, na které dále pokračují, čímž v grafu vzniká tvar písmene V (viz obrázek č. 28). Souhrnně tyto průběhy označujeme jako rozkolísané, jelikož v grafech dochází k výrazným výkyvům hodnot HNR.



Obrázek č. 26. Skupina 2: srovnání průběhů HNR typu „velbloud“ pro VC kontexty zastoupené v jednom případě frikativou a v druhém případě explozivou. V popisku jednotlivých grafů najdeme jméno nahrávky, ve které se dané spojení nacházelo, čas umístění manuálně korigované hranice – t0 a konkrétní spojení hlásek, které graf reprezentuje. Zelená linka označuje manuálně určenou hranici, červená linka reprezentuje automatickou hranici.

Na obrázku č. 26 můžeme sledovat průběhy HNR, v rámci kterých u konsonantů hodnoty HNR klesají, stoupají a opět klesají. V kontextu, ve kterém se vyskytuje frikativa, můžeme v oblasti přibližně kolem 25. ms za umístěním manuální hranice sledovat výrazný výkyv hodnot do oblasti kladných čísel. Graf nejprve sleduje v úseku 0 – 13 ms výrazné klesání hodnot HNR, které v oscilogramu odpovídá měnícímu se tvaru vlny a přibývajícimu šumu a koreluje s chybějící formantovou strukturou

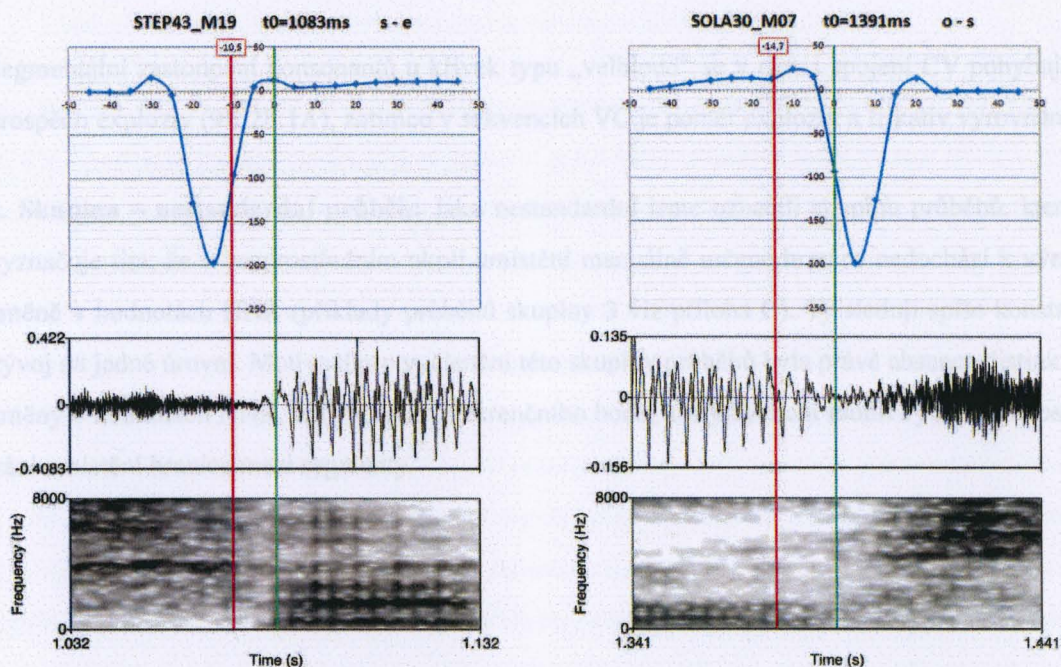
a slábnoucí nízkofrekvenční energií ve spektrogramu. Hodnoty HNR po tomto klesání opět rostou, a to v časovém úseku grafu 12 – 25 ms. V oscilogramu přitom můžeme pozorovat tvar vlny, která se opakuje, ale je výrazně zašuměná. V odpovídající části spektrogramu můžeme sledovat přítomnost výrazných šumových formantů, tj. zesílených pásem šumu. Nejspíše právě tyto spektrální projevy jsou zodpovědné za periodicitu, jež zde byla zjištěna za použití algoritmu autokorelace v programu *PRAAT*. V druhém grafu na obrázku č. 26 je tento průběh typu „velbloud“ reprezentován i spojením vokálu a explozivy [ot]. Výkyv v hodnotách HNR konsonantu se v oscilogramu projevuje jako opakující se vlna, jež pozvolna odeznívá. Ve spektrogramu můžeme sledovat pozvolna odeznívající formantovou strukturu, respektive přítomnost nízkofrekvenční energie, až 30 ms za umístěním manuální hranice (k přeznívání znělosti srov. kapitolu 2.4). Tento typ průběhu, ovšem pro spojení CV, znázorňuje obrázek č. 27.



Obrázek č. 27. Skupina 2: průběh HNR typu „velbloud“ pro CV kontext. V popisku grafu najdeme jméno nahrávky, ve které se dané spojení nacházelo, čas umístění manuálně korigované hranice – t0 a konkrétní spojení hlásek, které graf reprezentuje. Zelená linka označuje manuálně určenou hranici, červená linka reprezentuje automatickou hranici.

Ve spojení hlásek [po] na obrázku č. 27 koresponduje výkyv v konsonantických hodnotách HNR s aspiračním šumem, který vidíme ve spektrogramu v oblasti bezprostředně po explozi. V oscilogramu přitom v té samé oblasti pozorujeme 4 periody opakující se vlnu, která nejspíše navozuje zdání periodicity zodpovědné za výkyv v hodnotách HNR. V grafu koreluje se stoupajícími hodnotami HNR předznívání znělosti, které můžeme vidět ve spektrogramu jako tmavý pruh v nízkých frekvencích v oblasti mezi umístěním automaticky a manuálně určené hranice a které v oscilogramu odpovídá poměrně jednoduchému tvaru vlny.

Další vývoj hodnot HNR kopíruje tvar písmene V. Znázorňuje ho obrázek č. 28. V rámci těchto realizací dochází po předchozím poklesu hodnot v rámci konsonantu k opětovnému nárůstu HNR, přičemž se tyto hodnoty ustálí a dále pokračují konstantně na jedné úrovni. Ve spektrogramech zobrazujících spojení [se] a [os] (viz obrázek č. 28) můžeme u sykavek pozorovat poměrně prominentní pásma šumu, tzv. šumové formanty, které jsou nejspíše zodpovědné za naměřené hodnoty HNR, jež by dle našich úvah neměly odpovídat charakteru neznělých frikativ. V porovnání se spojením [as] na obrázku č. 26 zde již nepozorujeme překryv frikce a hlasu.

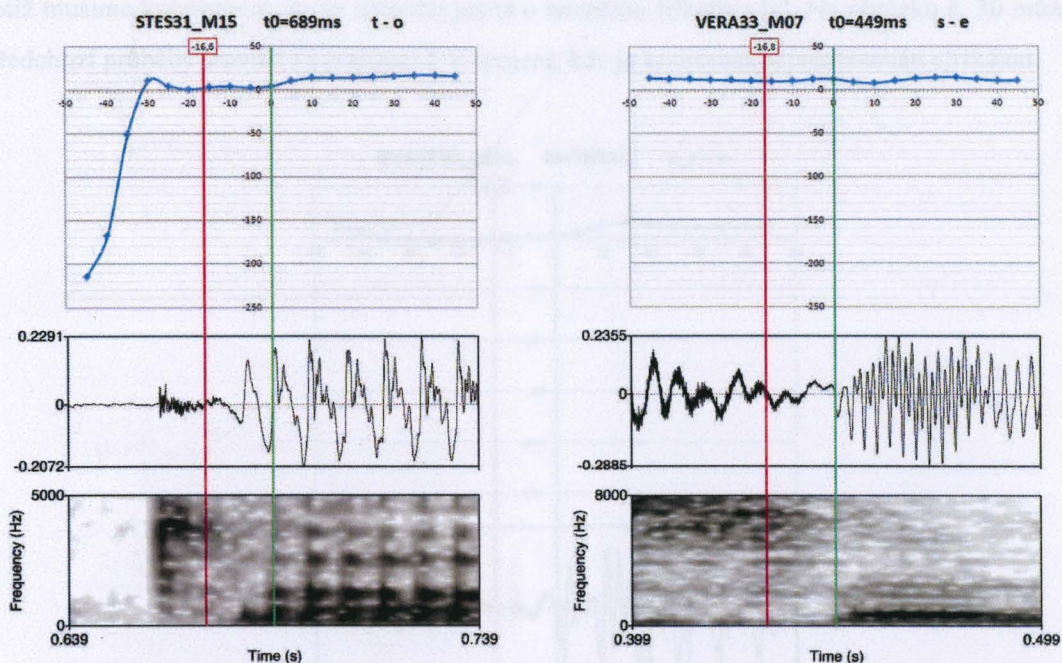


Obrázek č. 28. Skupina 2: srovnání průběhů HNR, které jsou podobné tvaru písmene V, pro CV a VC kontexty. V popisku jednotlivých grafů najdeme jméno nahrávky, ve které se dané spojení nacházelo, čas umístění manuálně korigované hranice – t0 a konkrétní spojení hlásek, které graf reprezentuje. Zelená linka označuje manuálně určenou hranici, červená linka reprezentuje automatickou hranici.

Co se týká zastoupení obou podtypů této skupiny, pohybuje se jejich četnost v poměru 2:1 ve prospěch průběhu typu „V“. Z hlediska segmentálního zastoupení konsonantů lze konstatovat, že se tento průběh skoro výhradně týká frikativ. Skupina spojení CV obsahuje z celkových 37 grafů 25 grafů ve tvaru písmene „V“. Z těchto grafů jich 20 vykresluje spojení frikativy s vokálem a zbylých 5 obsahuje jako konsonant afrikátu. V tomto případě můžeme afrikáty přiřadit k frikativám, jelikož se v jejich konečné fázi, kterou v těchto spojeních zkoumáme, objevuje šum jako u frikativ, takže lze říct, že ve všech 25 grafech se objevuje konsonant kontinuálního šumového charakteru. V rámci spojení VC jsme z 30 grafů přiřadili 20 grafů průběhu typu „V“, přičemž byl v 11 grafech konsonant zastoupen frikativou a v 9 grafech explozivou. Tady je nutno poznamenat, že všechny grafy obsahující explozivu byly do této skupiny zařazeny na základě dat, která nebyla úplně jednoznačná. To znamená, že jednotlivé grafy jako výsledky druhého měření, které zachycovalo interval 90 ms s časovým středem v místě manuálně určené hranice, neznázorňovaly očekávaný průběh křivky HNR. Obvykle v nich byl patrný razantní pokles hodnot HNR na začátku konsonantu, po kterém následoval jejich vzestup. Jelikož ani v jednom z těchto případů graf po daném výkyvu nezobrazoval další pokračování křivky, byly tyto realizace paušálně přiřazeny do podskupiny popisované na základě podobnosti grafu s tvarem písmene V.

Segmentální zastoupení konsonantů u křivek typu „velbloud“ se v rámci spojení CV pohybuje ve prospěch exploziv (9E:2F:1A), zatímco v sekvencích VC je poměr exploziv a frikativ vyrovnaný.

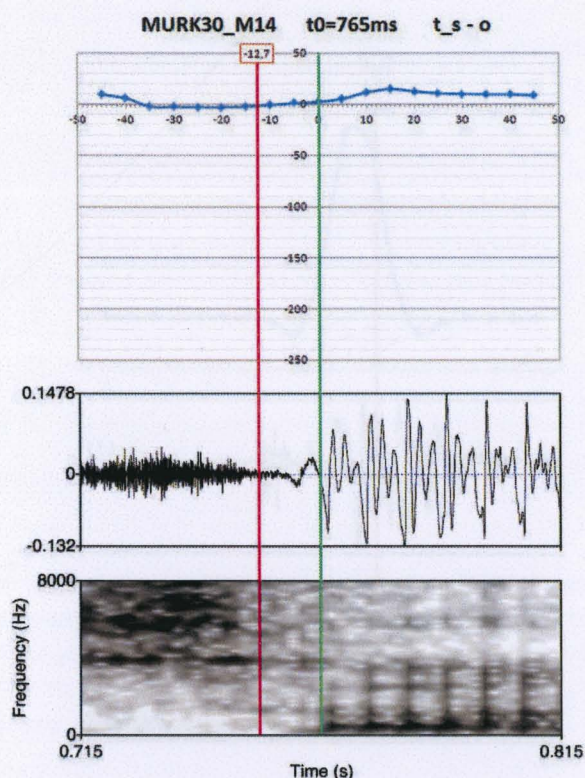
3. Skupina – nestandardní průběh: Jako nestandardní jsme označili skupinu průběhů, která se vyznačuje tím, že v bezprostředním okolí umístění manuálně určené hranice nedochází k výrazné změně v hodnotách HNR (příklady průběhů skupiny 3 viz příloha C). Ty sledují spíše konstantní vývoj na jedné úrovni. Motivací pro vyčlenění této skupiny průběhů byla právě absence distinktivní změny v hodnotách HNR, tedy absence referenčního bodu, který bychom mohli využít pro posuzování umístění hranice mezi segmenty.



Obrázek č. 29. Skupina 3: srovnání průběhů HNR pro spojení CV se segmentálním zastoupením konsonantu jednou jako exploziv a po druhé jako frikativy. V popisku jednotlivých grafů najdeme jméno nahrávky, ve které se dané spojení nacházelo, čas umístění manuálně korigované hranice – t0 a konkrétní spojení hlásek, které graf reprezentuje. Zelená linka označuje manuálně určenou hranici, červená linka reprezentuje automatickou hranici.

Na obrázku č. 29 můžeme sledovat příklady nestandardních průběhů křivky HNR pro kombinaci CV. V případě spojení hlásek [to] můžeme v grafu vidět razantní nárůst hodnot HNR v oblasti konsonantu, které se dále vyvíjejí konstantně v oblasti kladných čísel. Tato změna nicméně probíhá v oblasti –30 ms od místa určení manuální hranice a koresponduje se závěrovou částí explozivy, přičemž po explozi hodnoty HNR stoupají do oblastí kladných hodnot. Za tento nárůst HNR následovaný konstantním vývojem hodnot může podle spektrogramu aspirační šum přítomný v poslední fázi konsonantu a také harmonické komponenty, které v signálu přibývají v oblasti těsně před zeleně vyznačenou hranicí. Hodnoty HNR poté ještě jednou zaznamenají nárůst (i když v tomto případě už ne tak razantní), který se objeví spolu s přítomností plné formantové struktury vokálu. Ve spektrogramu spojení hlásek [se] na obrázku č. 29 můžeme pozorovat zvláštní realizaci frikativy [s], která podle tvaru křivky odpovídá nezvykle vysokým hodnotám HNR. Ty jsou skoro stejné jako u vokálu v dané realizaci, a to i přes to, že ve spektrogramu frikativy můžeme pozorovat přítomnost šumu ve vysokých frekvencích, který v oblasti vokálu nenajdeme. Znázorněné spojení hlásek [se] se vyskytovalo v rámci spojení slov „vlastně jsem“, přičemž byla při realizaci vynechána hláska [j]. Podle spektrogramu tak můžeme nabýt dojmu, že se jedná o případ sonorizace neznělé frikativy vlivem jejího segmentálního okolí, sluchový dojem nám ovšem tuto domněnku vyvrátí. Na základě poslechu

totiž musíme konstatovat, že se opravdu jedná o neznělou frikativu [s]. Na obrázku č. 30 můžeme předchozí průběhy srovnat i s realizací CV spojení, kde je konsonant reprezentován afrikátou.

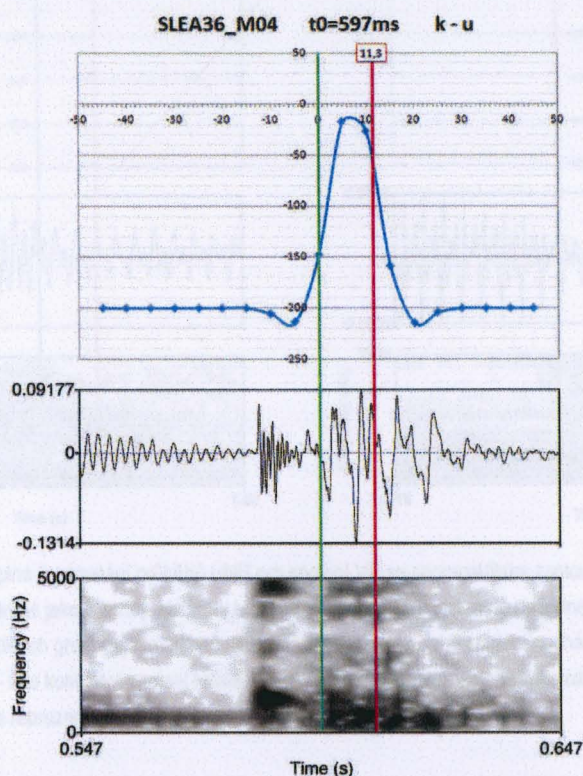


Obrázek č. 30. Skupina 3: průběh HNR pro CV realizaci, kde je konsonant reprezentován afrikátou. V popisku grafu najdeme jméno nahrávky, ve které se dané spojení nacházelo, čas umístění manuálně korigované hranice – t0 a konkrétní spojení hlásek, které graf reprezentuje. Zelená linka označuje manuálně určenou hranici, červená linka reprezentuje automatickou hranici.

Příklady nestandardního průběhu HNR v sekvenčních VC znázorňují obrázky č. 32 a 33, přičemž na

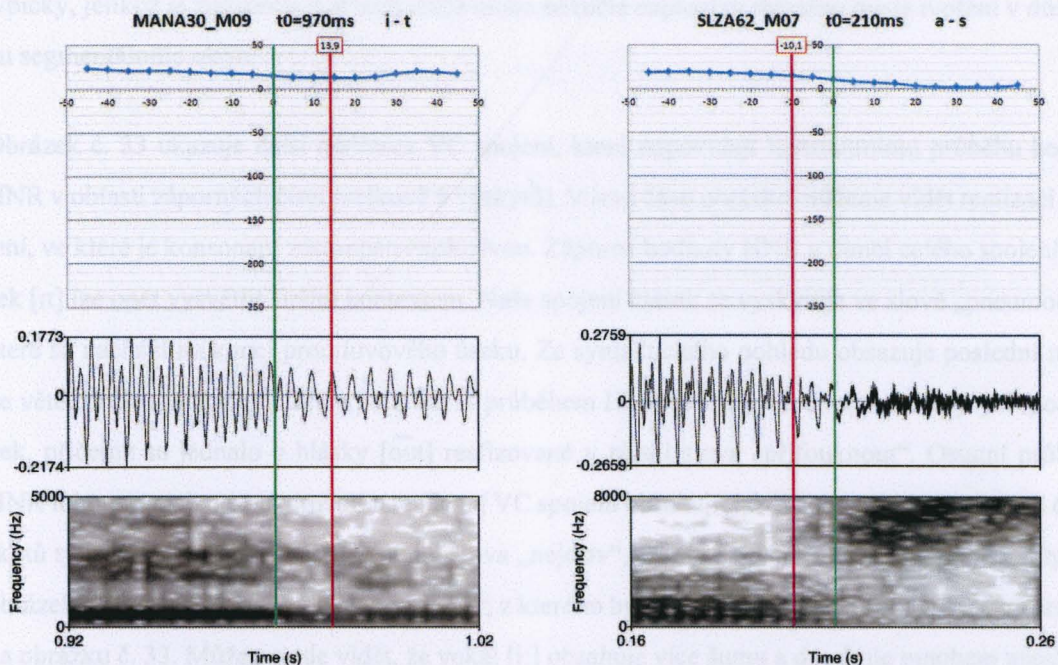
Jako další zvláštní průběhy HNR byly do této skupiny zařazeny krátké realizace vokálů. V těchto grafech (viz např. obrázek č. 31) můžeme pozorovat podobný vývoj hodnot HNR jako u některých grafů skupiny 2: na přechodu mezi konsonantem a vokálem hodnoty HNR prudce narůstají, ale poté záhy klesají opět do oblasti záporných čísel. Jak jsme již zmínili, je tento průběh křivky ovlivněn trváním vokalického segmentu. V našem případě dosahuje hláska [u] ve spojení [ku] na obrázku č. 31 délky 26 ms. S podobnými průběhy HNR jsme se setkali v rámci CV spojení dvakrát a v rámci VC spojení sedmkrát. Tvar křivky si přitom vysvětlujeme nejen na základě trvání segmentu, ale i v závislosti na použité metodě měření HNR, která k analýze využívá okno dlouhé 80 ms odpovídající minimálně

6 periodám signálu. Pro podobný algoritmus tak krátké realizace segmentů představují problém (viz obrázek č. 31, kde hlásce [u] na základě vymezených hranic přičítáme pouze 4 periody).



Obrázek č. 31. Skupina 3: krátká realizace vokálu. V popisku grafu najdeme jméno nahrávky, ve které se dané spojení nacházelo, čas umístění manuálně korigované hranice – t_0 a konkrétní spojení hlásek, které graf reprezentuje. Zelená linka označuje manuálně určenou hranici, červená linka reprezentuje automatickou hranici.

Příklady nestandardního průběhu křivky v sekvencích VC znázorňují obrázky č. 32 a 33, přičemž na obrázku č. 32 můžeme pozorovat konstantní vývoj hodnot HNR spíše v oblasti kladných čísel a na obrázku č. 33 v oblasti záporných čísel. Kladné průběhy byly ve spojeních VC početnější. Setkali jsme se s nimi v 71 případech z celkového počtu 87 nestandardních VC průběhů. Z toho obsahovalo 12 kontextů explozivu a zbylých 59 kontextů frikativu. To ukazuje na významnou převahu zastoupení konsonantu frikativou. Toto vysoké číslo také koresponduje s již zmíněnou tendencí frikativ kopírovat ve skupině VC kontextů spíše nestandardní průběh křivky HNR (nestandardní průběh u 66 frikativ ze 119 frikativ ve spojeních VC).

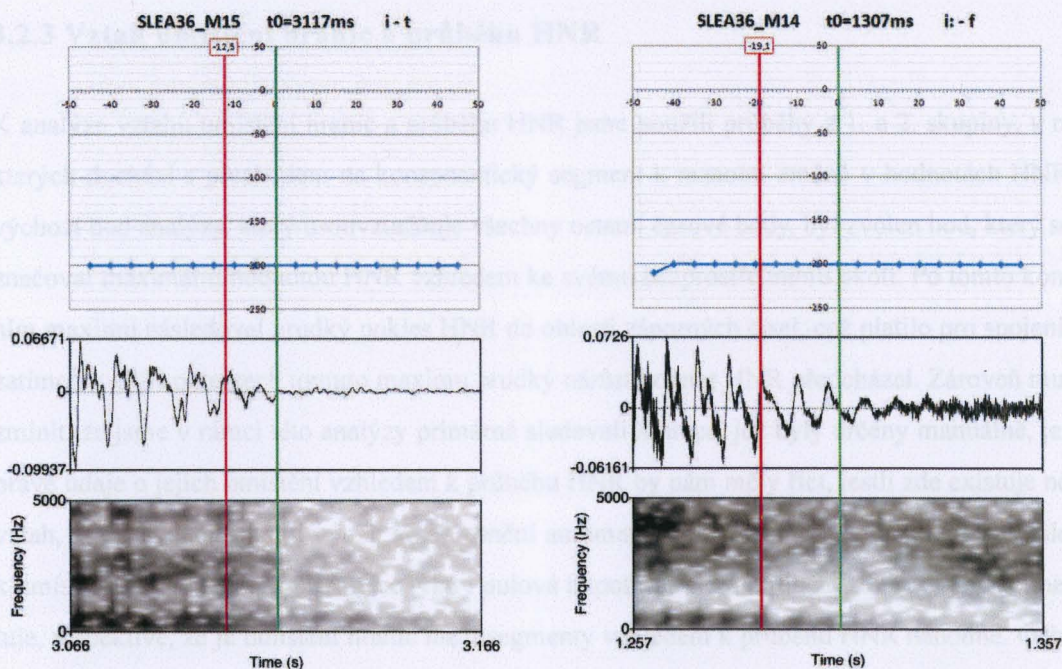


Obrázek č. 32. Skupina 3: srovnání průběhů HNR pro spojení VC se segmentálním zastoupením konsonantu jednou jako explozivou a po druhé jako frikativu. Hodnoty HNR se vyvíjejí spíše konstantně na jedné úrovni v oblasti kladných čísel. V popisku jednotlivých grafů najdeme jméno nahrávky, ve které se dané spojení nacházelo, čas umístění manuálně korigované hranice – t_0 a konkrétní spojení hlásek, které graf reprezentuje. Zelená linka označuje manuálně určenou hranici, červená linka reprezentuje automatickou hranici.

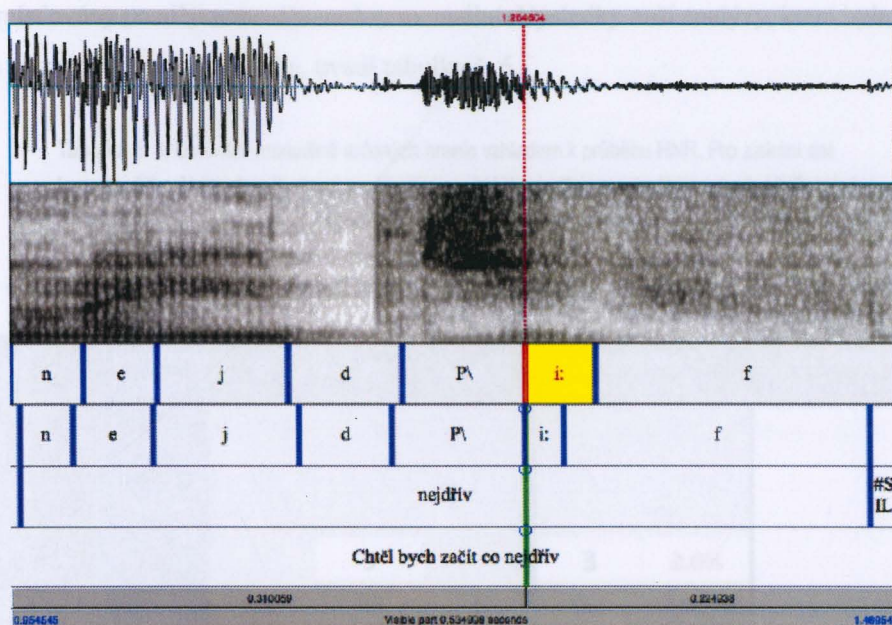
Obrázek č. 32 znázorňuje konstantní průběh hodnot HNR v oblasti kladných čísel pro spojení VC, kde je konsonant reprezentován explozivou a frikativou. Na základě naší analýzy můžeme konstatovat, že podobný průběh HNR lze považovat u frikativ za převládající. V oscilogramu frikativy [s] na obrázku č. 32 vidíme jistou periodicitu, která se zde objevuje spolu se šumem. V grafu průběhu HNR můžeme v konsonantické části vidět mírný pokles HNR v porovnání s jejími hodnotami v oblasti vokálu. Ty se nicméně pořád (zřejmě v důsledku zmíněné periodicity) pohybují v oblasti kladných čísel. Realizace s explozivou (viz levá část obrázku č. 32) dle oscilogramu a spektrogramu naznačuje, že se jedná spíše o znělou variantu konsonantu, tedy o hlásku [d]. Soudíme tak na základě přítomnosti znělosti ve spektrogramu, která koresponduje s opakující se periodou vlny, kterou znázorňuje oscilogram spojení [it]. Tato realizace vznikla v rámci spojení slov „vařit vodu“ jako výsledek znělostní asimilace, kterou způsobila hlásky [v]. Podobný vývoj hodnot HNR jsme sledovali celkem u 12 kontextů s explozivou, přičemž 5 z nich se objevilo právě ve spojení slov „vařit vodu“. Další realizace byly buď výsledkem znělostní asimilace (např. [at] jako [ad] v „například vztažné“, [i:t] jako [i:d] v „nebýt Laďových“ atd.), nebo realizace intervokální explozivy [p] v rámci spojení „skoro půlku“. Na základě těchto faktů lze usoudit, že daný průběh HNR není pro neznělé explozivy

typický, jelikož je způsoben realizací znělé místo neznělé explozivы stejného místa tvoření v důsledku segmentálního okolí.

Obrázek č. 33 ukazuje další realizace VC spojení, které odpovídají konstantnímu průběhu hodnot HNR v oblasti záporných čísel (celkově 9 výskytů). V levé části obrázku můžeme vidět realizaci spojení, ve které je konsonant zastoupen explozivou. Záporné hodnoty HNR v rámci celého spojení hlásek [rt] lze opět vysvětlit širším kontextem. Naše spojení hlásek se vyskytuje ve slově „pneumocyt“, které se nachází na konci promluvového úseku. Ze syntaktického pohledu obsazuje poslední místo ve větě. Další realizace hlásek s podobným průběhem HNR se vyskytla za splnění stejných podmínek, přičemž se jednalo o hlásky [out] realizované v rámci slova „přifouknout“. Ostatní průběhy HNR tohoto typu jsme pozorovali u realizací VC spojení obsahující frikativu, přičemž se všech 6 výskytů týkalo spojení hlásek [i:f] v rámci slova „nejdřív“, které se opět nacházelo na konci věty. Na obrázek č. 34 je spektrogram slova „nejdřív“, z kterého bylo vyjmuto spojení hlásek [i:f] znázorněno na obrázku č. 33. Můžeme zde vidět, že vokál [i:] obsahuje více šumu a dosahuje mnohem menší intenzity než vokál [e] nacházející se v první slabice daného slova (viz obrázek č. 34). U průběhů HNR pohybujících se konstantně v záporných číslech můžeme opět pozorovat závislost realizací hlásek na širším kontextu a zároveň můžeme konstatovat, že je frekvence výskytů těchto průběhů spíše nízká.



Obrázek č. 33. Skupina 3: srovnání průběhů HNR pro spojení VC se segmentálním zastoupením konsonantu jednou jako explozivou a po druhé jako frikativu. Hodnoty HNR se vyvíjejí spíše konstantně na jedné úrovni v oblasti záporných čísel. V popisku jednotlivých grafů najdeme jméno nahrávky, ve které se dané spojení nacházelo, čas umístění manuálně korigované hranice – t0 a konkrétní spojení hlásek, které graf reprezentuje. Zelená linka označuje manuálně určenou hranici, červená linka reprezentuje automatickou hranici.



Obrázek č. 34 (zpracovaný materiál – nahrávka SLEA36_M14). Realizace slova „nejdřív“, které se vyskytuje na konci věty. V další vrstvě pod spektrogramem, který znázorňuje frekvenční rozpětí 0 – 8 kHz, jsou naznačeny manuálně korigované hranice, ve vrstvě následující se nachází automatická segmentace.

3.2.3 Vztah umístění hranic a průběhu HNR

K analýze vztahu umístění hranic a průběhu HNR jsme použili průběhy z 1. a 2. skupiny, v rámci kterých dochází s přechodem na konsonantický segment k razantní změně v hodnotách HNR. Za výchozí bod analýzy, který usouvztažňuje všechny ostatní časové body, byl zvolen bod, který se vyznačoval maximální hodnotou HNR vzhledem ke svému bezprostřednímu okolí. Po tomto konkrétním maximu následoval prudký pokles HNR do oblasti záporných čísel, což platilo pro spojení VC, zatímco v CV kontextech tomuto maximu prudký nárůst hodnot HNR předcházel. Zároveň musíme zmínit, že jsme v rámci této analýzy primárně sledovali hranice, jež byly určeny manuálně, jelikož právě údaje o jejich umístění vzhledem k průběhu HNR by nám měly říct, jestli zde existuje nějaký vztah, který bychom mohli využít ke zpřesnění automatické segmentace. Průběhu HNR vzhledem k umístění hranic mezi segmenty se týká i nulová hypotéza, která tvrdí, že zde žádný vztah neexistuje, respektive, že je umístění hranic mezi segmenty vzhledem k průběhu HNR náhodné. Odhalení souvislosti mezi průběhem HNR a umístěním hranic mezi segmenty by tedy přispělo i k vyvrácení nulové hypotézy. Až v druhé řadě jsme se zabývali umístěním automatických hranic, a to především z hlediska jejich relativního umístění vůči manuálně korigovaným hranicím, tj. zda byly automatické

hranice umísťované později nebo dřív než ty manuální. Výsledky naší analýzy, která byla provedena v souladu se zmíněnými východisky, uvádí tabulka č. 6.

Tabulka č. 6. Umístění manuálně určených hranic vzhledem k průběhu HNR. Pro získání dat jsme použili průběhy 1. a 2. skupiny z klasifikace. MAX vyjadřuje maximální hodnotu HNR, která se objevuje těsně před výrazným poklesem hodnot HNR u VC nebo které předchází výrazný nárůst hodnot HNR u CV. Tento bod byl využit k usouvztažení ostatních časových bodů. Takže např. bod -5 se nachází 5 ms před výskytem maxima HNR atd.

	Umístění hranice	E	F	A	Celkem ve skupinách		Celkem položek
CV	víc						152
	-20						
	-15						
	-10		3		3	2.0%	
	-5	18	29	5	52	34.2%	
	MAX	16	42	9	67	44.1%	
	+5	11	13	2	26	17.1%	
	+10	2	1		3	2.0%	
	+15	1			1	0.7%	
	+20						
	víc						
VC	víc						204
	-20	5	4		9	4.4%	
	-15	13	1		14	6.9%	
	-10	14	5	2	21	10.3%	
	-5	19	3	2	24	11.8%	
	MAX	22	10	1	33	16.2%	
	+5	28	6	6	40	19.6%	
	+10	20	14		34	16.7%	
	+15	10	4	1	15	7.4%	
	+20	2	5	3	10	4.9%	
	víc	3	1		4	2.0%	

U analyzovaných spojení CV můžeme vidět, že 95,4 % umístění hranic se nachází v rozmezí 10 ms, tj. v časovém intervalu 5 ms před až 5 ms za hodnotou maxima HNR. V rámci spojení VC již výsledky nejsou tak vyhraněné a variabilita umístění hranic segmentů je vzhledem k bodu maxima HNR mnohem větší. Tento fakt si vysvětlujeme tím, že při přechodu z vokalického na konsonantický segment dochází k častějšímu výskytu přechodových oblastí, které jsou způsobené delší dobou

přibližování artikulátorů v porovnání s jejich oddálením, ke kterému dochází u CV spojení (Machač & Skarnitzl, 2009, str. 35 – 36). Pro VC sekvence jsme tedy zjistili, že početně nejvíc hranic se pohybuje v časovém rozmezí 10 ms, které se nachází v oblasti mezi maximem HNR a 10 ms po tomto maximu, přičemž střed tohoto intervalu tvoří časový bod +5 ms vzhledem k umístění tohoto maxima. Toto pásmo postihuje 52,5 % umístění manuálních hranic v rámci VC kontextů. Jelikož se u VC sekvencí vyskytuje větší rozptýl hodnot, rozšířili jsme časové rozmezí pro umístění hranice na 20 ms, přesněji na oblast –10 ms až +10 ms kolem hodnoty maxima HNR. Tento časový interval pak zahrnuje 74,5 % všech analyzovaných hranic ve spojeních VC. Výsledky této analýzy tedy potvrzují existenci určitých časových intervalů, které byly stanoveny v souladu s průběhem HNR a které zahrnují umístění většiny manuálně určených hranic. Právě na základě existence těchto časově omezených intervalů můžeme vyvrátit nulovou hypotézu. Umístění hranic s ohledem na průběh HNR tedy není náhodné, ale tato dvě kritéria spolu určitým způsobem korelují.

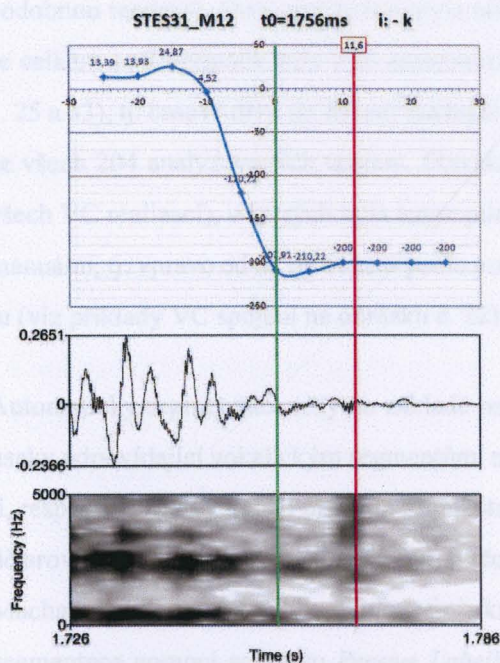
Podle další hypotézy jsou hranice segmentů umísťovány do oblasti kladných hodnot HNR, a to v závislosti na kritériu plné formantové struktury, které se v našem materiálu využívá při manuální segmentaci. Tuto hypotézu jsme formulovali hlavně z toho důvodu, abychom ověřili, zda jsou hranice mezi segmenty umísťovány spíše do oblasti, kde se hodnoty HNR vyvíjejí konstantně, respektive na začátek/konec oblasti s kladnými hodnotami HNR, nebo do oblasti přechodu z vokálního segmentu na konsonantický. Ten koresponduje s prudkým poklesem hodnot HNR do oblasti záporných čísel. Stejně relace platí i pro přechod z konsonantického segmentu na vokální, v rámci kterého dochází k růstu hodnot HNR. Tento růst se nicméně pořád odehrává v oblasti záporných čísel.

- Pro spojení CV se kladné hodnoty HNR vyskytují v časovém intervalu mezi maximem a +20 ms, který zahrnuje 63,8 % všech analyzovaných hranic v rámci CV spojení. Zde je ještě nutno poznamenat, že hranice umístěné v časovém bodu –5 ms před maximem HNR se z poloviny pohybovaly rovněž v kladných číslech. To znamená, že přiřadíme-li polovinu těchto realizací (26 z 52) k případům umístění hranice do oblasti kladných čísel, pracujeme až s 80,9 % hranic, které jsou umístěny do oblasti kladných hodnot HNR.
- V rámci spojení VC se v oblasti kladných čísel pohybují hranice u 52 % analyzovaných realizací, což představuje hranice nacházející se v naší analýze v časovém intervalu –20 ms až maximum HNR. V tomto případě jsme mohli opět pozorovat kladné hodnoty HNR i u hranic nacházejících se v časovém bodu +5 ms za hodnotou lokálního maxima HNR, celkem u 27 ze 40 hranic umís-

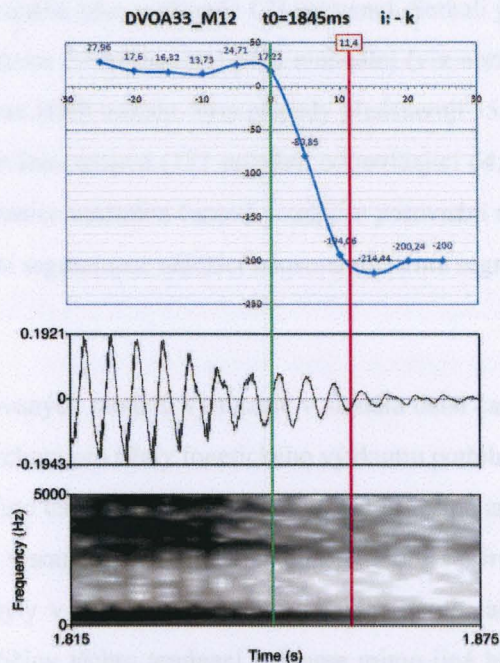
těných v tomto časovém bodě. Započteme-li i tyto hranice ke zmíněným 52 % hranic, zvýší se procento hranic umístěných do oblasti kladných čísel na 65,2 %.

Podíl hranic umístěných do oblasti kladných hodnot HNR v rámci VC kontextů (65,2 %) je znatelně nižší než u spojení CV (80,9 %), což dáváme do souvislosti s již zmíněným častým výskytem přechodových oblastí ve spojeních VC, v rámci kterých se hranice dle segmentační příručky Machač & Skarnitzl (2009) umísťují do temporálního středu přechodu. Ten již nemusí obsahovat harmonické komponenty, které by byly tak výrazné z hlediska své energie, což se odrazí v záporných hodnotách HNR (viz obrázek č. 34).

Dalším faktorem, který mohl zapříčinit tak velkou variabilitu v umístění hranic mezi segmenty u sekvencí VC, mohl být i způsob zpracování materiálu ze strany anotátora. Jak zmiňují Machač & Skarnitzl (2009, str. 13), neliší se výsledky segmentace jenom pro skupinu anotátorů, porovnáme-li je mezi sebou, ale výsledky segmentace se mohou lišit i u jednoho anotátora v čase. Sekvence VC tedy mohly být ovlivněny tímto faktorem více nežli spojení CV zejména proto, že se zde v důsledku delšího přibližování artikulátorů častěji objevují přechodové oblasti, jejichž segmentace je obtížnější. Dva různé případy umístění hranice mezi segmenty znázorňují obrázky č. 35 a 36. Na obrázku č. 35 je hranice umístěna v časovém bodu +15 ms po výskytu námi sledovaného maxima HNR, který koresponduje se zápornou hodnotou HNR. Na obrázku č. 36, který zobrazuje umístění hranice do časového bodu +5 ms vzhledem k výskytu maxima HNR, se harmonicita v oblasti umístění hranice pohybuje v kladných hodnotách. Po srovnání obou obrázků můžeme konstatovat, že hodnoty HNR v bodě umístění hranice jsou ovlivněny nejspíše realizací fonému /k/ (respektive délkou přechodového pásma mezi danými segmenty). U realizace [i:k] na obrázku č. 36 můžeme sledovat pozvolnější přechod mezi segmenty, v rámci kterého znělost výrazně přeznívá do konsonantického segmentu. Na obrázku č. 35 tomu tak není a nízkofrekvenční energie se ve spektrogramu konsonantu neobjevuje v tak velké míře. Dáme-li toto zjištění do spojitosti s délkou analyzačního okna, které se využívá pro měření HNR, můžeme vyvodit, že právě pozvolnější přechody obsahující znělost přeznívající do konsonantického segmentu nebo postupně odeznívající formantovou strukturu zapříčiňují určitou setrvačnost v hodnotách HNR, což má za následek umísťování hranic mezi segmenty do oblastí, kde se harmonicita pohybuje v kladných číslech.



Obrázek č. 35. Umístění manuální hranice mezi segmenty do oblasti záporných hodnot HNR, přesněji do časového bodu +15 ms po výskytu maxima HNR. V popisku jednotlivých grafů najdeme jméno nahrávky, ve které se dané spojení nacházelo, čas umístění manuálně korigované hranice – t_0 a konkrétní spojení hlásek, které graf reprezentuje. Zelená linka označuje manuálně určenou hranici, červená linka reprezentuje automatickou hranici.



Obrázek č. 36. Umístění manuální hranice mezi segmenty do oblasti kladných hodnot HNR, přesněji do časového bodu +5 ms po výskytu maxima HNR. V popisku jednotlivých grafů najdeme jméno nahrávky, ve které se dané spojení nacházelo, čas umístění manuálně korigované hranice – t_0 a konkrétní spojení hlásek, které graf reprezentuje. Zelená linka označuje manuálně určenou hranici, červená linka reprezentuje automatickou hranici.

Na základě naší analýzy můžeme říct, že hranice umístěné do oblasti kladných hodnot HNR převládají. Představují přitom ve spojeních CV 80,9 % a ve spojeních VC 65,2 % všech případů. Můžeme zde tedy pozorovat tendenci umísťovat hranice spíše do oblastí (respektive na začátky nebo konce těchto oblastí), kde se hodnoty HNR vyvíjejí konstantně v kladných číslech, nežli do oblastí, kde dochází k přechodu do záporných hodnot HNR. Ty korespondují buď s prudkým poklesem/nárůstem HNR, nebo s konstantním vývojem hodnot HNR v záporných číslech.

Další zajímavé zjištění nám poskytla analýza umístění automaticky určených hranic. U spojení CV se ze všech 152 realizací objevil pouze 1 případ, kdy byla automatická hranice umístěna až za manuální, tj. později v čase. Obecně jsme tedy mohli sledovat tendenci v umísťování automatických hranic do oblastí, kde byl přítomen aspirační šum nebo frikce neznělých obstruentů, případně tam zasahovala předznívající znělost. Nevyskytovala se tam ale ještě formantová struktura vokálu (viz příklady CV spojení na obrázku č. 22 a dále obrázky č. 24, 27, 29 a 30). V rámci spojení VC jsme pozorovali

podobnou tendenci, která zde opět nebyla tak vyhraněná jako v případě CV sekvencí. Setkali jsme se celkem se 73 případy, kdy byla automatická hranice umístěna ještě před manuální (viz obrázky č. 25 a 33), tj. časově dřív, do oblasti kladných hodnot HNR vokálu. Tyto případy představují 35,8 % ze všech 204 analyzovaných spojení. Převládaly ovšem spojení (131 položek odpovídající 64,2 % všech VC realizací), u kterých byla automatická hranice umístěna časově později v porovnání s tou manuální, tj. vpravo od ní, do oblasti podle manuální segmentace náležící konsonantickému segmentu (viz příklady VC spojení na obrázku č. 22).

Automatický segmentátor tedy na základě natrénovaných modelů vymezuje v signálu delší časové úseky odpovídající vokalickým segmentům, než bychom pro účely fonetického výzkumu potřebovali, respektive než vymezují pravidla segmentace. Tuto tendenci můžeme formulovat právě na základě srovnání manuálních hranic, jež byly určovány v souladu se segmentačními pravidly z příručky Machač & Skarnitzl (2009), s hranicemi, které byly v signálu umístěny na základě automatické segmentace pomocí softwaru *Prague Labeller*. Příčiny těchto tendencí můžeme mimo jiné hledat i v modelech, které při automatické segmentaci využívá *Prague Labeller*. Tyto modely jsou totiž závislé na zpracování materiálu, jež byl použit pro jejich trénování. To znamená, že nebyl-li materiál z trénovací databáze zpracován podle segmentačních pravidel, které využíváme v naší práci my, nemůžeme v rámci automatické segmentace očekávat výsledky, které budou lépe odpovídat našim očekáváním, respektive které budou odpovídat segmentaci respektující zmiňovaná segmentační pravidla. Materiál, jež byl použit k natrénování modelů, respektive jeho zpracování, tedy představuje také jeden z faktorů, které ovlivňují přesnost segmentace.

Naším cílem v této kapitole bylo ověřit, zda existuje vztah mezi umístěním hranice mezi segmenty a hodnotami HNR, který bychom potenciálně mohli využít ke zpřesnění automatické segmentace. V tomto ohledu tedy můžeme konstatovat, že pro CV kontexty existuje časový interval 10 ms (± 5 ms od námi sledovaného maxima HNR), v rámci kterého je umístěno 95,4 % manuálně určených hranic. Ty se v případě automatické segmentace pohybují v intervalu o délce 20 ms, který se vzhledem k maximu HNR nachází v oblasti -25 až -5 ms. Teoreticky by se identická oblast měla nacházet i v oblasti $+5$ až $+25$ ms od maxima HNR, ale vzhledem k tomu, že se v tomto pásmu nacházela pouze 1 ze 152 analyzovaných hranic, nepovažujeme tuto oblast z hlediska výskytu automatických hranic za významnou. Rozdělení a velikost těchto intervalů vyplývá z výběru našeho vzorku dat pro analýzu, který zohledňuje kritérium, aby se umístění automatické a manuální hranice lišilo o 10 až 20 ms v absolutní hodnotě. Z toho pak vyplývá, že v případě manuálních hranic, které se nacházely v bodě -5 ms vzhledem k maximu HNR, se umístění automatické hranice pohybovalo v rozmezí

od -25 do -15 ms od maxima HNR. Pro manuální hranice umístěné do bodu maxima HNR nebo do bodu $+5$ ms od tohoto maxima se automatické hranice pohybovaly v intervalech -20 až -10 ms a -15 až -5 ms od maxima HNR. Na základě této úvahy jsme tedy získali interval -25 až -5 ms od maxima HNR (potenciálně i $+5$ ms až $+25$ ms od maxima HNR), v rámci kterého by se měly pohybovat všechny automaticky určené hranice u analyzovaných spojení. Tyto údaje ovšem nejsou zcela směrodatné a uvádíme je zde pouze pro představu toho, v jak velké míře můžeme automatickou segmentaci zpřesnit na základě využití křivky HNR. Podobnou myšlenkovou linii jsme aplikovali i u sekvencí VC. Jelikož byl u této skupiny rozptýl dat větší, má interval, který obsahuje většinu ($74,5$ %) analyzovaných manuálně určených hranic, délku 20 ms a nachází se v oblasti ± 10 ms od výskytu maxima HNR. Automaticky určené hranice u těchto analyzovaných spojení se tedy teoreticky pohybují v rozmezí -30 ms až $+30$ ms od maxima HNR, což představuje souvislé pásmo 60 ms, v rámci kterého může umístění automatických hranic kolísat.

Zde je ještě vhodné připomenout, že se daná analýza umístění hranic týkala pouze 356 z celkově naměřených 1955 položek. To představuje $18,2$ % objemu všech dat. Nami sestavená klasifikace průběhů HNR (viz kapitola 3.2.1) se přitom zakládala na analýze 459 z celkového počtu 1955 položek, obsahovala tedy $23,5$ % z objemu všech naměřených dat. Z toho vyplývá, že u $76,5$ % ze všech naměřených dat, které nebyly zahrnuty do naší analýzy, se umístění hranic lišilo o méně než 10 ms (příčemž tento případ převažoval), nebo o více než 20 ms, což bylo nejspíš zapříčiněno tím, že text použitý pro nucené zarovnání hláskových hranic neodpovídal skutečně realizované výslovnosti. Ta mohla být ovlivněna například přítomností hezitace nebo přeřeků, substituce apod.

4. Závěr a diskuze

Výsledky, jež prezentuje tato práce, jsou zaměřeny na ověření využitelnosti harmonicity k indikování hranic mezi segmenty, a proto by dále mohly být využity ke zpřesnění automatické segmentace, kterou v současnosti ve Fonetickém ústavu FF UK zajišťuje automatický segmentátor *Prague Labeller*.

Práce vycházela z nahrávek (k materiálu srov. kapitola 3.1), které byly nejdříve segmentovány automaticky pomocí *Prague Labelleru* a následně v nich byla provedena manuální korekce umístění hranic podle segmentačních pravidel, jak je uvádějí Machač & Skarnitzl (2009). Ve zkoumaném materiálu jsme se zaměřili na spojení neznělých obstruentů s vokály, která se objevila buď jako sekvence CV, nebo VC. Pro kvalitativní analýzu průběhu HNR na hranicích segmentů jsme si vybrali vzorek dat, u kterého se lišilo umístění automatické a manuálně korigované hranice o 10 až 20 ms. Na základě tohoto postupu jsme získali 459 položek, které sloužily jako východisko pro posouzení průběhu křivky harmonicity na hranicích segmentů.

Dalším krokem v naší analýze bylo sestavení klasifikace průběhů HNR (viz kapitola 3.2.1), jež by vyvrátilo tvrzení nulové hypotézy, že průběh HNR je na hranicích segmentů náhodný. Identifikovali jsme celkem 3 skupiny průběhů HNR: standardní průběh (skupina 1), nestandardní průběh (skupina 3) a rozkolísaný průběh (skupina 2), který se pohyboval na pomezí obou předešlých skupin. Ve skupinách 1 a 2 docházelo v průběhu křivky harmonicity k prudkému poklesu nebo nárůstu hodnot HNR, které byly natolik distinktivní, že jsme je dále mohli využít pro zkoumání vztahu mezi umístěním hranic a průběhem křivky HNR mezi segmenty. U skupiny 3 se hodnoty HNR na hranicích segmentů vyvíjely spíše konstantně na jedné úrovni. Zjistili jsme, že nejpočetněji byla v klasifikaci průběhů zastoupena 1. skupina (skupina standardních průběhů), a to 68,5 % ve spojeních CV a 59,8 % ve spojeních VC. Početně druhou největší skupinou byla v rámci spojení CV skupina 2, která obsahovala 22 % všech CV realizací, zatímco ve spojeních VC byla na druhém místě skupina 3 s 29,9 %. Početně nejméně zastoupené skupiny průběhů tvořila skupina 3 u CV spojení (9,5 %) a skupina 2 u VC spojení (10,3 %). Stejně vztahy v četnosti zastoupení jsme objevili, když jsme zkoumali segmentální zastoupení konsonantů v rámci jednotlivých skupin průběhů u CV a VC sekvencí. U CV sekvencí klesala četnost zastoupení exploziv, frikativ a afrikát v pořadí: skupina 1, skupina 2, skupina 3. Ve VC kontextech sledovalo segmentální zastoupení konsonantů stejné tendence jako četnost zastoupení průběhů HNR v rámci jednotlivých skupin, ovšem pouze v případě exploziv a afrikát. U frikativ se projevila silná tendence sledovat spíše průběh 3. skupiny z klasifikace (a to v 55,5 % případů), až poté průběh skupiny 1 (ve 33,6 % případů) a nejmenší preference u VC spojení s konsonantem

reprezentovaným frikativou projevil průběh křivky harmonicity vůči skupině 2 (10,9 % případů). Zároveň jsme zjistili, že jsou pro skupinu 3 typické spíše průběhy křivky s konstantním vývojem hodnot HNR, které se pohybují v oblasti kladných čísel (viz kapitola 3.2.2). Celkem se tento trend objevil u 71 z 87 nestandardních VC průběhů, přičemž 59 z těchto 71 výskytů reprezentovalo kontexty, ve kterých se vyskytovala frikativa. Tento fakt je zcela v rozporu s naším očekáváním, jelikož, jak zmiňují Machač & Skarnitzl (2009, str. 40), měla by neznělá frikativa představovat čistý šum bez periodické složky, což by mělo odpovídat záporným hodnotám HNR. Nečekaný vývoj křivky HNR ve spojeních obsahujících frikativu, tedy vývoj hodnot HNR v oblasti kladných čísel, si vysvětlujeme zejména přítomností neobvykle prominentních šumových formantů, které s největší pravděpodobností způsobily zdání periodicity.

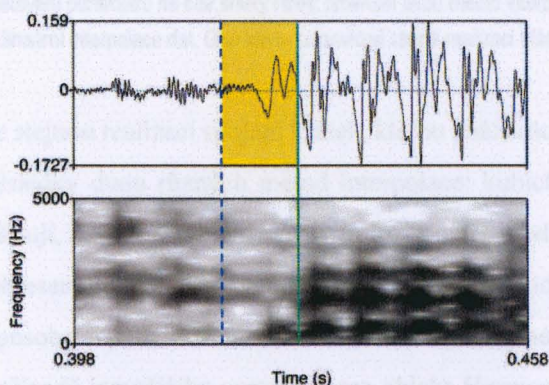
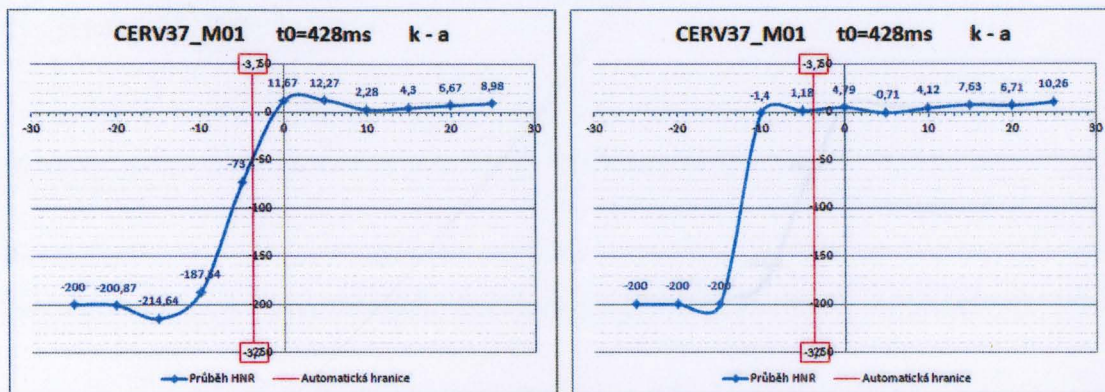
Posledním krokem v naší analýze bylo prozkoumání vztahu mezi umístěním manuální hranice a vývojem křivky HNR mezi segmenty (viz kapitola 3.2.3). Za tímto účelem byly využity průběhy 1. a 2. skupiny (celkem 356 položek). Zjistili jsme, že hranice mezi segmenty jsou umístěny spíše do oblasti s kladnými hodnotami HNR. Tento fakt koreluje s kritériem plné formantové struktury, které bylo využito při segmentaci vokálů. U spojení CV se v kladných číslech pohybovala hodnota HNR u 80,9 % realizací, ve spojeních VC u 65,2 % realizací. Dále jsme na základě naší analýzy stanovili konkrétní časové intervaly, v rámci kterých bylo umístěno početně nejvíc hranic. Tyto intervaly byly přitom vymezeny s ohledem na průběh HNR mezi segmenty. Pro spojení CV tento interval představoval délku 10 ms a nacházel se v rozmezí ± 5 ms od námi sledované hodnoty maxima HNR. V rámci tohoto intervalu bylo umístěno 95,4 % manuálně určených hranic. V případě automatické segmentace se přitom odpovídající hranice vyskytovaly v 20ms intervalu, který byl situován v oblasti -25 ms až -5 ms vzhledem k maximu HNR. Sekvence VC vykazovaly mnohem větší rozptyl v umístění hranic mezi segmenty. Ten si vysvětlujeme častým výskytem přechodových oblastí, které artikulačně odpovídají přibližování artikulátorů (Machač & Skarnitzl, 2009, str. 35 – 36) a které jsou zároveň s ohledem na manuální segmentaci dosti problematické. V případě VC sekvencí jsme mohli pozorovat umístění většiny (74,5 %) hranic v rámci časového intervalu o délce 20 ms, který se vzhledem k průběhu křivky HNR nacházel v oblasti ± 10 ms od námi sledované hodnoty maxima. V případě automatické segmentace se přitom odpovídající hranice analyzovaných spojení pohybovaly v rozmezí 60 ms, tj. v oblasti ± 30 ms kolem hodnoty maxima HNR.

Shrňme-li dílčí výsledky našich analýz, můžeme konstatovat, že harmonicita se projevila jako slibný indikátor hranic segmentů. K tomuto závěru jsme došli na základě identifikace 3 typů průběhů HNR na hranicích segmentů, přičemž průběh, který byl označen za standardní, také převládá v čet-

nosti svého zastoupení v rámci klasifikace. Tento průběh koresponduje s očekávanými vlastnostmi spojení neznělých obstruentů s vokály, a tím nám umožňuje stanovit hranici mezi segmenty. Tento závěr jsme zároveň ověřili v kapitole 3.2.3, kde jsme zkoumali souvislost mezi umístěním manuálně určených hranic a průběhem křivky HNR. Na základě srovnání dat jsme byli schopni stanovit časové intervaly o délce 10 ms (u spojení CV) respektive 20 ms (u spojení VC), které ale pořád představovaly zpřesnění v porovnání s výsledky automatické segmentace.

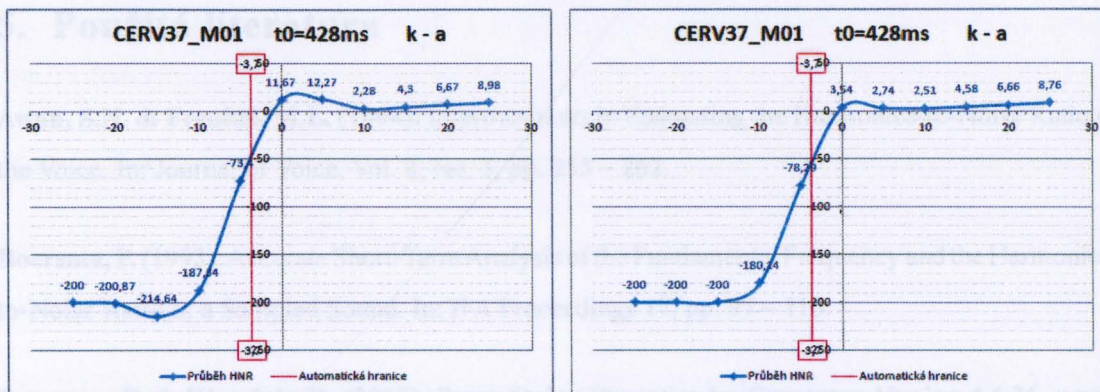
Tato práce prezentuje výsledky získané za použití standardního, tj. softwaru *PRAAT* doporučeného, nastavení, a to jak pro vytvoření objektu *Harmonicity*, tak i pro zjišťování hodnoty HNR v konkrétním čase. Pro implementaci harmonicity do procesu automatické segmentace řeči bude ovšem třeba najít takové nastavení parametrů, při kterém je úspěšnost určování hranic na základě průběhu křivky HNR největší. Jelikož tato práce zpracovává výsledky získané na základě jednoho nastavení parametrů, zůstává předmětem diskuze, do jaké míry a v jakém směru může nastavení jednotlivých parametrů ovlivnit konečné výsledky. Předmětem dalšího výzkumu pak bude prozkoumání úspěšnosti určování hranic mezi segmenty pro různá nastavení parametrů a stanovení optimální sady nastavení pro implementaci do procesu automatické segmentace.

Standardní nastavení používá při vytváření objektu *Harmonicity* časový krok 10 ms a pro odhad hodnoty HNR v konkrétním čase volí kubickou interpolaci dat. Jelikož nás zajímal vliv tohoto nastavení na výsledná data, provedli jsme v rámci diskuze pokusné měření, které srovnává dva různé časové kroky pro vytvoření objektu *Harmonicity* (viz obrázek č. 37) a dvě různé metody interpolace dat (viz obrázek č. 38).



Obrázek č. 37. Vliv nastavení parametrů na tvar křivky HNR: srovnání výsledků dvou nastavení pro vytvoření objektu *Harmonicity*. Vlevo nahoře standardní nastavení (časový krok 10 ms) s použitím kubické interpolace dat, vpravo nahoře jemnější nastavení (časový krok 1 ms) opět s použitím kubické interpolace dat. Obě křivky odpovídají stejné realizaci hlásek (CV). Dole odpovídající oscilogram a spektrogram: žlutě vyznačený interval odpovídá 10 ms a zelená linka odpovídá umístění manuální hranice.

Pokud zvolíme jemnější nastavení (tj. menší časový krok) pro vytvoření objektu *Harmonicity*, zůstává tvar křivky zachován (to se potvrdilo i pro průběhy z 2. a 3. skupiny klasifikace). Změny nastávají až v oblasti přechodu mezi hláskami, který je u křivky HNR v případě standardního nastavení pozvolnější. Žlutě vybarvené pásmo ve spektrogramu (viz obrázek č. 37) přitom ve standardním nastavení koreluje s prudkým nárůstem hodnot HNR, zatímco v nastavení s časovým krokem 1 ms tuto část signálu popisují již kladné hodnoty harmonicity. V tomto pásmu přitom dochází k předznívání znělosti a začíná se zde také formovat formantová struktura (viz spektrogram na obrázku č. 37), což v oscilogramu odpovídá jednoduchému tvaru vlny. V obou grafech na obrázku č. 37 je k získání hodnot HNR použita kubická interpolace, která ovšem v případě standardního nastavení časového kroku způsobuje větší výkyvy v tvaru křivky než v případě nastavení s menším časovým krokem.



Obrázek č. 38. Vliv nastavení parametrů na tvar křivky HNR: srovnání dvou metod interpolace dat. Vlevo kubická interpolace dat, vpravo lineární interpolace dat. Obě křivky odpovídají stejné realizaci hlásek (CV).

Na obrázku č. 38 vidíme stejnou realizaci spojení hlásek, kterou znázorňuje i obrázek č. 37, tentokrát ale můžeme srovnat výsledky dvou různých metod interpolace: kubické a lineární. Oba obrázky (č. 37 i č. 38) přitom indikují, že mírná zvlnění křivky (mírný pokles před prudkým nárůstem a mírný nárůst před prudkým poklesem hodnot) těsně před razantní změnou v hodnotách HNR, jak je popisuje kapitola 3.2.2, jsou způsobena použitím standardního nastavení, zejména pak kubické interpolace. Na obrázku č. 37 se v případě jemnějšího nastavení pro objekt *Harmonicity* tato zvlnění neprojeví v takové míře a stejně je tomu i na obrázku č. 38, kde jsou v případě použití lineární interpolace hodnot sledované výkyvy také menší.

V naší práci jsme prozkoumali využitelnost harmonicity v oblasti automatické segmentace řeči a dospěli jsme k závěru, že se harmonicitu prokázala jako dobrý indikátor hranic mezi segmenty. Podněty týkající se nastavení parametrů, které jsme předdeslali v diskuzi, mohou být tedy předmětem dalšího výzkumu, který se může zabývat např. tím, které nastavení lépe odpovídá fonetické realitě a/nebo tím, které nastavení je z hlediska automatické segmentace optimální a dosahuje nejlepších výsledků (respektive největší přesnosti) při určování hranic mezi segmenty.

Hansson, H.M., Stevens, K.N., Kim, H.-K. J., Chen, M.Y. & Siffra, J. (2001): Towards Models of Phonation. In: *Journal of Phonetics*, Vol. 29, pp. 451–480.

Heranová, J. (2009): *Harmonicitu českých segmentů*. Nepublikovaná písemná práce. Praha: Fonetický ústav FF UK.

Klinghoffer, K. & Norris, P. (1985): Quantitative Spectral Evaluation of Shimmer and Jitter. In: *Journal of Speech and Hearing Research*, Vol. 28, pp. 169–174.

5. Použitá literatura

Awan, S.N. & Frenkel, M.L. (1994): Improvements in Estimating the Harmonics-to-Noise Ratio of the Voice. In: *Journal of Voice*, Vol. 8, No. 3, pp. 255 – 262.

Boersma, P. (1993): Accurate Short-Term Analysis of the Fundamental Frequency and the Harmonics-to-Noise Ratio of a Sampled Sound. In: *IFA Proceedings 17*, pp. 97 – 110.

Boersma, P. & Weenink, D. (2007): Praat: Doing Phonetics by Computer, Version 4.6.36, www.praat.org.

Brockmann, M., Storck, C., Carding P.N. & Drinnan, M.J. (2008): Voice Loudness and Gender Effects on Jitter and Shimmer in Healthy Adults. In: *Journal of Speech, Language, and Hearing Research*, Vol. 51, pp. 1152 – 1160.

Cox, N.B., Ito, M.R. & Morrison, M.D. (1989): Technical Considerations in Computation of Spectral Harmonics-to-Noise Ratios for Sustained Vowels. In: *Journal of Speech and Hearing Research*, Vol. 32, pp. 203 – 218.

de Krom, G. (1993): A Cepstrum-Based Technique for Determining a Harmonics-to-Noise Ratio in Speech Signals. In: *Journal of Speech and Hearing Research*, Vol. 36, pp. 254 – 266.

Ferrand, C.T. (2002): Harmonics-to-Noise Ratio: An Index of Vocal Aging. In: *Journal of Voice*, Vol. 16, No. 4, pp. 480 – 486.

Gelfer, M.P. & Fendel, D.M. (1995): Comparisons of Jitter, Shimmer, and Signal-to-Noise Ratio from Directly Digitized Versus Taped Voice Samples. In: *Journal of Voice*, Vol. 9, No. 4, pp. 378-382.

Hanson, H.M., Stevens, K.N., Kuo, H.-K. J., Chen, M.Y. & Slifka, J. (2001): Towards Models of Phonation. In: *Journal of Phonetics*, Vol. 29, pp. 451 – 480.

Heranová, J. (2009): Harmonicita českých segmentů. Nепublikovaná písemná práce. Praha: Fonetický ústav FF UK.

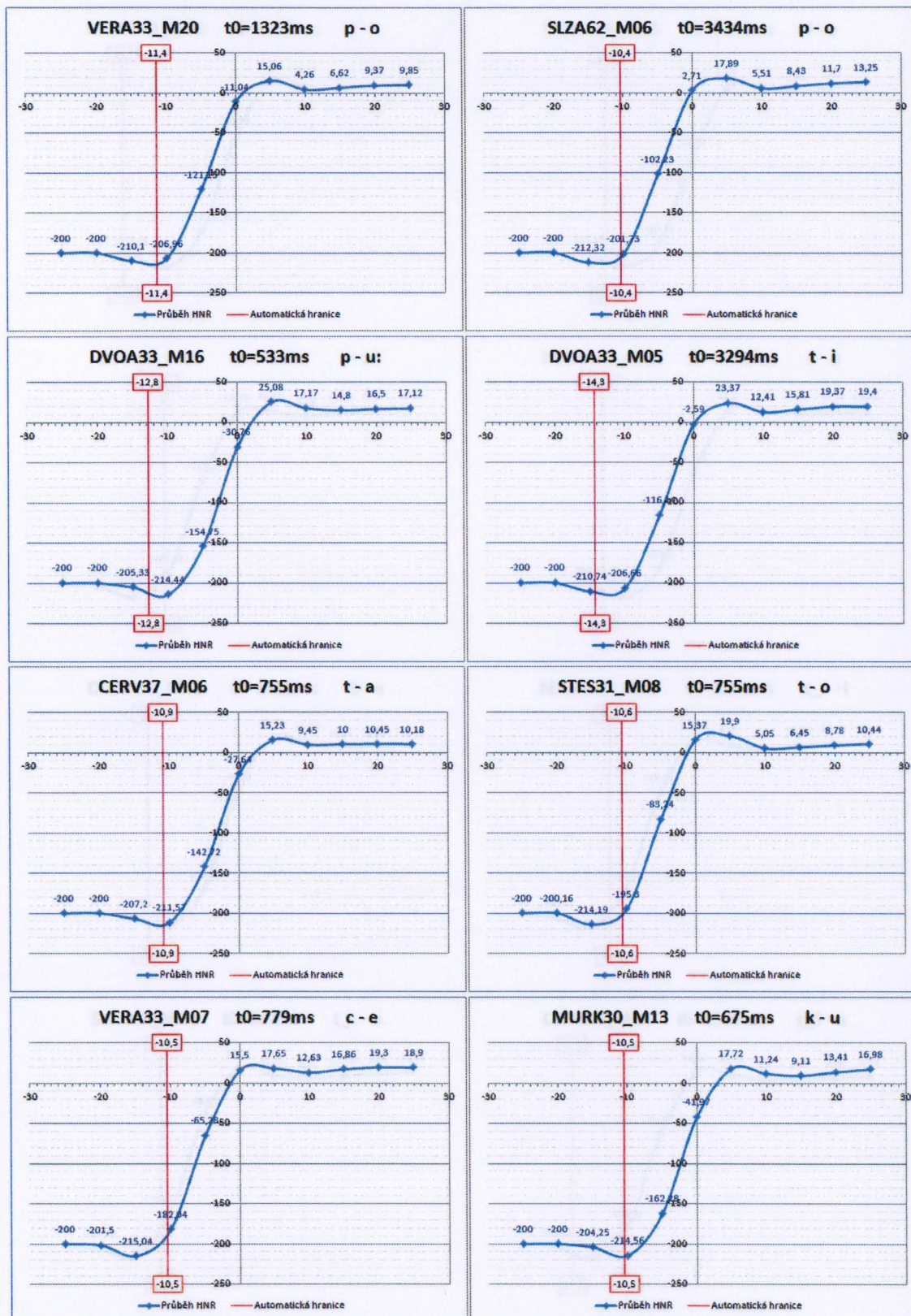
Klingholz, F. & Martin, F. (1985): Quantitative Spectral Evaluation of Shimmer and Jitter. In: *Journal of Speech and Hearing Research*, Vol. 28, pp. 169 – 174.

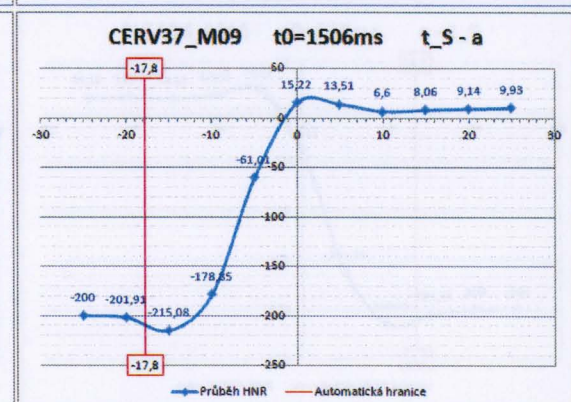
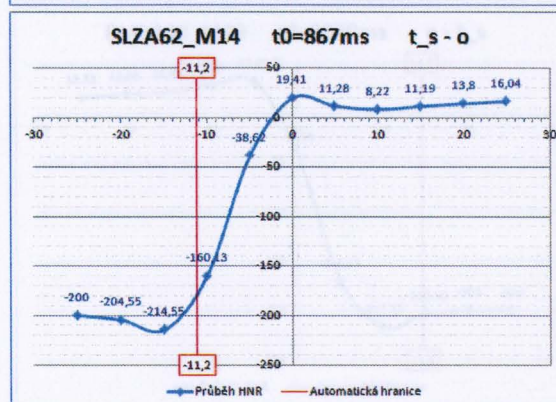
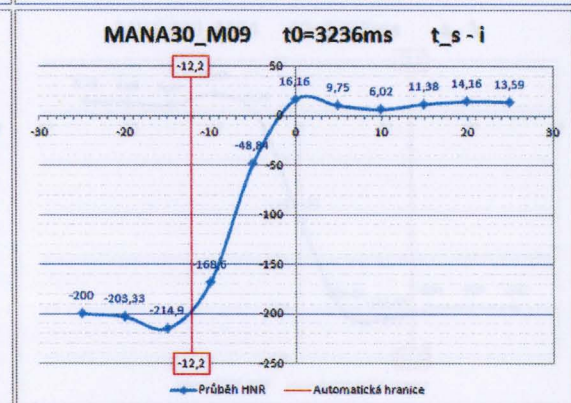
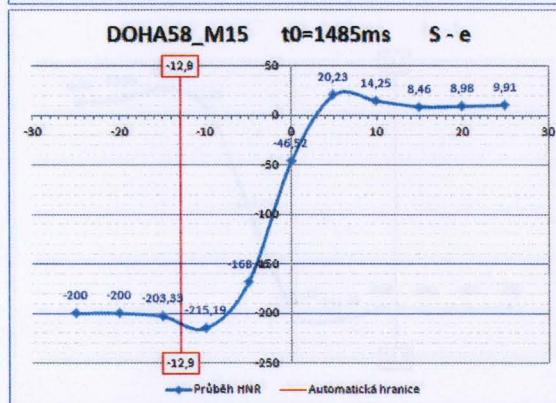
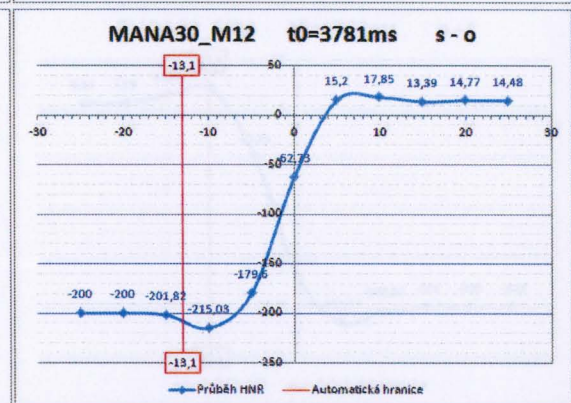
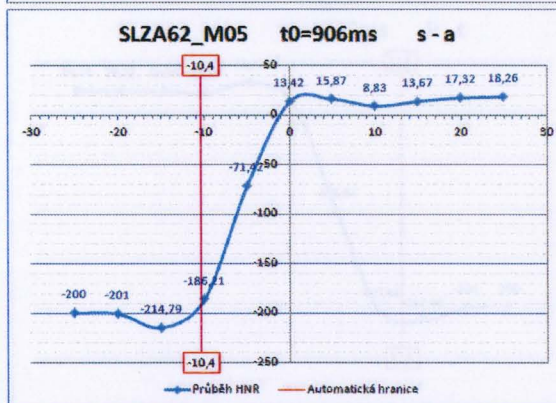
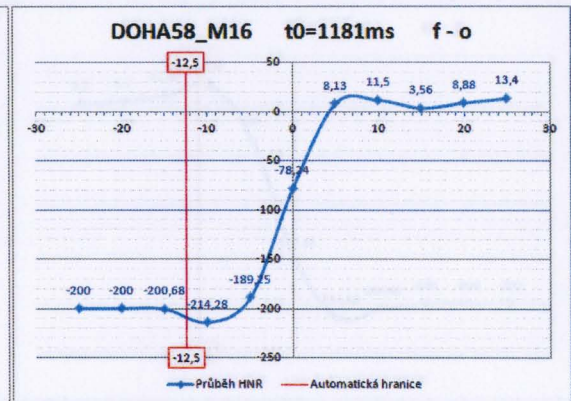
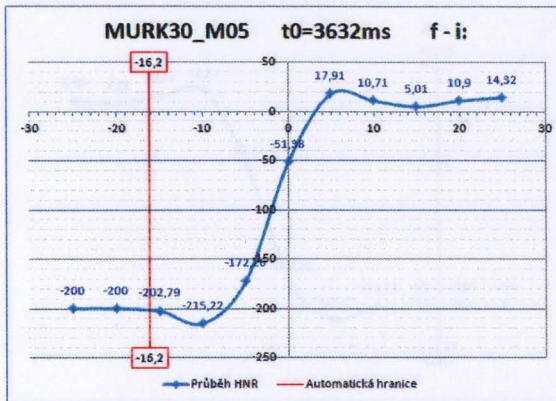
- Linville, S.E.** (2000): The Aging Voice. In: Kent, R. D. a Ball, M. J. (Eds.): Voice quality measurement. San Diego: Singular Publishing Group, pp. 359 – 372.
- Machač, P. & Skarnitzl, R.** (2009): Fonetická segmentace hlásek. Praha: Nakladatelství Epona.
- Möbius, B.** (2004): Corpus-Based Investigations on the Phonetics of Consonant Voicing. In: Folia Linguistica XXXVIII/1-2, Berlin: Mouton de Gruyter, pp. 5 – 26.
- Murphy, P.J.** (2006): Periodicity Estimation in Synthesized Phonation Signals Using Cepstral Harmonic Peaks. In: Speech Communication, Vol. 48, pp. 1704 – 1713.
- Palková, Z.** (1997): Fonetika a fonologie češtiny. Praha: Karolinum.
- Pollák, P.** (2008): Přednášky k rozpoznávání řeči v rámci předmětu „Speciální problémy fonetiky: Akustika“, Praha: Fonetický ústav FF UK.
- Pollák, P., Volín, J. & Skarnitzl, R.** (2007): HMM-Based Phonetic Segmentation in Praat Environment. In: Proceedings of the XIIth International Conference „Speech and computer – SPECOM 2007“, Moscow, pp. 537-541.
- Psutka, J., Müller, L., Matoušek, J. & Radová, V.** (2006): Mluvíme s počítačem česky. Praha: Academia.
- Qi, Y. & Hillman, R.E.** (1997): Temporal and Spectral Estimations of Harmonics-to-Noise Ratio in Human Voice Signals. In: Acoustical Society of America, Vol. 102, No. 1, pp. 537 – 543.
- Schötz, S. & Müller, C.** (2007): A Study of Acoustic Correlates of Speaker Age. In: Müller, C. (Ed.) (2007): Speaker Classification II, LNAI, Vol. 4441, Springer Verlag Berlin Heidelberg, pp. 1 – 9.
- Těšilová, M. & kolektiv** (1987): O češtině v číslech. Praha: Academia.
- Vaissière, J.** (1977): Phonological Use of the Larynx. In: LARYNX 1997, Marseille, France, June 16 – 18, pp. 115 – 125.
- Volín, J., Skarnitzl, R. & Pollák, P.** (2005): Confronting HMM-based Phone Labelling with Human Evaluation of Speech Production. In: Interspeech 2005, Lisboa.

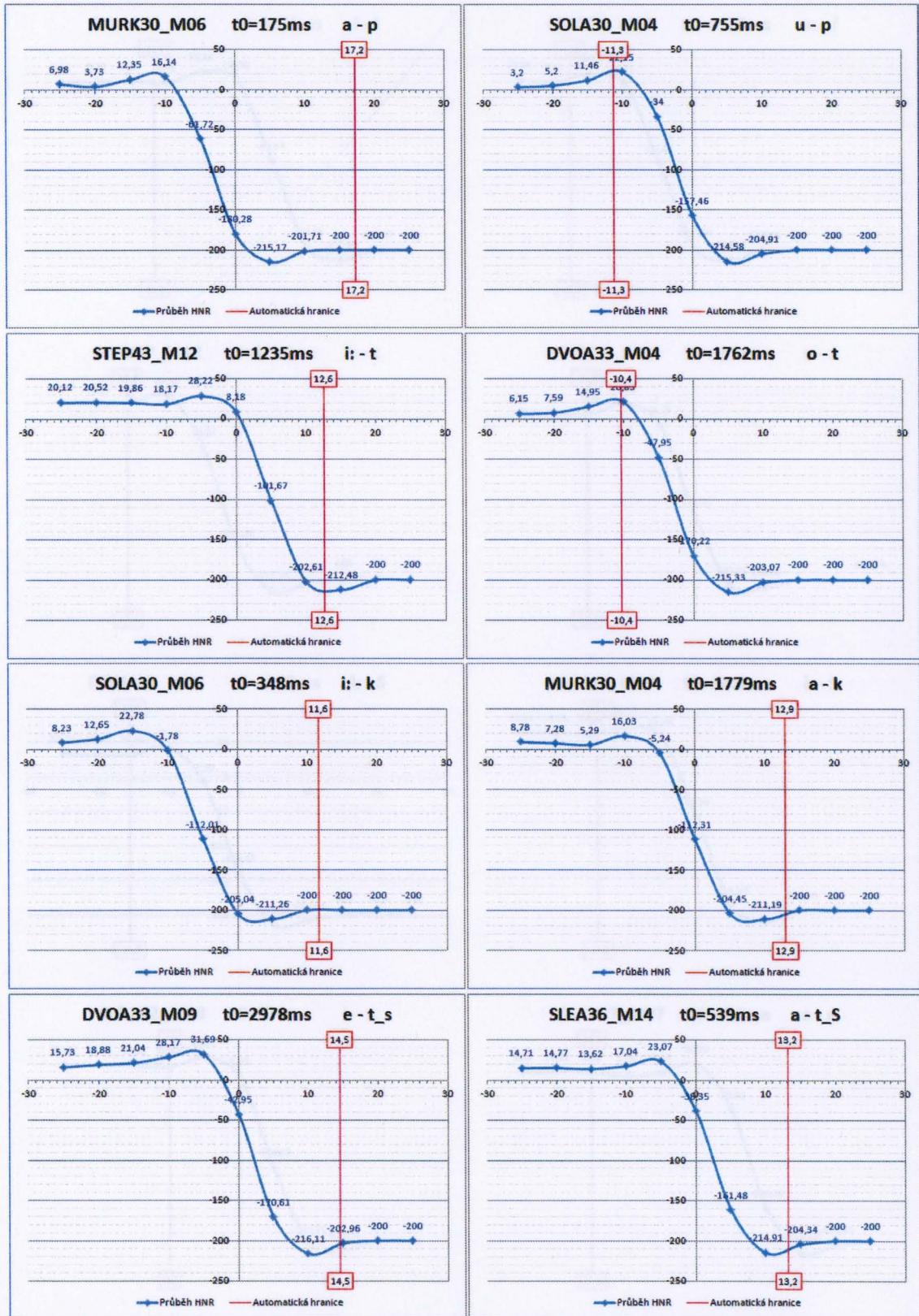
Yu, A.-T. & Wang, H.-C. (2006): New Speech Harmonic Structure Measure and its Applications to Speech Processing. In: Acoustical Society of America, Vol. 120, No. 5, pp. 2938 – 2949.

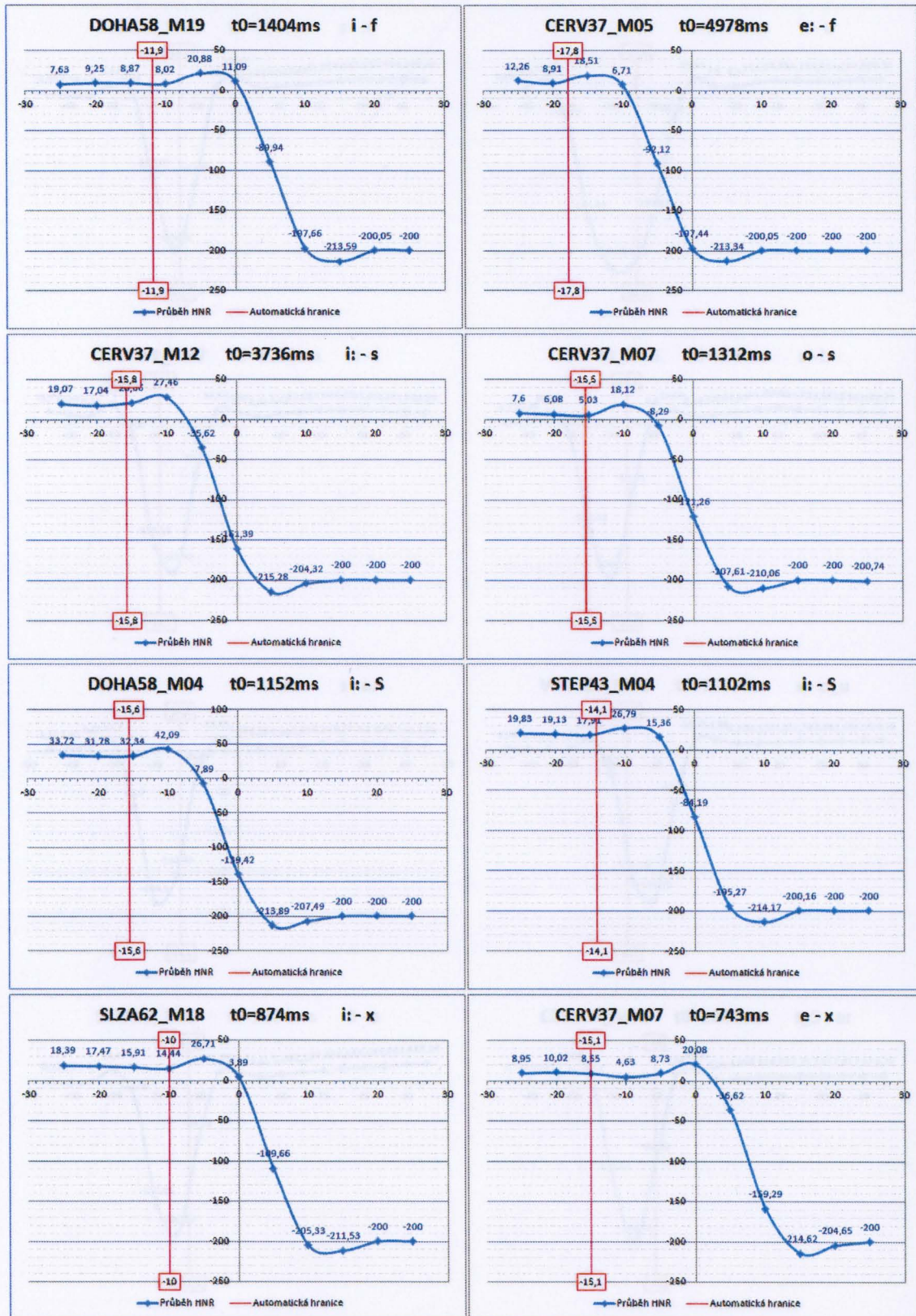
Yumoto, E., Gould, W.J. & Baer, T. (1982): Harmonics-to-Noise Ratio as an Index of the Degree of Hoarseness. In: Acoustical Society of America, Vol. 71, No. 6, pp. 1544 – 1549.

A. Příloha – grafy skupiny 1

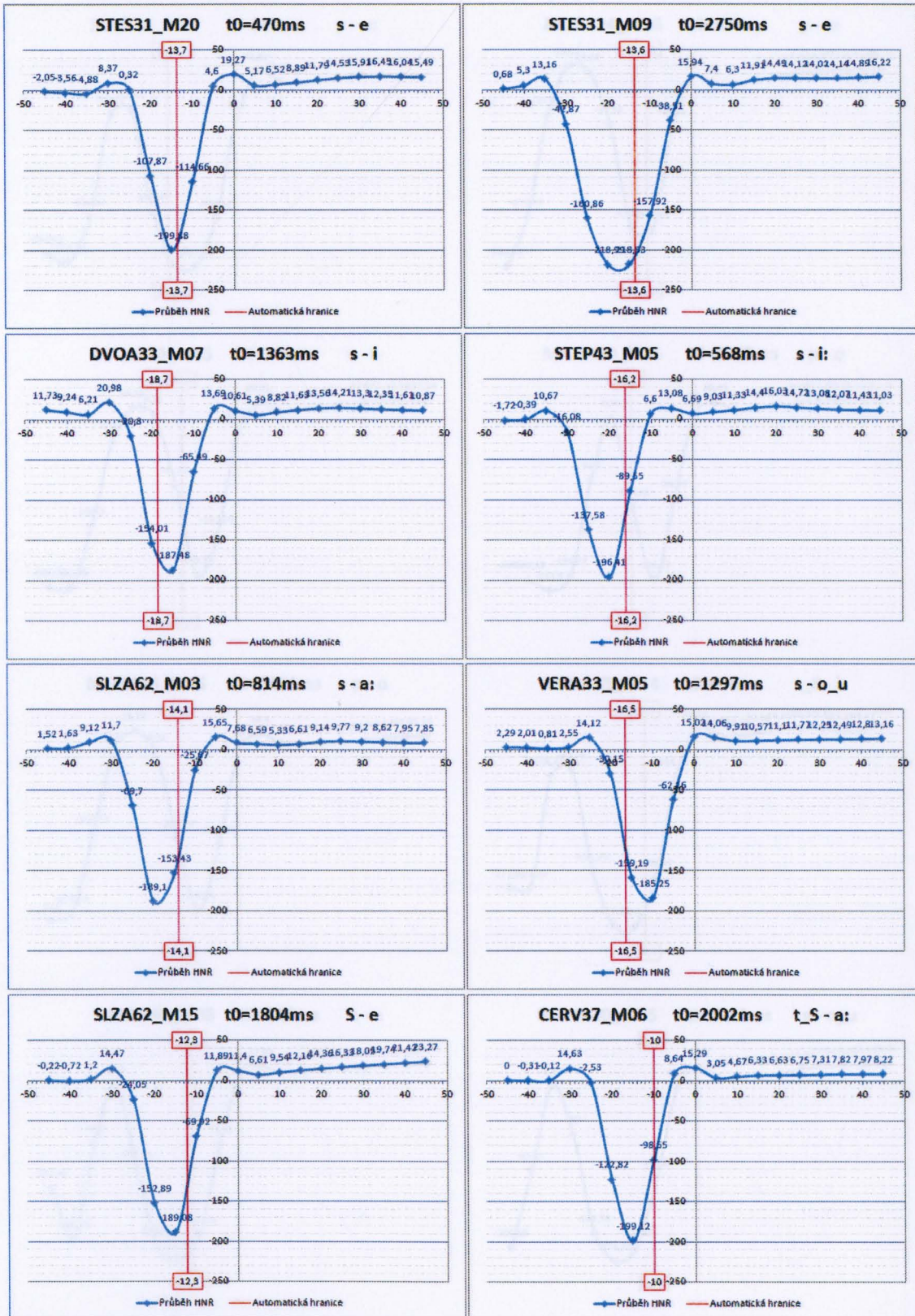


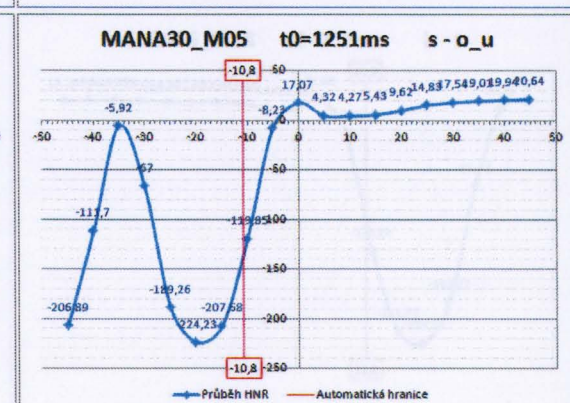
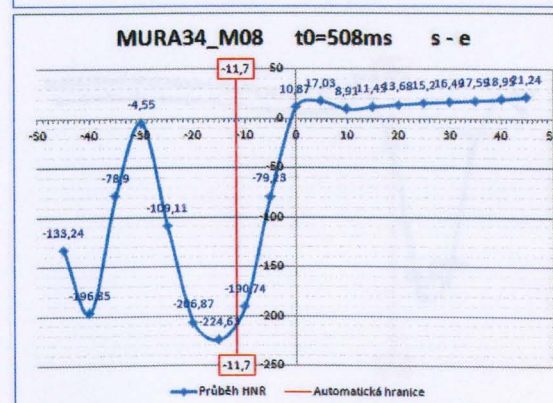
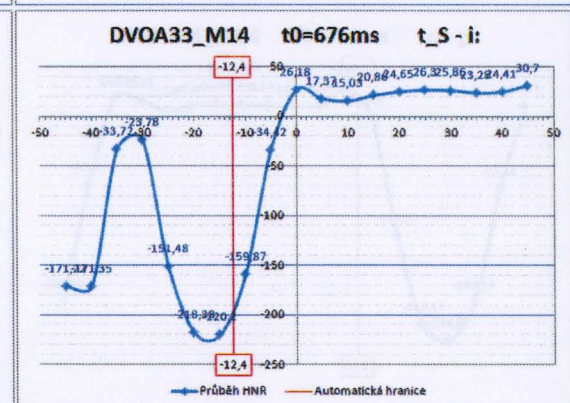
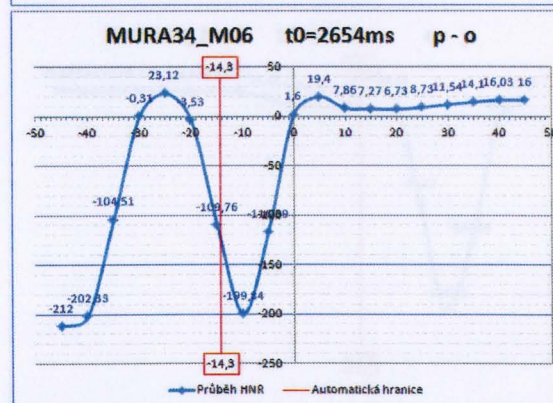
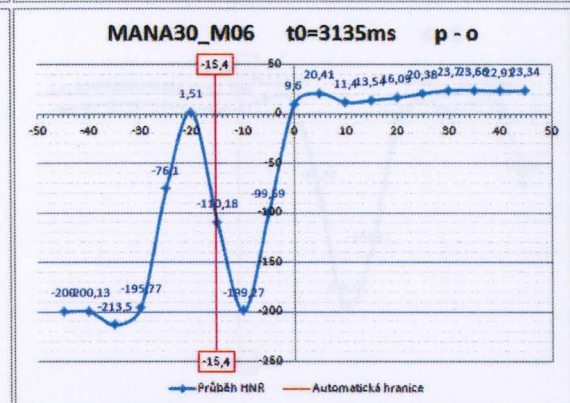
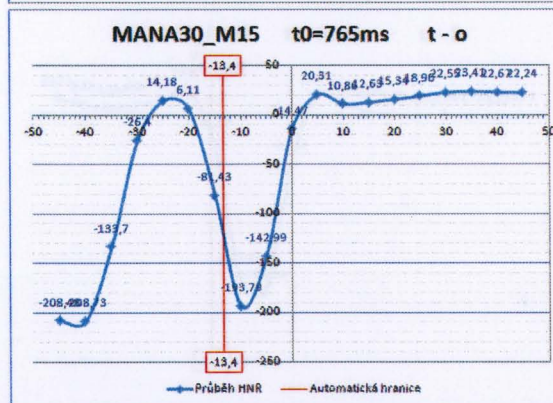
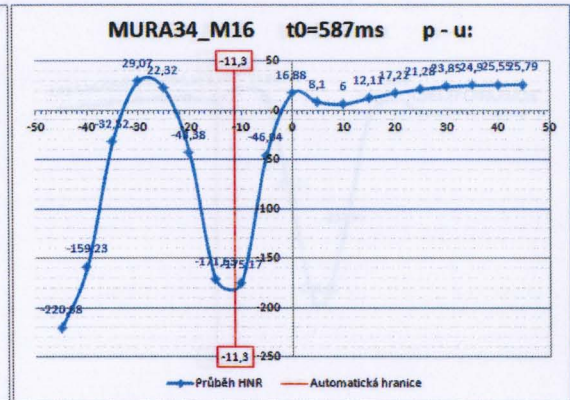
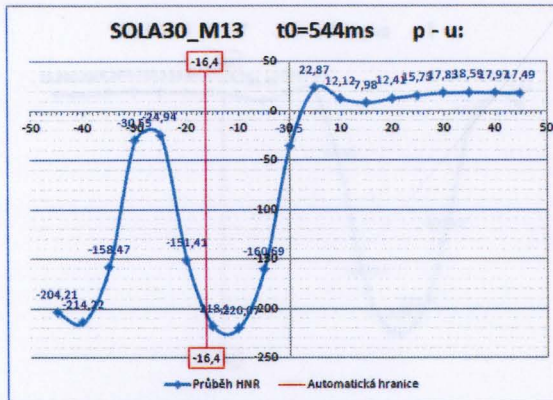


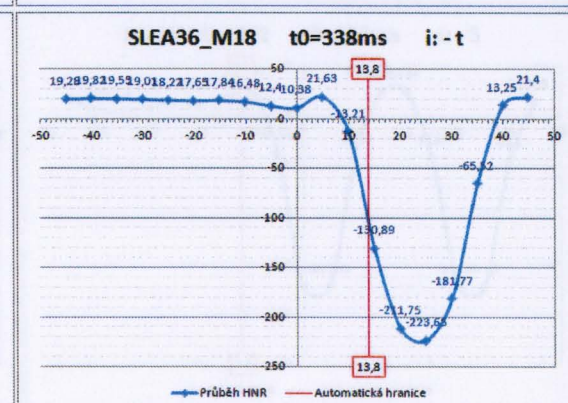
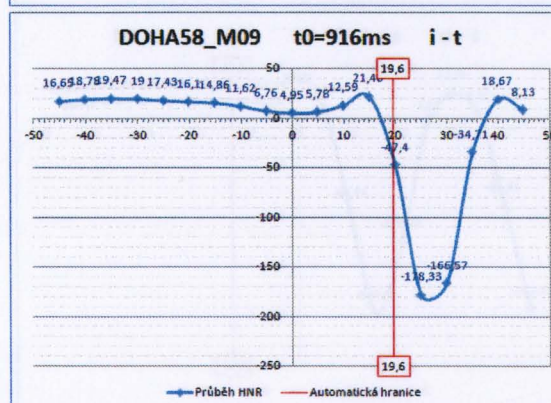
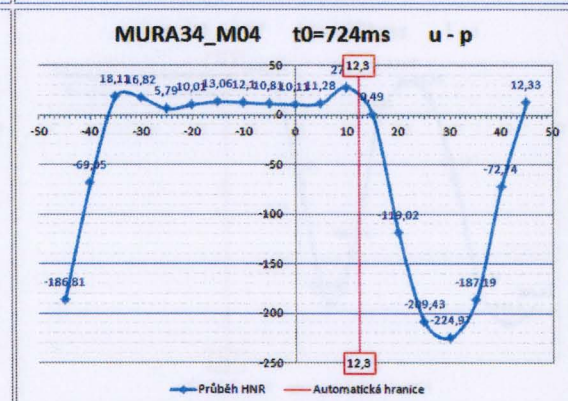
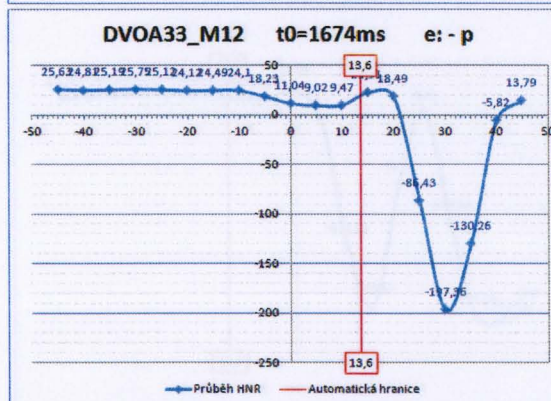
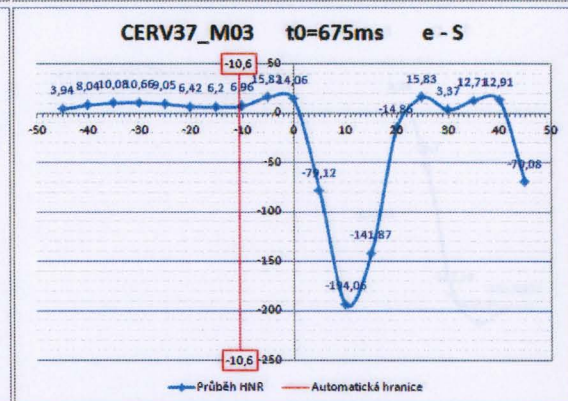
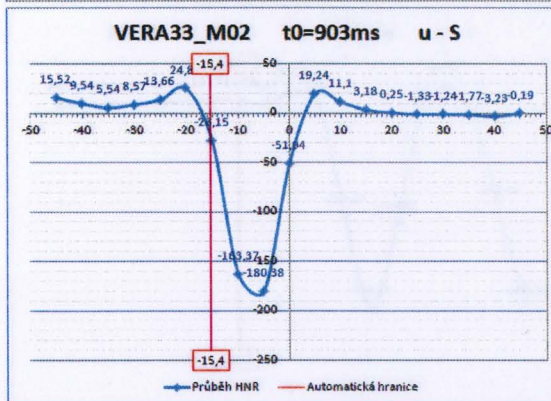
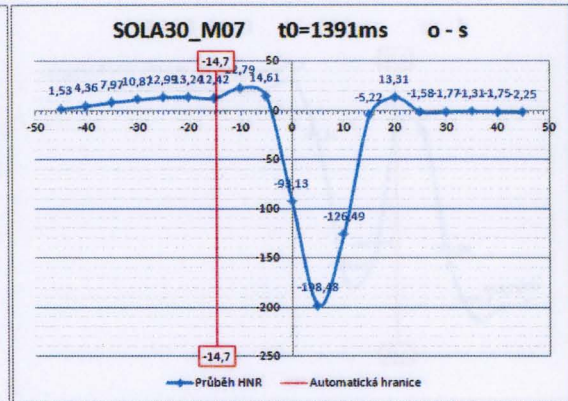
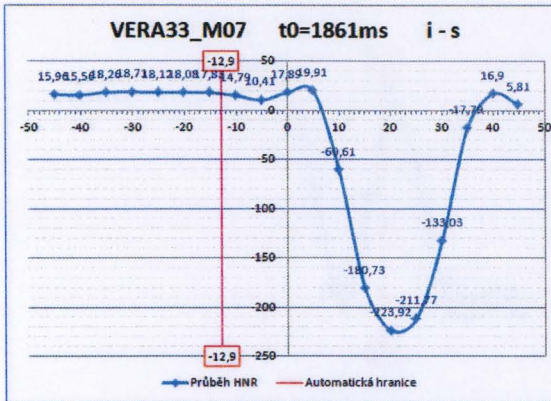




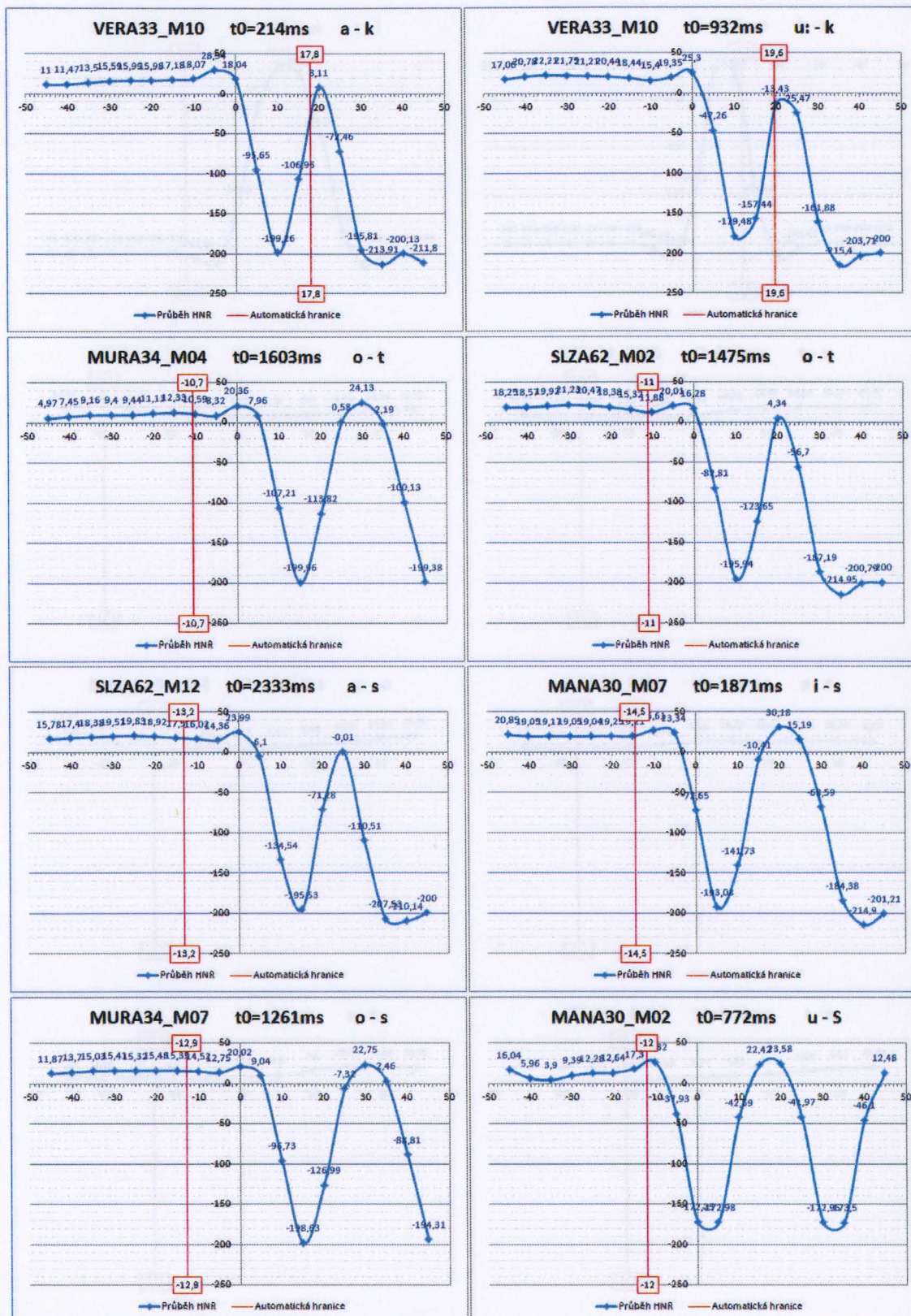
B. Příloha – grafy skupiny 2



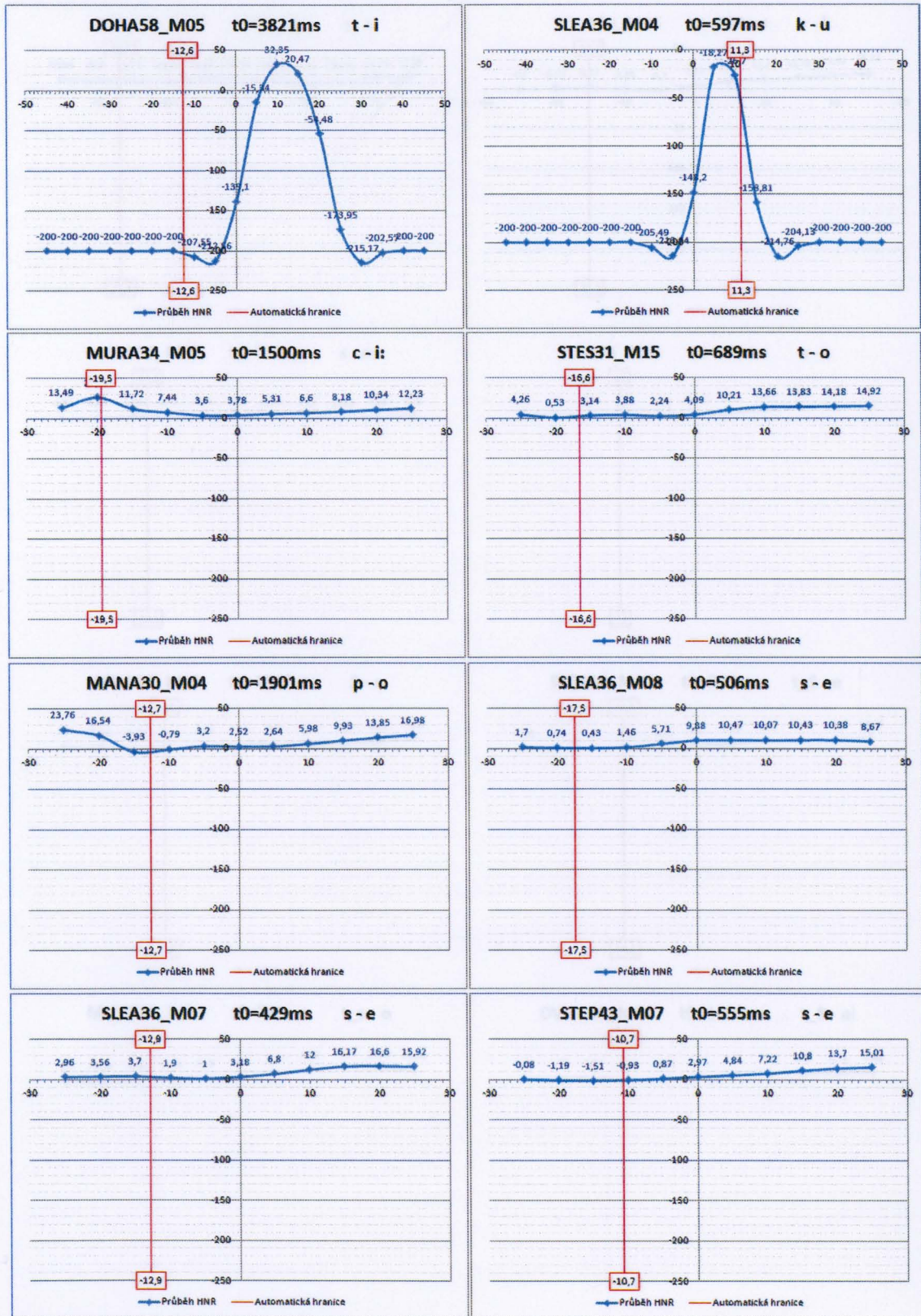


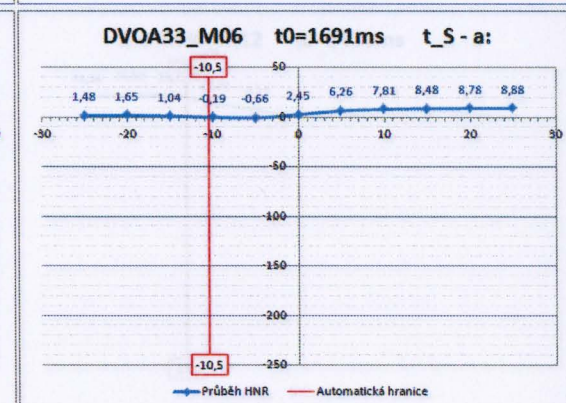
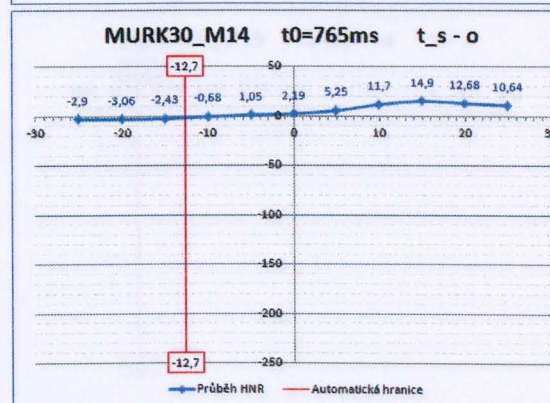
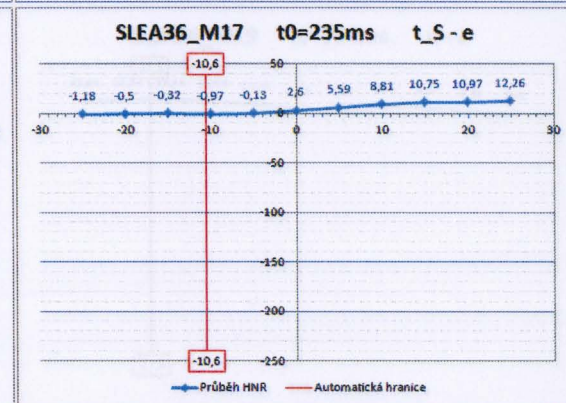
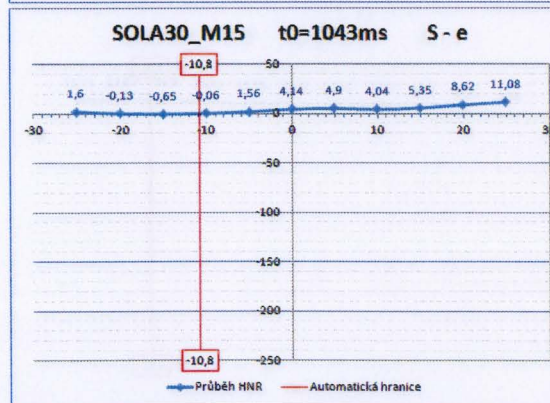
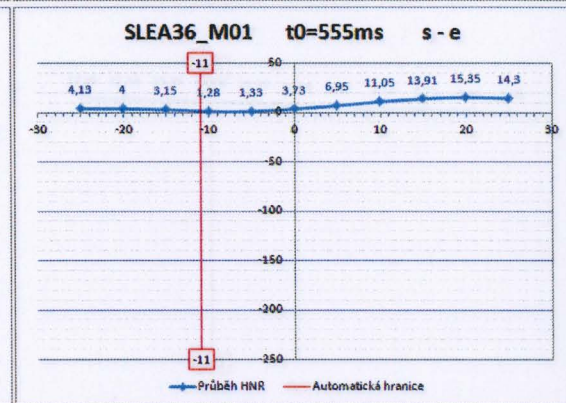
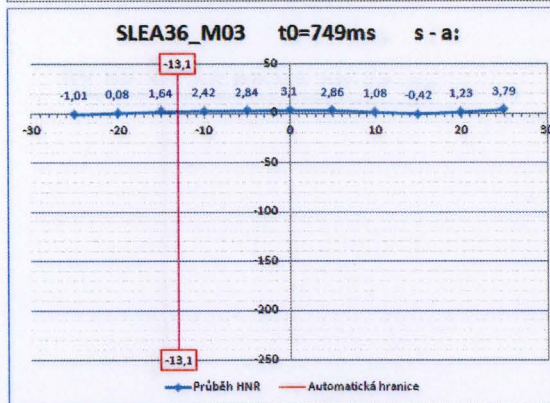
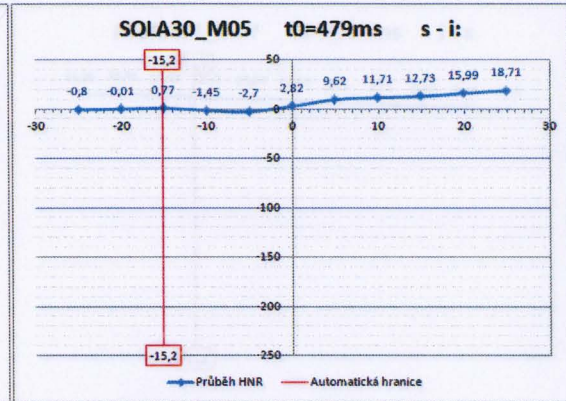
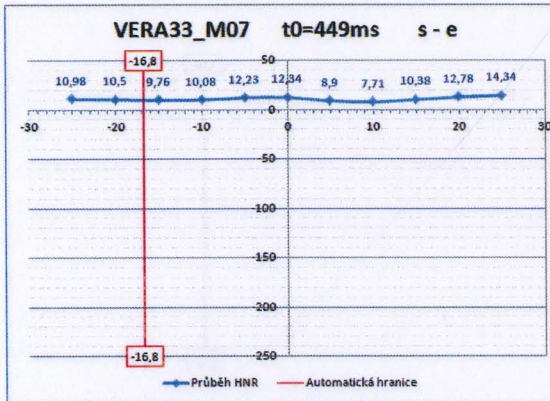


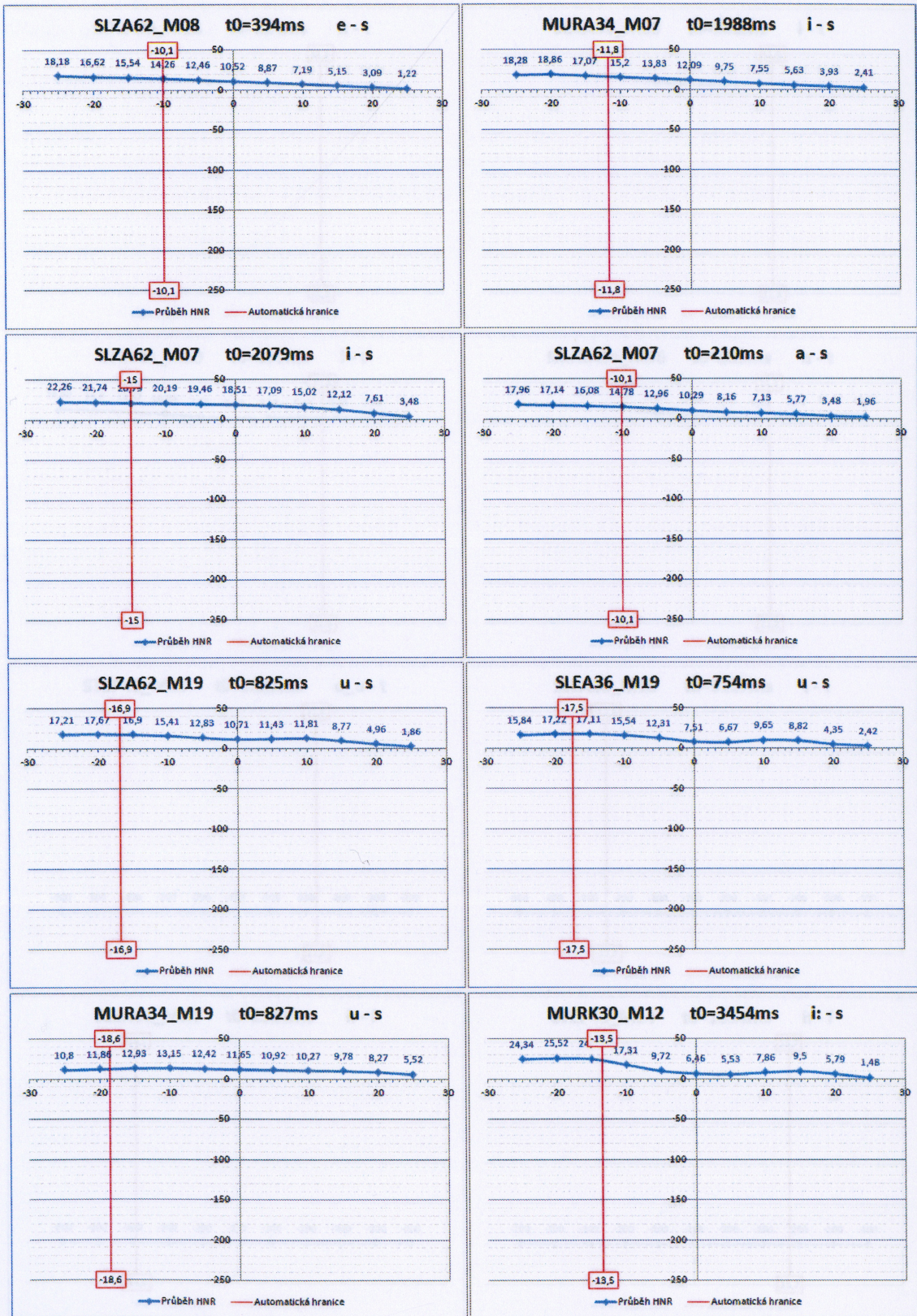
C. Příloha – grafy skupiny 3

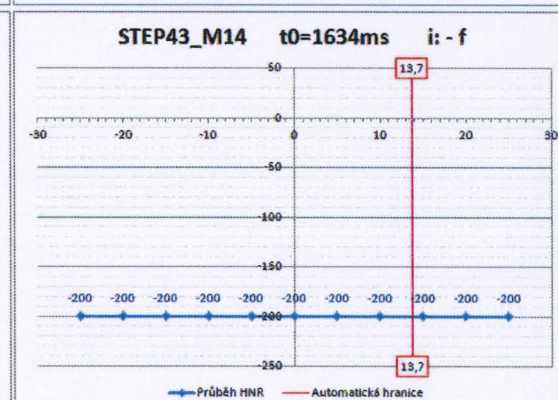
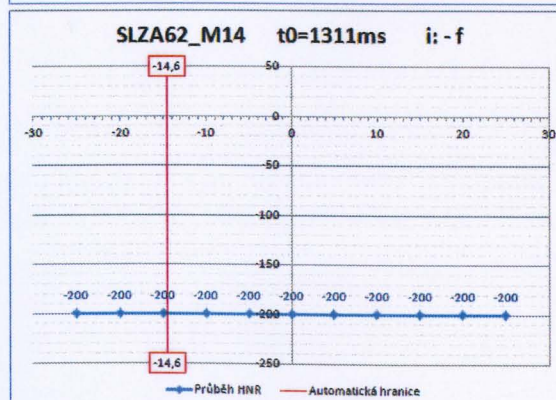
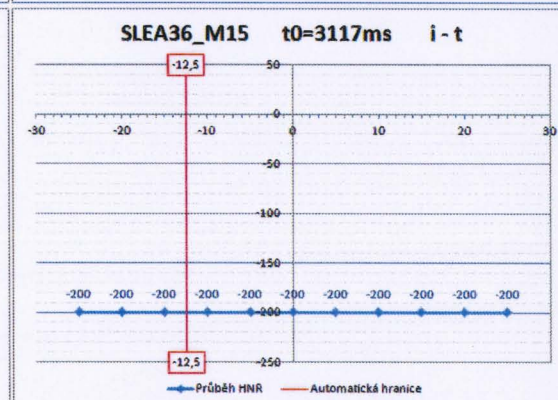
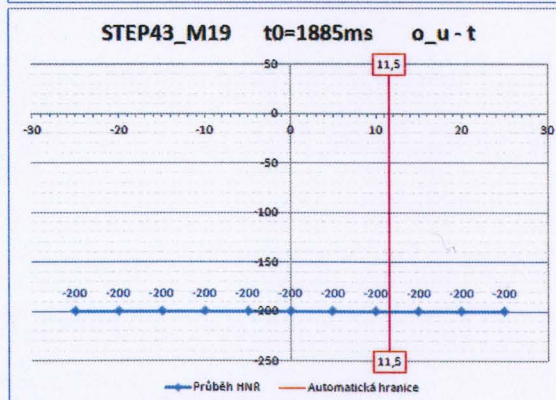
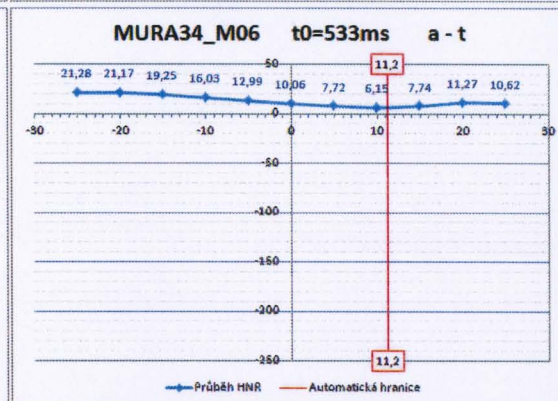
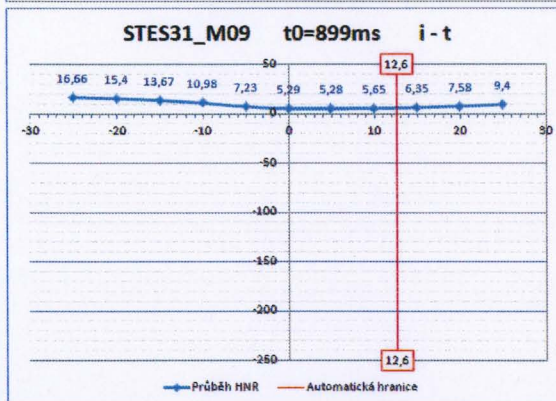
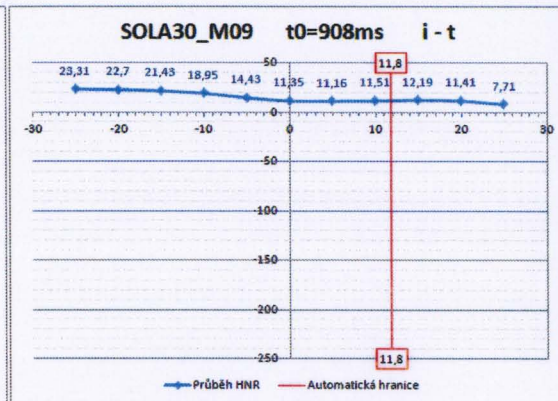
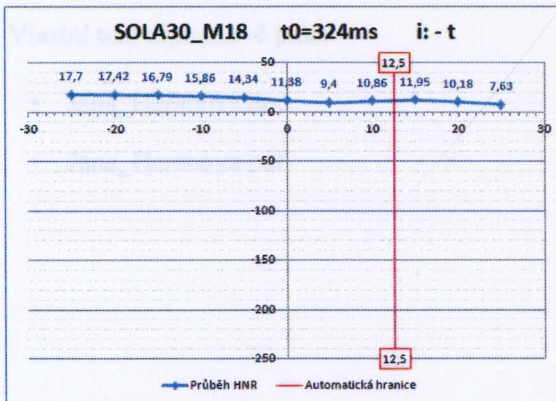


C. Příloha – grafy skupiny 3









D. Příloha – obsah přiloženého CD

Vlastní text diplomové práce:

- Jana_Heranova.doc
- Jana_Heranova.pdf