

**Pavel Machač
a Radek Skarnitzl**

Fonetická segmentace

HLÁSEK

- spektrogram
- formantová
struktura
- akustické ukazatele
- oscilogram
- přechodová oblast
- umístění hranice
- formantový sloupek
- hláskové kombinace
- shoda mezi
anotátory
- spektrální intenzita
- pečlivý poslech

EDICE ERUDICA

NAKLADATELSTVÍ EPOCHA

Vědecká redakční rada EDICE ERUDICA

prof. PhDr. František Mezihorák, CSc., Dr.h.c. – Univerzita Palackého v Olomouci, ČR

prof. PhDr. Erich Mistrík, CSc. – Univerzita Komenského Bratislava, SR

prof. ThDr. Jan B. Lášek – Univerzita Karlova v Praze, ČR

doc. PhDr. Zdeněk Novotný, CSc. – Univerzita Palackého v Olomouci, ČR

doc. PhDr. Miroslav Sapík, Ph.D. – Jihočeská univerzita v Českých Budějovicích, ČR

Mgr. Antonín Staněk, Ph.D. – Univerzita Palackého v Olomouci, ČR

prof. PhDr. Cyril Diatka, CSc. – Univerzita Konštantína Filozofa v Nitre, SR

doc. PhDr. Josef Oborný, Ph.D. – Univerzita Komenského Bratislava, SR

Dr. Małgorzata Świder – Instytut Historii Uniwersytetu Opolskiego, PL

prof. Dr. Andrew Burgess – University of New Mexico, American Academy of Religion, USA

PhDr. Martina Klicperová – Baker, CSc., mim. prof. – San Diego State University, USA

doc. PhDr. Naděžda Pelcová, CSc. – Univerzita Karlova v Praze, ČR

doc. PhDr. Nikolaj Demjančuk, CSc. – Západočeská univerzita v Plzni, ČR

doc. PaedDr. Vanda Hájková, Ph.D. – Univerzita Karlova v Praze, ČR

FONETICKÁ
SEGMENTACE
HLÁSEK



NAKLADATELSTVÍ
EPOCHA

Pavel Machač a Radek Skarnitzl

FONETICKÁ
SEGMENTACE
HLÁSEK

NAKLADATELSTVÍ EPOCHA

Poděkování

Rádi bychom vyjádřili svůj dík našemu kolegovi a příteli Janu Volínovi za jeho podnět k napsání této knihy i za cenné rady a předané zkušenosti. Chtěli bychom také poděkovat Janě Heranové a Lucii Ondruškové za užitečné poznámky k rukopisu.

Tato kniha byla napsána za podpory grantů MRTN-CT-2006-035561 (evropský grant Sound to Sense), VZ MSM 0021620825 a GACR 102/09/0989 (kapitoly 10 a 12).

Copyright © Pavel Machač a Radek Skarnitzl, 2009

Cover © Petra Süssová, 2009

Czech Edition © Nakladatelství Epoque, Praha 2009

ISBN 978-80-7425-031-6

Obsah

| | |
|--|-----------|
| 1. Úvod | 11 |
| 1.1. Proč potřebujeme hranice segmentů? | 11 |
| 1.2. Co rozumíme pod pojmem „hranice“?..... | 16 |
| 1.3. Fonetické rysy..... | 20 |
| 1.3.1. Vlastní fonetické rysy..... | 20 |
| 1.3.2. Nevlastní fonetické rysy..... | 21 |
| 1.3.3. Hranice segmentu a rozložení hláskových rysů.. | 22 |
| 1.4. Metodologické a terminologické poznámky | 23 |
| 2. Explozivy v intervokálníkové pozici | 27 |
| 2.1. Uvedení do artikulačních a akustických vlastností ... | 27 |
| 2.2. Vlastní fonetické rysy a základní pravidla pro segmentaci..... | 28 |
| 2.3. Další pravidla pro segmentaci | 32 |
| 2.4. Shrnutí | 39 |
| 3. Intervokálníkové frikativy | 40 |
| 3.1. Uvedení do artikulačních a akustických vlastností ... | 40 |
| 3.2. Vlastní fonetické rysy a základní pravidla pro segmentaci..... | 42 |
| 3.3. Další pravidla pro segmentaci | 44 |
| 3.4. /v/ a /h/ – „méně frikativní“ frikativy..... | 47 |
| 3.5. K segmentaci afrikát..... | 54 |
| 3.6. Shrnutí | 55 |
| 4. Intervokálníkové nazální konsonanty | 56 |
| 4.1. Uvedení do artikulačních a akustických vlastností ... | 56 |
| 4.2. Vlastní fonetické rysy a základní pravidla pro segmentaci..... | 57 |
| 4.3. Hranice v sekvenci vokál-nazála | 59 |
| 4.4. Hranice v sekvenci nazála-vokál | 61 |
| 4.5. Shrnutí | 66 |
| 5. Intervokálníkové vibranty | 67 |
| 5.1. Uvedení do artikulačních a akustických vlastností ... | 67 |
| 5.2. Vlastní fonetické rysy a základní pravidla pro segmentaci..... | 68 |

| | |
|--|------------|
| 5.2.1. Segmentace podle kmitu jazyka | 70 |
| 5.2.2. Rozšířená segmentace | 71 |
| 5.3. Další pravidla pro segmentaci | 72 |
| 5.4. Česká frikativní vibranta ř. | 75 |
| 5.5. Shrnutí | 77 |
| 6. Intervokálníkové aproximanty | 79 |
| 6.1. Uvedení do artikulačních a akustických vlastností . . | 79 |
| 6.2. Vlastní fonetické rysy a základní pravidla pro segmentaci. | 80 |
| 6.2.1. Akustický přístup | 80 |
| 6.2.2. Percepční přístup | 82 |
| 6.3. Další pravidla pro segmentaci | 84 |
| 6.4. Shrnutí | 90 |
| 7. Intervokálníkové laterální alveolární aproximanta | 92 |
| 7.1. Uvedení do artikulačních a akustických vlastností . . | 92 |
| 7.2. Vlastní fonetické rysy a základní pravidla pro segmentaci. | 93 |
| 7.3. Další pravidla pro segmentaci | 95 |
| 7.4. Shrnutí | 99 |
| 8. Obstruenty různého způsobu artikulace | 101 |
| 8.1. Uvedení do artikulačních a akustických vlastností . . | 101 |
| 8.2. Vlastní fonetické rysy a základní pravidla pro segmentaci. | 101 |
| 8.3. Další pravidla pro segmentaci | 104 |
| 8.4. Shrnutí | 107 |
| 9. Sekvence obstruent-likvida. | 108 |
| 9.1. Konsonantické skupiny s [l] | 108 |
| 9.2. Konsonantické skupiny s [r] | 113 |
| 9.3. Shrnutí | 114 |
| 10. Sekvence hlásek stejného způsobu artikulace | 115 |
| 10.1. Spojení dvou exploziv | 115 |
| 10.2. Spojení dvou frikativ | 119 |
| 10.3. Shrnutí | 123 |
| 11. Ráz před vokály na začátku slova. | 125 |
| 11.1. Ráz jako exploziva | 125 |

| | |
|---|------------|
| 11.2. Ráz jako třepená fonace | 128 |
| 11.3. Shrnutí | 131 |
| 12. Začátek a konec promluvy | 132 |
| 12.1. Iniciální hlásky | 132 |
| 12.2. Finální hlásky | 136 |
| 12.3. Shrnutí | 140 |
| 13. Závěr | 141 |
| 14. Použitá literatura | 144 |

Motto: Všechno má své meze,
alespoň do jisté míry.

1. Úvod

1.1. Proč potřebujeme hranice segmentů?

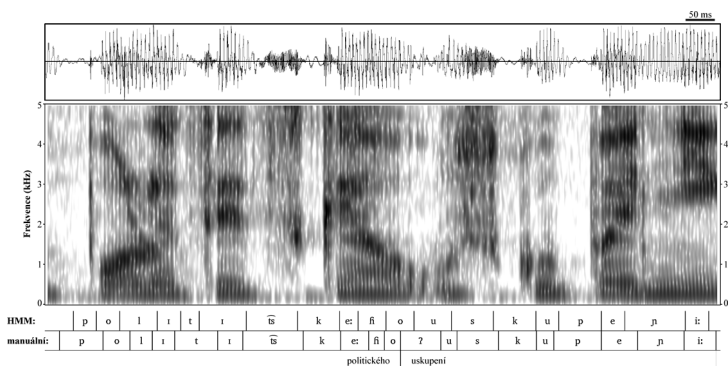
Konečným cílem veškerého fonetického bádání je porozumět struktuře řeči a jejím různým funkcím v komunikaci (Kohler, 2007). K odhalování této struktury se musíme pokusit najít smysluplný a obecně akceptovatelný způsob vymezení základních strukturních jednotek. Z hlediska praktického to znamená, že potřebujeme rozdělit akustický signál na diskrétní jednotky a spojit je s více či méně abstraktními fonetickými symboly. Rozměr takové jednotky obvykle závisí na povaze konkrétního výzkumného úkolu: můžeme mít zájem například o segmentaci hlásek, slov, přízvukových taktů, promluvových nebo nádechových úseků.

V této knize se zaměříme na segmentaci na úrovni hlásek. Bylo by možno namítnout (a s touto námitkou jsme se setkali), že znalost hranic segmentů není v řadě oblastí fonetického výzkumu potřebná. Je pravda, že některé specifické výzkumné úkoly pracují s jinými jednotkami nebo parametry. Domníváme se však, že znalost hranic segmentů je stále ještě nejuniverzálnější způsob jak přistupovat k řečovému materiálu. Anotace na úrovni jednotlivých segmentů jsou užitečné nejen pro výzkum vlastností řeči na úrovni segmentů (např. temporální vlastnosti, spektrální změny uvnitř hlásky), nýbrž také pro mnoho druhů úkolů souvisejících s takzvaným prozodickým výzkumem. Uvedme jen dva příklady: (1) ke zkoumání intonačních zákonitostí (nikoli pouze kontury F0) chceme znát temporální střed slabičného jádra; (2) výzkum rytmických vlastností jazyka se obvykle vztahuje k temporálnímu chování hlásek nebo jejich typů.

Je všeobecně známo, že jedna věta nebude nikdy vyslovena – viděno z objektivního, akustického hlediska – absolutně stejným způsobem. Bude se samozřejmě lišit, bude-li pronášena různými mluvčími, avšak ani týž mluvčí v témže komunikačním a sémantickém kontextu nepronese dvě zcela identické věty. Krátce řečeno, řeč je mimořádně variabilní fenomén. Smyslem fonetického bádání je najít ve variabilitě jistou stabilitu, invariantnost, protože pokud by jistý stupeň invariantnosti neexistoval, nemohla by řeč fungovat jako prostředek komunikace.

Invariantnost v řeči nelze odhalovat na základě zkoumání několika vět pronesených jedním mluvčím. K tomu je zapotřebí reprezentativní vzorek řečového materiálu, tedy rozsáhlý a strukturovaný korpus. Abychom mohli mluvit o skutečném **fonetickém korpusu**, musí být nahraná řeč zpracovávána jednotným způsobem. Pro naše účely toto zpracovávání neobnáší jen transkripci, nýbrž zejména segmentaci.

Ohraničování fonetických jednotek – ať už segmentů, nebo jiných – může probíhat dvěma způsoby, a to automaticky nebo manuálně. Bylo vyvinuto množství automatických nástrojů pro segmentaci, nejčastěji založených na HMM (např. Wester *et al.*, 2001; Kominek *et al.*, 2003; Pollák *et al.*, 2007). Bohužel dosud nejsou tyto metody pro fonetický výzkum dostatečně přesné a vyžadují manuální korekce. Segmentace vytvořená pomocí HMM a ručně opravená segmentace dvou slov je porovnána na obrázku 1.1 (příklad slouží jako ilustrace, rozdíly zde neanalyzujeme). Automatická segmentace pomocí HMM může být samozřejmě využívána pro hrubé naznačení hranic segmentů, avšak nikoli pro vyvozování lingvisticky interpretovatelných závěrů. To nás vede k přesvědčení, že pokud hodláme provádět skutečný fonetický výzkum, je lidský vstup při přípravě řečového korpusu nezastupitelný. V tomto případě to znamená manuální segmentaci.



Obr. 1.1. Srovnání segmentace generované pomocí HMM a ručně korigované segmentace dvou slov.

Jsme si pochopitelně vědomi nevýhod manuální segmentace. Jednak je časově náročná, takže příprava fonetického korpusu je věcí dlouhodobého úsilí, jednak je náročná na odborné znalosti anotátorů. Mnoho badatelů ji kritizuje jako od podstaty subjektivní, a proto nekonzistentní a nereprodukovatelnou (např. Wesenick & Kipp, 1996; Pitt *et al.*, 2005). Patrně každý, kdo si vyzkoušel ručně segmentovat nějaký úsek řeči, zaznamenal onu hořkou zkušenost, kdy se nebyl schopen rozhodnout kam hranici segmentu umístit. Častěji, než je nám libo, dospíváme k několika rozumným důvodům pro značně odlišné umístění hranice, případně nenalzáme žádná vodítka pro segmentaci. Nakonec učiníme rozhodnutí, a když se druhý den vrátíme ke stejnému problému, změním názor a hranici umístíme jinak. To znamená, že zásadní otázkou manuální segmentace je jak konzistentnost mezi různými anotátory, tak konzistentnost jednoho anotátora v čase.

Shoda mezi anotátory při manuální segmentaci byla podrobená zkoumání v různých studiích. Cosi *et al.* (1991, citováno v Pauws *et al.*, 1996) ve svém materiálu zjistili, že se přes 10 % hranic lišilo v umístění o více než 20 ms. Výsledky srovnání různých anotátorů ve studii autorů Pitt *et al.* (2005) hovoří

o průměrné odchylce v umístění hranic 16 ms, Wesenick & Kipp (1996) uvádějí odchylku přes 10 ms. Kvale & Foldvik (1991) segmentovali 748 hlásek na základě poměrně jednoduchých kritérií a zjistili, že 96,5 % hranic má odchylku nižší než 20 ms.

Před několika lety jsme se rozhodli, že se pokusíme nesrovnalosti mezi anotátory minimalizovat. Hodlali jsme zjistit, zda relativně jednoduchá pravidla, založená na (pokud možno) foneticky signifikantních událostech v akustickém kontinuu, mohou vést k vyšší shodě mezi anotátory. Formulovali jsme tedy pravidla pro určité kombinace hlásek: intervokálníkové explozivy, frikativy a nazály (Volín *et al.*, 2008). Průměrná odchylka mezi třemi anotátory se ukázala, jak zachycuje tabulka 1.1, jako podstatně nižší než ve výše zmíněné srovnatelné studii (Wesenick & Kipp, 1996).

| typ hranice | průměrná odchylka (ms) Wesenick & Kipp (1996) | průměrná odchylka (ms) Volín <i>et al.</i> (2008) |
|-----------------|--|--|
| vokál-exploziva | 12,0 | 1,8 |
| exploziva-vokál | 6,0 | 1,3 |
| vokál-frikativa | 8,0 | 3,0 |
| frikativa-vokál | 9,5 | 2,4 |
| vokál-nazála | 9,0 | 2,0 |
| nazála-vokál | 8,0 | 2,6 |

Tab. 1.1. Srovnání průměrné odchylky mezi anotátory ve studiích Wesenick & Kipp (1996) a Volín *et al.* (2008). Pro zjednodušení zde neuvádíme rozdíly mezi znělými a neznělými obstruenty.

Abychom mohli porovnat naše výsledky s těmi, k nimž dospěli autoři Cosi *et al.* (1991, citováno v Pauws *et al.*, 1996), uvádíme v tabulce 1.2 odchylky v umístění hranic podle stoupajícího rozsahu shody. Přestože se výsledky studie autorů Cosi *et al.*

patrně zakládají na kombinacích všech segmentů, je zřejmé, že pravidla pro segmentaci mohou rozdíly mezi anotátory výrazně snížit.

| rozsah shody | intervokalické explozivy | intervokalické frikativy | intervokalické nazály |
|--------------|--------------------------|--------------------------|-----------------------|
| = 0 ms | 53 % | 32 % | 43 % |
| < 3 ms | 82 % | 66 % | 74 % |
| < 6 ms | 96 % | 88 % | 91 % |
| < 9 ms | 98 % | 95 % | 96 % |
| < 15 ms | 99.4 % | 99 % | 98 % |

Tab. 1.2. Rozsah shody při segmentaci intervokalických exploziv, frikativ a nazál (podle studie Volín *et al.*, 2008).

Na základě těchto povzbuzujících údajů jsme se rozhodli formulovat podobná pravidla i pro segmentaci dalších hláskových kombinací a shromáždit je v této studii. Výsledek našeho snažení právě držíte ve svých rukou. Domníváme se, že existence pravidel umožní více lidem (včetně studentů) pracovat na vývoji fonetického korpusu tak, aby byl zaručen (alespoň do jisté míry) jednotný přístup k segmentaci. Tím se příprava korpusu urychlí, aniž by se snižovala spolehlivost segmentace. K reliabilitě našich výsledků z hlediska shody mezi anotátory se vracíme v závěrečné kapitole.

K pokusům o stanovení pravidel pro segmentaci docházelo již dříve, například tvůrci korpusu Buckeye publikovali online manuál pro segmentaci (Kiesling *et al.*, 2008). Jedná se o sadu psaných instrukcí, ovšem bez jakýchkoli ilustrací na základě spektrogramů nebo oscilogramů. Některé pokyny, podle našeho názoru, nejsou dostatečně podrobné. Naší snahou bylo specifikovat kritéria pro umísťování hranic co možná nejdůsledněji a doplňovat je vizuálními příklady.