

1. Úvod

1.1 Statistika v současném fonetickém výzkumu

Hledání rozumného, vyváženého přístupu k různým jevům reality bylo pro člověka vždy obtížné. Přitom je zřejmé, že vyhnout se extrémním stavům, např. stagnaci na jedné straně a chaosu na straně druhé, je žádoucí a že rozumný a vyvážený přístup se vyplácí. Tuto obecnou úvahu je možno velmi dobře uplatnit při hodnocení současného vztahu fonetiků ke statistice. Jakkoli je role statistických metod v lingvistickém výzkumu dnes již nezpochybnitelná, překvapivé množství fonetiků přistupuje ke statistice přinejmenším nepoučeně.

Z řady nepoučených přístupů vyberme dva typické příklady. Jeden se váže k názoru, že použití statistických metod je samo o sobě zárukou kvalitního výzkumu. Druhý se týká představy, že fonetika je specifickou lingvistickou disciplínou, která statistické metody nejenže nepotřebuje, ale měla by je odmítat jako zavádějící. První názor může být mimo jiné důsledkem logické chyby vzniklé vyvozováním. Je jisté, že pracuje-li fonetik s daty, provádí-li různá měření, sleduje-li reakce posluchačů na určité řečové podněty atd., nemůže svá pozorování zobecnit, pokud je statisticky neproověří. Aby se dozvěděl, zda jeho postřehy platí pouze pro daný vzorek materiálu, nebo mají platnost širší, nevystačí s pouhými procenty, aritmetickými průměry a tušením pravdy. Součástí kvalitního zpracování dat musí být třeba i jednoduchá statistická analýza. Problém je, že tato věta neplatí automaticky také obráceně. Představa, že

jakékoli použití statistické metody zaručuje kvalitní zpracování dat, je mylná. Samoučelně přidané nebo nesprávně vybrané statistické testy výzkum rozhodně nezvyšují.

Druhý extrém spojený s názorem, že statistika je pro fonetiku zbytečná, dnes již naštěstí ustupuje. Bývá obvykle nešťastným důsledkem toho, že statistické principy a postupy je obtížné vysvětlit jednoduše a názorně. Mnoho z těch, kteří k nim proto získají negativní postoj, pak namísto nových pokusů pochopit raději začne hledat důvody, proč je správné statistice se vyhýbat.

Samozřejmě existuje i celá řada dalších postojů, a je dokonce možno setkat se i s paradoxním přístupem, který je poučný, a přesto nesprávný. Jde o záměrné vyhýbání se statistice u jedinců, kteří vědí předem, že vzhledem k malému počtu měření, která provedli, a velkému rozptylu dat, který je ve fonetice poměrně častý, jsou jejich závěry bezcenné. Na to pochopitelně nechtějí nikoho upozorňovat.

Je ovšem nutno zmínit typ výzkumu, kde statistická analýza není potřebná. Jedná se o výzkum *kvalitativní*, tedy takový, který odhaluje, s čím vším je třeba při řešení problému počítat a které jevy a závislosti je možno s ohledem na předmět zájmu mapovat. Takový výzkum může sloužit např. k vytvoření nomenklatury problematiky a opírá se kromě důkladného studia související literatury také o detailní pozorování jednotlivin. Jeho cílem tedy není potvrzování hypotéz, ale jejich hledání. Závěry se pak nezobecňují, ale dále se prověřují ve výzkumu *kvantitativním*, který ukáže, jak spolehlivé jsou výsledky předpokládaných jevů, jakou míru mají jejich skutečné vlastnosti, jak pevné jsou vazby mezi nimi apod.

Spíše než plané úvahy o tom, který typ výzkumu je důležitější a významnější, je dobré se soustředit na to, jak se oba typy vzájemně doplňují, aby přispěly k efektivnějšímu bádání a spolehlivějšímu poznání reality. Obzvláště nepřijatelné by měly být snahy zneužívat nálepek kvantitativní nebo kvalitativní jako omluvu za badatelskou neschopnost či lenost. Po-

kud někdo jen náhodně proměřuje určité jevy bez předem formulované hypotézy nebo výzkumného plánu, pak nedělá nic užitečnějšího než jiný člověk, který se na základě rozboru jednoho dialogu pokouší postavit obecný popis nějakého řečového nebo jazykového jevu. V obou případech se jedná o výzkum nedostatečný a termíny kvantitativní ani kvalitativní mu lesku nedodají.

Jak kvantitativní, tak kvalitativní složka bádání se vzájemně doplňují a jsou nezastupitelné. Samozřejmě je možné, a dokonce poměrně obvyklé, věnovat se pouze jedné z nich, ovšem pak je nutno být členem nějakého buď reálného, nebo alespoň virtuálního týmu a mít jasno v tom, kdo doplní chybějící část výzkumu.

Volání, že hledat rozumný přístup ke statistickým metodám je naléhavým úkolem dneška, by mohlo vést k ukvapenému závěru, že se jedná o problém moderní. Ve skutečnosti je možno již ve 30. letech minulého století nalézt stopy velmi živé debaty na toto téma. Lze nalézt i velmi explicitní výzvy, aby se lingvisté, kteří dokládají svá tvrzení reálným materiálem, povinně zajímali o statistickou průkaznost svých dat (Zwirner a Zwirner, 1936, v Kohler, 2007). Citování autoři však jedním dechem také varují před samoúčelným proměřováním fyzikálních vlastností řečového signálu a opájením se tabulkami a numerickými údaji. Dokud budou badatelé např. vydávat měření kontury F0 v hertzech za „měření intonace“ a dokud nepochopí složitost vztahu mezi akustickou intenzitou řeči a vnímanou hlasitostí, pak ani statisticky významné nálezy nemohou mít žádný význam lingvistický. Budou naopak nebezpečné, neboť povedou k falešnému sebeuspokojení a často i k předstírání odbornosti tam, kde prostředky na poctivou lingvistiku odčerpává diletant či nabubřelec.

Aby se statistika stala účinnějším nástrojem lingvistického výzkumu, musí být lépe chápána. K tomuto poměrně triviálnímu postřehu je však nutno dodat, že aby byla statistika lingvisty lépe chápána, musí být také lépe prezentována, což se

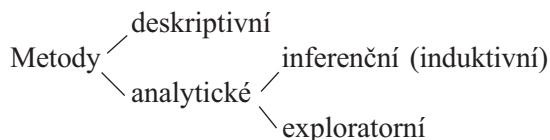
velmi často nedaří. Člověk, který je nadán pro práci s jazykem, nemusí být nutně nadán pro operace ve sféře matematické symboliky. Kognitivní psychologie, která se kromě řady jiných věcí zabývá také typologií představivosti a způsobů myšlení, rozlišuje mezi dovednostmi jazykovými a matematickými a nachází velmi často případy, kdy jednotlivci vykazují značnou převahu jedněch dovedností nad druhými. Lidé s různými typy představivosti a způsoby myšlení pak mnohdy jen obtížně komunikují o složitějších problémech. Není proto výjimkou, když se lingvista statistice instinktivně vyhýbá jako něčemu cizímu.

Určitou úlevu přinášejí dnešní výkonné počítače a statistické programy, které odbourávají úmornou dřinu s výpočty, tak typickou pro dřívější statistickou práci. Je to z valné části díky nim, že lingvisté používají statistické metody se stále větším sebevědomím. Rychlé procesory a pohodlné programy však přinášejí určité nebezpečí. I přes svou zdánlivou složitost a často ohromující ceny jsou stále jen nástroji, jejichž využití je limitováno poučeností uživatele. Samy o sobě zdaleka neumí vše, a především neumí za uživatele uvážit, zda má kvalitní data a jaké otázky si na základě svých dat může dovolit klást. Nezkušený badatel pak často spoléhá na dokonalost svého statistického softwaru a sbírá data, kterým příliš nerozumí. Přehlédnutí důležitých vztahů nebo jevů může ovšem výzkum zcela znehodnotit. Nepřehledná prezentace čehosi, co počítačový program vyprodukoval, pak možná udělá dojem na začátečníky, avšak u zkušených kolegů vzbudí nedůvěru.

Statistické metody je proto třeba studovat a díky studiu i pochopit. Je to jedna ze spolehlivých cest ke zkvalitnění výzkumu. Systematičtější uvažování nad daty a jejich zpracováním pak může přiblížit objevy, které se jinak naplánovat nedaří.

1.2 K obsahu a poslání knihy

Statistické postupy, které nacházíme ve fonetické literatuře, je možno typologizovat podle různých hledisek. Následující schéma přibližuje rozlišení hlavních typů podle účelu.



Takzvané deskriptivní statistice je věnována 3. kapitola této knihy. Při statistické deskripci jde především o úsporné, přehledné vyjádření důležitých údajů o datových souborech. Její postupy pomáhají např. odhalit typické trendy v datech a atypické případy či skupiny případů. Specificky fonetickým problémem je nutnost rozlišovat, kdy je deskriptivně statistický údaj použit tradičním způsobem, tedy k popisu sledovaného datového souboru, a kdy je použit jako parametr nějaké fonetické jednotky (např. spektrální momenty aperiodických hlásek, intonační rozpětí u promluvových úseků nebo směrodatná odchylka průměrného trvání vokalického intervalu pro větu jako rytmickou jednotku). Kromě 3. kapitoly se však deskriptivní statistice věnujeme spíše okrajově. Není to proto, že by snad byla méně významná – v praxi se s ní setkáváme ze všeho nejčastěji. Její postupy a koncepty jsou však natolik známé a pochopitelné, že jim v této monografii není třeba věnovat mnoho místa.

Je tedy zřejmé, že se zaměřujeme především na statistiku analytickou. Kapitoly 5 až 8 jsou věnovány statistice inferenční, která zahrnuje metody určené k takzvanému testování hypotéz, tedy vyčíslování statistické významnosti vztahů nalezených ve vzorcích dat. V souvislosti s touto problematikou se v praxi opět setkáváme s některými nesprávnými přístupy.

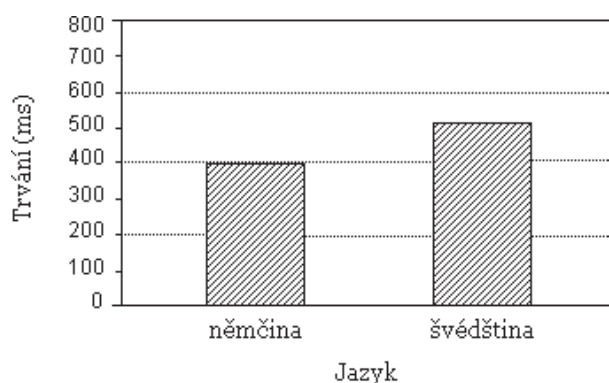
Jak upozorňuje myslitel a pedagog Edward de Bono (a před ním mnoho jiných), jednou z nejčastějších chyb v lidském myšlení je unáhlené zobecňování (de Bono, 1991). Lidé často dělají závěry, aniž by je měli podloženy dostatečným množstvím pozorování. Vezměme si hypotetický příklad experimentu, v němž se zjišťuje, zda posluchači poznají, která exploziva byla editací zvuku odstraněna ze slabiky [pe] nebo [ke]. Jestliže 12 ze 16 posluchačů uvede správně, o kterou explozivu šlo, mohlo by se zdát, že výsledek je jasný. Inferenční statistika však dokáže vyčíslit, že pokud budou posluchači hádat jen tak nazařbůh, je téměř osmiprocentní pravděpodobnost, že při jednom pokusu se 16 náhodně vybranými posluchači jich 12 bude hádat stejně. Podle zavedených konvencí tento výsledek není statisticky významný – jako hranice ještě přijatelné pravděpodobnosti náhodného výsledku bylo stanoveno 5 %. Odtud známých $p < 0,05$, používaných obvykle v sociálních, a tedy i ve fonetických vědách.

Podobně nevyhověl kritériu statistické významnosti rozdíl nalezený Botinisem, Gawronskou, Bannertem a Sandblomem (2003, s. 1307), kteří měřili trvání pauz ve čtených rozhlasových zprávách a porovnávali situaci v němčině a švédštině. Část výsledků je zachycena na obr. 1.1. Průměrné trvání pauz v německém materiálu bylo zhruba 400 ms, zatímco ve švédském něco málo přes 500 ms (přesný údaj autoři neuvádějí). Tento rozdíl byl však vzhledem k masivnímu rozptylu dat a k velikosti vzorku shledán nevýznamným.

Podmínky je samozřejmě možné ještě zpřísnit a stanovit, že se nebude zobecňovat, pokud pravděpodobnost náhodného výsledku p nebude menší než 0,01, nebo dokonce 0,001. Inferenční statistika tedy zjišťuje, zda je vzhledem k počtu případů, které byly prozkoumány, možno činit závěry obecnější platnosti. Základním výstupem jejích metod je údaj o pravděpodobnosti omylu při zobecňování.

Protipólem ignorace metod inferenční statistiky je jejich povýšení na absolutní důkaz. Tento přístup je ovšem také ne-

správný. Představa, že něco může být vědecky prokázáno, je ošidná a většinou vysloveně laická. Věda jakožto systematické poznávání světa dochází pouze k více nebo méně pravděpodobným závěrům. Statistická významnost znamená jediné to, že bylo dosaženo určité arbitrárně stanovené hranice pravděpodobnosti náhody pro danou výzkumnou situaci. Metody inferenční statistiky neříkají nic o logice navrženého výzkumu, o bezchybnosti provedeného sběru informací nebo o rozumnosti formulované hypotézy.



Obř. 1.1 Průměrné trvání pauz ve čtených rozhlasových zprávách. Srovnání německého a švédského vzorku; měřeno v milisekundách. (Volně dle Botinise et al., 2003.)

V souvislosti s jejich užíváním je proto potřeba neustále zdůrazňovat, že cenou za pomíjivé chvilkové sebeuspokojení nad „objevem“ v nepořádném výzkumu je diskreditace vědního oboru. To je u humanitních věd poměrně reálné nebezpečí a souvisí nakonec i s prostředky, které je společnost ochotna na humanitní výzkum vydávat!

(Záměrnou formulací paradoxu upozorňuje na nezbytnost používání metod inferenční statistiky v kvantitativním výzkumu Colin Robson [1973]. Říká, že bez statistické analýzy je možno se obejít v případech, kdy má výzkumník naprostou

jistotu zcela jednoznačných vztahů mezi proměnnými. U takovýchto jednoznačných vztahů však stěží budou formulovány nějaké výzkumné úkoly.)

Třetí skupinou metod, jimž se v této knize věnujeme, jsou analytické metody exploratorní. Naleznete je od 9. kapitoly dále. Jedná se o složitější postupy pomáhající odhalovat skryté souvislosti a vzájemné vazby mezi jevy u méně přehledných souborů dat. Zde je třeba si uvědomit, že složitost metody nesouvisí nutně s její exaktností. Stejně jako komplikované, málo srozumitelné popisy nejsou vědecktější než ty dobře srozumitelné, nejsou ani složité metody nutně spolehlivější. Exploratorní metody umožňují hledat strukturu v datech tam, kde na první pohled není zjevná, avšak pouze zřídka dovolují jednoznačnou interpretaci výsledků. Jedním z poslání této knihy je tedy ukázat, co je podstatou těchto metod, a jak jsou nebo mají být používány ve fonetickém výzkumu, aby se uživatel vyhnul chybám.

Ty části knihy, které nebyly jmenovány, jako např. 2. a 4. kapitola, terminologický rejstřík a podobně, mohou být považovány za pomocné. Spíše než o úplnost šlo při jejich koncipování o účelnost. Představení základních pojmů a konceptů je tedy podřízeno základnímu tématu monografie, jímž je využití statistických metod v současném fonetickém výzkumu.

Při prezentaci jednotlivých metod jsou využity tři typy dat. Za prvé jde o materiál ze studií publikovaných v domácí i zahraniční literatuře. Za druhé jsou analyzována data původní, většinou dosud nepublikovaná, jejichž zdrojem je autorův vlastní výzkum. Nakonec jsou zde i data hypotetická, která jsou využívána především pro základní explikaci statistických principů. U všech dat je jednoznačně uveden jejich zdroj. Prameny neuvádíme pouze u příkladů špatně provedených studií, chybného využití metod a jiných negativních jevů. Zde jsme připraveni referenci vydat ve zvláště odůvodněných případech formou osobní korespondence. Důvodem je skutečnost, že chybující badatel si před vydáním této knihy mohl svou chybu

uvědomit a čestně ji deklarovat. Pak by si nezasloužil, aby jeho jméno bylo nadále s danou chybou spojováno.

Prameny dále neuvádíme u obecně známých vzorečků, které lze např. nalézt i ve středoškolských učebnicích matematiky. Tyto vzorečky jsou dnes již považovány za společný intelektuální majetek lidstva a není zvykem citovat materiál, v němž jsou k nalezení. V souvislosti se vzorci je také vhodné poznamenat, že v této monografii neslouží k dosazování a nejsou určeny k memorování. Jsou uváděny pouze tam, kde usnadňují pochopení principů rozebíraných metod. Proto také neotiskujeme varianty, které sloužily k usnadnění ručních výpočtů a z hlediska výkladu jsou irelevantní. Z důvodů úspory místa je oproti konvencím zjednodušován např. symbol součtu. Tam, kde nemůže dojít k omylu, často píšeme pouze

$$\sum x_i \text{ namísto plného } \sum_{i=1}^n x_i .$$

Z fonetické literatury posledních desetiletí je jasné, že statistické metody již mají své pevné místo ve výzkumné praxi. Ať tato kniha přispěje k tomu, aby neustále narůstající míra jejich zapojení do výzkumu byla provázena také celkovým zkvalitněním jejich využívání.

2. Základní pojmy

2.1 Základní a výběrový soubor

Při fonetickém výzkumu často odhalujeme určité skutečnosti týkající se zvuku řeči v nějak omezeném řečovém materiálu a snažíme se je zobecnit pro popis zvukové stránky jazyka. Odhadujeme tak vlastně realitu na základě poznání nějaké její části. Tu množinu objektů, kterou chceme systematicky a obecně popsat, nazýváme *základním souborem* nebo též *populací*. Naproti tomu konkrétní objekty, které máme k dispozici a které zkoumáme přímo, označujeme jako *výběrový soubor* neboli *vzorek* populace.

Původní česká terminologie *základní soubor* – *výběrový soubor* je elegantnější a pro člověka s citem pro mateřštinu přijatelnější. Novější terminologie *populace* – *vzorek* je svázaná s anglickou tradicí a umožňuje snadnější přechod od domácí literatury k literatuře mezinárodní. V této knize budou uvedené termíny používány střídavě, abychom reflektovali rozmanitost situací, v nichž se ve fonetické literatuře můžeme se statistikou setkávat.

Pokud je použit termín *populace*, je třeba pamatovat, že se nevztahuje pouze na živočišnou říši. Označuje se jím množina jakýchkoli prvků, které jsou předmětem výzkumného zájmu. Může tedy znamenat množinu všech rodilých mluvčích českého jazyka, všech pětiletých dětí přistěhovalců v Rakousku, ale také množinu všech zjišťovacích otázek v ostravské češtině nebo množinu všech předělů mezi odstavci při četbě souvislého textu. Pro účely výzkumu si z této populace (základního

souboru) vybíráme vzorek (výběrový soubor) a samozřejmě usilujeme o to, aby zjištěné skutečnosti měly obecnější platnost.

Metodologie výzkumu se pravidly pro výběr vzorku zabývá dosti podrobně. V této souvislosti se často připomíná princip náhodnosti a princip reprezentativnosti. Oba principy jsou spolu úzce spjaty. Pokud má totiž výběrový soubor reprezentovat soubor základní, pak z něj musí být vybrán zcela náhodně. Každý prvek základního souboru musí mít při výběru stejnou šanci stát se součástí souboru výběrového. Pokud tedy má být zkoumána např. čeština v semispontánních dialogích, je nutno mít na paměti, že vzorek vysokoškolských studentů filologie z Karlovy univerzity v Praze nereprezentuje všechny uživatele češtiny, protože např. důchodce z Nového Jičína neměl možnost být do tohoto vzorku vybrán. Čím menší je náhodnost výběru, tím obezřetněji je třeba zobecňovat výsledky bádání. Náhodný vzorek studentů filologie z Karlovy univerzity umožňuje, striktně vzato, zobecňovat pouze pro základní soubor studentů filologie z Karlovy univerzity.

Samotné statistické metody počítají vždy s tím, že výběr byl proveden správně, a proti špatně provedeným výběrům nemají žádnou pojistku. Na druhou stranu se ovšem počítá s tzv. *výběrovou chybou*. To je v rámci inferenční statistiky velmi důležitý koncept. I při správně provedeném výběru se totiž různé parametry vzorku neshodují přesně s parametry populace. Náhodně vybraný a reprezentativní vzorek může při určité obvyklé variabilitě jevů vykazovat např. jinou průměrnou hodnotu sledované vlastnosti nebo jiný rozptyl hodnot, než jaký vyazuje základní soubor. Výběrová chyba je tedy rozdíl mezi určitým ukazatelem v základním souboru a v souboru výběrovém. Protože však parametry základního souboru obvykle nemůžeme pro jeho velikost zjistit, musíme výběrovou chybu odhadovat (viz dále oddíl 5.1).

Podívejme se nyní na jeden reálný případ, který demonstruje dosti rozšířený, avšak nesprávný postup nakládání s populací

a vzorkem. Jistí výzkumní pracovníci dostali za úkol popsat vývoj dětské řeči z akustického hlediska za účelem zkvalitnění jejího automatického rozpoznávání. Zaměřili se na čtyři věkové skupiny a každou z nich měli zastoupenou dvěma nebo třemi dětmi. Aby ve své výzkumné zprávě vůbec mohli porovnávat různé parametry a zobecňovat své nálezy, museli se vyrovnat s tím, že je velmi problematické, aby míra vlastnosti zjištěná u dvou subjektů nebyla ve srovnání s mírou vlastnosti u tří jiných subjektů označena statistickým testem jako dílo náhody. Tento problém se často obchází tak, že výzkumník nepovažuje za jednotlivý případ daného mluvčího, ale každou jeho promluvu. Je jasné, že má-li 90 promluv určitou míru nějaké vlastnosti a 60 jiných promluv míru odlišnou, pak určení nenáhodnosti tohoto rozdílu již není vyloučeno. Zmínění pracovníci však pozapomněli, že test významnosti rozdílu (v jejich případě analýza rozptylu – ANOVA) jim potvrdil pouze to, že budou-li nadále sbírat promluvy od těch několika dětí, které si vybrali, pak u nich budou s velkou pravděpodobností nacházet podobné rozdíly, jako u původního vzorku promluv. V žádném případě však tento rozdíl nemohou zobecňovat na celou populaci dětí daného věku. Vyberou-li si jinou dvojici či trojici dětí, mohou se výsledky úplně obrátit. Obecně tedy platí, že testy statistické významnosti rozdílu dovolují ze vzorku nějakých dětí usuzovat na populaci takovýchto dětí, a ze vzorku promluv dané dvojice či trojice dětí usuzovat na populaci promluv těchto dvou nebo tří dětí. Žádný test však neumožňuje zobecňovat ze vzorku určitých případů na populaci úplně něčeho jiného! Zcela konkrétně je pak jasné, že pokud byly vybrány dvě tříleté děti, u nichž bylo zjištěno průměrné F_0 255 Hz, a tři čtyřleté děti s průměrným F_0 285 Hz, pak test statistické významnosti tohoto rozdílu, v níž je jednotlivým případem promluva některého z dětí, zřejmě umožní předpoklad, že i ostatní promluvy oněch dvou tříletých dětí budou převážně hlubší než promluvy získané od dané trojice dětí čtyř-

letých. Zmínění výzkumníci však chybně formulovali závěry ve smyslu obecných rozdílů mezi tříletými a čtyřletými dětmi.

Podobným problémům se obvykle vyhnou ti badatelé, kteří si při plánování svého výzkumu dají práci s formulováním hypotéz. Ve statistice má zvláštní význam tzv. *nulová hypotéza*, která předpokládá, že různé vztahy mezi proměnnými nebo rozdíly mezi soubory jsou čistě náhodné. Statistický test pak zcela adresně nulovou hypotézu buď potvrdí, nebo nepotvrdí. Je-li nulová hypotéza pečlivě formulována, nebývá obtížné odvodit, jaká zobecnění si badatel může dovolit a jaká ne.

I přes maximální snahu o reprezentativnost výběrového souboru však nikdy nedosáhneme absolutní jistoty v tom, že námi nalezené vztahy platí i v základním souboru. Vždy půjde jen o odhad v mezích nějaké pravděpodobnosti. Aby badatel při své práci na tuto skutečnost nezapomínal, existují užitečné konvence ohledně zápisů hodnot, které rozdíl mezi výběrovým a základním souborem neustále připomínají. Hodnoty týkající se výběrového souboru (v.s.) mají být zapisovány malými písmeny latinské abecedy, zatímco jim odpovídající hodnoty základního souboru (z.s.) písmeny velkými anebo malými písmeny řecké abecedy. Tak např.:

n = počet případů ve výběrovém souboru	N = počet případů v základním souboru
\bar{x} = aritmetický průměr hodnot ve v.s.	μ = aritmetický průměr hodnot v z.s.
s = směrodatná odchylka od průměru ve v.s.	σ = směrodatná odchylka od průměru v z.s.

S ohledem na správné využívání statistických metod je ještě nutno krátce se zmínit o jednom aspektu výběrových souborů. Týká se vztahu porovnávaných skupin dat nebo měření. (Pojem měření používáme pro stanovování hodnot jednotlivých proměnných, tedy i pro sběr odpovědí v dotazníku.) Jak již vyplývá z povahy výzkumu, výběrový soubor má obvykle

poskytnout nějaké informace o jedné a téže skutečnosti, ale za změněných podmínek. Chceme znát např. hodnotu intonačního rozpětí u čtených a u spontánních projevů, míru koartikulační u mužů a u žen, trvání vokálu pod přízvukem a v nepřízvukné pozici nebo počet dysfluencí u střízlivých a u alkoholem intoxikovaných mluvčích. Výběrový soubor je pak nutno buď rozdělit na dvě skupiny, z nichž každá poskytne informace o jedné z podmínek, nebo se soubor nedělí a měření se provedou nejprve za jedné a později za druhé z podmínek. U první z možností se pak hovoří o *nezávislých měřeních* (také *nezávislých výběrech*), zatímco v druhém případě se mluví o *závislých měřeních* (také *opakovaných měřeních* nebo *spárovaných výběrech*). Do kategorie závislých měření patří i ty případy, kdy pracujeme např. s jednovaječnými dvojčaty nebo s dvojčeci respondentů vybranými tak, aby členové jednotlivých párů vykazovali shodnou míru nějakých relevantních vlastností. Z hlediska metod, jimiž se budeme zabývat, není rozdíl mezi opakováním a spárováním podstatný, zatímco rozdíl mezi nezávislými a závislými měřeními je nutno věnovat pozornost.

Příkladem nezávislých výběrů může být zjišťování artikulačních zvyklostí u profesionálních a neprofesionálních mluvčích. Volín a Skarnitzl zjišťovali, zda jsou asimilace před /v/ přes hranici slova stejně časté u obou skupin mluvčích (Volín a Skarnitzl, 2006). Je zřejmé, že výsledky v jedné skupině neměly žádný přímý vliv na výsledky v druhé skupině. Pokud by však byl tento problém asimilací zkoumán u osob před požitím a po požití alkoholu, jednalo by se o opakovaná měření a bylo by vhodné používat metody určené pro spárované výběry. O této problematice se ještě zmíníme později.

2.2 Typy proměnných

2.2.1 Proměnné podle matematických vlastností

Proměnnou v kontextu statistických metod rozumíme určitou vlastnost, která je předmětem našeho zájmu a již se sledované případy v souboru mohou navzájem lišit, neboť nabývá alespoň dvou různých hodnot. Tato vlastnost bývá v české statistice označována tradičně také jako *znak*. Nesprávné určení typu proměnných, s nimiž pracujeme, může dramaticky ovlivnit výsledky našeho bádání, neboť je v přímé souvislosti s výběrem metod statistického zpracování dat. Některé metody jsou vhodné pouze pro určité typy proměnných, a jsou-li použity pro typy jiné, mohou se stát zdrojem omylů.

Primárně je vhodné dělit proměnné podle toho, kolik informací je v nich z výpočetního hlediska obsaženo. Některé vlastnosti např. pouze odkazují na příslušnost ke skupině nebo na zařazení do kategorie. Takovéto vlastnosti se obvykle vyjadřují nějakým názvem, a proto se jim říká proměnné *nominální*. Určitý případ v našich datech tak může vykazovat např. znak žena, hlasatel, neprofesionální mluvčí, zadní středová samohláska, afrikáta, melodém ukončující klesavý, promluva s dysfluencí apod. Z výpočetního hlediska se u tohoto typu informace uplatňuje pouze rovnost nebo nerovnost. Daný případ buď afrikátou je, nebo není. Nominální proměnné je možno číselně zakódovat, tím se ovšem nijak nemění jejich povaha. Označíme-li si např. melodém ukončující klesavý číslicí 1, melodém ukončující stoupavý číslicí 2 a melodém neukončující stoupavý číslicí 3, neměli bychom první dva melodémy odečítat od třetího a spekulovat o tom, že rozdíl mezi nimi je 0. Stejně matoucí by byla i informace týkající se výběrového souboru padesáti mluvčích o tom, že jejich průměrné pohlaví je 0,5. Je lépe jasně specifikovat, že máme k dispozici 25 žen a 25 mužů a pro jednoduchost zpracování kódujeme jedno

z pohlaví číslicí 1 a druhé číslicí 0. Nominální proměnná tedy klasifikuje, tj. zařazuje ke kategoriím objektů.

K nominálním proměnným mají blízko *řadové proměnné*. Ty se sice také týkají zařazení do určitých kategorií, ale s tím, že zjištěné hodnoty je možno ještě seřadit podle kritéria „je větší než“ nebo „je menší než“. Řadové proměnné však také neumožňují kvantifikaci rozdílů mezi dosaženými hodnotami. Např. hodnocení artikulačních stylů podél osy *silně pokleslý – lehce nedbalý – neutrální – pečlivý – hyperartikulační* je příkladem využití řadové proměnné. Neumožňuje nám ovšem stanovit, o kolik je nedbalý artikulační styl horší než neutrální nebo pečlivý, neboť vzdálenost mezi jednotlivými stupni na ose není konstantní a exaktně měřitelná. Také při seřazení souhlásek podle sonority víme, že nazály jsou sonornější než frikativy, ale nemůžeme říci o kolik, protože sonorita není prozatím definována jako přesně měřitelná fyzikální veličina, ale pouze jako relativní vlastnost.

Ve fonetickém výzkumu často pracujeme také s vlastnostmi, které odkazují na pozici případu na určité škále exaktně měřitelných hodnot a namísto kódování číslicí je můžeme vyjádřit skutečným číslem udávajícím míru dané vlastnosti. To jsou proměnné *numerické*. Např. frekvence nějakého zvuku může být udána v hertzech (Hz), trvání pauzy v milisekundách (ms), intenzita ražené souhlásky v místě exploze v decibelech (dB). Numerické proměnné ovšem nemusí být vázány na striktně fyzikální veličiny. Také hodnoty hlasitosti v sonech (son), určení hladiny hlasitosti ve fonech (Ph) nebo výška formantu v erbech (ERB) či jiných percepčních jednotkách jsou numerickými proměnnými. Důležité je, aby hodnoty dané proměnné bylo možno měřit v nějakých jasně definovaných jednotkách.

U numerických proměnných je dále praktické uvažovat o dělení na *intervalové* a *poměrové*. Intervalové proměnné umožňují zjišťovat rozdíly mezi hodnotami, avšak ne jejich poměry, neboť mají arbitrárně stanovenou nulu. Např. ve frek-

venčním pásmu 50–500 Hz, v němž se pohybuje intonace lidské řeči, je výhodné pracovat s půltóny (ST), jak experimentálně prokázal Nolan (2003). Právě zde je však dobré si uvědomit, že určování půltónů počítá vždy s nějakým arbitrárně stanoveným počátkem. Široce rozšířený analyzační program Praat (viz např. Boersma a Weenink, 2005) přepočítává měření v hertzech na půltóny podle vzorce:

$$ST = \frac{12 \cdot LN (Hz / 100)}{LN (2)},$$

kde ST je hodnota v půltónech, LN je přirozený logaritmus (tj. takový, jehož základem je Eulerovo číslo) a značka Hz zastupuje převáděnou hodnotu v hertzech. Číslo 100 ve vzorci reprezentuje právě onu arbitrárně stanovenou nulu. Vzorec však funguje stejně spolehlivě, zadáme-li do něj místo čísla 100 třeba číslo 60. I v tomto případě bude převádět hodnoty v hertzech na hodnoty v půltónech, jen hodnota 0 ST se bude vztahovat k nižší frekvenci. Je tedy zřejmé, že naměříme-li za určitých podmínek pro jednoho mluvčího třeba výšku nějakého intonačního vrcholu v hodnotě 4 ST a pro jiného mluvčího v hodnotě 8 ST, nelze hovořit o tom, že druhý mluvčí dosáhl dvakrát větší výšky. V okamžiku, kdy stanovíme nulu jinak, se poměr mezi oběma výškami změní. Co však bude platit vždy, je skutečnost, že druhý mluvčí dosáhl hodnoty o čtyři půltóny vyšší (viz tab. 2.1).

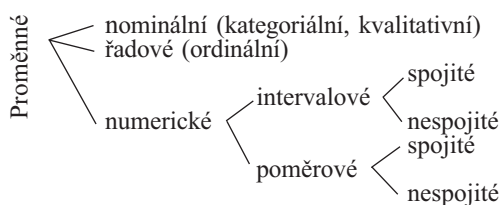
Počátek	Mluvčí 1	Mluvčí 2	Poměr	Rozdíl
0 ST = 100 Hz	4 ST	8 ST	1 : 2	4 ST
0 ST = 80 Hz	7,86 ST	11,86 ST	1 : 1,5	4 ST
0 ST = 60 Hz	12,84 ST	16,84 ST	1 : 1,3	4 ST

Tab. 2.1 Hypotetické měření výšky určitého intonačního vrcholu u dvou mluvčích na intervalové numerické škále. Při změně umístění počátku škály se mění poměr mezi měřeními, nikoli však jejich rozdíl (interval mezi nimi).

Problematika intervalové povahy pŕltónů je zejména pro intonology velice důležitá, a proto se jí ještě budeme zabývat. Nyní se však vraťme k typům proměnných a zmiňme se o takových, které se vztahují ke stupnicím s nějakou objektivně stanovitelnou absolutní nulou. Těmto proměnným se říká poměrové, neboť je u nich možno zjišťovat nejen o kolik, ale také kolikrát je jejich hodnota u jednoho případu vyšší než u druhého. Taková je situace třeba u trvání vyjádřeného v milisekundách. Máme-li například samohlásku o trvání 70 ms a jinou, která trvá 175 ms, můžeme říci nejen, že ta druhá je o 105 ms delší, ale také že je delší dva a půlkrát.

Již jen pro úplnost ještě dodejme, že numerické proměnné je možno také dělit na *spojité* a *nespojité*. Spojité proměnné mohou nabývat libovolných hodnot v rámci jim vymezeného intervalu. Takovými jsou např. artikulační rychlost, výška základního tónu, trvání segmentu nebo šířka formantového pásma. Nespojité proměnné nabývají pouze vybraných hodnot v rámci jim vymezeného intervalu. Patří mezi ně třeba počet slabik ve slově, počet taktů v promluvovém úseku, složitost slabičné kody udávaná v souhláskových segmentech, množství úspěšně rozpoznávaných slov v percepčním testu apod.

Obecně se ve statistice v souvislosti s nominálními a řadovými proměnnými mluví o neparametrických metodách a v souvislosti s numerickými proměnnými o metodách parametrických. Situace je ve skutečnosti poněkud složitější, jak bude patrné dále v souvislosti s problematikou tzv. normálního rozdělení. Zjednodušený přehled typů proměnných je uveden na obr. 2.1.



Obr. 2.1 Zjednodušené shrnutí typologie proměnných podle výpočetně relevantní informace v nich obsažených.

2.2.2 Proměnné podle role ve zkoumaných vztazích

Jiný úhel pohledu na proměnné nabízí otázka, jak vlastně tyto proměnné vlastnosti fungují v rámci reality, kterou zkoumáme. Zde se nejčastěji, i když ne výhradně, pracuje s termíny proměnná *nezávislá* a *závislá* (někdy též *nezávisle proměnná* a *závisle proměnná*). Tak jako je určení správného typu proměnných z hlediska jejich matematických vlastností důležité pro správný výběr statistické metody, je určení role proměnných ve zkoumaných vztazích zásadní pro správné dosazení hodnot do výpočetních vzorců, popřípadě správné zadání hodnot do statistického počítačového programu.

Termíny *závislá* a *nezávislá* proměnná se primárně vztahují k experimentálnímu výzkumu, kde vyšetřujeme nějaký vztah příčiny a následku. *Nezávislá* proměnná bývá v takových případech spojována s příčinou. Je to ta vlastnost, kterou experimentátor ovlivňuje, aby zjistil, dostaví-li se nějaký systematický následek, tj. bude-li se měnit *závislá* proměnná, neboli vlastnost, jejíž závislost na obměňované vlastnosti se má ověřit. Víme například, že řečový signál je akusticky poměrně bohatý a uplatňuje se v něm princip redundance neboli nadbytečnosti. Může nás proto zajímat, bez kterých frekvenčních pásem v řečovém signálu je možno se spíše obejít a která naopak nesou nejvíce informace. V experimentu tedy budeme testovat, jak se bude měnit index srozumitelnosti, tj. nějaký ukazatel správného rozpoznávání řeči (viz např. Ptáček, 1993, s. 19), v závislosti na různých polohách a šířkách odfiltrovaných pásem. Srozumitelnost řeči zde pro nás bude závislou proměnnou: bude záviset na poloze a šířce odfiltrovaného pásma. Poloha a šířka pásma, které můžeme měnit, jsou proměnnými *nezávislými*.

Z matematického hlediska by se závislost jedné proměnné na jiné dala zapsat následovně:

$$y = f(x)$$

Tento zápis říká, že vlastnost y je nějakou funkcí vlastnosti x , nebo že hodnota y závisí na hodnotě x , a pomocí funkce f můžeme z hodnoty x vypočítat hodnotu odpovídajícího y . Ve fonetickém výzkumu většinou nebývá funkce f přesně matematicky specifikována a odkazuje nás pouze do oblasti pravděpodobností. Často totiž pracujeme s proměnnými, které nevstupují do dostatečně uchopitelných číselných vztahů. Například výška prvních dvou samohláskových formantů může mimo jiné souviset i s geografickým původem mluvčího. Tak třeba lidé v některých regionech České republiky používají zavřenější [i] a [e], jinde jsou tyto hlásky otevřenější. Zavřenost se projevuje nižšími hodnotami prvního formantu a u samohlásek přední řady i vyššími hodnotami druhého formantu. Nebylo by však zřejmě nikterak výhodné hledat matematickou závislost hodnot F1 a F2 přímo na zeměpisné šířce a délce, i když matematický aparát tuto možnost zajiště poskytuje. Daleko obvyklejší je uvažovat o tom, že obyvatelé Hodonínska budou mít s velkou pravděpodobností u předních samohlásek nižší F1 než obyvatelé Prahy. Pro potřeby ladění artikulačního řečového syntetizéru se pak třeba můžeme snažit konkrétní rozdíly blíže specifikovat.

Příklad regionálních výslovnostních variant naznačuje jednu z možných nesnází při dělení proměnných na nezávislé a závislé. Fonetický výzkum je velmi často neexperimentální. Může být založen na prostých pozorováních, kdy žádnou z proměnných úmyslně nemanipulujeme. V řadě případů je pak vhodnější mluvit o korelačním výzkumu a o souvislostech, spíše než závislostech, neboť tvrzení, že zeměpisná šířka a délka je příčinou jiného artikulačního nastavení mluvidel, je přinejmenším zkratkovité. (Více o tom v 8. kapitole věnované korelacím.)

Zkušený badatel ví, že termíny závislá a nezávislá proměnná jsou formální a není vždy vhodné brát je doslova. (Snad i proto je v některých typech úloh nezávislá proměnná nazývána skupinovou/grupovací proměnnou, faktorem nebo expli-

kační/explanatorní proměnnou.) Je však užitečné vědět, že data, která zachycujeme a která jsou primárním předmětem našeho zájmu, se nazývají závislou proměnnou, zatímco vlastnosti, které by mohly mít nějakou zajímavou souvislost s variabilitou oněch dat, se nazývají proměnnými nezávislými. Odhalení skutečných kauzálních vztahů pak překračuje rámec vlastního statistického zpracování výsledků. Pokud nám při pohledu na sluchové pole člověka postačí závěr, že poloha hladiny hlasitosti pro nějaký zvuk závisí na jeho výšce, je to v pořádku. Nesmíme však zapomínat, že sama výška není příčinou tohoto jevu. Tu musíme hledat ve fyziologických vlastnostech sluchového aparátu člověka.

Z uvedeného také vyplývá, že stejná veličina může být jednou nezávislou a podruhé závislou proměnnou podle toho, jak je zadán výzkumný úkol. Vezměme třeba finální intonační minimum, tj. výšku zakončení klesavé kadence nějaké prozodické jednotky. Odhadujeme ji např. z F0 posledních znělých vzorků řečového signálu. Ve výzkumu Sluijterové a Terkena (1993) šlo mimo jiné o to, zda se hodnoty tohoto finálního intonačního minima mění podle pozice v odstavci. V tomto výzkumu je tedy finální intonační minimum závislou proměnnou a pozice v odstavci proměnnou nezávislou. Budeme-li však zjišťovat, zda pocit ukončenosti promluvy závisí na výšce finálního intonačního minima jako Cohen a jeho spolupracovníci (Cohen, Collier a t'Hart, 1982, s. 264), pak musíme manipulovat minimem jako nezávislou proměnnou a zaznamenávat data o pocitu ukončenosti jakožto hodnoty závislé proměnné.

Zvláštním druhem nezávislých proměnných, které také ovlivňují námi sledovaný jev, jsou proměnné vnější a proměnné rušivé neboli náhodné. Zamysleme se nad nimi v následujícím oddílu v souvislosti s problematikou chyb.

2.3 Typy chyb

Působení jiných proměnných než plánovaných není při výzkumu žádoucí. Přesto je jasné, že např. nastavení experimentálních podmínek tak, aby se v nich uplatňovaly pouze sledované nezávislé a závislé proměnné, je velmi obtížné a v humanitních vědách často zcela nemožné. Složitost sledovaných jevů s sebou přináší i spolupůsobení velkého množství dalších skutečností, které nějakým způsobem ovlivňují námi sledovanou závislou proměnnou. Vzhledem k typu chyb, které do výzkumu zanášejí, je dělíme na *vnější* a *rušivé*. Vnější proměnné jsou vlastně přehlédnuté či opomenuté nezávislé proměnné, které definovatelně fungují ve zkoumaném vztahu. Rušivé proměnné jsou různé náhodné vlivy, které často neumíme ani postihnout a jejichž působení je nejednotné. Z hlediska řešení výzkumného úkolu pak vnější proměnné způsobují tzv. *systematické chyby*, tj. takové, které posouvají výsledek určitým směrem od výsledku správného. Rušivé proměnné způsobují tzv. *náhodné chyby*. Působí totiž všemi směry a jejich vliv se proto nescítá. Většinou ovšem zvyšují variabilitu dat, takže vlastně správný výsledek zamlžují.

Četné nepatrné vlivy, které neumíme zachytit myšlenkou, pojmenováváme jako náhodu. Protože je samo chování takové náhody kvantifikovatelné, metody indukční statistiky nám pomáhají se s nimi do určité míry vypořádat. To ovšem neznamená, že bychom se neměli snažit je poznat a popřípadě je při výzkumu usměrňovat nebo eliminovat. Mnohem důležitější však je zabývat se možnými systematickými chybami, které mohou výsledek znehodnotit. Jestliže je např. značkování korpusu zadáno pracovníkovi, který chápe hranice likvid a vokálů jinak, než jak s tím počítá logika daného výzkumu, pak je zřejmé, že výsledky budou nějakým směrem posunuty. Podobný účinek má také již výše zmíněná nesprávná metodologie výběru vzorku. Pokud např. vyzvu skupinu lidí, aby poskytla dobrovolníky za účelem testování percepce intonačních jevů,

pak je dost pravděpodobné, že se přihlásí pouze ti, kteří mají pro vnímání intonace určité nadání. Výsledek je v důsledku toho systematicky ovlivněn.

Z oblasti percepčních experimentů vůbec nejspíše pocházejí ty nejmarkantnější případy systematických chyb. Při přípravě percepčních podnětů se např. někdy zapomíná na frekvenci výskytu dané jednotky. Již téměř legendárním se stal výzkum jisté skupinky badatelů, kteří chtěli mapovat problémy japonských studentů angličtiny s rozlišováním bokové aproximanty [l] a postalveolární aproximanty [ɭ]. U páru podnětů *look* – *rook* jim vycházelo jednoznačné upřednostňování [l] na úkor [ɭ]. Tento výsledek však neodrážel vlastnosti samotných hlásek. Většina studentů totiž slovo *rook* vůbec neznala, a proto se jim při vlastním experimentu percepčně aktivovalo pomaleji než slovo *look*, což v situaci binárního rozhodování mezi *rook* a *look* vedlo k jeho potlačení.

Systematické chyby mohou vstoupit do hry také při zadávání experimentu. Jestliže např. osobnost experimentátora respondenty demotivuje a vzbuzuje neochotu podílet se na výzkumu, pak se může stát, že výsledek bude mylně odkazovat na neschopnost posluchačů rozpoznávat určité kontrasty.

Jiným příkladem systematické chyby je tzv. *efekt pořadí*. Je známo, že vnímání řeči je vázáno na kontextové kontrasty. V percepčním testu, kde pracujeme s izolovanými řečovými jednotkami, se však jedna položka stává kontextem druhé. Vnímání každé následující může být ovlivněno tou předchozí. Pokud test nezadáme v několika verzích s náhodným pořadím nebo nerušíme vliv předchozí položky nějakým jiným, neutrálním zvukem, systematická chyba se snadno objeví.

Efekt pořadí se ovšem může objevit i v produkčních výkonech. Představme si třeba Myersův experiment, v němž se srovnávaly intonační charakteristiky řeči při různé hlasitosti (Myers, 1999). Respondenti měli za úkol číst určitý text nejprve nahlas, poté normálně a nakonec potichu. Kromě osobních představ respondentů o kvalitách nahlas, normálně a potichu, které

můžeme označit jako rušivé proměnné, zde zůstává otázka, zda právě pořadí úkolů nějak systematicky neovlivnilo výkon mluvčích. Budeme-li po hlasitém čtení bezprostředně vyžadovat tzv. čtení normální, pak je možné, že kvůli okamžitému kontrastu normální nebude. Myers však zajisté s tímto problémem počítal a nepožadoval, aby jeho mluvčí produkovali jednotlivé verze textu okamžitě jednu za druhou. Navíc, jak popisuje, důkladně instruoval své pokusné osoby, aby při hlasitém čtení četly tak, jako by osoba, ke které promlouvají, byla ve vedlejší místnosti. Poté měly stejné věty číst normálně, jako by mluvily k někomu v téže místnosti, kdo sedí naproti u stolu. Nakonec četly věty co nejtíšeji, jako by mluvily k člověku sedícímu přímo vedle nich, ale bez šepotu.

Literatura věnovaná plánování experimentu přináší různé metody, jak se systematickým chybám vyhýbat. Nejobtížnější je samozřejmě riziko systematické chyby vůbec postřehnout. Pro správné modelování zkoumaného vztahu je pak důležité systematickou chybu buď ze vztahu vyřadit, nebo ji převést na chybu náhodnou. Je také možné zkoumat její zdroj jako další nezávislou proměnnou nebo ji zavést jako konstantu (např. sjednocením jak v experimentálních, tak v kontrolních podmínkách).

Jiný pohled na chyby při statistickém zpracování se vztahuje k nakládání s hypotézami a mírou nejistoty ohledně zobecnování výsledků, která je přítomna i v případě sebelépe připraveného a sebepečlivěji provedeného výzkumu. Zkušené badatelé vědí, jak důležité je formulovat hypotézy o vztazích, které chceme prozkoumat. Jak jsme již uvedli výše a jak budeme i dále opakovaně demonstrovat, zvláštní místo patří tzv. *hypotéze nulové*, která říká, že hledaný vztah neexistuje a že rozdíly, které by ho snad mohly indikovat, způsobila pouze náhoda. Proti nulové hypotéze pak stojí nějaká *hypotéza pracovní* neboli *alternativní*. Ta hledaný vztah předpovídá. Pokud je možné, že by hledaný vztah měl více podob, pak je také více pracovních hypotéz. Volín (2006) kupříkladu zkoumal gravi-

tační trendy v intonaci britských rozhlasových hlasatelů. Jedna z nulových hypotéz, s nimiž pracoval, říkala, že vytvoří-li se regresní přímky intonačních trendů pro všechny věty, pak některé budou klesat, jiné stoupat, nebo budou víceméně rovné, ale žádný z těchto trendů nebude v převaze. Z odpovídajících alternativních hypotéz pak byla nejdůležitější ta, která počítala s převahou klesajících (deklinujících) trendů.

Statistické testy mají umožňovat rozhodování o zamítnutí hypotéz. V této souvislosti se hovoří o *chybě 1. druhu* a *chybě 2. druhu*. Pokud je nesprávně zamítnuta nulová hypotéza a badatel tvrdí, že odhalil vztah, který ve skutečnosti neexistuje (tj. je jen nespolehlivou souhrou náhod), pak se jedná o chybu prvního druhu. Pokud by došlo k chybě opačné, tedy nulová hypotéza by zamítnuta nebyla a badatel by nabyl přesvědčení, že hledaný vztah neexistuje, i když by se ve skutečnosti pouze neprojevil v jeho datech, jednalo by se o chybu druhého druhu.

Při statistických analýzách je možné nastavovat testy tak, aby se snížilo riziko buď té, či oné chyby. Z povahy věci však vyplývá, že snižuje-li se riziko jedné, zvyšuje se riziko té druhé. Zde je pak nutno vzít v úvahu praktické důsledky obou možných omylů a rozhodnout se pro ten, který způsobí menší škody. Dejme tomu, že chceme prozkoumat určitou logopedickou vadu z hlediska jejího negativního dopadu na hodnocení mluvčího jeho okolím. Nulová hypotéza bude formulována v tom smyslu, že daná logopedická vada nemá žádný vliv na sympatie nebo antipatie okolí vůči nositeli této vady. Chyba 1. druhu by znamenala mylný závěr, že vada mluvčího znevýhodňuje, tzn. zbytečně by byly vynakládány prostředky na odstraňování něčeho, co ve skutečnosti nijak neškodí. Chyba 2. druhu by vedla k ignorování skutečnosti, která negativně ovlivňuje kvalitu života mluvčího. Je zřejmé, že ekonom, který se zabývá hledáním úsporných opatření, se bude chtít vyhnout spíše chybě 1. druhu. Naproti tomu člověk s humanis-

tickým citěním, který se zabývá harmonií v mezilidských vztazích, bude považovat chybu 2. druhu za škodlivější.

Ještě zřetelnější je tento problém ve forenzní fonetice, která se uplatňuje při vyšetřování kriminálních činů nebo závažných nehod. Jedním z nejčastějších úkolů forenzního fonetika je identifikace mluvčího z nějakých zvukových záznamů. Na základě podrobných analýz nahrávek musí rozhodovat, zda je hlas pachatele totožný s hlasem podezřelého. Nulová hypotéza bude tvrdit, že veškeré rozdíly mezi oběma hlasy jsou způsobeny pouze náhodnými vlivy, a že tedy pocházejí z jednoho zdroje. Chyba 1. druhu by znamenala, že nulová hypotéza bude zamítnuta, i když by neměla být. Jinými slovy dojde k chybnému odmítnutí shody a podezřelý i pachatel jsou mylně považováni za dvě různé osoby. Při chybě 2. druhu by naopak došlo k chybné identifikaci: podezřelý i neznámý by byli mylně považováni za tutéž osobu. Právní systém civilizovaných zemí v těchto případech považuje chybu 2. druhu za závažnější, neboť vede k potrestání nevinného.

Riziko chybných závěrů je možno snížit pouze pečlivou, trpělivou prací, důkladným výběrem vzorků dostatečné velikosti, opakováním experimentů a měření a především souběhem výsledků vícera různých šetření.

2.4 Pravděpodobnost

Koncept pravděpodobnosti je ve statistice konceptem klíčovým. Matematické postižení pravděpodobnosti, nebo z jiného úhlu pohledu matematické postižení náhody, zaměstnává úvahy lidí odedávna. V okamžiku, kdy se setkáváme s nějakými proměnlivými výsledky naší činnosti, zajímá nás, nakolik je tato proměnlivost u dalších výsledků téže činnosti předpovědatelná. Dejme tomu, že se zaměříme na nízkou samohlásku [a] v češtině a za účelem modelování chceme zjistit její akustické vlastnosti. Už v projevu jediného mluvčího zjistíme, že

akustické vlastnosti jednotlivých realizací samohlásky [a] se od sebe překvapivě liší. Některé příčiny této variability jsme schopni postihnout, řada drobných, různými směry působících vlivů však zůstane nepostižena. Každopádně zjistíme, že to, čemu říkáme samohláska [a], není z akustického hlediska *jev jistý* (tj. takový, který je možno jednoznačně popsat a spoléhat se, že budeme-li ho chtít vyvolat, vždy dosáhneme přesně specifikovaného výsledku), ale *jev náhodný* (tj. takový, jehož vlastnosti můžeme jen více nebo méně přesně předpovídat, ale tyto vlastnosti se budou pouze s určitou pravděpodobností vyskytovat v nějak vymezeném pásmu možností).

Pro statistiku bylo objevem zásadní důležitosti zjištění, že fungování náhody se přece jenom řídí určitými principy a tyto principy je možno konceptualizovat natolik, že jsou prakticky využitelné. Disciplína, která se těmito principy zabývá, se nazývá *stochastika*, někdy též metonymicky pravděpodobnostní počet (anebo přesněji *nauka o počtu pravděpodobnosti*).

Počet pravděpodobnosti je se statistikou natolik spjat, že se občas setkáváme i s názory, že jde vlastně o dvě stránky téhož (Swoboda, 1977, s. 29). Pro potřeby našeho výkladu není nutno čistě teoretickou otázku vztahu statistiky a pravděpodobnostního počtu řešit. Je však několik aspektů pravděpodobnostní problematiky, které je užitečné zmínit.

Prvním z nich je význam malého p udávaného při prezentaci výsledků některých základních statistických testů. Pokud není uvedeno jinak, tento symbol odkazuje na pravděpodobnost chyby 1. druhu. Signalizuje vyčíslení rizika omylu ve smyslu nesprávného zamítnutí nulové hypotézy. Představme si např. problém vztahu spektrálního profilu hlásek a pozice v metrické struktuře promluvy. Ve fonetickém výzkumu byla již několikrát nadnesena myšlenka, že různé hlásky mohou vykazovat různý poměr převahy energie v různých pásmech svého spektra v závislosti na prozodických poměrech, v nichž se právě nalézají (viz též Sluijterová a van Heuven, 1996). Určitá

sonora by třeba mohla mít stejnou polohu formantů jak v akcentované, tak v neakcentované formě, avšak lišila by se intenzitními poměry v jednotlivých pásmech spektra. Nulová hypotéza by byla formulována v tom smyslu, že veškeré rozdíly mezi akcentovanými a neakcentovanými sonorami z hlediska spektrálního profilu jsou náhodné. Provedená měření by však rozdíly mezi dvěma skupinami sonor zachytila a test statistické významnosti těchto rozdílů by pro ně našel dejme tomu $p = 0,038$. Tomuto výsledku je třeba rozumět tak, že existuje zhruba čtyřprocentní riziko (přesněji 3,8 %), že nalezené rozdíly jsou pouhou vlastností vzorku, v němž byly naměřeny. Z 96 % je tedy pravděpodobné, že tyto rozdíly skutečně existují i v základním souboru, pro který chceme zobecňovat.

Uvedený příklad také naznačil, jakým způsobem bývá pravděpodobnost kvantifikována. Při řešení různých úloh pracujeme s hodnotami od 0 do 1. Nulová pravděpodobnost znamená, že daný jev nikdy nenastane, zatímco $p = 1$ signalizuje, že daný jev nastane vždy. Vypočtené hodnoty pravděpodobnosti je možno také při interpretacích převést na procenta tak, že je násobíme stem. Jestliže pak chceme vědět, zda se v nějaké otázce mýlíme, či ne, a zjistíme, že $p = 0,5$, pak je zřejmé, že na základě dat, která máme k dispozici, je možno činit závěry pouze s tím, že existuje padesátiprocentní riziko omylu (což z praktického hlediska znamená, že bychom závěry činit neměli).

Kolikaprocentní riziko omylu je přijatelné, bývá určováno arbitrárně s ohledem na povahu řešeného problému. Tomuto rozhodování se říká nastavování *hladiny významnosti* (značené malým řeckým písmenem α). U správně metodologicky prováděného výzkumu se hladina významnosti α určuje ještě před výpočtem p . Jinými slovy to, jaké riziko omylu nám daný problém dovoluje nést, má být určeno a priori, tj. dříve, než ho početně určíme. Na základě empirické zkušenosti byla přijata konvence, že je-li p nižší než 0,05, hovoříme o významném výsledku, je-li p nižší než 0,001, hovoříme o výsledku vysoce

významném. Řada problémů řešených v humanitních vědách dovoluje také pro p zhruba v oblasti 0,05 až 0,08 hovořit o výsledku okrajově významném. Je ovšem nutno mít stále na paměti, že jakkoli mohou uvedená čísla činit na laika hluboký dojem a vzbuzovat jeho respekt, jedná se o arbitrárně zvolené hodnoty a pětiprocentní riziko chyby 1. druhu není o nic „teoreticky správnější“ než riziko čtyřprocentní.

Dále je užitečné vědět, že nastavíme-li před experimentem příliš nízké α , snižujeme sice riziko chyby 1. druhu, ale zvyšujeme riziko chyby 2. druhu. Budeme-li tedy chtít příliš velkou jistotu, že hledaný vztah existuje, pak se může stát, že bude-li slabší, mylně ho prohlásíme za neexistující. Takováto chyba může vést k poškození jinak cenného výzkumného projektu. Proto je vždy důležité testovat hypotézy opakovaně a různými způsoby a nespoléhat pouze na jeden experiment, popřípadě na jeden zdroj dat.

Další, již méně obecnou oblastí, v níž se ve fonetických vědách setkáváme s pravděpodobností, je oblast fonotaktiky, která se zabývá pozičním a kombinačním uspořádáním řečových jednotek. Pro řešení různých problémů v řadě oblastí fonetického výzkumu je nutno znát, v jakých pozicích uvnitř jednotek vyššího řádu a v jakých vzájemných kombinacích se ty které řečové jednotky vyskytují. Je například známo, že anglický souhláskový systém má 24 prvků. Pouze 22 z nich se ale může vyskytnout ve slabičné préture. V préturách se nevyskytují /ŋ/ a /z/. Ani pravděpodobnost výskytu oněch 22 souhlásek však není čistě náhodná. Ve slabičné préture nalezneme daleko častěji /k/ než /g/, podobně /s/ je častější než /z/. Obecně je možno říci, že poziční fonotaktická pravděpodobnost znělých obstruentů v préture anglické slabiky je vyšší než pravděpodobnost obstruentů znělých.

U kombinační fonotaktické pravděpodobnosti si zase povšimneme, že vytváří-li oněch 22 souhlásek v préturách dvoučlenné skupiny, pak by výsledkem mělo být 22^2 (tj. 484) možností (popřípadě $22 \times 21 = 462$ možností, odmítneme-li

gemináty). Ve skutečnosti není z těchto potenciálních kombinací využito ani deset procent.

Při zkoumání např. principů osvojování řeči se pak v souvislosti s fonotaktickými pravděpodobnostmi hovoří o *hustotě okolí*. Ta vyčísluje množství podobných jednotek (např. slov) vzhledem k jednotce sledované. Tak třeba české slovo *dát* /da:t/ má poměrné husté okolí /ba:t, va:t, ka:t, la:t, ra:t, sa:t/.

Jiným zajímavým příkladem práce s pravděpodobnostmi byly výzkumy MacNeilage, Davisové a dalších, týkající se různých fonotaktických preferencí v dětské řeči. Badatelé srovnávali teoreticky náhodný výskyt určitých sekvencí s reálným výskytem těchto sekvencí v jazyce úplném (tj. v kanonické formě) a v dětské řeči (MacNeilage et al., 2000, viz též Davis a MacNeilage, 2003 a MacNeilage a Davis, 2003). U deseti zkoumaných jazyků bylo například zjištěno, že reduplikace jednoduchých slabik se vyskytují hluboko pod hranicí náhody. Naopak jsou doloženy případy disimilací, které reduplikace odstraňují (např. *purpur* → *purple* v angličtině). Naproti tomu u dětí jsou reduplikace vysoko nad hranicí náhody, a to jak ve žvatlání, tak v prvních slovech (z češtiny známe třeba *autobus* → *abubus*, *pomeranč* → *poraranč*).

V dalším výkladu se termín pravděpodobnost bude neustále, znovu a znovu objevovat, neboť, jak bylo uvedeno na počátku tohoto oddílu, z hlediska statistiky se jedná o koncept zcela zásadní.