

## Longitudinální korpus raného vývoje češtiny

*Manuál pro přepisovatele*

*Anna Chromá, Filip Smolík, Klára Matiasovitsová*

Tento manuál vznikl jako výtah z komplexního, pravidelně aktualizovaného manuálu pro přepis ve formátu CHAT [MacWhinney, B. (2000). The CHILDES Project: Tools for Analyzing Talk. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates].

- Lze očekávat poměrně vysoký výskyt nesrozumitelných nebo špatně srozumitelných úseků. Je potřeba nalézt kompromis mezi předčasným vzdáním snahy o porozumění a zbytečně obětovaným časem. Ani jedno není žádoucí. Neposlouchejte určitý úsek (orientačně) více než pětkrát (pokud je důvodem nesrozumitelnost).
- Název přepisu je přesně stejný jako název nahrávky či datapointu.
- Vždy pracujte s **poslední verzí** manuálu. Při jeho čtení věnujte prosím důkladnou pozornost příkladům, které jsou nedílnou součástí.
- Přepisujte vždy v textovém editoru (doporučujeme Notepad++) a dodržujte formát souboru (**.txt**) a kódování znaků (**UTF-8**).

## PRAVIDLA PŘEPISU

### Základní rámec přepisu: hlavička a ukončení

V hlavičce se používají řádky uvozené znakem @. Uvádí se seznam osob na nahrávce, jazyk, datum nahrávky, atd.

@Begin

@Languages: ces

@Participants: CHI Viktor Target\_Child, MOT Mother

formát řádku Participants: ZKRATKA Jméno Role *jméno být nemusí, role se vybírají ze stanovené nabídky*

@Transcriber: TERBIN

Na konec přepisu patří řádek označující konec:

@End

V přepisu nikde **nesmí zůstat prázdné řádky**.

Za každou dvojtečkou řádkového úvodu následuje **přesně jeden tabulátor**.

### Typy řádků v přepisu a jejich zakončení

Každý segment promluvy má svůj vlastní řádek. Každý řádek v přepisu začíná řídicím znakem, který indikuje, o jaký typ řádku se jedná. **Hlavní řádek**, na kterém se zapisuje, co bylo ve skutečnosti proneseno, začíná hvězdičkou. Všechny **vedlejší řádky**, na kterých se zaznamenávají vysvětlivky k pronesenému, příp. chyby (viz dále), začínají znakem pro procento %. Po tomto úvodním znaku následuje třípísmenný kód označující mluvčího nebo typ řádku (v případě hlavních řádků velká písmena, v případě vedlejších malá písmena), dále dvojtečka, tabulátor a vlastní přepis. **Každý řádek musí být zakončen a před koncovým znakem vždy musí být mezera** (viz znovu níže).

\*CHI: musím přinést polštářek .

%pho: muším pineš poutáše .

Vedlejší řádky se vážou k nejbližšímu předcházejícímu hlavnímu řádku a dovysvětlují ho. Řádek začínající **%pho** upřesňuje fonologickou podobu pronesené výpovědi. Tento řádek vždycky končí tečkou, bez ohledu na intonaci. Každý hlavní řádek s verbálním obsahem je vždy následován řádkem pho, ať se zápis na hlavním řádku a pho řádku liší či nikoliv.

Řádek **%com** je univerzální a umožňuje komentovat situaci – **obsahuje komentáře vztahující se pouze a přímo k dané výpovědi**.

Řádek **%err** blíže specifikuje chybu označenou v hlavním řádku hvězdičkou v hranatých závorkách (viz dále).

\*CHI: zejtra zvednul [\*] .  
%pho: žejtja žvednul .  
%err: zvedne .

## Ortografický přepis

Přepis je **na hlavním řádku** zaznamenáván ortograficky. To znamená, že pro zápis rekonstruujeme z fonologického hlediska slovníkovou, ortografickou formu daného slova, tj. nebereme ohled na mnoho výslovnostních odchylek, např. nevyslovuje-li dítě ještě *r*, přesto ho zaznamenáváme, nepíšeme tedy *motolka*, ale *motorka*. Systematické chyby či zvláštnosti jedince uvádíme v řádcích %pho nebo %com.

**Velká písmena píšeme pouze u vlastních jmen, nikoli na začátku výpovědi.**

Šva **na fonologickém řádku** (%pho) značíme @, např. p@ší (**nesmí se vyskytnout na hlavním řádku**).

Morfologické zvláštnosti (např. vynechané nebo modifikované koncovky) naopak ponecháváme! Stejně tak ponecháváme silně zkomolená slova. Podrobněji k těmto záležitostem viz níže v sekci *Lexikální a morfologické zvláštnosti*.

## Jak tedy zacházet s fonetickými redukcemi apod.?

1. *Redukce souhláskových skupin uvnitř a na začátku, proteze, změny kvantity*

**na hlavní řádek neznačíme – zapisujeme ortograficky**

\*MOT: takhle si zavážeš tkaničku .

%pho: takle si zavážeš kaničku .

\*MOT: to jsem nevěděla .

%pho: to sem nevěděla .

\*MOT: tam půjdu až za nějakej čas .

%pho: tam pudu až za ňákej čas .

\*MOT: nevím .

%pho: nevim .

\*MOT: s tím nic nenadělám takhle od oka .

%pho: s tim nic nenadělám takle vod voka .

= jevy typické pro "dospělou" češtinu X jevy zvláštní pro dané dítě (viz sekce *Lexikální a morfologické zvláštnosti*):

\*CHI: oblečeme žamko@c .

%pho: obečeme žamko .

%com: žamko = pyžamko .

2. *Redukce a další odchylky v sufixech*

**na hlavní řádek značíme, nestandardizujeme**

\*MOT: s tím nemůžem takhle hejbat .

%pho: s tim nemůžem takle hejbat .

\*MOT: s těma holčičkama už si nebudeš hrát ?

%pho: s těma holčičkama už si nebudeš hrát .

3. *y > ej / é > í/y:*

na hlavní řádek značíme, nenahrazujeme

\*MOT: zejtra se vrátí mladej od sousedů .

%pho: zejtra se vrátí mladej vod sousedů .

4. *Způsob zápisu vybraných jednotlivostí:*

Kontrakce *jsi*

\*MOT: přišels [: přišel jsi] pozdě .

%pho: přišels pozdě .

Demonstrativa *tohle, tohlenco* apod. s d-realizací (*todle, todlencto*)

\*MOT: na tohlencto prosím\_tě nesahej .

%pho: na todlencto simtě nešahej .

Nepisovný kondicionál

(a)bysem, (a)bysi, (a)bysme, (a)byste

Prézens slovesa *být*

bez ohledu na výslovnost zapisujeme na hl. řádek ,j': *jsem, jsi, je, jsme, jste, jsou*

neplatí pro tvar *seš* – vždy bez ,j'

,&eee' zvuk zastavení po chybě a před opravou (něco tako „teda“ nebo „vlastně“)

,hm' souhlas 1

,mhm' souhlas 2 (dvouslabičný)

,emem' nesouhlas (dvouslabičný)

,nee' dvouslabičné s rázem

,(s) sebou'

## Kódy na hlavním řádku

### 1. "Neslova" a nesrozumitelné úseky

Nesrozumitelné úseky značíme jako xxx. Pro označení fragmentů slov (chybných nebo přerušených začátků) použijeme symbol &, který připojíme na začátek. Stejný symbol se použije také pro označení hezitačních zvuků a dalších „pazvuků“, které nejsou fonologicky kódovatelné a zřejmě nenesou žádný význam. **Hezitační zvuk konzistentně zapisujeme jako &em.** Kromě &em je obvyklé také hezitační **&to.** Podobnou povahu má i **&eee** (viz výše). Jejich významné prodloužování a zvláštní modulace je možné komentovat pomocí řádku %com.

\*CHI: xxx .

%pho: xxx .

\*CHI: bílý &em onp@c .  
%pho: bílý em onp .  
%com: onp = význam nejasný .  
\*CHI: &p prstík .  
%pho: p p@stík .

#### *Rozhodování mezi xxx a &*

V případě, že jde o komplexnější výpověď, a pouze určitá její část má být zapsána jako neúplná nebo nesrozumitelná, je dobré se zabývat otázkou: "Funguje daná výpověď syntakticky a sémanticky i bez nesrozumitelné části?" Jde o to, že výpověď, která je jakkoliv neúplná (obsahuje xxx nebo končí znakem pro neúplnou výpověď), se standardně vyřazuje z analýz. To se ale netýká fragmentů se znakem &.

- pokud spíše **ne**, zaznamenám pomocí **xxx** (tím signalizuji: "chybí tam podstatný syntaktický element, kterému nerozumím")
- pokud **ano**, zaznamenám pomocí **&** a přibližné hláskové rekonstrukce (&s / &nel ... tím signalizuji, že jde pravděpodobně jen o fragment nějakého nedokončeného/spolknutého slova nebo možná o zcela nejazykový zvuk)
- případně je také možné daný úsek/repliku rozčlenit vhodně na samostatné výpovědi, aby některá část zůstala jako srozumitelná a úplná

## **2. Zakončení řádku: nedokončené výpovědi a překryvy**

Každý řádek musí mít koncový znak. Před koncovým znakem je vždy mezera. Vedlejší řádky jsou VŽDY zakončeny tečkou. Nikdy nesmí být více koncových znaků na jednom řádku. Jakmile vložíte jeden koncový znak, musíte ukončit výpověď.

Na hlavních řádcích kromě tečky, otazníku a vykřičníku, používáme i kódy pro nedokončené výpovědi:

**+/.**  znamená přerušeni jednoho mluvčího jiným mluvčím.

**+<** použijeme pro překrývající se výpovědi: tento kód připojíme na začátek výpovědi, která začala dříve, než předcházející skončila. Za kódem následuje mezera. Kód pro překryv nemusí být vždy doprovázen kódem pro přerušeni, protože první mluvčí se nemusel vstoupením druhého mluvčího nechat ve své výpovědi přerušit.

**+...**  používáme pro sebepřerušeni mluvčího, které není následované sebeopravou, tj. následuje pauza a potom zcela nová výpověď téhož nebo jiného mluvčího, anebo tentýž mluvčí bezprostředně začne zcela novou výpověď (všimněte si rozdílu oproti kódu [//], viz níže).

\*CHI: takhle &ši +/.  
%pho: takle ši .  
\*MOT: Adélko, na to se nesahá .  
%pho: Adélko na to se nešahá .

\*MOT: tak kam se posadí tahle panenka ?  
%pho: tak kam se posadí tadle panenka .  
\*CHI: +< posadí 0se tady .  
%pho: posadí tady .

\*CHI: &tona my jsme +...  
%pho: tona my šme .  
\*CHI: mami, kde je ?  
%pho: mami de je .  
%com: vrací se k hledané ponožce .

\*MOT: tak mu +...  
%pho: tak mu .  
\*MOT: počkej !  
%pho: počkej .  
\*MOT: ohnem mu ty nohy .  
%pho: ohnem mu ty nohy .

### 3. Opakované začátky a sebeopravy

Další kódy se používají pro situaci, kdy mluvčí začne nějakou výpověď vícekrát, případně kdy se sám opraví. Jestliže se opakuje, použijeme kód <> [/], jestliže se opraví, pak použijeme <> [//].

\*MOT: tak <co mu teda> [//] jaký mu dáváš pleničky ?  
%pho: tak co mu teda jaký mu dáváš pleničky .  
\*CHI: já mu dávám <ještě> [/] ještě jednu pempřsku .  
%pho: já mu dávám eště eště jedu pep@sku .  
\*CHI: <až až až> [/] až tamtu poslední doděláme +...  
%PHO: až až až až tamtu posení doděuáme .  
  
\*CHI: <&tona my jsme vystupovali> [//] asi prvníma dveřma jsme vystupovali .  
%pho: tona my šme vyštupovaui asi p@níma dvema sme vykupovaui .  
  
\*MOT: a <chodíte> [//] jezdíte tam autobusem ?  
%pho: a chodíte jezdíte tam autobusem ?

Jiné případy opakování značíme jako [x počet\_opakování]. Mezi "x" a číslem je mezera.

\*MOT: <počkej> [x 3] .  
%pho: počkej počkej počkej .

#### Sebeopakování

- pokud dochází k opakování na začátku výpovědi nebo fráze, tak jde obvykle o **opakované začátky** (false starts, zakoktání) značené pomocí <> [/]
- pokud k němu dochází na konci nebo se týká celé výpovědi, tak jde o opakování značené pomocí [x počet]. Každé **opakování celé výpovědi** značíme jako [x počet].

Ne všechno, co se na povrchu jeví jako opakování, je skutečně opakování. Totéž slovo může být bezprostředně po sobě použito v referenci k odlišným konceptům, s odlišnou funkcí. Např. slovo "tohle" ve výpovědi "Tohle, tohle a tohle." může referovat ke třem různým hračkám. **Pokud jsme to schopni interpretovat, pak v takovém případě neznačíme opakování, nýbrž zapíšeme všechna slova.**

#### 4. Pauzy a prodloužení

K označení pauzy mezi slovy slouží znak (.), ale v principu je lepší začít při jakékoli (delší) pauze novou výpověď. Je také možné značit pauzy mezi slabikami, jde o znak ^, který se bez mezer vloží mezi slabiky. Prodlužování jednotlivých slabik je možné zapisovat pomocí dvojtečky, opět bez mezer vložené na konec prodloužené slabiky.

Se značením prodlužování slabik i mezer mezi slabikami spíše šetříme a užíváme ho v markantních případech (významnost případu určujeme s ohledem na typické vyjadřování dané osoby). V případě, že dítě často vkládá mezi slabiky pauzy (váhá nebo se zajímá), okomentujeme (%com) to u konkrétní výpovědi nebo v obecném komentáři (@Comment) kdekoliv během přepisu (viz dále).

\*CHI: my jsme jeli do lékárny pak (.) za tej^den .

%pho: my šme jeli do lékány pak (.) ža tej den .

\*MOT: he:le@i, táta .

%pho: héle táta .

##### Dělení na výpovědi

- Dělení na výpovědi se nemusí krýt s členěním dialogů na repliky (replika = co řekne jeden mluvčí, než ho vystřídá jiný mluvčí).
- Dělíme podle klesavé i stoupavé koncové intonace a podle smyslu výpovědi, u velmi malých dětí často i po jednotlivých slovech.
- Zároveň ale platí, že na jednom hlavním řádku jsou **max 2 souřadně spojené hlavní věty ve slučovacím poměru** (bez spojky nebo s "a").
- Pokud se bezprostředně opakuje celá výpověď (byť by to bylo jen jedno slovo), pak značíme opakování [x 2] (viz výše).

Dělení na výpovědi je důležité, protože délka výpovědi je důležitým obecně využívaným měřítkem stupně jazykového rozvoje, a to zejména zpočátku, dokud jsou výpovědi spíše krátké.

#### 5. Neverbální jevy, komentáře a vysvětlivky

##### 5.1 Začlenění komentářů do hlavního řádku

&=smích	neverbální hlasový projev; nekoinciduje s řečí (vložíme přímo na místě, kde se jev odehrál) další časté kódy tohoto typu: &=smrkání; u velmi malých dětí: &=námaha, &=žádost
&=imit:auto	takto můžeme zapsat imitaci nejrůznějších zvuků (za dvojtečku bez mezery napíšeme jednoduše vlastními slovy, co se imituje)
<scope> [?]	zaznamenání nejistoty přepisu

\*MOT: a jakou &=smích barvu má tahle ?

%pho: a jakou barvu má tahle .

%com: z nahrávky není zřejmé, čemu se směje .  
\*CHI: <žlutou> [?] .  
%pho: žuu .  
\*MOT: žlutou, ano .

### 5.2 „Nulový“ zápis

V případě, že se osoba v dané chvíli nějak hlasově projevuje, ale nic neříká, můžete zvolit následující způsob zápisu s nulou. Při každém takovém použití nuly je nutné vždy připojit komentář (%com). Takový řádek nesmí obsahovat nic kromě nuly. **Tento zápis je univerzálně využitelný pro kašel, smích, u mladších dětí pro žvatlání a podobné vokalizace bez komunikační funkce apod.**

\*CHI: 0 .  
%com: mumlání .

### 5.3 Využití řádků hlavičkového typu

@New Episode signalizuje předěl na nahrávce, je bez dvojtečky a bez obsahu  
@Situation: popis situace, objeví se hned v hlavičce, ale taky během přepisu, pokud dojde k předělu  
@Comment: cokoli dalšího, co je potřeba sdělit (od %com se liší tím, že se zde jedná o informace obecnější povahy, které se nevztahují jen k jedné konkrétní výpovědi)

### 5.4 Využití vedlejších řádků

%com: komentář, který se týká přímo dané výpovědi, má univerzální využití.  
%err: využíváme pro objasnění chyb (viz *dále*).  
%add: specifikuje adresáta dané výpovědi; **využijeme vždy, pokud se nahrávky účastní více než dvě osoby** (může být velmi důležité pro přehlednost)  
%tim: časový údaj, kdy výpověď začala. Referenčním bodem je 00:00 = začátek nahrávky. Nepoužívá se často, ale může posloužit k upozornění na výraznější pauzy.

Pokud něco neprobíhá úplně synchronizovaně s řečí, ale před nebo po ní, je lepší využít obecný řádek @Comment.

**U komentářů a situačních popisů dbejte na to, aby byl přepis společně s těmito komentáři co nejsrozumitelnější pro osobu, která nemá k dispozici nahrávku. Tedy s nimi opravdu nemusíte šetřit.**

\*MOT: já vám to tady takhle připravím, pane doktore, a můžete číst .  
%pho: já vám to tady takle připravím pane doktore a můžete číst .  
@Comment: MOT odejde z pokoje .  
CHI: jednou ráno Geppetto chtělo mít Pinocchia [\*] .  
%pho: jednou jáno Dzepeto chtělo mít Pinocchia .  
%err: chtělo = chtěl .  
%com: CHI ,předčítá' podle knihy .



Řádek %com se využívá velmi často a velmi univerzálně. Jeho nejdůležitějším užitím je dovysvětlení intonace, resp. funkce výpovědi (která jinak bez nahrávky není zřejmá). Může to být důležité zejména u jednoslovných dětských výpovědí, které mohou dle intonace a doprovodných zvuků a gest mít funkci čistě popisnou, anebo žádací, a často také jen opakovací (dítě záměrně opakuje, učí se, zpřesňuje svou výslovnost).

\*CHI: &aj korálek .  
%pho: aj kouajak .  
%com: drží sáček, tónem a pokýváním hlavou žádá o otevření sáčku a vyndání korálků .

\*CHI: &te korálek .  
%pho: te koueuek .  
%com: komentuje vybraný korálek a ukazuje na něj .

\*FAT: a ještě tenhle dáš na tu šňůrku ?  
%pho: a eště tenle dáš na tu šňůrku .  
%com. ukazuje na jiný nenavlečený korálek ležící před CHI .  
\*CHI: šňůrku .  
%pho: čupšu .  
%com: opakuje .  
\*FAT: šňůrku .  
%pho: šňůrku .  
%com: opakuje a upřesňuje .

### 5.5 Básničky, písničky, názvy

Pokud se na nahrávce objeví recitace básniček, zpívání nebo jen zlomky takových naučených textů, je potřeba je většinou vyřadit z analýzy. Takovou sekvenci vložíme do špičatých závorek a označíme ji komentářem jako "zpěv" nebo jako "básničku". Jako "básničku" označujeme i písňové texty, pokud nejsou zpívané, ale jen přeřikávané. V případech, kdy jsou takové úseky rozsáhlé, např. se na nahrávce zpívá úplně celá písnička, nemusíte ji přepisovat a vložte pouze vysvětlující komentář. Někdy je možné sekvenci normálně ponechat v textu bez závorek, pokud se text básničky/písničky jeví hodně aktualizovaný nebo nějak vědomě analyzovaný.

Z víceslovných názvů (a také frází typu prosím\_tě, pane\_jo) uděláme pro automatickou analýzu jedno slovo pomocí podtržíttek.

\*CHI: <bába leze do bezu> [=! básnička] .  
%pho: bába leze do bezu .  
\*CHI: a on zpívá <Hajdom\_hajdom\_tydlidom> [x 2] .  
%pho: a on zpívá hajdom hajdom tydlidom hajdom hajdom tydlidom .  
\*MOT: zpívá Hajdom\_hajdom\_tydlidom, no .  
%pho: zpívá hajdom hajdom tydlidom no .

## 6. Lexikální a morfologické zvláštnosti

Speciální dětské výrazy (bez zjevné morfologické inovace) značíme na konci slova @c, neologismy se zjevnou morfologickou inovací značíme jako @n, citoslovce @i. **Citoslovce je třeba důsledně rozlišovat od odpovídajících neverbálních jevů: &=smích vs. chichi@i nebo &=povzdech vs. uh@i.** Chichi@i nebo uh@i se v přepisu velmi pravděpodobně vůbec neobjeví.

Další speciální kódy využíváme pro zvláštní užívání citoslovcí v raných stádiích řeči. Pro užití citoslovce jako predikátu použijeme kód @z:ip, pro citoslovce v nominální pozici @z:in, v atributivní pozici @z:ia. Tyto funkce kódujeme pouze v jednoznačných kontextech, nikdy je tedy nekódujeme v jednoslovných výpovědích (toto kódování se častěji vyskytne u dospělých).

Zejména u mladších dětí v jednoslovném období je velmi časté užívání kódu @c. Za hlavním řádkem obsahujícím slovo s tímto kódem vždy musí následovat řádek %com, který vysvětlí význam slova s @c, příp. obsahuje poznámku ‚význam nejasný‘.

\*CHI: pojd', něco ti ukázám@n .

%pho: pod', něco ty ukázám .

%com: ukázám = ukážu .

\*CHI: já se tam takhle nakloňuju@n .

%pho: já se tam takle naklonuju .

%com: nakloňuju = nakláním .

\*MOT: co že jí dáme ?

%pho: co že jí dáme .

\*CHI: dáme jí za zvířeto@n opičku .

%pho: dáme jí za zvířeto opicku .

%com: zvířeto = zvíře .

\*CHI: hopla@i .

%pho: hopla .

\*MOT: já@i, tadydle je nějaký robot .

%pho: já tady je nějaký robot .

\*MOT: uděláme ham@z:ip ?

%pho: uděláme ham .

\*MOT: podívej, to je brm\_brm@z:in !

%pho: podívej to je brmbrm .

%com: brm\_brm = auto .

\*MOT: tady je oheň, to je au\_au@z:ia .

%pho: tady je oheň to je auau .

\*CHI: to není holčička .

%pho: to není oučiča .

\*SIT: a kdo je to ?

%pho: a kdo je to .

\*CHI: to je kluk .  
%pho: toe tuť .  
  
\*SIT: co je tohle ?  
%pho: co je todle .  
\*CHI: &e yvuladlo@c .  
%pho: e yvuladlo .  
%com: umyvadlo .

#### Standardizace zkomolenin vs. speciální dětské tvary

Jako @c se mohou označovat jednak formy, které si dítě úplně vymýšlí (říká např. *bu* pro vodu), jednak i formy, které jsou výraznými zkomoleninami oproti cílové dospělé formě (*yvuladlo*). V mnoha případech jsou odchylky mírné a chceme je na hlavním řádku raději standardizovat na dospělou formu. Abyste to mohli udělat, tak:

1. **musíte vědět, co daná forma v daném kontextu znamená** (často totiž můžete jako @c označit i formy, které jsou pro dítě zjevně slovem, ale ani rodiče třeba netuší, co mají znamenat)
2. **musí být zachovaná gramatická část** (koncovka/sufix: *-dlo*, resp. *-o* u *yvuladlo*)  
dále přihlížejte k následujícím faktorům:
3. počet slabik
4. pořadí slabik
5. zachování vokálů
6. pokud jsou konsonantické odchylky, měli byste si je dokázat odůvodnit (chybějící nebo jinak tvořený konsonant = zjednodušení artikulace; je dobré se odrazit od vytvoření seznamu zkomolenin daného dítěte pro danou nahrávku a podívat se, které odchylky dělá pravidelně – pravděpodobně to bude stačit jednou u prvního přepisu)

Pokud **nejsou** kritéria ve větším počtu splněna, ponecháme na hlavním řádku realizaci dítěte označenou jako @c). Kritéria mají částečně sestupnou důležitost (1. a 2. jsou nejdůležitější). Viz příklady na předchozí stránce: např. *tuť* > *kluk* bylo standardizováno, protože dané dítě substituovalo *t/ť*>*k* pravidelně.

**Když s dětskou formou operuje i rodič, opravuje dítě, nebo naopak přistupuje na jeho tvar, pak tuto formu vždy zachovejte s kódem @c i na hlavním řádku!**

## 7. Chyby

V hlavním řádku na chybu upozorňujeme hvězdičkou v hranatých závorkách [\*], kterou umístíme na konec výpovědi před koncový znak. Chybu pak zaznamenáme/okomentujeme na řádku %err. Pokud chyba spočívá ve vynechání slova, napíšeme ho už na hlavní řádek, přitom na začátek připojíme nulu (v tom případě nepoužíváme hvězdičku) – chybějící slovo jde nicméně často těžko takto jednoznačně rozřešit, proto stanovíme seznam takto kódovaných vynechaných kategorií, prozatím: 2. osoba pomocného slovesa být; zvrtná zájmena a další enklitika; předložky.

Chyby můžou někdy splývat s morfologickými zvláštnostmi. Lze říct, že chyby obecně značíme méně často a **rozhodně neznačíme jako chyby nespisovné jevy běžné mluvy**. Typickými chybami jsou koncovky nesprávného pádu/osoby/atd. nebo syntaktické odchylky. Jako @n (viz sekce 6. *Lexikální a morfologické zvláštnosti*) naopak označíme koncovku odpovídajícího pádu/osoby/atd., **ale z jiného ohýbacího typu** (např. u sloves záměna koncovek 1sg prez -u vs. -ám).

**Zkomoleniny neznačíme jako chyby**, ale jako speciální dětské tvary pomocí kódu @c (viz výše).

\*CHI: my chodíme jenom normální cestu [\*] .

%pho: myd chodíme jenom nojmálný testu .

%err: cestou .

\*CHI: já ti ukážu moje kladívko .

%pho: já ti ukázu moe kadívko .

\*MOT: potom, jo ?

\*CHI: já ti ho ukážu do postele [\*] .

%pho: já ti ho ukázu do postele .

%err: přinesu ti ukázat .

### 8. Anonymizace a pseudonymizace

Jméno hlavního dítěte je nahrazováno pseudonymem, nejlépe ihned v průběhu přepisu. Pro každý pseudonym vytvoříme i set příslušných hypokoristik, které následně v celém korpusu používáme konzistentně (např. Báro = Sáro; Barunko = Sárinko). Jméno je nutné vhodným způsobem nahradit i na řádce %pho. Jména ostatních lidí ponecháváme. Příjmení, adresy apod. se na hlavním a %pho řádce nahradí kódem zzz a v komentáři se doplní, jaký typ informace je vynechán. Pokud si nejste jistí, jak s daným údajem naložit, označte ho kódem @an, abyste se k němu mohli snadno vrátit později.

\*CHI: já se jmenuju Adam zzz .

%pho: já se menuu Adam zzz .

%com: vynechané příjmení .

### 9. Hlavní rozdíly proti verzi 1.0

- nově zavedená mezera před koncovým znakem

- básničky a písničky se nespojují podtržítka, pouze se uzávorkují a označí

- nepoužíváme kód <scope> [=! komentář] (kromě básniček)

- u všech kódů s hranatými závorkami (kromě [\*]) je nutné vyznačit rozsah pomocí špičatých závorek i

v jednoslovných případech

- nepoužíváme kód ,yyy'

- hvězdička pro označení chyb se umísťuje na konec výpovědi

- nekódujeme větný důraz