**Collecting and Analyzing Sign Language Data: Video Requirements an Use of Annotation Software**

**Sběr a analýza dat ve znakovém jazyce – náležitosti videozáznamu a využití anotačního softwaru**

- technický pokrok jde stále dopředu – je nutné sledovat novinky

- text nezachází do technických detailů, ale soustředí se na to, jak správně natočit video – korektní a použitelná data odpovídající výzkumným otázkám, anotovat specifické výzkumné otázky (efektivní a využitelné kódování)

- analýzu aktuálního užití ZJ, primárních jazykových zdrojů – ve formě videonahrávky (ta by měla odpovídat výzkumné otázce)

- materiál vyžaduje transkripci a kódování, kterou usnadňují sofistikované anotační softwary (ELAN, Anvil, iLex, LEXUS aj.)

**Sběr jazykových dat**

- technická a obsahová kvalita videa

- video musí být ve formátu, který umožňuje následnou analýzu

- data dvojího typu – spontánní vyprávění (přirozená data) x elicitace na základě různých stimulů – podnětů (elicitovaná data)

- sběr dat většinou nejde zopakovat za stejných podmínek

- sběr dat vyžaduje kontakt s komunitou a musí probíhat ve spolupráci s neslyšícími výzkumníky

Přirozená data

- spontánní produkce v přirozených podmínkách

- aktuální úzus neovlivněný požadavky výzkumníků – vhodné pro porozumění interakci a dynamice přirozené komunikace

- vhodné při výzkumu osvojování ZJ, interakce dítě – pečovatel (např. při longitudinálních sledováních)

**Elicitovaná data**

- výhodou je možnost kontroly, zachycení specifické domény výzkumu (přirozená data negarantují přítomnost sledovaných jevů)

- různá forma elicitovaných materiálů (obrázky, krátká videa, komiksy, příběhy – nelingvistické formáty x lingvistické formáty – psaná slova, věty většinového MJ – nebezpečí strukturního ovlivnění!)

- elicitovaná vyprávění a popisy musí být produkovány pro neslyšícího adresáta, ne pro kameru nebo výzkumníka – projev bude přirozenější – adresát má nějaký úkol (např. najít obrázek), to motivuje znakujícího k popisu

- pro kvantitativní a statistickou analýzu je nutné mít porovnatelná data od více znakujících, je to nutné také k posouzení toho, zda data reprezentují aktuální úzus

- využitelné také pro crosslingvistické studie

- korpus by měl kombinovat oba typy dat

- data musí přinášet odpovědi na položenou otázku

- pozor na uměle vzniklá data vlivem procesu sběru

**Technické aspekty sběru dat**

- jak sbírat vysoce kvalitní archivovatelná data

- vybavení: moderní kamery a stativy (dbát na výběr techniky, konzultovat s odborníky)

- naučit se s vybavením zacházet

- mít dostatek místa na data

- natáčet s kamerou zapojenou do elektřiny nebo mít dostatečně nabitou baterii

- přenést data do úložiště v počítači, zálohovat data

- důležitý je program na zpracování dat

- video by mělo být ve formátu použitelném pro SW, který chceme používat (MOV, AVI, MPEG) – pozor na formáty!

**Natáčení – prostředí, podmínky, technika**

- není to pouhé zmáčknutí nahrávání – REC

- kamera musí být na stativu – je stabilní a statická, lépe je nešvenkovat, nezoomovat – nebezpečí rozmazaných, pohyblivých, neanalyzovatelných obrázků (+ přitahuje to nechtěnou pozornost respondentů)

- natáčecí situaci je třeba předem připravit – rozmístěný, označit místa na podlaze, kde budou participanti, elicitační materiál atd.

- participanti musí být dobře vidět s dostatkem místa okolo těla, aby ruce nevylézaly za záběru

- pravidlo: spodní strana záběru na kolenou ke středu stehna, nad hlavou nechat prostor velikosti hlavy

- znakující by měl být zabrán z mírného bočního úhlu, ne úplně zepředu, mírně shora (kamera mírně nakloněná dolů) – lepší hloubka a rozlišitelnost videa

- pokud natáčíme 2 participanty na 1 kameru, sedí naproti sobě, napůl natočeni ke kameře, napůl k sobě

- lze využít více kamer, na každého respondenta jedna, nebo kamera pouze pro nemanuální komponenty atd.

- pohled shora – stativ s kamerou umístíme na stůl, nebo kameru namontujeme na mikrofonový stojan (na příčné rameno)

- při použití více kamer je pro následnou synchronizaci dobré využití vizuální+zvukový signál na začátku natáčení, např. tlesknutí

**Archivování a přístupnost videonahrávek**

- ihned po natáčení uložit video do počítače a pojmenovat

- zpracovat data, rozstříhat na segmenty – např. jednotlivé úkoly, lépe se zpracovává, nejsou problémy s pamětí

- pozor na smazání! Zálohovat! – mít alespoň jednu zálohu!!

- co nejrychleji vytvořit metadata – zásadní pro archivaci a přístup k datům – měly by obsahovat informace o čase, datu, místě natáčení, obsah nahrávky, participanti, údaje o nich

- metadata musí být propojitelná s videem skrze jméno souboru – důležité používat konzistentní, systematické, popisné názvy, aby bylo rozpoznatelné i později

- v průběhu času je třeba aktualizovat formát, aby nebyl zastaralý

- důležité uchovávat data ZJ – nemá psanou podobu

**Software pro anotování dat ZJ**

- anotace – označování prvků jazykových dat, která chceme popisovat, analyzovat, kvantifikovat atd.

- elementy mohou být na různé úrovni: fonologické, morfol., syntakt., sémant., pragmat., aj. označování může být deskriptivní nebo analytické

- moderní SW – časové propojení přímo s korespondujícími částmi videa

- je možné dělat i samostatnou anotaci (tabulka, excel), ale lepší je propojení s videem

- pře anotováním je třeba mít jasnou představu, co chceme anotovat – kódovací schéma přizpůsobené výzkumu – vytvoření schématu kódování

- analýza musí umožnit vytěžit všechny informace, které potřebujeme pro analýzu a kvantifikaci – konkrétní příklady s. 64

+ stručný popis systému ELAN