

# GOVORJENA BESEDILA IN KORPUS SLOVENSKEGA JEZIKA

*Simona Kranjc*

Filozofska fakulteta v Ljubljani

Aškerčeva 2

1000 Ljubljana

tel.: 061-1769-232, fax: 061-1259-337

Simona.Kranjc@guest.arnes.si

## POVZETEK

Govorjena besedila so v slovenskem prostoru malo raziskana, kar je predvsem posledica značilnosti te vrste besedil. Na podlagi pilotne analize otroškega govora bomo ob obravnavi osebnih deiksivov, to je izrazov s kazalno vlogo, skušali pokazati razliko med govornimi in zapisanimi besedili in s tem utemeljiti potrebo po načrtovanem zbiranju govornih besedil in njihovo vključitev v korpus slovenskega jezika.

## ABSTRACT

So far spoken texts have been rather poorly researched by Slovene experts, which is to be seen as connected with the very nature of such texts. On the basis of a pilot analysis of child speech, which takes into particular consideration personal deictics (i.e. expressions having a demonstrative function), the paper will attempt to show differences between spoken and written texts and argue that systematic collection of spoken texts and their inclusion in the Slovene language corpus would be necessary.

## 1 GOVORJENA BESEDILA IN KORPUS SLOVENSKEGA JEZIKA

Govorjena besedila so bila v zgodovini jezikoslovnega raziskovanja jezika ves čas nekoliko zapostavljena. Vzrok takemu stanju gre iskati v naravi prenosnika, to je govora, in v razvoju tehnike, ki podpira zbiranje gradiva. Poleg tega pa so dodatno težavo predstavljale tudi same značilnosti govornega besedila. Ko namreč človek govori, je njegovo besedilo takoj na voljo poslušalcu in zato nima možnosti popravljanja ali preoblikovanja. Nasprotno pa ima pri zapisanih besedilih tvorec vedno vsaj nekaj trenutkov časa, da besedilo preveri in popravi. Za pogovor so značilni prekrivajoči se govor, ponavljanje, premori, zapolnjevalci vrzeli, napačni starti, samokorekture. Vse te značilnosti so kazalci kognitivnega procesa na različnih ravneh, in sicer od načrtovanja teme pogovora do iskanja besed in strukturiranja izrekov. Tekočnost govora se je pogosto raziskovala kot dimenzija pragmatičnega obnašanja. Označena je bila kot ustrezna ali neustrezna. Tudi preskoki in premori so pojavi, ki so odvisni od konteksta, ki se spreminja z vsako

konverzacijsko udeležbo. Interakcija med sogovorcema je strukturno organizirana glede na določene socialne oziroma družbene konvencije, kar pomeni, da se jo lahko preučuje neodvisno od psiholoških ali drugih atributov udeležencev. Veliko osnovnih, ne pa vedno tudi očitnih, lastnosti konverzacijske organizacije se ujema s predpostavkami teorije vljudnosti (Levinson, Brown 1987), kar podpira trditev, da je pogovor zgrajen v skladu z družbenimi konvencijami. Kontekstualno vezana konverzacijska udeležba je lahko razumljena le z referenco na kontekst, v katerem je izrečena, predvsem je tu mišljen predhodni kontekst, hkrati pa tudi gradi nov kontekst za tisto, kar ji bo sledilo. Posledica primarnosti konteksta je, da je pomen izreka produkt interakcijskega delovanja. Vse to je v nasprotju s Searlovo teorijo govornega dejanja, ki daje prednost izolirani povedi, to nato analizira glede na njene skladenjske in semantične lastnosti. S stališča novejših pristopov h govornemu diskurzu je pomanjkljivost Searlovega in Griceovega pristopa v tem, da se ukvarjata ločeno s pomenom, ki ga je govorec hotel izraziti, in s poslušalčevim razumevanjem. Schlegloff (1987) je poudaril, da morajo biti izreki primarno razumljeni s pomočjo reference na mesto v nizu; relevantno vprašanje, ki se pri tem postavlja, pa je, kako ljudje skupaj usmerjajo potek pogovora in izpeljujejo probleme, kot se pojavljajo. Vprašanje, kdo kdaj govori, je treba raziskovati skozi sistem menjavanja vlog. Ko ena oseba govori, druga ne le posluša, marveč z različnimi jezikovnimi in nejezikovnimi sredstvi kaže razumevanje in strinjanje ali nestrinjanje. Že sama odsotnost teh znakov sprejemanja je znak, da se pogovor ne odvija normalno.

Najpopolnejšo sliko spontanega govornega besedila je dalo dnevniško zapisovanje.<sup>1</sup> Vendar je tako delo precej težavno, saj je zapisovalec omejen z možno hitrostjo zapisovanja in opazovanja več stvari hkrati. Veliko težav je postalo manjših z razvojem avdiovizualne tehnike. Korak naprej je najprej pomenilo snemanje z magnetofonom (ob tem si je raziskovalec še vedno moral

<sup>1</sup>Dnevniško zapisovanje je znano predvsem iz študij otroškega govora, kjer so raziskovalci, predvsem starši ali stari starši, zapisovali vse, kar je otrok rekel in storil. Popolno sliko je tako zapisovanje nudilo tudi zato, ker so poleg govora zapisovali še okoliščine in sobesedilo, v katerem so se vršila govorna dejanja.

zapisovati opombe o nejezikovnem sobesedilu in okoliščinah, ki so spremljale govorno dejanje), nato pa snemanje s kamero. Slednje nudi skoraj popolno informacijo o govornem dejanju in njegovem kontekstu, manjkajo le še socialnokulturne informacije o življenju ožje in širše družbene skupnosti, v kateri živi govorec.

Z razvojem tehnike zbiranja gradiva pa se je pojavil nov problem - transkripcija. Raziskovalci so se znašli pred vprašanjem, kaj in kako prepisati s traku. Sprašujemo pa se tudi, ali je mogoče pričakovati razvoj take programske opreme, ki bo omogočala strojni prepis govora v digitalno obliko. Da bi bilo tako opremo mogoče ustvariti, je verjetno prej treba ustvariti bazo gradiva.

Govorjena besedila predstavljajo večino besedil v komunikacijskih procesih, tako v spontanah situacijah kot tudi v vnaprej načrtovanih, na primer medijih. In zakaj ni dovolj, če se ustvari baza gradiva iz zapisanih besedil? Ker se zaradi opisanih značilnosti med obema vrstama gradiva kaže velika razlika. Za primer si oglejmo razlike med rabo osebnih deiksivov ali izrazov s kazalno vlogo v govorjenih in zapisanih besedilih.

Deiksivi ali izrazi s kazalno vlogo so posebne oblike v jeziku, ki služijo kot most med jezikovnimi oznakami, ki so usmerjene v mišljenje, in tistimi, ki se odražajo v govoru. Teorije o razumevanju jezika, med njimi N. Chomsky s pretvorbno slovnico, so ostale na razlagalnem pristopu misli in mišljenja, ki predpostavlja pomen "egocentričnega" subjekta. Podobno so se teorije o pomenu (semantika) predvsem ukvarjale z naravo abstraktnih entitet, kot na primer propozicije, in kot osnovno enoto analize vzele poved. Domnevali so celo, da predmet semantičnih raziskav ni neposredno naravni jezik. Sodobna psiholingvistika skuša razjasniti problematičen odnos med mišljenjem (kognitivne dejavnosti, ki potekajo v individuumu) in jezikom. Povečalo se je zanimanje za funkcijski pristop: kaj jezik pomeni v kontekstu. Pri tej želji, razjasniti pomen jezika v kontekstu, je študij deiktičnih izrazov ena od najbolj obetajočih poti.

Deiksivi so slovnični izrazi za skupino besed in besednih zvez, ki vzpostavljajo interakcijo med izrekom in kontekstom, v katerem je nastal. Razumemo jih lahko le z referenco na situacijo, v kateri so bili uporabljeni. Opravljajo pa lahko enkodiranje vrste različnih prvin okoliščin izrekanja izrekov. "Deiktični izrazi služijo za usmerjanje poslušalčeve pozornosti na prostorske in časovne vidike sporočanjeve situacije, ki so pogosto kritični za interpretacijo. Služijo kot središčna točka skladišnih, semantičnih in pragmatičnih vidikov jezika." (Lyons, 1977: 637 [1].) Izrazi in besedne zveze, kot so *jaz/ti*, *to/tisto*, *tukaj/tam*, *pred/za*, *včeraj/naslednji teden*; *priti/iti*; *pod/med* ..., so vsi deiktični izrazi. Mednje pa se uvrščajo tudi izrazi spoštovanja, ki označujejo status ali položaj osebe, ki se nanjo obračamo.

Pojmovanje deiksivov je za psihologijo pomembno iz več vzrokov, in sicer

1) predstavljajo most med mišljenjem in govorom;

2) ležijo na površini med razumevanjem jezika in socialno interakcijo;

3) zanimivi so z vidika učenja, ker nudijo veliko možnosti za raziskavo, kako se otroci učijo razumeti, da je perspektiva drugih različna od njihove (otroci imajo težave pri izrazih, kot so *tu/tam*, *to/tisto*; mnogi te težave razlagajo s stopnjo otrokovega kognitivnega razvoja);

4) četrti razlog zanimanja pa izvira iz nasprotujočih si mnenj o tem, ali obstaja univerzalna slovnica. Chomsky je trdil, da je edini način, da se razloži način učenja jezika pri majhnem otroku, ta, da predpostavljamo genetične predispozicije za to funkcijo. Ta nativistični vidik lahko najdemo tudi v sodobnih delih S. Pinkerja (1994) in Karmiloff-Smith (1983).

Ko govorci v komunikacijskem procesu menjavajo vloge, se spreminja tudi deiktični center. Kanonični kontekst izrekov je namreč egocentričen, kar pomeni, da govorec vse gleda s svojega stališča. Postavljen je na ničto točko prostorsko-časovnih koordinat (Lyons, 1993: 638 [1]). To točko imenujemo deiktični kontekst (Lyons, 1993: 638 [1]) ali deiktično središče (Levinson, 1983: 64 [2]). Govorec je središčna oseba, središčni čas je čas, v katerem govorec izreka svoj izrek, središčni prostor je prostor, v katerem se govorec nahaja v času izrekanja izreka, središče diskurza je točka, na kateri se nahaja govorec v času izrekanja, socialno središče pa je socialni status ali položaj govorca. Obstajajo sicer tudi konkretni govorni položaji, v katerih se deiktično središče pomakne k drugim udeležencem, Lyons jih imenuje deiktične projekcije (1993: 636-640 [1]). Deiktično središče si lahko predstavljamo kot štiridimenzionalni prostor z govorcem v središču. Dodati bi bilo treba še eno dimenzijo, in sicer tisto, ki bi kazala socialni status govorca in njegov odnos do drugih udeležencev komunikacijskega procesa. Ko se vloge zamenjajo, se spremenijo tudi koordinate deiktičnega središča, središče se postavi v novo časovno, prostorsko in socialno določeno točko, ki je spet odvisna od tokratnega govorca. Taka slika deiktičnega središča in njegovih sprememb se zdi zapletena. Da taka v resnici tudi je, kažejo težave, ki jih imajo otroci pri učenju deiktičnih izrazov.

Slovnična kategorija osebe je odvisna od predstav o udeleženskih vlogah in njihove gramatikalizacije v posameznih jezikih. Prvi osebi je namenjena najpomembnejša vloga, ta je tista, ki govori, njej nasprotna je druga oseba, ki ji je namenjen diskurz prve osebe; tretja oseba pa zaznamuje vse ostale osebe, ki niso neposredno vključene v komunikacijski proces, ker se nanje prva oseba neposredno ne obrača. Takoj, ko bi se nanje obrnila, bi namreč spremenile svoj status in bi postale druga oseba. Tretja oseba ima poseben status tudi v zaimenskem sistemu. Lyons (1993: 639 [1]) meni, da prvoosebni in drugoosebni zaimki kot taki sploh niso pomembni. Veliko jezikov (med njimi tudi slovenščina) jih v določenih sklonih sploh ne izraža (v imenovalniku je

osebni zaimsek v slovenščini v nepoudarjenih položajih izpuščen, in sicer tako prvoosebni, kot drugoosebni in tretjeosebni), ker je informacija o osebi izražena v glagolski končnici. Jasno pa je tudi, da se govorec večinoma nanaša nase, ko izreka neki izrek. Do kakšnih nesporazumov lahko pripelje spreminjajnje deiktičnega središča, ilustrira naslednja židovska zgodba. "Melamed (hebrejski učitelj) je ugotovil, da je spet pozabil doma svoje udobne copate in je ponje poslal študenta s sporočilom za svojo ženo. V sporočilu je pisalo: "Pošlji mi tvoje copate s tem fantom!" Ko je študent vprašal, zakaj je napisal "tvoje" copate, je melamed odgovoril: "Hja! ^e bi napisal "moje" copate, bi prebrala "moje" copate in bi poslala svoje. Kaj naj počnem z njenimi copati? Tako sem napisal "tvoje" copate, prebrala bo "tvoje" in mi poslala moje." (Levinson, 1983: 68 [2].)

Podobne težave imajo tudi otroci, ko se učijo jezika. V tem procesu jim odrasli lahko pomagamo, lahko pa jim stvari otežujemo tudi na tak način, da v svojem govoru deiksivov ne uporabljamo pravilno.

Čeprav se osebni deiksivi odražajo neposredno v slovnčni kategoriji osebe, se moramo zavedati različnih udeleženskih vlog in načina njihove gramatikalizacije. Tako moramo recimo ločevati med govorcem in virom izreka, sprejemnikom in ciljem, poslušalcem ali opazovalcem in naslovnikom ali ciljem. Primer iz opombe o otrocih, ki so vsi Špele, kaže ravno nasprotno - na neločevanje. Prva oseba tako predstavlja govorca (G), druga vključuje naslovnika (N), tretja pa izključuje tako govorca kot naslovnika.

Oseba	Udeleženska vloga
prva oseba	+ G
druga oseba	+ N
tretja oseba	- G, - N

Primer iz gradiva bo pokazal, s katerimi sredstvi otroci, stari približno dve leti in pol, izražajo osebo. Gradivo je bilo posneto 11. 11. 1992.

1. Rok: Jes te bom griznu. Bom griznu. Griznu bom.
2. Sara: Bejš. Sara boji. (Pojdi stran, bojim se.)
3. Sarah: Lej. Glej.
4. Rok: Ne.
5. Alen: Hov, hov, hov.
6. Blaž: Ja.
7. Sarah: Glej.
8. Sara: Stan pejt. Stran.
9. Marcel: Torba strgana.
10. Sara: Sara boji. Nagaja. Si rumen. Boji kužka. Boji se, boji se kužka.

Vlogo osebnih deiksivov opravljajo osebni zaimsek za prvo osebo ednine jaz, ki ga otrok izreče tudi v imenovalniku, kjer je sicer v slovenščini v nepoudarjenih položajih ničta kategorija. Še v isti vlogi ga potem v dveh naslednjih izrekih izpusti, vendar je referenca na naslovnika izražena v glagolski končnici. V drugi vlogi

pa deklica referenco z govorcem, to je s samo seboj, vzpostavlja z lastnim imenom v tretji osebi ednine. Neustrezno slovnično sredstvo, ki ga uporabi, je v njenem govoru pogosto, je pa posledica govornega zgleda iz okolja. Isti način vzpostavljanja reference z govorcem kasneje v isti propozicijski sekvenci uporabi še nekajkrat. Z vzpostavljanjem reference z naslovnikom nimajo težav, pri čemer uporabljajo predvsem glagolske končnice (ukazi bejš, pejt) za drugo osebo ednine. Tretja oseba pa se pojavlja tudi v svoji primarni vlogi, torej kot udeležene opazovalec, ki ni ne govorec ne naslovnik (Torba strgana), ampak kot informacija.

Uporaba deiksivov je samo eden od elementov, ki kaže, kako različna so med seboj govorjena in zapisana besedila. Poimenovanj referentov z izrazi, ki nimajo lastne posebne reference, je v govorjenih besedilih več kot v zapisanih, ker je v prvih informacija dopolnjena še z nejezikovnimi sredstvi. Da je razlik med obema vrstama besedil še veliko več, bi lahko pokazale primerjave med besedili. Te bodo mogoče v večji meri šele tedaj, ko bodo poleg zapisnih besedil na voljo tudi podatkovne baze govorjenih besedil.

## 2 SKLEP

Govorjena besedila se zaradi narave komunikacijskega procesa, v katerem nastajajo, razlikujejo od zapisanih besedil. Poleg jezikovnih sredstev na njihovo razumevanje v veliki meri vplivajo tudi nejezikovna sredstva, predvsem kontekst. Ta značilnost pa do neke mere verjetno pogojuje tudi razvrstitev jezikovnih sredstev. Kot pomembna razlika med zapisanimi in govorjenimi besedili se kaže uporaba deiksivov, to je izrazov s kazalno vlogo, ki splošno referenco vzpostavijo šele v kontekstu. Zato jih je v zapisanih besedilih manj kot v govorjenih.

Vključevanje govorjenih besedil v korpus slovenskega jezika bi pomenilo prvi korak v njihovo intenzivnejše raziskovanje, hkrati pa bi omogočalo tudi premišljen način načrtovanja jezika v različnih govornih položajih, predvsem v javnih formalnih govornih položajih, na primer medijih.

## 3 VIRI

- [1] Lyons, John (1993). *Semantics 1-2*. Cambridge: Cambridge University Press.
- [2] Levinson, Stephen C., Penelope Brown (1992). *Immanuel Kant among the Tenejapans: Anthropology as Empirical Philosophy*. (Working paper No. 11). Nijmegen: Cognitive Anthropology Research Group at the Max Planck Institute for Psycholinguistics, 29 str.
- [3] Bach, Kent (1994). *Thought and Reference*. Oxford: Clarendon Press, Oxford University Press.
- [4] Forrester, Michael A. (1996). *Psychology of language. A critical introduction*. London - Thousand Oaks - New Delhi: Sage Publications.

- [5] Fox, (1993). *Discourse structure and anaphora. Written and conversational English.* (Cambridge studies in linguistics 48). Cambridge: Cambridge University Press.
- [6] Hofmann, Thomas R. (1993). *Realms of Meaning. An Introduction to Semantics.* (Learning about Language). London & New York: Longman, 339 str.
- [7] Jackendoff, Ray (1994). *Patterns in the mind. Language and Human Nature.* (Basic books). New York: A Division of Harper Collins Publishers, Inc., 246 str.
- [8] Kranjc, Simona (1995). *Razvoj govora predšolskih otrok.* (Magistrska naloga). Mentorica: prof. dr. Breda Pogorelec, somentorica: prof. dr. Ljubica Marjanovič Umek. Ljubljana: Univerza v Ljubljani, Filozofska fakulteta, 145 str.
- [9] Kranjc, Simona (1996/97). Govorjeni diskurz. *Jezik in slovstvo* 42, 7, 307-320.
- [10] Kranjc, Simona (1998). *Skladnja otroškega govora od prvega do tretjega leta.* (Doktorska disertacija). Mentorica: prof. dr. Breda Pogorelec, somentorica: prof. dr. Ljubica Marjanovič Umek. Ljubljana: Univerza v Ljubljani, Filozofska fakulteta, 190 str.