

UDK 811.163.6'322'34

Darinka Verdonik

Fakulteta za elektrotehniko, računalništvo in informatiko Univerze v Mariboru
darinka.verdonik@um.si

OZNAČEVANJE DIALOŠKIH DEJANJ V KORPUSIH

SHEME za označevanje dialoških dejanj v govornih korpusih so v svojo zasnovo vključevale teorijo govornih dejanj, vendar jo vsaka po svoje prilagajale in razširjale. Posledica tega je, da pogosto nesistematično mešajo različne vrste oznak. Cilj prispevka je raziskati, katere vrste pragmatičnega pomena označujejo različne sheme in na kakšne probleme v zvezi s tem naletimo v praksi, hkrati s tem pa ponovno preučiti, kaj vse delamo v diskurzu.

Ključne besede: dialoška dejanja, govorna dejanja, korpusna pragmatika

Schemes for annotating dialogue acts in spoken corpora have incorporated pragmatic speech act theory; however, they have also significantly adapted and expanded the original categorisation of illocutionary acts. Consequently, an unsystematic mixture of different types of annotations has arisen. The aim of this article is to investigate what types of pragmatic meaning are annotated in different schemes and what problems appear in the annotation. At the same time, the question of what kind of things we do in discourse is addressed.

Keywords: dialogue act, speech act, corpus pragmatics

1 Uvod

Govorna dejanja¹ predstavljajo eno temeljnih enot komunikacije. Levinson (2016) trdi, da so temelj pogovora in pogovor osrednji način jezikovne rabe, za Searla (1972) pa je bila ena temeljnih funkcij jezika, da izraža govorna dejanja. Tudi če se zdijo take izjave nekoliko pretirane, pa je nedvomno, da gre za enega osrednjih pojmov pragmatičnega jezikoslovja.

Govorna dejanja gledajo na izjave ne samo kot na nosilce pomena, ampak kot na enote jezika, ki povzročijo neki učinek pri naslovniku oz. opravijo neko aktivnost. Lahko rečemo, da so enote jezika z energijskim nabojem, in energija, ki jo nosijo, se prek komunikacije prenaša na druge in vpliva na njihovo delovanje.

Po objavi Austinove teorije in Searlovi nadgradnji le-te je vse do osemdesetih let dvajsetega stoletja vladalo zanjo veliko zanimanja. Z nekaj zamika se je to zanimanje preneslo tudi v slovenski prostor, zlasti po zaslugi Olge Kunst Gnamuš (1984), ki jo je podrobno obravnavala in pri nas širila znanja o njej. Kot ugotavlja Levinson (2016), pa je po dobrem desetletju razcveta v osemdesetih zanimanje zanjo začelo upadati, še preden bi bila temeljna vprašanja, povezano z njo, zares razrešena. Koliko tipov

¹ V članku uporabljamo termin 'govorna dejanja', kadar se nanašamo na rabo tega pojma v pragmatiki in izvorni teoriji. V kontekstu jezikovnih tehnologij in korpusne pragmatike uporabljamo termin 'dialoška dejanja'.

govornih dejanj obstaja, ali so kulturno vsaj delno specifična, kako se izražajo, kakšna je povezava med jezikovno podobo in funkcijo govornih dejanj, kako jih v jezikovni rabi prepoznamo – na ta in druga vprašanja še vedno nimamo zanesljivih odgovorov. Po eni strani torej vplivna in močna, po drugi pa v mnogih podrobnostih nerazrešena teorija se je tako vpisala na seznam temeljnih znanj jezikoslovnih in drugih sorodnih študijskih programov, hkrati pa izpisala iz seznama atraktivnih raziskovalnih tem.

Tako je tudi v slovenskem prostoru bolj izjemoma predmet raziskovalnega zanimanja. Še največ pozornosti ji posveti Žagar (2018), ki pristopi k temi predvsem s filozofskega vidika. Podrobno raziskuje zgodovino teorije govornih dejanj, njen nastanek, razvoj, cepitev in zlom ter možnost povezave s sorodnimi teorijami. V drugem delu se osredotoči na problem performativnosti in izpostavi prispevek Stanislava Škrabca na tem področju. Škrabcu kot utemeljitelju teorije performativnosti na Slovenskem je posvečena tudi monografija (Grgič, Žagar 2004).

V slovenističnem jezikoslovju je teorija govornih dejanj najpogosteje oživiljena, ko se avtorji lotevajo problema besedilnih skupin. M. Grosman (1997) uporabi govorna dejanja kot izhodišče za opredelitev literarne kritike, S. Hudej (2002) pa za delitev besedilnih vrst na šest skupin glede na tvorčev namen. Tudi M. Krajnc Ivič (2020) ugotavlja, da se v učbenikih in učnih načrtih ilokucijska vloga ali želen perlokucijski učinek pogosto uporabljata za definiranje besedilnih vrst.

Podrobnejša analiza posameznih vrst govornih dejanj, kar je ena od plodnih raziskovalnih niš (Hancher 1979; Blum-Kulka idr. 1989; Trosborg 1995; García 2019, če navedemo samo nekatere), je v slovenistiki redka. M. Nidorfer Šiškovič (2004) na kratko predstavi osnovne funkcijske sestavine zahteve kot vrste govornega dejanja. Bolj podrobno analizo predstavi M. Krajnc Ivič (2010), ki se ukvarja z govornim dejanjem opravičevanja, raziskuje njegove vzorce in opozarja na večfunkcijskost govornih dejanj, s katerimi se (med drugim) opravičujemo.

Posebno področje raziskav, ki se naslanjajo (tudi) na teorijo govornih dejanj, se je razvilo v navezi z razvojem tehnologij, ki omogočajo bodisi govorno komunikacijo s strojem bodisi stroj posreduje v govorni komunikaciji med ljudmi. Govorna dejanja so bila prepoznana kot zelo obetavna informacija, ki bi stroju pomagala učinkoviteje procesirati človeški govor. Osrednji cilj je bil označiti transkribirano govorno gradivo z informacijami, ki bodo pomagale stroju procesirati pomen komunikacije. Žal je ta empirični preizkus vedno znova pokazal številne težave teorije in njeno nepripravljenost na terensko uporabo. Tako se je označevanje govornih – dialoških dejanj v korpusih začelo razvijati po svoje, mimo izvirne teorije, žal pa pri tem namesto zapiranja odprtih problemov kvečjemu odpiralo še dodatna vprašanja.

Predmet našega zanimanja bo prav to, danes zelo aktualno (Georgila idr. 2009; Král, Cerisara 2014; Weissner 2014; Bunt idr. 2018), a hkrati konfuzno in nadvse težavno

področje – označevanje dialoških dejanj v korpusih.² V drugem poglavju bomo na kratko pregledali, kako se je razvijalo znotraj korpusne pragmatike in tehnoloških ved, v tretjem poglavju bomo argumentirali in predstavili svoj zorni kot na problematiko ter nakazali, kako k problemu pristopimo, v četrtem poglavju predstavljamo rezultate lastne analize in poglobitve v problematiko ter v petem poglavju sintezo spoznanj iz analize.

2 Korpusna pragmatika in dialoška dejanja

Povezovanje pragmatike in korpusnega jezikoslovja poteka iz dveh smeri. Na eni strani se v korpusnem jezikoslovju razvija teorija semantične prozodije (Stewart 2010; v slovenskem prostoru piše o tej temi izčrpno Šorli v tisku), ki raziskuje načine, kako se skozi leksikalno rabo realizira pragmatični, implicitni pomen. Ta za razvoj teorije izredno zanimiva in plodna veja je osredotočena zlasti v analizo leksikogramatike, frezeologije in idiomov, torej na enote, manjše od izjave.

Po drugi strani se je pragmatika začela aktivno spraševati, kako nove možnosti analize jezikovnega gradiva, ki jih omogočajo računalniška orodja in jezikovni viri, uporabiti za svoje (tradicionalne) raziskovalne cilje. Začele so se posebne številke pragmatičnih revij, konferenc in monografij (Romero-Trillo 2008; Felder idr. 2011; Taavitsainen idr. 2014; Aijmer, Rühlemann 2015), posvečene temu vprašanju. Ob tem se je rodilo tudi novo poimenovanje, korpusna pragmatika.³ Vidno mesto v njej je zasedla tudi tematika govornih dejanj.

Govorna dejanja pa so, še preden je pragmatika začela iskati sinergijo s korpusnimi metodami, našla pot v jezikovne tehnologije in tam dobila ime dialoška dejanja in hkrati delno novo razumevanje pojma. Označevanje dialoških dejanj se je uvajalo v prvih sistemih dialoga v področno omejenih situacijah, kot so komunikacija z avtomobilom, načrtovanje potovanja z vlakom, poizvedovanje po poti, dogovarjanje termina sestanka in načrtovanja potovanja, odgovarjanje na uporabniška vprašanja itd.

Kasneje se je pokazala potreba po splošno veljavni, generični shemi za označevanje. Eden prvih in hkrati najvplivnejših poskusov je bila shema DAMSL (Allen, Core 1997). Ta označuje dialoška dejanja na štirih različnih ravneh, imenovanih dimenzije. Prav večdimenzionalnost sheme je bila najbolj problematizirana, saj je bilo posledično mogočih več kot 4 milijone različnih kombinacij oznak, kar zahteva ogromno učno množico, obenem pa je bilo pogosto uporabljenih le okoli 200 različnih kombinacij. Zato so shemo DAMSL preoblikovali v SWBD-DAMSL (Jurafsky idr. 1997), ki vsebuje 42 medsebojno izključujočih se oznak, uporabljenih za označevanje telefonskih

² Raziskovalno delo, predstavljeno v prispevku, je financirano s strani ARRS v okviru projekta HUMANIPA – Fuzija verbalnih in neverbalnih signalov za naslednjo generacijo inteligentnih komunikacijskih vmesnikov (J2-1737).

³ Zelo ohlapno in široko opredeljeno se korpusna pragmatika nanaša na pragmatične raziskave, ki pri svojem delu uporabljajo korpusno gradivo in korpusnojezikoslovne postopke. Bolj natančno se bo njen pomen oblikoval skozi področja raziskav, ki se bodo razvijala pod njenim okriljem.

pogovorov na temo izvajanja naloge. Nadaljnja transformacija je sledila v MRDA (Dhillon idr. 2004), z razlogom, da se SWBD-DAMSL ne prilega ustrezno na nov tip diskurza, tj. sestanke. Posamezne oznake SWBD-DAMSL so izpustili, nekatere dodali ter ponovno omogočili kopičenje več različnih oznak za isto izjavo. Shema MRDA je predstavljala izhodišče za shemo MALTUS (Clark, Popescu-Belis 2004), kjer so bile oznake zreducirane, nekoliko posplošene in omogočajo le okrog 500 različnih kombinacij. DAMSL je bila osnova tudi v mnogih drugih poskusih označevanja.

Povsem ločen poskus označevanja dialoških dejanj predstavlja shema AMI (AMI 2005), ki so jo uporabili za označevanje diskurza delovnih sestankov. Shema je bila redko opažena v literaturi, čeprav izvira iz teorije govornih dejanj, je bila uporabljena na velikem obsegu podatkov in je dobro dokumentirana.

Shema DIT (Bunt 1994) predstavlja začetke posebne smeri označevanja dialoških dejanj. Nadgrajena je bila v DIT++ in uporabljena kot izhodišče za definiranje standarda ISO na področju označevanja dialoških dejanj (ISO 24617-2 2012).

Vsi navedeni poskusi označevanja dialoških dejanj so se razvijali primarno pod okriljem tehnološkega razvoja. Pod večjim vplivom jezikoslovja so zametki označevanja dialoških dejanj nastajali v okviru iniciative EAGLES (Leech idr. 1998). Na podlagi teh priporočil je bil predstavljen prvi poskus taksonomije dialoških dejanj (Leech, Weisser 2003), ki ga je potem Weisser razvijal naprej v shemo DART (Weisser 2019b).

Večina korpusov z oznakami dialoških dejanj obstaja za angleški jezik. Posamezni obsegajo tudi več kot 100 ur govora, mnogi pa samo po nekaj deset ur. Gradivo, ki ga označujejo, so večinoma študijski oziroma poustvarjeni diskurzi, pogosto na temo posredovanja informacij, izvajanja rezervacij, opravljanja nalog ali delovnih sestankov. Avtentično gradivo je uporabljeno izjemoma. Za neangleške jezike redkeje naletimo na korpus dialoških dejanj. Za slovenski jezik je bil konec 2019 razvit prvi enourni korpus dialoških dejanj GORDAN (Verdonik 2020). Seznam korpusov z oznakami dialoških dejanj je predstavljen v Verdonik idr. (2020).

3 Cilji in metodologija

V teoriji govornih dejanj je ilokucijsko dejanje opredeljeno kot dejanje vplivanja. To, kar izrečemo oz. zapišemo, ima neko vplivajnsko silo, namero, da pri naslovniku/ih nekaj doseže.

Različne označevalne sheme dialoških dejanj so ene bolj, druge manj eksplicitno izhajale iz teorije govornih dejanj, vendar jo vsaka po svoje prilagajale in razširjale oz. kot pravi Jurafsky (2003), so izhodiščno pojmovanje ilokucijskega dejanja bistveno razširile in označevale različne vrste komunikacijskih funkcij, ki jih izjava lahko opravlja. Zaradi tega se je namesto izvirnega termina govorno (oz. ilokucijsko) dejanje začel uporabljati termin dialoško dejanje. Toda v nasprotju z ilokucijskim je pojem

dialoškega dejanja, kot ugotavlja Jurafsky (2003), žal precej dvoumen. Pogosto pomeni kombinacijo ilokucijskega dejanja in semantične sile izjave ali dejanje, ki se navezuje na diskurzno/dialoško funkcijo, brez jasne opredelitve, kaj točno naj bi slednja bila. Posledica tega je, da različne sheme za označevanje dialoških dejanj mešajo različne vrste oznak, večinoma precej nesistematično in nekoherentno.

Vprašanje, katere vrste informacij o diskurzu označujejo sheme za označevanje dialoških dejanj, je tako vse prej kot enostavno, hkrati pa ima bistveno večji domet, kot se zdi na prvi pogled – skozi analizo shem in izkušenj z označevanjem namreč hkrati raziskujemo, *kaj vse v diskurzu delamo*. To pa je zelo sorodno izhodiščnemu vprašanju teorije govornih dejanj (Austin 1990: 85): »[K]oliko smislov je v tem, da je reči *kaj napraviti* kaj, oziroma, da napravimo kaj, *ko* rečemo kaj, in da napravimo kaj celo *tako, da* rečemo kaj.« Eksplicitno se tega zavedo tudi avtorji standarda ISO 24617-2 (2012), ki pri definiranju dimenzij dialoških dejanj izhajajo iz ponovnega premisleka o tem, »da mora udeleženec v diskurzu opravljati veliko stvari« (Bunt 2006: 921).

Cilj tega prispevka je raziskati, katere vrste pragmatičnega pomena označujejo različne sheme za označevanje dialoških dejanj in na kakšne probleme v zvezi s tem naletimo v praksi, pri označevanju avtentičnega gradiva. Vrsto pragmatičnega pomena pri tem razumemo kot skupino oznak, ki označujejo izjave iz neke skupne pragmatične perspektive, na primer: ilokucijska sila, informacije o kontekstu, vljudnostne strategije, kataforično-anaforični učinki, ton izjave ipd. Skozi to skušamo storiti korak naprej v razumevanju, *kaj vse v diskurzu delamo*.

Metodološko postopamo v analizi na način, da v prvem koraku analize (poglavje 4) izhajamo iz tega, kako so oznake kategorizirane in opredeljene v samih shemah, v diskusiji (poglavje 5) pa reintrepretiramo vrste oznak skupno za vse analizirane sheme.

Področno specifičnih označevalnih shem, ki so namenjene razvoju posameznih aplikacij, kot je poizvedovanje po informacijah o železniškem prometu (TRAINS) ali dogovarjanje za termin sestanka (Verbmobil), v analizo ne vključujemo, saj za splošno uporabo niso primerne. Med splošno veljavnimi shemami pa ločimo tri smeri, ki jih zajamemo v analizo:

- teorija govornih dejanj in sheme, ki izhajajo iz nje: SPICE-Ireland (Kirk 2013), DART (Weisser 2019b), AMI (AMI 2005)
- DAMSL (Allen, Core 1997) in njene izpeljanke, zlasti SWB-DAMSL (Jurafsky idr. 1997), MRDA (Dhillon idr. 2004) in MALTUS (Clark, Popescu-Belis 2004)
- ISO 24617-2 (2012) in prehodnice (DIT in DIT++)

4 Predstavitev in analiza označevalnih shem

V nadaljevanju izbrane označevalne sheme dialoških dejanj predstavimo, analiziramo oznake in na kratko povzamemo izkušnje iz označevanja.

4.1 Teorija govornih dejanj

Izhodiščno kategorizacijo ilokucijskih dejanj ponudi že Austin (1990), ki pa tudi sam omeni, da ni z vsemi kategorijami enako zadovoljen. Searle (1979) meni, da je največja pomanjkljivost Austinove taksonomije, da ni razvidno nobeno konsistentno načelo, na katerem bi temeljila. Zato naredi klasifikacijo na novo in loči:

- asertive: tvorec predstavi stvari take, kot so (velja pravilo resničnosti/neresničnosti)
- direktive: poskusi tvorca, da naslovnika pripravi do tega, da nekaj naredi
- komisive: tvorec se zaveže k nekemu prihodnjemu dejanju
- ekspresive: tvorčev odnos, občutja, duševna stanja do tega, kar je bilo izraženo v propoziciji
- deklarative: spremenijo realnost

Pri uporabi Searlove kategorizacije na avtentičnem gradivu so avtorji označevalnih shem DAMSL in DART (Allen, Core 1997; Weisser 2019b) navajali različne razloge, zakaj ni ustrezna, čeprav večinoma niso ustrezno argumentirani. Ob označevanju obsežnega korpusa SPICE-Ireland tako Kirk (2013) ugotavlja, da je za robustno označevanje velike količine raznovrstnega gradiva ravno Searlova taksonomija najbolj primerna. Vseeno Kirk (2013) doda štiri nove kategorije dejanj, s katerimi se ne Austin ne Searle ne ukvarjata, ob empiričnem delu z avtentičnim gradivom pa se jim ne moremo izogniti. To so:

- za konverzacijo relevantne nedoločne enote: metadiskurzni izrazi, kot so *ja, mhm, okej, v redu*
- družbeni izrazi: vljudnostne fraze, kot so pozdravi
- pragmatično nerazumljive enote: nedokončani ali opuščeni deli izjav
- ključni: izjave, ki so humorne, sarkastične, ironične, kjer izjavitelj ne misli resno tega, kar reče, in kjer normalni pogoji uporabe ne veljajo (Kirk 2013: 216)

Searlovo taksonomijo ilokucijskih dejanj smo preskusili na vzorcu iz korpusa GORDAN (Verdonik 2020). Pomanjkljivosti, ki smo jih ugotovili, so treh vrst: (1) v interakciji obstaja del jezikovnega materiala, ki mu ni mogoče pripisati neke prepoznavne ilokucijske sile in posledično ni mogoče označiti celotnega gradiva; (2) pogosto je lahko z isto izjavo izraženih več ilokucijskih dejanj ali pa vsaj ni mogoče jasno prepoznati enega točno določenega; (3) definirani razredi ilokucijskih dejanj so za korpusno analizo pregrobi, podroben seznam ilokucijskih dejanj pa nikoli ni bil izdelan, je vprašljivo, ali je končen, je zelo zahteven za interpretacijo in za korpusno označevanje preveč podroben.

Označevalna shema DART vsebuje dolg abecedni seznam 162 različnih dialoških dejanj. Ta naj bi bila definirana tako, da »označujejo dejanja, ki jih lahko izvedemo verbalno, in da oznake kar najbolj ustrezajo tistim, ki bi jih uporabili v vsakdanjem življenju, ko opisujemo ista dejanja« (Weisser 2019a: 11). Oznake sami ločimo na tiste,

ki se nanašajo na vsebinski oz. propozicijski del diskurza, in preostale, kar približno sovpadajo z delitvijo oznak v standardu ISO 24617-2 (2012).

Pri oznakah DART, ki jih uporabljamo za označevanje vsebinskega dela diskurza, opazimo precejšen odstop od ilokucijskih dejanj. Searle (1979: 368) opozarja, da ne smemo mešati ilokucijskih glagolov z ilokucijskimi dejanji. Dve od oznak sheme DART sta primera, ki ju v tem kontekstu Searle eksplicitno navaja kot ilokucijska glagola, ki ne izražata ilokucijskega dejanja: odgovoriti 'answer' in predlagati 'suggest'. Klasična ilokucijska dejanja, kot so ponuditi 'offer', opravičiti 'apologise', izjaviti 'state', so na seznamu oznak DART redka. Večina oznak bolj kot ilokucijsko dejanje povzema semantično silo izjave (npr. izraziti 'express', uvajati 'initiate', razložiti 'elaborate', dopolniti 'complete', pojasniti 'clarify', nasprotovati 'contradict', naštetih 'enumerate'), pri čemer so nekatere oznake pretirano podrobne (npr. prositi za informacijo o trajanju 'reqInfoDuration'). Zdi se, da so oznake DART precej stihijsko definirane in da nikakor niso zaprt nabor.

Oznake DART, s katerimi označimo nevsebinski del diskurza, interpretirajo (zelo poenostavljeno) nekatera metadiskurzna dejanja, ki se nanašajo večinoma na upravljanje poteka diskurza, procese tvorjenja, procese interpretacije in interakcijske procese: dokončati 'complete', samopopravi 'correctSelf' itd. Oznaka črkovati 'spell' pa označuje celo način, kako je nekaj izrečeno. Vljudnostne strategije označujejo oznake, kot sta posloviti se 'bye', pozdraviti 'greet', zahvaliti se 'thank'.

Ob preskusu sheme na korpusu GORDAN se pokaže, da so številne oznake dvoumne in izmenljive. Prav tako je opazno nesorazmerje med izredno podrobnimi oznakami na eni strani, kot je npr. izraziti upanje 'expressHope', in zelo splošnimi na drugi strani, kot je npr. izjaviti 'state', ki se nanaša na prenos informacije ali vedenja.

Shema AMI (2005) eksplicitno navaja, da označuje govorčev namen v izjavi. Pri tem loči štiri osrednje kategorije namena: izmenjava informacij (posredovanje ali poizvedovanje), akcije (direktivi, komisivi ter poizvedovanje v zvezi z direktivi in komisivi), komentiranje (angl. assess) in družbena dejanja (kjer ločijo samo pozitivna in negativna). Dodane so še tri posebne kategorije: oporni signali, zavlačevanje in fragment, z namenom, da se lahko označi celotno gradivo, eksplicitno pa avtorji AMI opozorijo, da te tri kategorije ne označujejo tvorčevega namena.

Shema AMI je enostavno uporabljiva v praksi, vendar je v primerjavi z drugimi shemami zelo groba in posledično nizko informativna.

4.2 DAMSL in njene izpeljanke

Allen in Core (1997) uporabljata termin komunikacijska dejanja, ne dialoška dejanja. Izjava kot temeljna enota označevanja, torej kot nosilka dejanja, je v shemi DAMSL definirana na podlagi teorije govornih dejanj in temelji na analizi tvorčevih namenov

(Allen, Core 1997: 2). Izjavi se pripisujejo različne vrste pomena, ki so kategorizirane v štiri dimenzije (Allen, Core 1997), opisane v nadaljevanju.

Prva dimenzija, komunikacijski status, označi, ali je izjava razumljiva in dokončana oz. ali gre za samogovor. Na tem nivoju torej shema DAMSL loči tiste dele diskurza, ki za razumevanje diskurza niso posebej relevantni, od tistih, ki so.

Druga dimenzija, informativna raven, zelo splošno in abstraktno označi vsebino izjave. Pri tem empirično izhaja iz diskurzov, katerih namen je izvajanje določene naloge (npr. rezervacija termina, pridobitev informacij ipd.). Če njene kategorije interpretiramo na bolj splošni ravni, ločuje izjave, ki so na neki način metadiskurzne (upravljajo diskurz), in tiste, ki so prozodijske in izvajajo neko akcijo oz. nalogo. Nekaterih ne morejo jasno ločiti v nobeno od kategorij, zato jih označijo kot 'ostale'. Izjave, ki so povezane z izvajanjem akcije/naloge, ločijo na take, ki prispevajo k napredovanju te akcije (imajo namero), in take, ki se nanašajo na dogovarjanje o tem, kako akcijo uspešno izvesti (upravljanje izvajanja akcije).

Tretja dimenzija, neodzivi oz. naprej usmerjene funkcije, označuje, kako izjave vplivajo na prihodnje dogajanje/prihodnji dialog. Pri tem izhaja DAMSL iz Searlove klasifikacije ilokucijskih dejanj, vendar označevanje ne poteka z vidika tvorca, ampak iz (neke vrste) perlokucije, tj. kakšen učinek ima izjava na nadaljevanje dialoga (Allen, Core 1997: 9), pri čemer dopušča večfunkcionalnost. V tej dimenziji DAMSL loči:

- izjave, ki omogočajo preverjanje resničnostnih pogojev (je res/ni res) in ustrezajo asertivom po Searlu
- izjave, ki od naslovnika pričakujejo neko informacijo
- izjave, ki skušajo vplivati na naslovnikovo dejanje v prihodnosti in ustrezajo direktivom po Searlu
- izjave, ki zavezujejo tvorca, da bo nekaj naredil v prihodnosti, in ustrezajo komisivom po Searlu
- konvencionalne fraze, kot so pozdravi
- eksplicitni performativ
- vzklike
- drugo

Četrta dimenzija, odzivi oz. nazaj usmerjene funkcije, označuje, kako se izjava navezuje na predhodni diskurz, pri čemer se označi tako vrsta tega navezovanja kot tisti del diskurza, na katerega se navezuje. Loči:

1. odzive, ki se nanašajo na izražanje (ne)strinjanja
2. odzive, ki se nanašajo na izraze (ne)razumevanja
3. podajanje informacij kot odgovor
4. razmerja med informacijami v izjavah, ki pa niso podrobneje razvita, ampak omogočajo prosto mesto za uporabnike sheme, da jih definirajo po svoje

Ključni problem sheme DAMSL pri označevanju je njena večdimenzionalnost, ki po eni strani omogoča nevzdržno veliko množico različnih kombinacij oznak, po drugi strani pa številne kombinacije v praksi niso nikoli uporabljene. Večini dejanj tako od možnih štirih dimenzij pripišemo po dve dimenziji.

V projektu WS97 Language Modeling so shemo DAMSL preoblikovali, tako da so več dimenzij združili v eno, izpustili oznake, ki niso bile dovolj pogosto ali sploh ne uporabljene, ter dodali nekatere nove, od katerih mnoge konceptualno temeljijo na teoriji bližnjih parov konverzacijske analize. Tako je nastala shema SWBD-DAMSL (Jurafsky idr. 1997). Njen namen je bil omogočiti hiter in enostaven način označevanja za govorni korpus Switchboard, ki bi pomagal pri razvoju avtomatskega razpoznavanja govora. K izbranim oznakam sheme DAMSL tako shema SWBD-DAMSL doda nekaj novih, ki so naslednjih vrst:

- komu je izjava namenjena: udeležencem v diskurzu, samemu sebi (samogovor, retorična vprašanja) ali nekomu tretjemu, ki ni udeleženec potekajočega diskurza
- kdo je pravi avtor izjave: shema doda oznake za dobesedni govor
- opazen dodatek je podrobna specifikacija vprašanj in odgovorov na različne oblikovno-pomenske tipe, kot so »da odgovor«, »ne odgovor«, »odgovor da plus dodatek« oz. »da-ne vprašanje«, »K-vprašanje«, »odprto vprašanje«, »ali-vprašanje« itd.
- v sklopu odzivov, ki se nanašajo na izražanje (ne)razumevanja, se začnejo dodajati oznake za izraze t.i. odnosa (angl. attitude): hvaležnost, sočutje, zmanjševanje pomena, zadržek (hedges)

Nabor konvencionalnih fraz se razširi z zahvalo, sprejemanjem zahvale in opravičilom.

Ključni problem označevanja s shemo SWBD-DAMSL je njena eklektičnost: s tem ko združi različne, sicer jasno razmejene dimenzije pragmatičnega pomena iz sheme DAMSL v enodimenzionalne oznake, izgubi notranjo koherenco.

V projektu ICSI so delali s posnetki delovnih sestankov. Kot izhodišče za označevanje dialoških dejanj so vzeli shemo SWBD-DAMSL, vendar ponovno nekaj oznak izpustili, nekaj pa dodali in novo shemo poimenovali MRDA (Dhillon idr. 2004).

Najdemo naslednje nove skupine oznak:

- menjavanje vlog: prevzem vloge (kot taki so običajno označeni začetni metadi-skurzni izrazi, kot so *torej, ja, no, zdaj* ...) ali zadrževanje vloge (izrazi, kot sta *eee* ali *v bistvu*)
- menjavanje tem: posebne oznake za izjave, ki bodisi začenjajo bodisi končujejo neko temo
- preverjanje razumevanja: tvorec preveri, ali je pravilno razumel informacije oz. ali ga naslovnik posluša, razume, mu sledi, npr. z diskurzivnimi označevalci *a ne?*, *a veš*
- argumenti: k odzivom sheme DAMSL je dodana oznaka za izjave, ki dodatno argumentirajo ali branijo neko stališče tvorca

- humor: dodana je oznaka za humor

Za označevanje s shemo MRDA velja podobno kot za DAMSL, notranje teoretske koherence, kaj z njo označujemo, ni. Ker ponovno omogoča kopičenje različnih oznak eno na drugo, se hkrati vrne problem prevelike množice možnih kombinacij oznak.

Schema Maltus (Clark, Popescu 2004) izhaja iz sheme MRDA, s tem da omejuje nabor možnih kombinacij med različnimi oznakami MRDA. Ne dodaja nobenih novih vrst oznak.

4.3 ISO 24617-2

Standard ISO 24617-2 je nadgradnja DIT++, slednja pa je nastala na podlagi shem DIT in DAMSL. Izhaja iz DAMSL-ove informativne ravni, vendar loči samo med dvema osrednjima sklopoma: na eni strani so komunikacijske oz. metadiskurzne izjave, na drugi strani izjave, ki se navezujejo na izvajanje akcije/naloge. S tem standard ISO 24617-2 na neki način sledi znani delitvi v uporabnem jezikoslovju na metadiskurz in propozicijsko vsebino (Hyland 2005; Pisanki Peterlin 2011).

Že pri shemi DAMSL smo izpostavili, da avtorji opozarjajo na to, da se v diskurzu izvajajo hkrati različne vrste dejanj in posledično označujejo dialoška dejanja z vidika več različnih dimenzij. Avtorji ISO 24617-2 se navezujejo na predhodne razprave (Bunt 2006), v katerih na novo razpravljajo o problemu, katere vrste komunikacijskih aktivnosti udeleženci diskurza opravljajo med diskurzom, in ugotavljajo, da to nikakor ni samo izvajanje določenih (vsebinskih) nalog ali dejanj, ampak tudi komunikacija o tem, ali in kako se razumejo, signaliziranje izmenjavanja vlog, pridobivanje časa za tvorjenje, vzdrževanje stika, izražanje pozornosti ipd. Te vrste dejanj ISO 24617-2 opredeli kot različne vrste diskurznih dimenzij, ki so naslednje:

- naloga
- odziv o lastnem razumevanju, pozornosti ali mnenju
- preverjanje sogovornikovega razumevanja, pozornosti ali mnenja
- menjavanje vlog
- upravljanje časa oz. pridobivanje časa za tvorjenje ali interpretacijo
- strukturiranje diskurza (menjavanje tem, začenjanje, zaključevanje)
- upravljanje družbenih razmerij (upoštevanje konvencij družbenega obnašanja)
- upravljanje lastne komunikacije (npr. samopopravljanja)
- upravljanje sogovornikove komunikacije (popravljanje ali dopolnjevanje sogovornika)

Od navedenih dimenzij lahko vse razen prve opredelimo kot metadiskurzne. Problem je, da ISO 24617-2 pri tem ne upošteva, da se praviloma mnoge navedene dimenzije izvajajo hkrati, saj je metadiskurz notorično večfunkcijski.

Druga raven označevanja pomena po ISO 24617-2 so komunikacijske funkcije. Tiste, ki so vezane na dimenzijo naloge, torej propozicijske vsebine, se delijo v štiri tematske sklope:

- iskanje informacij
- podajanje informacij
- komisivi (tvorec se zaveže k nekemu prihodnjemu dejanju)
- direktivi (tvorec skuša pripraviti naslovnika do nekega dejanja)

Komunikacijske funkcije, ki so vezane na metadiskurz, so bolj kot ne specifične za vsako posamezno dimenzijo, navedeno zgoraj, in jo samo podrobneje opišejo.

Dodatno predvideva ISO 24617-2 tudi možnost, da se k dialoškim dejanjem dodajajo neobvezni atributi. Ti so treh vrst (ISO 24617-2 2012: 86–87):

- oznake pogojnosti: ali tvorec presoja, da obstajajo določeni pogoji za izvedbo dejanja
- oznake zanesljivosti: koliko je tvorec prepričan o pravilnosti informacije
- oznake sentimenta: tvorec izraža svoje čustveno stanje oz. odnos do vsebine ali do sogovornika

Označevanje z ISO 24617-2 kaže, da je seznam dimenzij in pripadajočih komunikacijskih funkcij poenostavljen in ne upošteva v zadostni meri večfunkcijskosti metadiskurza, seznam oznak za propozicijsko vsebino oz. dimenzijo naloge pa preskromen in neustrezen za vse vrste informacij, ki jih izražamo skozi diskurz. Dopolnjevanje z atributi za pogojnost, zanesljivost in sentiment je tako nujno, da vsaj nekoliko natančneje označimo vrsto dialoškega dejanja.

4.4 Empirični preskus označevanja

Pregled in analiza označevalnih shem potrjujeta in podrobno razgrneta problematiko označevanja dialoških dejanj na korpusnem gradivu. Vidimo, da ni enotnega pristopa niti enotnega razumevanja, kaj naj bi dialoška dejanja v korpusnem gradivu, od vrstice do vrstice, dejansko označevala. Zdi se, da je število različnih vrst dejanj, ki jih skozi diskurz izvajamo, veliko večje, kot razlaga teorija govornih dejanj, zlasti na račun raznih metadiskurznihih aktivnosti med komunikacijo, pri čemer pa je enoznačno označevanje zaradi večfunkcijskosti teh dejanj tako rekoč nemogoče.

Za jasnejšo predstavo problematike v tem poglavju na kratkem izseku pogovora prikažemo označevanje dialoških dejanj po vseh analiziranih shemah. Odsek je iz pogovora, v katerem se družinski prijatelji dogovarjajo o skupnih smučarskih počitnicah. Vzet je iz govornega korpusa GOS (Verdonik idr. 2013), tako da se vsaka od izjav lahko tudi posluša na spletni povezavi [korpusa](#). Prostor nam ne omogoča več kot analizo le nekaj začetnih izjav pogovora.

Govorka A, izjava 1: *no v glavnem lej*

Govorka A, izjava 2: *gljeh dones sem dobila talele za smučarijo ne*

Govorka A, izjava 3: *taku sam zarad cen da videš ta del kamer bi mi šli*

Govorka A, izjava 4: *če te zanima*

Govorec B, izjava 5: *ja seveda mi smo že čist tako rekoč se že pripravljamo*

Vsako izjavo analiziramo z vidika izraženih ilokucijskih/dialogskih dejanj, najprej z lastno interpretacijo, nato pa po vseh analiziranih shemah. Oznake opišemo, če v shemi niso same po sebi dovolj opisne. Nekatero izjavo dodatno razdelimo z znakom +, saj lahko po posameznih shemah pomenijo več osnovnih enot za označevanje. Interpretacije zaradi omejitve prostora predstavimo samo za tri izjave.

Izjava 1: *no v glavnem glej*

- interpretacija: z izjavo naredi tvorka uvod v temo, pritegne pozornost sogovornika in se sklicuje naprej na to, kar bo sledilo
- kategorija po Searlu: ne sodi v nobeno od kategorij
- oznaka po DART: začetek nove faze dialoga
- oznaka po AMI: ni posebne primerne oznake
- oznake po DAMSL: dimenzija informativne ravni, upravljanje komunikacije; dimenzija naprej usmerjenih funkcij, ni primerne oznake
- SWBD-DAMSL: ni primerne oznake
- MRDA: prevzem vloge
- ISO 24617-2: dimenzija 'strukturiranje diskurza', funkcija 'strukturiranje interakcije'

Izjava 2: *gljeh dones sem dobila talele za smučarijo + ne*

- interpretacija: poda relevantno informacijo o preteklem dogodku + poudari, pritegne pozornost, ohranja stik s sogovorniki
- kategorija po Searlu: asertiv
- DART: izjava + semantično prazen element
- AMI: informiranje
- DAMSL: dimenzija informativne ravni, izvajanje aktivnosti; dimenzija naprej usmerjenih funkcij, asertivna izjava
- SWBD-DAMSL: izjava, ki ne izraža mnenja
- MRDA: izjava + izjava – vprašanje tipa 'tag'
- ISO 24617-2: dimenzija 'naloga', funkcija 'informirati' + dimenzija 'preverjanje sogovornikovega razumevanja', funkcija 'pozivanje k odzivu'

Izjava 5: *ja seveda + mi smo že čist + tako rekoč se že pripravljamo*

- interpretacija: potrditev predloga + popravljena enota + poda informacijo o dogajanju
- kategorija po Searlu: komisiv + ni oznake + asertiv

- DART: odgovor + opuščeno + izjava
- AMI: informiranje
- DAMSL: dimenzija nazaj usmerjenih funkcij, 'sprejeti' + komunikacijski status, 'opuščno' + dimenzija informativne ravni, 'izvajanje aktivnosti'; dimenzija naprej usmerjenih funkcij, 'zaveza'
- SWBD-DAMSL: da-odgovor + opuščeno + zaveza
- MRDA: izjava/sprejeti + opuščeno + izjava/zaveza
- ISO 24617-2: dimenzija 'naloga', funkcija 'potrditev' + dimenzija 'upravljanje lastne komunikacije', funkcija 'umik' + dimenzija 'naloga', funkcija 'obljuba'

Iz primera vidimo, da v praksi ni niti enotnega razumevanja, kaj predstavlja osnovno enoto označevanja, niti kako to enoto interpretirati v smislu ilokucijskih oz. dialoških dejanj niti kako podrobno oz. na drugi strani posplošeno to interpretacijo izvesti.

5 Vrste pragmatičnega pomena v dialoških korpusnih oznakah

Ob označevanju korpusnega gradiva se problematika govornih dejanj razpre v širši perspektivi in povezana s sorodnimi teorijami ali področji, ki k diskurzu pristopajo iz drugačnih izhodišč, a se soočajo s podobnimi fenomenami (npr. konverzacijska analiza in uporabo jezikoslovje). Sinteza rezultatov analiz pokaže nadvse zanimivo sliko različnih vrst pragmatičnega pomena, pa čeprav je v shemah ponekod bolj, drugod manj sistematično in/ali podrobno zajet v oznakah. Na podlagi analiz v poglavju 4 opredelimo oznake v analiziranih shemah v naslednje kategorije:

1. Komunikacijski status izjave:
 - a. ali je izjava nerelevantna za vsebino diskurza (samopopravljanja, samogovor)
 - b. ali je izjava opuščena
2. Kontekstne informacije o izjavi:
 - a. kdo je pravi tvorec (dobesedni in poročani govor)
 - b. komu je namenjena (ne sogovorniku, ampak tvorcu samemu ali nekemu zunaj diskurza)
3. Kataforično-anaforične in deiktične vezi: se izjava navezuje naprej ali nazaj v diskurzu, na vsebino, na naslovnika ...
4. Ilokucijska sila:
 - a. tvorec se zaveže k aktivnosti
 - b. tvorec skuša naslovnika pripraviti k aktivnosti
 - c. tvorec podaja informacije
 - d. tvorec poizveduje po informacijah
 - e. tvorec izraža mnenje, odnos, sentiment

5. Preverjanje in izražanje strinjanja ali nestrinjanja
6. Zadržki (angl. hedges) oz. prepričanost o pravilnosti informacij
7. Oblikovno-semantične lastnosti izjave: »da odgovor«, »ne odgovor«, »odgovor da plus dodatek« oz. »da-ne vprašanje«, »K-vprašanje«, »odprto vprašanje«, »ali-vprašanje« itd.
8. Semantična sila: zlasti pri shemi DART se kaže težavnost ločevanja med pragmatičnim in semantičnim pomenom
9. Ton izjave: humoren, satiričen ...
10. Vljudnostne konvencije: fraze pozdravljanja, poslavljanja, predstavljanja, zahvaljevanja, opravičevanja ipd.
11. Metadiskurzne funkcije izjave:
 - a. signaliziranje o poteku diskurza (menjavanje vlog, prehodi med temami, uvajanje novih tem, začetek, konec diskurza)
 - b. signaliziranje o uspešnosti sporazumevanja
 - c. signaliziranje procesov tvorjenja in interpretacije
 - d. nezaznana ostane medosebna metadiskurzna funkcija, čeprav je sicer iz uporabnega jezikoslovja dobro znana.
12. Način, kako je izjava podana: vzkliki, črkovanje ...

6 Diskusija in zaključek

Kaj smo skozi analizo ugotovili o tem, kaj vse v diskurzu delamo? Potrjuje se ločnica med dvema osrednjima kategorijama, ki ju iz teoretskih osnov systemske funkcijske slovnice izpelje že uporabno jezikoslovje: propozicijska vsebina na eni in metadiskurz na drugi strani. Čeprav iz uporabnega jezikoslovja vemo, da besedila ni mogoče enoumno ločiti na eno in drugo, so v govorjenem diskurzu, zlasti v vsakdanjem pogovoru, vendarle mnoge izjave izrazito takšne, da ne prispevajo k propozicijski vsebini, in za te izjave je vedno znova problem, kako jih označiti iz vidika dialoških dejanj.

Pri teh izjavah je tudi največ razlik v označevanju. Ključni problem je njihova več-funkcijskost. Kakor hitro jih začnemo strogo razporejati v posamezne metadiskurzne razrede, smo jih nepopolno označili; če dopuščamo kopičenje več različnih oznak, se zapletemo v mrežo subjektivnega in nedoslednega označevanja, kjer je neka metadiskurzna izjava lahko skoraj karkoli.

Propozicijska vsebina se v praksi označevanja pogosto razporeja v nekoliko drugačne razrede, kot jih je opredelil Searle (1979). Komisivi in direktivi so bili dokaj

široko sprejeti in prevzeti, vendar večkrat združeni v eno skupino – aktivnosti. Od te skupine se ločuje druga velika skupina, izmenjava informacij, kjer se vzpostavi ločnica med dajanjem informacij in poizvedovanjem po informacijah. Veliko nedoslednosti, različnih označevalnih modelov in nejasnosti je v zvezi s skupino izjav, v katerih tvorec bodisi izraža svoja mnenja, odnos, sentiment bodisi sprašuje tvorca o tem. Deklarativi pa v večini shem niso bili prevzeti, verjetno je razlog predvsem ta, da v korpusnem gradivu niso bili prisotni.

Poleg ilokucijske sile in metadiskurznihi funkcij smo opredelili še več specifičnih kategorij, od katerih so mnoge zelo relevantne za celotno sliko o pomenih, posredovanih v interakciji (komunikacijski status; kontekstne informacije; anaforično-kataforične in deiktične vezi; vljudnostne konvencije ipd.), nekatere odražajo problematiko ločevanja med semantiko in pragmatiko (semantična sila; oblikovno-semantične lastnosti; zadržki ipd.), druge pa problematiko izločanja besedila iz njegovega večmodalnega okolja, v katerem nastopa v govoru (način, kako je izjava podana; ton ipd.).

VIRI IN LITERATURA

- Karin AIJMER, Christoph RÜHLEMANN, 2015: *Corpus Pragmatics: A Handbook*. Cambridge: Cambridge University Press.
- James ALLEN, Mark CORE, 1997: *Draft of DAMSL: Dialog Act Markup in Several Layers*. [Na spletu](#).
- AMI, 2005: *Guidelines for Dialogue Act and Addressee Annotation Version 1.0*. [Na spletu](#).
- John L. AUSTIN, 1990: *Kako napravimo kaj z besedami*. Ljubljana: Studia Humanitatis.
- Shoshana BLUM-KULKA, Juliane HOUSE, Gabriele KASPER (ur.), 1989: *Cross-Cultural Pragmatics: Requests and Appologies*. Nordwood: Ablex.
- Harry BUNT, 1994: Context and Dialogue Control. *Think Quarterly* 3. 19–34.
- Harry BUNT, 2006: Dimensions in dialogue act annotation. *Zbornik konference International Conference on Language Resources and Evaluation, LREC' 06*. Genova: ELRA, ELDA, ILC. 919–24.
- Harry BUNT, 2009: The DIT++ taxonomy for functional dialogue markup. *Towards a Standard Markup Language for Embodied Dialogue Acts: Proceedings*. Ur. Dirk Heylen, Catherine Pelachaud, Roberta Catizone, David Traum. Budimpešta. 36–48.
- Harry BUNT, Volha PETUKHOVA, Andrei MALCHANAU, Alex FANG, Kars WIJNHOFEN, 2018: The DialogBank: dialogues with interoperable annotations. *Language Resources and Evaluation* 53. 213–49.
- Alexander CLARK, Andrei POPESCU-BELIS, 2004: Multi-level Dialogue Act Tags. *Zbornik konference 5th SIGdial Workshop on Discourse and Dialogue at HLT-NAACL*. Cambridge: Association for Computational Linguistics. 163–70.
- Rajdip DHILLON, Sonali BHAGAT, Hannah CARVEY, Elisabeth SHRIBERG, 2004: *Meeting Recorder Project: Dialog Act Labeling Guide*. ICSI Technical Report TR-04-002. [Na spletu](#).

- Ekkehard FELDER, Marcus MÜLLER, Friedemann VOGEL (ur.), 2011: *Korpuspragmatik: Thematische Korpora als Basis diskurslinguistischer Analysen*. Berlin, Boston: Walter de Gruyter.
- Ernesto Wong GARCÍA, 2019: Toward a pragmatic account and taxonomy of evaluative speech acts. *Pragmatics* 29/1. 107–31.
- Kallirroi GEORGILA, Oliver LEMON, James HENDERSON, Johanna D. MOORE, 2009: Automatic annotation of context and speech acts for dialogue corpora. *Natural Language Engineering* 15/3. 315–53.
- Matejka GRGIČ, Igor Ž. ŽAGAR, 2004: *Čas in dejanje v jeziku: oblikovanje performativne teorije na Slovenskem*. Ljubljana: Založba CF.
- Meta GROSMAN, 1997: Kritika kot govorno dejanje. *Slavistična revija* 45/1–2. 269–92.
- Michael HANCHER, 1979: The classification of cooperative illocutionary acts. *Language in Society* 8. 1–14.
- Sonja HUDEJ, 2002: Uspešnost uresničevanja tvorčevega namena v šestih besedilnih vrstah. *Slavistična revija* 50/1. 61–81.
- Ken HYLAND, 2005: *Metadiscourse: Exploring Interaction in Writing*. London, New York: Continuum.
- ISO 24617-2, 2012: *Language resource management – Semantic annotation framework (SemAF): Part 2: Dialogue acts*. Reference number ISO 24617-2:2012(E). Geneva.
- Andreas H. JUCKER, Irma TAAVITSAINEN, Jukka TUOMINEN (ur.), 2014: *Diachronic Corpus Pragmatics*. Amsterdam: John Benjamins.
- Dan JURAFSKY, 2003: Pragmatics and computational linguistics. *The Handbook of Pragmatics*. Ur. Laurence R. Horn, Gregory Ward. Oxford: Blackwell Publishing Ltd.
- Daniel JURAFSKY, Elisabeth SHRIBERG, Debra BIASCA, 1997: *Switchboard SWBD-DAMSL shallow-discourse-function annotation. Coders manual, draft 13*. University of Colorado at Boulder & +SRI International. [Na spletu](#).
- John M. KIRK, 2013: Beyond the Structural Levels of Language: An Introduction to the SPICE-Ireland Corpus and its Uses. *After the storm: Papers from the Forum for Research on the Languages of Scotland and Ulster Triennial Meeting*. Ur. Janet Cruickshank, Robert McColl Millar. Aberdeen: Forum for Research on the Languages of Scotland and Ireland. 207–32.
- Mira KRAJNC IVIČ, 2010: Jezikovno-vedenjski vzorci opravičevanja v slovenščini. *Slavistična revija* 58/1. 133–47.
- Mira KRAJNC IVIČ, 2020: Obravnava besedil: Merila za razlikovanje med besedilno vrsto in besedilnim tipom. *Slavistična revija* 68/1. 55–71.
- Pavel KRÁL, Christophe CERISARA, 2014: Automatic dialogue act recognition with syntactic features. *Language Resources and Evaluation* 48. 419–41.
- Olga KUNST GNAMUŠ, 1984: *Govorno dejanje – družbeno dejanje: Komunikacijski model jezikovne vzgoje*. Ljubljana: Pedagoški inštitut pri Univerzi Edvarda Kardelja Ljubljana.
- Geoffrey LEECH, Martin WEISSER, Andrew WILSON, Martine GRICE, 1998: *LE-EAGLES-WP4-4: Integrated Resources Working Group Survey and Guidelines for the Representation and Annotation of Dialogue*. [Na spletu](#).
- Geoffrey LEECH, Martin WEISSER, 2003: Generic speech act annotation for task-oriented dialogues. *Zbornik konference 2003 Corpus Linguistics Conference*. Ur. Dawn

- Archer, Paul E. Rayson, Andrew Wilson, Tony McEnery. Lancaster University: Centre for Computer Corpus Research on Language Technical Papers. 441–46.
- Stephen C. LEVINSON, 2016: Speech acts. *Oxford Handbook of Pragmatics*. Ur. Yan Huang. [Na spletu](#).
- Mojca NIDORFER-ŠIŠKOVIČ, 2004: Kako izrekamo zahteve. *Moderno v slovenskem jeziku, literaturi in kulturi: zbornik predavanj*. Ur. Marko Stabej. Ljubljana: Center za slovenščino kot drugi/tuji jezik pri Oddelku za slovenistiko Filozofske fakultete. 182–86.
- Agnes PISANSKI PETERLIN, 2011: *Metabesedilo med dvema kulturama*. Ljubljana: Univerza v Ljubljani, Znanstvena založba Filozofske fakultete.
- Jesús ROMERO-TRILLO (ur.), 2008: *Pragmatics and Corpus Linguistics: A Mutualistic Entente*. Berlin, New York: Mouton de Gruyter.
- John R. SEARLE, 1972: Chomsky's Revolution in Linguistics. *The New York Review of Books*. 29. junij 1972.
- John R. SEARLE, 1979: *Expression and Meaning: Studies in the Theory of Speech Acts*. Cambridge: Cambridge University Press.
- Dominic STEWART, 2010: *Semantic prosody: a critical evaluation*. London, New York: Routledge.
- Mojca ŠORLI, v tisku: *Semantična proizvodnja: Leksikalni in besedilno-diskurzivni vidiki*.
- Anna TROSBORG, 1995: *Interlanguage Pragmatics: Requests, Complaints, and Apologies*. Berlin: Mouton de Gruyter.
- Darinka VERDONIK, 2020: *The GORDAN scheme for dialogue act annotation - version 1.0: technical report*. [Na spletu](#).
- Darinka VERDONIK, Iztok KOSEM, Ana ZWITTER VITEZ, Simon KREK, Marko STABEJ, 2013: Compilation, transcription and usage of a reference speech corpus: the case of the Slovene corpus GOS. *Language Resources and Evaluation* 47/4. 1031–48.
- Darinka VERDONIK, Andrej ŽGANK, Simona MAJHENIČ, Izidor MLAKAR, 2020: *Govorni, dialoški in multimodalni jezikovni viri: pregled stanja*. Vsebinsko poročilo. [Na spletu](#).
- Martin WEISSER, 2014: Speech Act Annotation. *Corpus Pragmatics: A Handbook*. Ur. Karin Aijmer, Christoph Rühlemann. Cambridge University Press. 84–113.
- Martin WEISSER, 2019a: The DART annotation scheme: form, applicability & application. *Studia Neophilologica* 91/2. 131–53.
- Martin WEISSER, 2019a: *The DART Taxonomy v. 3*. [Na spletu](#).
- Igor Ž. ŽAGAR, 2018: *Od performativa do govornih dejanj*. Ljubljana: Pedagoški inštitut.

SUMMARY

Speech acts have been recognized as an interesting topic not merely in pragmalinguistics, but also in speech technologies. When developing the first systems for communication between computers and humans, researchers have found that annotating speech acts is useful. However, when applying the original speech act theory (i.e., the concept of illocutionary acts) to recorded speech data, for various reasons significant changes, adaptations, and broadenings of the original concept have been made. Simultaneously, the term has changed from speech acts into dialogue acts. However, this (partially) new concept is unfortunately also very ambiguous.

A number of dialogue act annotation schemes were developed, domain-oriented as well as generic, but each of them followed a different internal logic and different comprehension of the notion of dialogue act. The aim of this paper is to analyse what types of pragmatic meaning different annotation schemes annotate and what problems we encounter when using these schemes for annotating authentic data.

Based on our analyses of different lines of dialogue act annotation schemes, we have defined a number of different levels of pragmatic meaning that these schemes implicitly annotate: communicative status (e.g., abandoned sentences, self-repairs); contextual information (e.g., reported speech, self-talk); cataphoric-anaphorical, and deictic bonds (e.g., forward looking functions; backward looking functions); illocutionary force (e.g., offer, commit, instruct); expressing agreement; hedges; morphological-semantic characteristics (e.g., or-question; yes answer plus addition); semantic force (e.g., express assumption, state require information); framing (e.g., humorous; satiric); politeness strategies (e.g., greeting, thanking, apologizing); metadiscourse functions (e.g., turn-taking, signalling understanding, hesitating); and mode of speaking (e.g., exclaiming, spelling).