

Multi-CAST

Sanzhi Dargwa
list of referents

Diana Forker
University of Jena

May 2019
v1.0

Nils Norman Schiborr
University of Bamberg



ARC CENTRE OF EXCELLENCE FOR
THE DYNAMICS OF LANGUAGE



Australian Government
Australian Research Council



University of Bamberg

DFG

Multi-CAST

Multilingual Corpus of
Annotated Spoken Texts

Citation for this document

Forker, Diana & Schiborr, Nils N. 2019. Multi-CAST Sanzhi Dargwa list of referents. In Haig, Geoffrey & Schnell, Stefan (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts*. (multicast.aspra.uni-bamberg.de/) (date accessed)

Citation for the Multi-CAST collection

Haig, Geoffrey & Schnell, Stefan (eds.). 2015. *Multi-CAST: Multilingual corpus of annotated spoken texts*. (multicast.aspra.uni-bamberg.de/) (date accessed)

The Multi-CAST collection has been archived at the *University of Bamberg*, Germany, and is freely accessible online at multicast.aspra.uni-bamberg.de/.

The entirety of Multi-CAST, including this document, is published under the *Creative Commons Attribution 4.0 International Licence* (CC BY 4.0), unless noted otherwise. The licensing terms can be reviewed online at creativecommons.org/licenses/by/4.0/.

Multi-CAST Sanzhi Dargwa list of referents v1.0 last updated 11 May 2019
This document was typeset by NNS with X_YL^AT_EX and the *multicast3* class (v3.0.9045).

Contents

1	Preface to the list of referents	1
2	Lists of referents	2
2.1	<i>asabali</i>	2
2.2	<i>bazhuk</i>	3
2.3	<i>dragon</i>	4
2.4	<i>kurban</i>	5
2.5	<i>mill</i>	6
2.6	<i>patima</i>	7
2.7	<i>ramazan</i>	8
2.8	<i>tape</i>	10

1 Preface to the list of referents

This document collects lists of the indexed referents in version 1905 (from May 2019) of the Multi-CAST Sanzhi Dargwa corpus. The information listed here are also available in a TSV file format from the Multi-CAST website. Please refer to the RefIND annotation guidelines (Schiborr et al. 2018) for details on the annotations.

Semantic classes of referent

hum	<i>human</i>	human beings and anthropomorphized non-humans
anm	<i>animate</i>	animals
inm	<i>inanimate</i>	inanimate physical objects
bdp	<i>body part</i>	body parts of human beings and animals
mss	<i>mass</i>	non-individuable masses like water or sand
loc	<i>location</i>	physical locations and places
tme	<i>time</i>	points or periods of time
abs	<i>abstract</i>	emotions, thoughts, ideas, speech, etc.

Relations between referents

- < *set member of* (partial co-reference)
is member of a previously introduced referent containing multiple entities
- > *includes* (split antecedence)
contains multiple entities previously established as separate referents
- M *part-whole*
is part of another entity or mass that was previously established

References

Schiborr, Nils N. & Schnell, Stefan & Thiele, Hanna. 2018. *RefIND — Referent Indexing in Natural-language Discourse: Annotation guidelines (v1.1)*. Bamberg / Melbourne: University of Bamberg / University of Melbourne. (<http://multicast.aspra.uni-bamberg.de/#annotations>) (Accessed 2019-03-08).

2 Lists of referents

2.1 *asabali*

<i>index</i>	<i>label</i>	<i>description</i>	<i>class</i>	<i>relations</i>	<i>notes</i>
0000	<i>people</i>		hum		
0001	<i>the war</i>		abs		
0002	<i>years of 0001</i>		tmp		
0003	<i>speaker</i>		hum		
0004	<i>name of 0003</i>		abs	M 0003	
0005	<i>record book</i>		inm		
0006	<i>village head</i>		hum		
0007	<i>the year 1953</i>		tmp		
0008	<i>Stalin</i>		hum		
0009	<i>cars</i>		inm		
0010	<i>trains</i>		inm		
0011	<i>mountains</i>		loc		
0012	<i>sound of 0010</i>		inm		
0013	<i>there</i>		loc		
0014	<i>sheep</i>		anm		
0015	<i>kolkhoz</i>		abs		
0016	<i>army</i>		abs		
0017	<i>enlistment office</i>		inm		
0018	<i>application</i>		inm		
0019	<i>train</i>		inm		
0020	<i>army officials</i>		hum		
0021	<i>0003 and other recruits</i>		hum	> 0003	
0022	<i>Voronezh</i>		loc		
0023	<i>a captain</i>		hum		
0024	<i>soldiers</i>		hum		
0025	<i>hospital</i>		inm		
0026	<i>a corridor in 0025</i>		loc		
0027	<i>commissariat</i>		loc		
0028	<i>Dargi people</i>		hum	< 0021	
0029	<i>canteen</i>		loc		
0030	<i>porridge</i>		mss		
0031	<i>half of 0030</i>		mss	M 0030	
0032	<i>other half of 0030</i>		mss	M 0030	
0033	<i>eyes of 0003</i>		bdp	M 0003	
0034	<i>a train</i>		inm		
0035	<i>army officials</i>		hum	> 0023	
0036	<i>Kursk</i>		loc		
0037	<i>Kursk station</i>		loc		
0038	<i>car of 0035</i>		inm		
0039	<i>barracks</i>		loc		
0040	<i>the Russian language</i>		abs		

0041	people	hum	
0042	guardhouse	inm	
0043	guard company	hum	
0044	people	hum	
0045	mine	loc	
0046	volunteers	hum	
0047	mother of 0003	hum	
0048	miners	hum	> 0003, 0044
0049	three Daghestani	hum	> 0003; < 0044
0050	two Russians	hum	< 0044
0051	finger of 0003	bdp	M 0003
0052	an engine	inm	
0053	hospital	inm	
0054	home of 0003	loc	
0055	money	inm	
0056	mine officials	hum	
0057	people	hum	
0058	prosecutor	hum	
0059	Urkaraq	loc	
0060	house of 0003	inm	
0061	documents	inm	
0062	Sanzhi	loc	
0063	driving course	loc	
0064	workers	hum	
0065	wine factory	inm	
0066	wine	inm	
0067	story	abs	
0068	people	hum	
0069	here	loc	
0070	events	abs	
0071	service	abs	

2.2 bazhuk

<i>index</i>	<i>label</i>	<i>description</i>	<i>class</i>	<i>relations</i>	<i>notes</i>
0000	a boy		hum		
0001	father of 0000		hum		
0002	a flock of sheep		anm		
0003	a place		loc		
0004	trees		inm		
0005	home of 0000		loc		
0006	words of 0001		abs		
0007	apple tree		inm	< 0004	
0008	apples		inm	M 0007	
0009	bazhuk		hum		
0010	apples		inm	M 0007	

0011	<i>a sack</i>	inm	
0012	<i>a place</i>	loc	
0013	<i>bush</i>	loc	
0014	<i>house of 0009</i>	loc	
0015	<i>a pot</i>	inm	
0016	<i>stones</i>	inm	
0017	<i>a box</i>	inm	
0018	<i>little finger of 0009</i>	bdp	M 0009
0019	<i>water in 0015</i>	mss	

2.3 *dragon*

<i>index</i>	<i>label</i>	<i>description</i>	<i>class</i>	<i>relations</i>	<i>notes</i>
0000	<i>a village</i>		loc		
0001	<i>a family</i>		hum		
0002	<i>a girl</i>		hum	< 0001	
0003	<i>dogs</i>		anm		
0004	<i>food for 0003</i>		inm		
0005	<i>people</i>		hum		
0006	<i>eating people</i>		abs		
0007	<i>villagers</i>		hum		
0008	<i>one villager</i>		hum	< 0007	
0009	<i>teeth of 0002</i>		bdp	M 0002	
0010	<i>brother of 0002</i>		hum	< 0001	
0011	<i>other villagers</i>		hum		
0012	<i>house of 0001</i>		inm		
0013	<i>a mouse</i>		anm		
0014	<i>house of 0013</i>		loc		
0015	<i>story</i>		abs		
0016	<i>other village</i>		loc		
0017	<i>three trees</i>		inm		
0018	<i>one of 0017</i>		inm		
0019	<i>village dogs</i>		anm	> 0003	
0020	<i>whistling</i>		abs		
0021	<i>second of 0017</i>		inm		
0022	<i>third of 0017</i>		inm		
0023	<i>little finger of 0002</i>		bdp	M 0002	
0024	<i>father of 0010</i>		hum	< 0001	
0025	<i>mother of 0010</i>		hum	< 0001	
0026	<i>belly of 0002</i>		bdp	M 0002	
0027	<i>a girl</i>		hum		
0028	<i>sound made by 0010</i>		abs		

2.4 kurban

<i>index</i>	<i>label</i>	<i>description</i>	<i>class</i>	<i>relations</i>	<i>notes</i>
0000	(speaker)		hum		
0001	cousin of 0000		hum		
0002	audience		hum		
0003	sister of 0000		hum		
0004	forty-day vigil		abs		
0005	0000 and family		hum	< 0000, 0001	
0006	home of 0000		inm		
0007	family of 0000		hum	< 0005	
0008	a friend of 0000		hum		
0009	Cilibki village		loc		
0010	village head		hum		
0011	head of the district administration		hum		
0012	such		abs		
0013	0000 and 0008		hum	> 0000, 0008	
0014	such		abs		
0015	agronomist		hum		
0016	expert on livestock		hum		
0017	veterinarian		hum		
0018	people of 0009		hum		
0019	such		abs		
0020	time to go there		tmp		
0021	0008 and others		hum	> 0008	
0022	family except 0001		hum	< 0005	
0023	brother Ramazan		hum	< 0005	
0024	qulhu		inm		
0025	everyone		hum	> 0000, 0001, 0005, 0008, 0023	
0026	office of 0011		loc		
0027	husband		hum		
0028	wife		hum		
0029	other villagers in 0009		hum		
0030	other engineers in 0009		hum		
0031	other mechanics in 0009		hum		
0032	other villagers in		hum		

	0009		
0033	one of 0027 and 0028	hum	< 0027, 0028
0034	villagers of 0009	hum	> 0029, 0030, 0031, 0032
0035	circumstances	abs	
0036	that problem	abs	
0037	the fifth day	tmp	
0038	Noghay district	loc	
0039	a woman from 0009	hum	
0040	someone else	hum	
0041	story	abs	
0042	story	abs	
0043	another sister of 0000 (?),	hum	
0044	everyone	hum	> 0000, 0005, 0008, 0023
0045	0001 and 0039	hum	> 0001, 0039
0046	people in 0009	hum	> 0029, 0030, 0031, 0032, 0034

2.5 mill

<i>index</i>	<i>label</i>	<i>description</i>	<i>class</i>	<i>relations</i>	<i>notes</i>
0000	a mill		inm		depicted reality
0001	the Sanzhi people		hum		
0002	grain		inm		
0003	the first mill		inm		
0004	the first location		loc		
0005	channel		inm		
0006	millhouse		inm		
0007	a man		hum		
0008	Allah		hum		
0009	one of 0001		hum	< 0001	
0010	two boys		hum	< 0001, 0034	
0011	water		mss		
0012	sound of 0011		abs		
0013	a windmill		inm		depicted reality
0015	the second mill		inm		

0016	<i>big river</i>		inm		
0017	<i>banks of 0016</i>		loc	M 0016	
0018	<i>channel</i>		inm		
0019	<i>water</i>		mss		
0020	<i>other villagers</i>		hum		
0021	<i>qadis</i>		hum		
0022	<i>mullahs</i>		hum		
0023	<i>one of 0021, 0022</i>		hum	< 0021, 0022	
0024	<i>bride</i>		hum		
0025	<i>nose of 0024</i>		bdp	M 0024	
0026	<i>millstone</i>		inm	M 0015	
0027	<i>hammer</i>		inm		
0028	<i>part of 0015</i>		inm	M 0015	
0029	<i>heart of 0015</i>		inm	M 0015	
0030	<i>water</i>		mss		
0031	<i>the third place</i>		loc		
0032	<i>the third mill</i>		inm		
0033	<i>a bride</i>		hum		depicted reality
0034	<i>builders</i>		hum	< 0001; > 0009	
0035	<i>that</i>		abs		
0036	<i>water</i>		mss		
0037	<i>people and bride</i>		hum	> 0001, 0024	

2.6 *patima*

<i>index</i>	<i>label</i>	<i>description</i>	<i>class</i>	<i>relations</i>	<i>notes</i>
0000	<i>Patima</i>		hum		
0001	<i>family of 0000</i>		hum	> 0000	
0002	<i>house of 0001</i>		inm		
0003	<i>a nut</i>		inm		
0004	<i>chimney of 0002</i>		inm	M 0002	
0005	<i>sisters of 0000</i>		hum	< 0001	
0006	<i>one of 0005</i>		hum	< 0005	
0007	<i>another of 0005</i>		hum	< 0005	
0008	<i>forest</i>		loc		
0009	<i>nuts</i>		inm		
0010	<i>a sack</i>		inm		
0011	<i>a wolf</i>		anm		
0012	<i>story</i>		abs		
0013	<i>0000 and 0011</i>		hum	> 0000, 0011	
0014	<i>entrance of 0002</i>		loc		
0015	<i>a fox</i>		anm		
0016	<i>story</i>		abs		
0017	<i>deeds</i>		abs		

0018	0015 and 0011	anm	> 0011, 0015	
0019	<i>a place</i>	loc		
0020	<i>urine</i>	inm		bridging-contained: reference to light verb complement
0021	<i>ice</i>	inm		
0022	<i>a fork</i>	inm		
0023	<i>a spade</i>	inm		
0024	<i>belly of 0011</i>	bdp		
0025	0000 and 0005	hum	> 0000, 0005	
0026	<i>mattress</i>	inm		
0027	<i>barn</i>	loc		
0028	<i>cotton</i>	inm		
0029	<i>door</i>	inm	M 0002	

2.7 ramazan

<i>index</i>	<i>label</i>	<i>description</i>	<i>class</i>	<i>relations</i>	<i>notes</i>
0000	<i>a journalist</i>		hum		
0001	<i>Tawlu</i>		hum		
0002	<i>an article</i>		abs		
0003	<i>surname of 0001</i>		abs		
0004	<i>such</i>		abs		
0005	<i>passport of 0001</i>		inm		
0006	<i>such</i>		abs		
0007	<i>a story like this</i>		abs		
0008	<i>stories like this</i>		abs		
0009	<i>the most beautiful place</i>		abs		
0010	<i>the world</i>		loc		
0011	<i>(speaker)</i>		hum		
0012	<i>Latvia</i>		loc		
0013	<i>fruit</i>		inm		
0014	<i>forests</i>		loc		
0015	<i>roads in 0012</i>		inm		
0016	<i>lorry of 0011</i>		inm		
0017	<i>towns in 0012</i>		loc		
0018	<i>traffic lights</i>		inm		
0019	<i>two people</i>		hum		
0020	<i>three people</i>		hum		
0021	<i>end of the road</i>		loc		
0022	<i>house in 0012</i>		loc		
0023	<i>pieces of asphalt</i>		inm		
0024	<i>chewing gum</i>		inm		
0025	<i>asphalt</i>		inm		
0026	<i>(impersonal)</i>		hum		

	<i>reference)</i>		
0027	<i>the letter D</i>	inm	
0028	<i>middle of licence plate</i>	loc	
0029	<i>front of the lorry</i>	loc	
0030	<i>that trick</i>	abs	
0031	<i>a Russian woman</i>	hum	
0032	<i>0011 and others</i>	hum	> 0011
0033	<i>torso of 0011</i>	bdp	M 0011
0034	<i>gold</i>	inm	
0035	<i>dressing table</i>	inm	
0036	<i>audience</i>	hum	
0037	<i>rings</i>	inm	
0038	<i>bracelets</i>	inm	
0039	<i>all things</i>	inm	
0040	<i>money</i>	inm	
0041	<i>wealth of 0031</i>	inm	
0042	<i>circumstances</i>	abs	
0043	<i>lock</i>	inm	
0044	<i>first neighbour</i>	hum	
0045	<i>second and third neighbour</i>	hum	
0046	<i>our place</i>	loc	
0047	<i>face of 0011</i>	bdp	M 0011
0048	<i>conscience of 0011</i>	abs	
0049	<i>room with 0035</i>	loc	
0050	<i>things not belonging to 0011</i>	inm	
0051	<i>being accused of theft</i>	abs	
0052	<i>a match</i>	inm	
0053	<i>light of 0052</i>	inm	
0054	<i>a cigarette</i>	inm	
0055	<i>passers-by</i>	hum	
0056	<i>cleanliness</i>	abs	
0057	<i>ashes</i>	inm	
0058	<i>like that</i>	abs	
0059	<i>such a place</i>	loc	
0060	<i>the worst place</i>	loc	
0061	<i>Azerbaijan</i>	loc	
0062	<i>every town</i>	loc	
0063	<i>Soviet Union</i>	loc	
0064	<i>a racket</i>	abs	
0065	<i>a scandal</i>	abs	
0066	<i>diesel fuel</i>	mss	
0067	<i>a hill</i>	loc	
0068	<i>petrol station</i>	inm	

0069	<i>a bucket of fuel</i>		inm	
0070	<i>attendant at 0068</i>		hum	
0071	<i>money for 0069</i>		inm	
0072	<i>a no entry sign</i>		inm	
0073	<i>other people at 0068</i>		hum	
0074	<i>one of 0073</i>		hum	< 0073
0075	<i>another of 0073</i>		hum	< 0073
0076	<i>another bucket</i>		inm	
0077	<i>such</i>		abs	
0078	<i>what 0011 had done</i>		abs	
0079	<i>0070 and 0073</i>		hum	> 0070, 0073
0081	<i>like that</i>		abs	
0082	<i>place like 0081</i>		loc	
0083	<i>events</i>		abs	
0084	<i>event</i>		abs	
0085	<i>Estonia</i>		loc	
0086	<i>a person</i>		hum	
0087	<i>Estonians</i>		hum	
0088	<i>Icari people</i>		hum	
0089	<i>market</i>		loc	
0090	<i>bus</i>		inm	
0091	<i>a town</i>		loc	
0092	<i>Latvians</i>		hum	
0093	<i>people who do not speak Latvian</i>		hum	
0094	<i>Moldova</i>		loc	
0095	<i>home of 0011</i>		loc	
0096	<i>0001 and 0000</i>		hum	> 0001, 0000
0097	<i>this kind of thing</i>		abs	
0098	<i>that</i>		abs	

2.8 *tape*

<i>index</i>	<i>label</i>	<i>description</i>	<i>class</i>	<i>relations</i>	<i>notes</i>
0000	<i>Kajakent</i>		loc		
0001	<i>speaker and friends</i>		hum		
0002	<i>Abdulhalik</i>		hum	< 0000	
0003	<i>Isakadi</i>		hum	< 0000	
0004	<i>(speaker)</i>		hum	< 0000	
0005	<i>work of 0000</i>		abs		
0006	<i>survey</i>		abs		
0007	<i>roundabout</i>		inm		
0008	<i>one side of 0007</i>		loc	M 0007	

0009	<i>a place</i>	loc	
0010	<i>0000 and 0002</i>	hum	> 0000, 0002
0011	<i>here</i>	loc	
0012	<i>car</i>	inm	
0013	<i>shop</i>	inm	
0014	<i>assistant</i>	hum	
0015	<i>tape</i>	inm	
0016	<i>one roll of 0015</i>	inm	
0017	<i>mouth of 0002</i>	bdp	M 0002
0018	<i>surveyors</i>	hum	

Multi-CAST

Multilingual Corpus of Annotated Spoken Texts



multicast.aspra.uni-bamberg.de/

