
Supplementary Material

FUNCTION	LANGUAGE	SOURCE
(imperfective) aspect, causation	Chapacuran family	Birchall (2018)
thought, intention ('internal awareness'), perception, explaining human action, naming, hearsay evidential	Ainu	Bugaeva (2008)
Complementizer, Topic marker, Conjunction, Hearsay evidentiality, Irrealis marker, Clause-final discourse marker	Sinitic family	Chappell (2012)
cognition ('think'), intention/volition, purpose, cause/reason, hearsay evidentiality	Sanzhi	Forker (2019)
subordination, purpose	Besleney Kabardian	Ershova (2012)
intention	Wari'	Everett (2008)
Naming, Reported evidence, Illocution reinforcement and related discourse functions, Similarity and manner, Internal awareness, Intention, proximative, and future, Deontic modality and indirect causation, Clause linkage, Multipurpose auxiliary, Nominal identification, listing	African languages (Nilo-Saharan, Khoisan, Afroasiatic and Niger-Kordofanian families)	Güldemann (2008)
Rhetorical highlighting, causal relations, awareness attribution, specification, concepts	Aguaruna	Larson (1978)
auxiliary/light verb, complementiser, comparison, concessive, conditional, discourse particle, enumerative particle, hearsay evidential, perceptual/mirative evidential; causation, metalinguistic marker (naming), purpose, reason, Report marker, topic	Siberian languages (various families)	Matić and Pakendorf (2013)
say, think, want, do	Warrwa, Nyulnyul	McGregor (2014)
purpose, lest, pretense, mistaken thought, inception, attempting, listing	Usan	Reesink (1993)
thought, intention, causation, naming	Ungarinyin	Rumsey (1990); Spronck (2015)
Causal conjunction, Embedded questions, Purpose conjunction, Naming-labelling, Evidential, "Deliberately", Onomatopoeic expressions, Question word complementizer, Conditional, Expletive	Tibeto-Burman family	Saxena (1988)
purpose, wish, intention, causation, warning	Kwaza	van der Voort (2002)

Table S1. A summary of functions from language-specific studies

1.2 Appendix B

Language Name	Genus	Language Family	Macro Area	Source
Balanta [bjt]	Bak	Niger-Congo	Africa	Creissels, Denis; Biaye, Séckou (2016) Le balant ganja. Phonologie, morphosyntaxe, liste lexicale, textes
Bena [bez]	Bantu	Niger-Congo	Africa	Morrison, Michelle Elizabeth (2011) A reference grammar of Bena
Fongbe [fon]	Kwa	Niger-Congo	Africa	Lefebvre, Claire; Brousseau, Anne-Marie (2002) A grammar of Fongbe
Goemai [ank]	West Chadic	Afro-Asiatic	Africa	Hellwig, Birgit (2011) A Grammar of Goemai
Gula [glu]	Bongo-Bagirmi	Central Sudanic	Africa	Nougayrol, Pierre (1999) Les parlers gula. Centrafrique, Soudan, Tchad. Grammaire et lexique
Gumer [sgw]	Semitic	Afro-Asiatic	Africa	Völlmin, Sascha (2017) Towards a Grammar of Gumer. Phonology and Morphology of a Western Gurage Variety
Kakabe [kke]	Western Mande	Mande	Africa	Vydrina, Alexandra (2017) A corpus-based description of Kakabe, a Western Mande language: prosody in grammar
Kotoko [aal]	Biu-Mandara	Afro-Asiatic	Africa	Sean David Allison (2012) Aspects of a Grammar of Makaray Kotoko
Luwo [lwo]	Nilotic	Eastern Sudanic	Africa	Storch, Anne (2014) A Grammar of Luwo: An anthropological approach
Maale [mdy]	North Omotic	Afro-Asiatic	Africa	Amha, Azeb (2001) The Maale Language
Tamashek [taq]	Berber	Afro-Asiatic	Africa	Heath, Jeffrey (2005) A grammar of Tamashek (Tuareg of Mali)
Saaho [ssy]	Lowland East Cushitic	Afro-Asiatic	Africa	Tajebe, Esayas (2015) Descriptive Grammar of Saaho
Samba Leko [ndi]	Samba-Duru	Niger-Congo	Africa	Fabre, Gwenaëlle (2002) Le Samba Leko, Langue Adamawa du Cameroun
Supyire [spp]	Senufo	Niger-Congo	Africa	Carlson, Robert (1994) A Grammar of Supyire
Arrernte [arr]	Central Pama-Nyungan	Pama-Nyungan	Australia	Wilkins, David P. (1989) Mparntwe Arrernte (Aranda)
Bininj Gun-Wok [gup]	Gunwinygic	Gunwinyguan	Australia	Evans, Nicholas (2003) Bininj Gunwok: a pan-dialectal grammar of Mayali, Kunwinjku and Kunz

continued on next page

Kayardild [gyd]	Tangkic	Tangkic	Australia	Evans, Nicholas (1995) A Grammar of Kayardild
Kurrama [vku]	Western Pama-Nyungan	Pama-Nyungan	Australia	Hill, Peter M. (2011) Morphology and Sentence Construction in Kurrama: A Language of the Pilbara Region of Western Australia
Kuuk Thaayorre [thd]	Northern Pama-Nyungan	Pama-Nyungan	Australia	Gaby, Alice R. (2006) A Grammar of Kuuk Thaayorre
Nyulnyul [nyv]	Nyulnyulan	Nyulnyulan	Australia	McGregor, William B. (2011) The Nyulnyul language of Dampier Land, Western Australia
Worrora [wro]	Worroran	Worroran	Australia	Clendon, Mark (2014) Worrora: a language of the north-west Kimberley coast
Yuwaalaraay and Gamilaraay [kld]	Southeastern Pama-Nyungan	Pama-Nyungan	Australia	Giacon, John (2014) A grammar of Yuwaalaraay and Gamilaraay:
Basque [eus]	Basque	Basque	Eurasia	Hualde, José I. & Jon Ortiz de Urbina (2003) A Grammar of Basque
Belorussian [bel]	Slavic	Indo-European	Eurasia	Toledo, Kristian Roncero (2019) A typological approach to West Polesian Morphology and Syntax
Bulo Stieng [sti]	Bahnaric	Austro-Asiatic	Eurasia	Bon, Noëllie. (2014) Une grammaire de la langue stieng, langue en danger du Cambodge et du Vietnam
Bunan [bfu]	Bodic	Sino-Tibetan	Eurasia	Widmer, Manuel (2017) A Grammar of Bunan
Catalan [cat]	Romance	Indo-European	Eurasia	Hualde, José Ignacio (1992) Catalan
Darai [dry]	Indic	Indo-European	Eurasia	Dubi Nanda Dhakal (2012) Darai Grammar
Hinuq [gin]	Avar-Andic-Tsezic	Nakh-Daghestanian	Eurasia	Forker, Diana (2013) A grammar of Hinuq
Icelandic [isl]	Germanic	Indo-European	Eurasia	Thráinsson, Höskuldur (2007) The Syntax of Icelandic
Lao [lao]	Kam-Tai	Tai-Kadai	Eurasia	Enfield, Nick J. (2007) A Grammar of Lao
Madurese [mad]	Malayo-Sumbawan	Austronesian	Eurasia	Davies, William D. (2010) A Grammar of Madurese
Meithei [mni]	Kuki-Chin	Sino-Tibetan	Eurasia	Chelliah, Shobhana L. (1997) A Grammar of Meithei

continued on next page

Pashto [pst]	Iranian	Indo-European	Eurasia	Boyle David, Anne (2014) Descriptive Grammar of Pashto and its Dialects
Sichuan Yi [iii]	Burmese-Lolo	Sino-Tibetan	Eurasia	Gerner, Matthias (2013) A Grammar of Nuosu
Thangmi [thf]	Mahakiranti	Sino-Tibetan	Eurasia	Turin, Mark (2012) A Grammar of the Thangmi Language. With an Ethnolinguistic Introduction to the Speakers and Their Culture
Uyghur [uig]	Turkic	Altaic	Eurasia	Abdurehim, Esmael (2014) The Lopnor dialect of Uyghur
Coatlan-Loxicha Zapotec [zps]	Zapotecan	Oto-Manguean	North America	Beam de Azcona, Rosemary Grace (2004) A Coatlan-Loxicha Zapotec grammar (Mexico)
Columbia-Wenatchi [col]	Interior Salish	Salishan	North America	Willett, Marie Louise (2003) A Grammatical Sketch of Nxa'arnxcin (Moses-Columbia Salish)
Cree [crk]	Algonquian	Algic	North America	Wolvengrey, Arok Elessar (2011) Semantic and Pragmatic Functions in Plains Cree Syntax
Creek [mus]	Muskogean	Muskogean	North America	Martin, Jack B. (2011) A Grammar of Creek (Muskogee)
Hidatsa [hid]	Core Siouan	Siouan	North America	Park, Indrek (2012) A Grammar of Hidatsa
Huehuetla Tepehua [tee]	Totonacan	Totonacan	North America	Kung Smythe, Susan (2007) A Descriptive Grammar of Huehuetla Tepehua
Itzaj Maya [itz]	Mayan	Mayan	North America	Hofling, Charles Andrew (2000) Itzaj Maya Grammar
Jamul Tiipay [dih]	Yuman	Hokan	North America	Miller, Amy (2001) A Grammar of Jamul Tiipay
Lushootseed [lut]	Central Salish	Salishan	North America	Zahir, Zalmai (2018) Elements of Lushootseed Grammar in Discourse
Mixtec Magdalena Peñasco [xtm]	Mixtecan	Oto-Manguean	North America	Erickson de Hollenbach, Elena (2013) Gramática del mixteco de Magdalena Peñasco (Sa'an Nñuu Savi)
Nahuatl Tlaxcala [nhn]	Aztecan	Uto-Aztecan	North America	Flores Nájera, Lucero (2019) La Gramática de la cláusula simple en el Náhuatl de Tlaxcala
Northern Paiute [pao]	Numic	Uto-Aztecan	North America	Thornes, Timothy Jon (2003) A Northern Paiute Grammar with Texts

continued on next page

Pomo [peq]	Pomoan	Hokan	North America	Walker, Neil Alexander (2013) A Grammar of Southern Pomo: An Indigenous Language of California
Seneca [see]	Northern Iroquoian	Iroquoian	North America	Chafe, Wallace (2015) A Grammar of the Seneca Language
Sierra Popoluca [poi]	Mixe-Zoque	Mixe-Zoque	North America	Boudreault, Lynda (2018) A Grammar of Sierra Popoluca
Slave [xsl]	Athapaskan	Na-Dene	North America	Rice, Keren (1989) A Grammar of Slave
Sochiapan Chinantec [cso]	Chinantecan	Oto-Manguean	North America	Foris, David Paul (2000) A Grammar of Sochiapan Chinantec
Yup'ik [esu]	Eskimo	Eskimo-Aleut	North America	Miyaoka, Osahito (2012) A Grammar of Central Alaskan Yupik (CAY)
Adang [adn]	Greater Alor	Timor-Alor-Pantar	Papunesia	Haan, Johnson Welem (2001) The Grammar of Adang: A Papuan Language Spoken on the Island of Alor East Nusa Tenggara, Indonesia
Aghu [ahh]	Awju-Dumut	Trans-New-Guinea	Papunesia	Heuvel, Wilco van den (2016) Aghu: annotated texts with grammatical introduction and vocabulary lists
Cebuano [ceb]	Greater Central Philippine	Austronesian	Papunesia	Tanangkingsing, Michael (2009) A Functional Reference Grammar of Cebuano
Dom [doa]	Chimbu	Trans-New-Guinea	Papunesia	Tida, Syuntarô (2006) A grammar of the Dom language: A Papuan Language of Papua New Guinea
Hua [ygr]	Eastern Highlands	Trans-New-Guinea	Papunesia	Haiman, John (1980) HUA: A Papuan Language of the Eastern Highlands of New Guinea
Ilocano [ilo]	Northern Luzon	Austronesian	Papunesia	Galvez Rubino, Carl Ralph (1997) A Reference Grammar of Ilocano
Kalam [kmh]	Madang	Trans-New-Guinea	Papunesia	Pawley, Andrew & Bulmer, Ralph (2011) A dictionary of Kalam with ethnographic notes
Kambera [xbr]	Central Malayo-Polynesian	Austronesian	Papunesia	Kalmer, Marian (1998) A Grammar of Kambera
Kelabit [kzi]	North Borneo	Austronesian	Papunesia	Hemmings, Charlotte (2016) The Kelabit Language, Austronesian Voice and Syntactic Typology
Kilmeri [kih]	Border	Border	Papunesia	Gerstner-Link, Claudia (2018) A Grammar of Kilmeri

continued on next page

Korafe-Yegha [kpr]	Binanderean	Trans-New-Guinea	Papunesia	Farr, Cynthia J. M. (1999) The interface between syntax and discourse in Korafe: a Papuan language of Papua New Guinea
Ma Manda [skc]	Finisterre-Huon	Trans-New-Guinea	Papunesia	Pennington, Ryan (2016) A grammar of Ma Manda: A Papuan language of Morobe Province, Papua New Guinea
Makasar [mak]	South Sulawesi	Austronesian	Papunesia	Jukes, Anthony. (2020) A grammar of Makasar: A language of South Sulawesi, Indonesia
Mali [gcc]	Baining	Baining	Papunesia	Stebbins, Tonya N. (2011) Mali (Baining) grammar: a language of the East New Britain Province, Papua New Guinea
Mehek [nux]	Tama Sepik	Sepik	Papunesia	Hatfield, Adam (2016) A grammar of Mehek
Menya [mcr]	Angan	Trans-New-Guinea	Papunesia	Whitehead, Carl R. (2004) A Reference Grammar of Menya, an Angan Language of Papua New Guinea
Nias [nia]	Northwest Sumatra-Barrier Islands	Austronesian	Papunesia	Brown, Lea (2001) A Grammar of Nias Seltan
Oksapmin [opm]	Oksapmin	Oksapmin	Papunesia	Loughnane, Robyn (2009) A Grammar of Oksapmin
Pendau [ums]	Celebic	Austronesian	Papunesia	Quick, Phil (2007) A grammar of the Pendau language of central Sulawesi, Indonesia
Savosavo [svs]	Savosavo	Solomons East Papua	Papunesia	Wegener, Claudia (2012) A Grammar of Savosavo
Siar-Lak [sjr]	Oceanic	Austronesian	Papunesia	Frowein, Friedel Martin (2011) A Grammar of Siar, An Oceanic Language of New Ireland Province
Tidore [tvo]	North Halmaheran	West Papuan	Papunesia	van Staden, Miriam (2000) Tidore: A Linguistic Description of a Language of the North Moluccas
Waigeo [wgo]	South Halmahera -West New Guinea	Austronesian	Papunesia	Arnold, Laura. (2018) A grammar of Ambel: An Austronesian language of Raja Ampat, west New Guinea
West Coast Bajau [bdr]	Sama-Bajaw	Austronesian	Papunesia	Miller, Mark T. (2007) A Grammar of West Coast Bajau

continued on next page

Yeri [yev]	Wapei-Palei	Toricelli	Papunesia	Wilson, Jennifer (2017) A Grammar of Yeri a Toricelli language of Papua New Guinea
Aguaruna [agr]	Jivaroan	Jivaroan	South America	Overall, Simon E. (2017) A Grammar of Aguaruna (Iiniá Chicham)
Apinayé [apn]	Ge-Kaingang	Macro-Ge	South America	Cunha de Oliveira, Christiane (2005) The language of the Apinajé People of Central Brazil
Araona [aro]	Tacanan	Tacanan	South America	Emkow, Carola (2006) A Grammar of Araona, an Amazonian language of Northwestern Bolivia
Baure [brg]	Bolivia-Parana	Arawakan	South America	Danielsen, Swintha (2007) Baure: An Arawak Language of Bolivia
Hup [jup]	Nadahup	Nadahup	South America	Epps, Patience (2008) A grammar of Hup
Kakataibo [cbr]	Panoan	Panoan	South America	Zariquiey, Roberto (2018) A Grammar of Kakataibo
Kokama-Kokamilla [cod]	Tupi-Guaraní	Tupian	South America	Vallejos Yopán, Rosa (2010) A Grammar of Kokama-Kokamilla
Kotiria (Wanano) [gvc]	Tucanoan	Tucanoan	South America	Stenzel, Kristine (2013) A Reference Grammar of Kotiria (Wanano)
Kulina [cul]	Arauan	Arauan	South America	Dienst, Stefan (2014) A Grammar of Kulina
Mapudungun [arn]	Araucanian	Araucanian	South America	Smeets, Ineke. (2008) A grammar of Mapuche
Murui Huitoto [huu]	Huitoto	Huitotoan	South America	Wojtylak, Katarzyna Izabela (2017) A Grammar of Murui (Bue): A Witotoan language of Northwest Amazonia
Nanti [cox]	Pre-Andine Arawakan	Arawakan	South America	Michael, Lev David (2008) Nanti evidential practice: Language, knowledge, and social action in an Amazonian society
Nivaclé [cag]	Matacoan	Matacoan	South America	Fabre, Alain (2016) Gramática de la lengua Nivacle (familia Mataguayo, Chaco Paraguayo)
Northern Emberá [emp]	Choco	Choco	South America	Mortensen, Charles Arthur (1999) A Reference Grammar of Northern Embera Languages
Tariana [tae]	Inland Northern Arawakan	Arawakan	South America	Aikhenvald, Alexandra Y. (2003) A Grammar of Tariana, from Northwest Amazonia

continued on next page

Totoro [ttk]	Barbacoan	Barbacoan	South America	Gonzales Castaño, Geny (2019) Una gramática de la lengua namtrik de Totoró: Lengua barbacoa hablada en los Andes colombianos
Trio [tri]	Cariban	Cariban	South America	Carlin, Eithne B. (2004) A Grammar of Trio, a Cariban Language of Suriname
Yagua [yad]	Peba-Yaguan	Peba-Yaguan	South America	Payne, Doris L. & Payne, Thomas E. (1990) Yagua
Yanomám [wca]	Yanomam	Yanomam	South America	Perri Ferreira, Helder (2017) Yanomama Clause Structure
Yauyos Quechua [qux]	Quechuan	Quechuan	South America	Shimelman, Aviva (2017) A grammar of Yauyos Quechua

REFERENCES

- Birchall, J. (2018). Historical change in reported speech constructions in the Chapacuran family. *Journal of Historical Linguistics* 8(1), 7–30.
- Bugaeva, A. (2008). Reported discourse and logophoricity in Southern Hokkaido dialects of Ainu. *Gengo Kenkyu* 133, 31–75.
- Chappell, H. (2012). Say in Sinitic: From *Verba Dicendi* to attitudinal discourse markers. In J. van der Auwera and J. Nuyts (Eds.), *Grammaticalization and (inter-)subjectification*, pp. 81–110. Wetteren: Universa Press.
- Ershova, K. (2012). Reported speech and reportative grammaticalization in Beslenny Kabardian. In *Proceedings of the Second Central European Conference in Linguistics for Postgraduate Students*, pp. 71–87.
- Everett, D. (2008). Wari' intentional state constructions. In R. D. Van Valin, Jr. (Ed.), *Investigations of the Syntax–Semantics–Pragmatics Interface*, pp. 381–409. Amsterdam/Philadelphia: John Benjamins.
- Forker, D. (2019). Reported speech constructions in Sanzhi Dargwa and their extension to other areas of grammar. *Sprachwissenschaft* 44(2).
- Güldemann, T. (2008). *Quotative Indexes in African Languages: A Synchronic and Diachronic Survey*. Berlin: Mouton de Gruyter.
- Larson, M. L. (1978). *The Functions of Reported Speech in Discourse*. Dallas: Summer Institute of Linguistics.
- Matić, D. and B. Pakendorf (2013). Non-canonical SAY in Siberia. *Studies in Language* 37(2), 356–412.
- McGregor, W. B. (2014). The 'say, do' verb in Nyulnyul, Warrwa, and other Nyulnylan languages is monosemic. In K. Robering (Ed.), *Events, Arguments and Aspects: Topics in the Semantics of Verbs*, pp. 301–327. Amsterdam/Philadelphia: John Benjamins.
- Reesink, G. P. (1993). 'Inner speech' in Papuan languages. *Language and Linguistics in Melanesia* 24, 217–225.
- Rumsey, A. (1990). Wording, meaning and linguistic ideology. *American Anthropologist* 92(2), 346–361.
- Saxena, A. (1988). On syntactic convergence: The case of the verb 'say' in Tibeto-Burman. In *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, pp. 375–388.
- Spronck, S. (2015). *Reported speech in Ungarininyin: grammar and social cognition in a language of the Kimberley region, Western Australia*. Ph. D. thesis, The Australian National University.

van der Voort, H. (2002). The quotative construction in Kwaza and its (de-) grammaticalisation. In M. Crevels, S. van de Kerke, S. Meira, and H. van der Voort (Eds.), *Current Studies on South American Languages Indigenous Languages of Latin America*, Number 3, pp. 307–328.