

Table 1. The 210 word meaning categories collected in the Austronesian Basic Vocabulary Database.

Adjectives	bad/evil, big, cold, correct/true, dirty, dry, dull/blunt, good, heavy, long, narrow, new, old, painful/sick, rotten, sharp, short, shy/ashamed, small, thick, thin, warm, wet, wide
Animals	bird, dog, egg, fish, louse, mosquito, rat, snake, spider, worm (earthworm)
Body Parts	back, belly, blood, bone, breast, ear, eye, feather, hair, hand, head, intestines, leg/foot, liver, mouth, neck, nose, shoulder, skin, tail, tongue, tooth, wing
Colors	black, green, red, white, yellow
Directions	above, at, below, far, in/inside, left, near, right
Numbers	one, two, three, four, five, six, seven, eight, nine, ten, twenty, fifty, one hundred, one thousand
People	child, father, he/she, husband, I, man/male, mother, name, person/human being, they, thou, we, wife, woman/female, you
Plants	branch, flower, fruit, grass, leaf, root
Other	all, and, how?, if, no/not, other, that, this, what?, when?, where?, who?
Other Nouns	ash, cloud, day, dust, earth/soil, fat/grease, fire, fog, house, lake, lightning, meat/flesh, moon, needle, night, rain, road/path, rope, salt, sand, sea, sky, smoke, star, stick/wood, stone, thatch/roof, thunder, water, wind, woods/forest, year
Verbs	to bite, to blow, to breathe, to burn, to buy, to chew, to choose, to climb, to come, to cook, to count, to cry, to cut/hack, to die/be dead, to dig, to dream, to drink, to eat, to fall, to fear, to flow, to fly, to grow, to hear, to hide, to hit, to hold, to hunt, to kill, to know/be knowledgeable, to laugh, to lie down, to live/be alive, to open/uncover, to plant, to pound/beat, to say, to scratch, to see, to sew, to shoot, to sit, to sleep, to sniff/smell, to spit, to split, to squeeze, to stab/pierce, to stand, to steal, to suck, to swell, to swim, to think, to throw, to tie up/fasten, to turn, to vomit, to walk, to work, to yawn

The core database schema of the ABVD is shown in Figure 1. The table *languages* stores information about each language. This includes the name of language (“language”), the data source information (“author”), and the name of the person entering the data (“typedby”). The field “silcode” contains the ISO 639-3 language identification code. This identification code provides a way of linking each language to broader information about the language at other resources such as SIL International’s *Ethnologue* (<http://www.ethnologue.com>, Gordon, 2005), or the *World Atlas of Language Structures* (<http://www.wals.info>, Haspelmath et al. 2005). The “classification” field stores the current

classification of the language, obtained from the *Ethnologue* classification. Finally, there is a text field (“notes”) for any extra information about the language. The *languages* table is linked to the *locations* table via a one-to-many foreign key. The *locations* table stores geographical information (“latitude”, “longitude”) about each language to enable plotting of the languages on a map. The *resources* table is also linked to the *languages* table via foreign key. This table contains links to other relevant websites about a language, such as *Wikipedia* entries, homepages of research groups investigating the language, or other online resources like dictionaries.