

APPENDIX L: Sepedi Core Vocabulary List

Core word list with frequency of occurrence, commonality score, and part of speech classification

root word	Number of occurrences	Frequency per mille	Commonality	Part of speech	Content/function words
ke@x (i)	985	50.9465191	6	Subject concord	function word
cn (child's name)	901	46.6018413	6	Proper noun	content word
o (you)	777	40.1882694	6	Subject concord	function word
a@x (she, he, they)	442	22.8612806	6	Subject concord	function word
e (it)	440	22.7578359	6	Subject concord	function word
nna (i)	432	22.3440571	6	Absolute pronoun	content word
mo@k (here)	337	17.4304334	6	Demonstrative particle/pos	function word
ke@oo (is, are)	332	17.1718217	6	Copulative particle	function word
ba@x (they)	300	15.5167063	6	Subject concord	function word
wa@x (you)	294	15.2063722	6	Subject concord	function word
bona (see)	291	15.0512051	6	Verb	content word
se (she, he, it)	288	14.8960381	6	Subject concord	function word
ee (yes)	285	14.740871	6	Interjection	function word
tlo (will, shall)	285	14.740871	6	Future morpheme	function word
ka@q (inside, into, in)	252	13.0340333	6	Locative particle	function word
le@p (and, with)	237	12.258198	6	Connective particle	function word
tla@j (come)	229	11.8444192	6	Verb	content word
re@x (we)	214	11.0685838	6	Subject concord	function word
a@n	197	10.1893038	6	Negative morpheme	function word
re@j (say)	188	9.72380263	6	Verb	content word

root word	Number of occurrences	Frequency per mille	Commonality	Part of speech	Content/function words
ye (this)	181	9.36174615	6	Demonstrative particle/pos	function word
ngwala (write)	165	8.53418848	6	Verb	content word
wena (you)	155	8.01696493	6	Absolute pronoun	function word
ya@w (of)	150	7.75835316	6	Possessive concord	function word
go@m (you)	149	7.70663081	6	Object concord	function word
go@t (to)	148	7.65490845	5	Infinitive prefix	function word
ka@z (my, mine)	140	7.24112962	6	Possessive pronoun	function word
di (they)	139	7.18940726	6	Object concord	function word
go@x (it)	138	7.13768491	6	Subject concord	function word
nyaka (want, search)	128	6.62046136	6	Verb	content word
ka@qq (with, by means of)	126	6.51701665	6	Instrumental particle	function word
ka@x (i)	118	6.10323782	6	Subject concord	function word
akere (isn't it)	111	5.74118134	6	Interjection	function word
le@x (you, she, he, it)	111	5.74118134	6	Subject concord	function word
mo@m (her, him)	110	5.68945898	6	Object concord	function word
gore (so)	108	5.58601428	6	Conjunction	function word
tše (these ones)	108	5.58601428	6	Demonstrative particle/pos	function word
ria (do, make)	102	5.27568015	6	Verb	content word
fa (give)	101	5.22395779	6	Verb	content word
ko (i am going to)	99	5.12051309	6	Future morpheme	function word
re@m (us)	91	4.70673425	6	Object concord	function word

root word	Number of occurrences	Frequency per mille	Commonality	Part of speech	Content/function words
swara (hold)	87	4.49984483	6	Verb	content word
gago (yours)	85	4.39640012	6	Possessive pronoun	function word
nto (a thing)	84	4.34467777	6	Noun	content word
mang (who)	83	4.29295542	6	Noun	content word
botša (tell)	82	4.24123306	6	Verb	content word
fela (finish)	81	4.18951071	6	Verb	content word
kae (where)	80	4.13778835	6	Adverb	content word
kua (over there)	80	4.13778835	6	Locative particle	function word
bea (put, place)	79	4.086066	6	Verb	content word
so (like this)	79	4.086066	5	Preposition	function word
tšea (take)	79	4.086066	6	Verb	content word
ja (eat)	74	3.82745423	6	Verb	content word
tša (of)	73	3.77573187	5	Possessive concord	function word
wa@w (of)	73	3.77573187	6	Possessive concord	function word
motho (person)	72	3.72400952	6	Noun	content word
aowa (no)	71	3.67228716	4	Interjection	function word
ahh	68	3.5171201	6	Interjection	function word
tseba (know)	68	3.5171201	6	Verb	content word
ya@x (she, he, it, they)	68	3.5171201	6	Subject concord	function word
še (here it is)	67	3.46539774	6	Demonstrative copulative pos	function word
tloga (leave)	67	3.46539774	4	Verb	function word
gape (again)	66	3.41367539	6	Adverb	content word

root word	Number of occurrences	Frequency per mille	Commonality	Part of speech	Content/function words
haeh	66	3.41367539	5	Interjection	function word
rubba (erase)	66	3.41367539	6	Verb	content word
tšhela (pour)	66	3.41367539	5	Verb	content word
a@xx	65	3.36195304	6	Present tense morpheme	function word
bo (it)	65	3.36195304	6	Subject concord	function word
ema (stop)	65	3.36195304	6	Verb	content word
la (you)	64	3.31023068	6	Subject concord	function word
tlalea (report)	64	3.31023068	5	Verb	content word
mmemo (teacher)	63	3.25850833	6	Noun	content word
ge (when, while)	59	3.05161891	6	Conjunction	function word
tla@u (shall, will)	59	3.05161891	6	Future morpheme	function word
no (only, just)	56	2.89645185	6	Aspectual prefix	function word
we	56	2.89645185	6	Interjection	function word
be	55	2.84472949	6	Auxiliary Verb	function word
eng (what)	54	2.79300714	6	Noun	content word
kwa (hear, feel, taste, smell)	54	2.79300714	6	Verb	content word
tn (teacher's name)	54	2.79300714	6	Proper noun	content word
le@m (you, it)	50	2.58611772	6	Object concord	function word
dira (do, make)	48	2.48267301	5	Verb	content word
ae	47	2.43095066	6	Interjection	function word
gešu	47	2.43095066	6	Possessive pronoun	function word

root word	Number of occurrences	Frequency per mille	Commonality	Part of speech	Content/function words
gona (there)	47	2.43095066	6	Absolute pronoun	function word
bedi (two)	46	2.3792283	6	Adjective	content word
ga@w (of)	46	2.3792283	5	Possessive concord	function word
mmm	45	2.32750595	5	Interjection	function word
ne (have)	45	2.32750595	6	Copulative verb	function word
ntšha (take out)	43	2.22406124	5	Verb	content word
dula (sit)	42	2.17233888	6	Verb	content word
ngwe	42	2.17233888	6	Adjective	content word
itia (hit)	40	2.06889418	5	Verb	content word
mama (mother)	40	2.06889418	4	Noun	content word
meetse (water)	40	2.06889418	5	Noun	content word
šo (here she/he is)	40	2.06889418	4	Demonstrative copulative pos	function word
ngwana (child)	39	2.01717182	5	Noun	content word
pela (haste)	38	1.96544947	4	Noun	content word
sepela (walk)	38	1.96544947	5	Verb	content word
tee (one)	38	1.96544947	6	Adjective	content word
tjo	38	1.96544947	6	Interjection	function word
rena (we)	37	1.91372711	6	Absolute pronoun	function word
kgopela (ask)	35	1.8102824	6	Verb	content word
ya@j (go)	35	1.8102824	6	Verb	content word
owo (okay)	34	1.75856005	6	Adjective	content word
ra (we)	34	1.75856005	5	Subject concord	function word

root word	Number of occurrences	Frequency per mille	Commonality	Part of speech	Content/function words
hmm	33	1.7068377	5	Interjection	function word
mara (but)	33	1.7068377	6	Conjunction	function word
nke (as if)	33	1.7068377	6	Conjunction	function word
puku (book)	33	1.7068377	6	Noun	content word
fase (down)	32	1.65511534	6	Noun	content word
yey	32	1.65511534	5	Interjection	function word
adima (borrow)	31	1.60339299	5	Verb	content word
gae (home)	31	1.60339299	5	Noun	content word
le@k (this)	31	1.60339299	6	Demonstrative particle/pos	function word
maan	31	1.60339299	5	Interjection	function word
nnyane (small, few)	31	1.60339299	6	Adjective	content word
tsena (enter, get in)	31	1.60339299	6	Verb	content word
yo (go and, go to)	31	1.60339299	5	Aspectual prefix	function word
ga@q (at the place of)	30	1.55167063	6	Locative particle	function word
mo@q (on)	30	1.55167063	6	Locative particle	function word
one@cs	30	1.55167063	3	Adjective	content word
sa@w (of)	30	1.55167063	5	Possessive concord	function word
sa@x (she, he, it)	30	1.55167063	6	Subject concord	function word
ebile (then)	29	1.49994828	6	Conjunction	function word
fora (decieve)	29	1.49994828	6	Verb	content word
namela (climb)	29	1.49994828	5	Verb	content word
bapala (play)	28	1.44822592	5	Verb	content word

root word	Number of occurrences	Frequency per mille	Commonality	Part of speech	Content/function words
kgona (can)	28	1.44822592	6	Verb	content word
le@o (is, are, am)	28	1.44822592	5	Copulative verb	function word
why@cs	28	1.44822592	4	Conjunction	function word
yena (her/him)	28	1.44822592	6	Absolute pronoun	function word
gabotse (well)	27	1.39650357	5	Adverb	content word
khalara (colour)	27	1.39650357	4	Noun	content word
nnang (refuse)	27	1.39650357	6	Interjection	function word
red@cs	27	1.39650357	6	Adjective	content word
topa (pick up)	27	1.39650357	5	Verb	content word
tšwa (come out/from)	27	1.39650357	6	Verb	content word
ile (went)	26	1.34478121	6	Verb	content word
lena (yours)	26	1.34478121	5	Possessive pronoun	function word
tshela (jump over)	26	1.34478121	5	Verb	content word
yah	26	1.34478121	5	Interjection	function word
ga@n	25	1.29305886	5	Negative morpheme	function word
kgale (long ago)	25	1.29305886	6	Noun	content word
sa@tt (still)	25	1.29305886	5	Aspectual prefix	function word
school-peke (schoolbag)	25	1.29305886	6	Noun	content word
sekolo (school)	25	1.29305886	6	Noun	content word
te	25	1.29305886	6	Reflexive pronoun	function word
apara (wear)	24	1.24133651	5	Verb	content word

root word	Number of occurrences	Frequency per mille	Commonality	Part of speech	Content/function words
go@q (at, from, to)	24	1.24133651	6	Locative particle	function word
ke@j	24	1.24133651	6	Auxiliary verb	function word
lebelela (look for/at)	24	1.24133651	4	Verb	content word
monna (man, male person)	24	1.24133651	3	Noun	content word
phaka	24	1.24133651	4	Verb	content word
sa@n	24	1.24133651	5	Negative morpheme	function word
baitše (they said)	23	1.18961415	4	Verb	content word
ganijwale	23	1.18961415	4	Adverb	content word
ka@pp (at, on)	23	1.18961415	5	Temporal particle	function word
nka (i can/may)	23	1.18961415	6	Modal verb	function word
blue@cs	22	1.1378918	4	Adjective	content word
bula (open)	22	1.1378918	6	Verb	content word
duma (wish)	22	1.1378918	6	Verb	content word
gana (refuse)	22	1.1378918	6	Verb	content word
ka@j (never)	22	1.1378918	4	Auxiliary verb	function word
napile (then)	22	1.1378918	3	Auxiliary Verb	function word
two@cs	22	1.1378918	5	Adjective	content word
aga	21	1.08616944	3	Interjection	function word
iša (take to)	21	1.08616944	6	Verb	content word
jwang (how)	21	1.08616944	6	Adverb	content word
ma	21	1.08616944	5	Interjection	function word

root word	Number of occurrences	Frequency per mille	Commonality	Part of speech	Content/function words
pencil@cs	21	1.08616944	6	Noun	content word
swana (the same)	21	1.08616944	6	Verb	content word
tafola (table)	21	1.08616944	6	Noun	content word
bolela (talk/speak)	20	1.03444709	6	Verb	content word
gafa (mad)	20	1.03444709	4	Verb	content word
kgwatha (touch)	20	1.03444709	6	Verb	content word
oe (that one)	20	1.03444709	5	Demonstrative particle/pos	function word
sesi (sister)	20	1.03444709	6	Noun	content word
simba (chips)	20	1.03444709	5	Noun	content word
tšhaba (run away, escape)	20	1.03444709	5	Verb	content word
yellow@cs	20	1.03444709	5	Adjective	content word
yona (it)	20	1.03444709	6	Absolute pronoun	function word
ba@k (these ones)	19	0.98272473	6	Demonstrative particle/pos	function word
bitša (call)	19	0.98272473	6	Verb	content word
ka@u (can, could)	19	0.98272473	5	Potential morpheme	function word
kgetha (choose)	19	0.98272473	3	Verb	content word
reng (say what)	19	0.98272473	5	Verb	content word
tsamo (go and)	19	0.98272473	5	Auxiliary verb	function word
ba@j (furthermore, and)	18	0.93100238	4	Auxiliary verb	function word
bowa (come)	18	0.93100238	5	Verb	content word
letsogo (arm)	18	0.93100238	5	Noun	content word

root word	Number of occurrences	Frequency per mille	Commonality	Part of speech	Content/function words
reka (buy)	18	0.93100238	5	Verb	content word
šia (here they are)	18	0.93100238	4	Verb	content word
tjebanna	18	0.93100238	6	Interjection	function word
tšhelete (money)	17	0.87928002	5	Noun	content word
ba@m (them)	16	0.82755767	6	Object concord	function word
lahla (throw away)	16	0.82755767	5	Verb	content word
moka (all, every)	16	0.82755767	5	Noun	content word
rata (love, like)	16	0.82755767	5	Verb	content word
šala (stay, remain)	16	0.82755767	4	Verb	content word
sega (laugh)	16	0.82755767	6	Verb	content word
sekotlelo (plate)	16	0.82755767	5	Noun	content word
senya (destroy)	16	0.82755767	6	Verb	content word
apola (undress)	15	0.77583532	6	Noun	content word
dilo (things)	15	0.77583532	6	Noun	content word
neh	15	0.77583532	5	Interjection	function word
ngwatha (break off a piece)	15	0.77583532	4	Verb	content word
nwa (drink)	15	0.77583532	4	Verb	content word
tala (green/blue)	15	0.77583532	5	Adjective	content word
gabo (of their family)	14	0.72411296	4	Communal possessive pronoun	function word
ke@y (by, with)	14	0.72411296	5	Agentive particle	function word
lehono (today)	14	0.72411296	4	Noun	content word

root word	Number of occurrences	Frequency per mille	Commonality	Part of speech	Content/function words
maaka (lies)	14	0.72411296	6	Noun	content word
ntšhi (many, a lot)	14	0.72411296	4	Adjective	content word
torowa (draw)	14	0.72411296	5	Verb	content word
wo (this one)	14	0.72411296	4	Demonstrative particle/pos	function word
abuti (brother)	13	0.67239061	3	Noun	content word
bjale (now)	13	0.67239061	5	Adverb	content word
faka (put in)	13	0.67239061	3	Verb	content word
koko (grandmother)	13	0.67239061	5	Noun	content word
maabane (yesterday)	13	0.67239061	4	Noun	content word
phaphuši (classroom)	13	0.67239061	4	Noun	content word
rela (say)	13	0.67239061	5	Verb	content word
šapo (okay)	13	0.67239061	5	Adjective	content word
seeta (shoe)	13	0.67239061	4	Noun	content word
tharo (three)	13	0.67239061	5	Adjective	content word
wa@j (fall)	13	0.67239061	4	Verb	content word
ai	12	0.62066825	5	Interjection	function word
chaya	12	0.62066825	4	Verb	content word
ga@yy	12	0.62066825	4	Hortative particle	function word
grade@cs	12	0.62066825	3	Noun	content word
hlatswa (wash)	12	0.62066825	4	Verb	content word
righte (right)	12	0.62066825	4	Noun	content word
baba (spicy, pain)	11	0.5689459	4	Verb	content word

root word	Number of occurrences	Frequency per mille	Commonality	Part of speech	Content/function words
ene (and)	11	0.5689459	4	Conjunction	function word
fihla (arrive)	11	0.5689459	6	Verb	content word
goba (or)	11	0.5689459	6	Conjunction	function word
hwetša (find)	11	0.5689459	6	Verb	content word
ntle (out, outside)	11	0.5689459	4	Noun	content word
number@cs	11	0.5689459	5	Noun	content word
robala (sleep)	11	0.5689459	5	Verb	content word
seng (which/who is not)	11	0.5689459	5	Copulative verb	function word
skhaftin (lunchbox)	11	0.5689459	3	Noun	content word
thoma (start, begin)	11	0.5689459	5	Verb	content word
botse (beauty)	10	0.51722354	4	Noun	content word
ditšhila (dirt)	10	0.51722354	4	Noun	content word
eban	10	0.51722354	4	Interjection	function word
haah	10	0.51722354	3	Interjection	function word
hlogo (head)	10	0.51722354	4	Noun	content word
mfanaka (my boy)	10	0.51722354	4	Proper noun	content word
pn (name of place)	10	0.51722354	4	Proper noun	content word
šupa (point)	10	0.51722354	5	Verb	content word
three@cs	10	0.51722354	5	Adjective	content word
tllaseng (classroom)	10	0.51722354	3	Noun	content word