

UTUMIZI WA KONGOO KATIKA KUBAINISHA VIRAI VITENZI VYA KISWAHILI VYENYE HADHI YA KIKAMUSI

Seleman S. Sewangi

1.0 Utangulizi

Hakuna kitu kinachomkwaza mtumiaji wa kamusi kama kukutana na neno lenye maana nyingi halafu akashindwa kupata maana anayoitafula kwa ajili ya matumizi yake (Sinclair 1996:x; *tafsiri yetu*).

Matarajio ya mtumiaji wa kamusi, hasa anayejifunza lugha au anayetafsiri, ni kupata visawe na maana za maneno katika miktadha mbalimbali wa matumizi ya lugha. Hata hivyo, matarajio hayo mara nyingi huwa hayafikiwi hasa pale inapohusu visawe na maana za virai katika kamusi. Kwa mujibu wa Sinclair (2006) tatizo hili linatokana na mkabala wa kimapokeo wa utunzi wa kamusi ambao kwa kiasi kikubwa huhusisha maana katika kamusi na neno mojamoya. Kwa hali hiyo, maana zinazohusu virai au miambatano ya maneno huingizwa tu kwenye kamusi kwa kuhusishwa na istilahi au nahau na kwa kawaida huwa hazipewi umuhimu mkubwa. Pamoja na tatizo hilo la kimkabala, bila shaka tatizo la ubainishaji wa virai au miambatano ya maneno yenye maana kwa ajili ya kuingizwa kwenye kamusi pia ni kikwazo kingine.

Makala haya yana lengo la kuonyesha namna ya kutatua tatizo hilo la ubainishaji wa virai kwa ajili ya kamusi kwa kutumia mbinu za kikompyuta, hasa za kutumia kongoo. Mbinu za kikongoo zinahusu utumiaji wa lugha iliyokusanywa na kuhifadhiwa kwenye kompyuta kwa ajili ya kuwakilisha lugha fulani na kutumika kama chanzo cha data katika utafiti wa lugha hiyo.

Baada ya utangulizi, sehemu ya pili ya makala haya inatoa ufafanuzi mfupi kuhusu dhana ya msamiati wa kikamusi. Aidha inadokeza kuhusu kanuni mbili muhimu za kushughulikia msamiati wa kamusi: kanuni ya uteuzi huru na kanuni ya unahau kama zilivyofafanuliwa na Sinclair (1987). Sehemu ya tatu inabainisha vyanzo vya msamiati wa kamusi kwa kuzingatia mikabala miwili, pamoja na mbinu zake za uteuzi wa msamiati wa kamusi: mkabala wa kimapokeo na mkabala wa kikongoo. Sehemu ya nne inatoa picha ya jumla kuhusu uteuzi wa msamiati katika kamusi za Kiswahili. Sehemu ya tano inafafanua kuhusu mwelekeo wa utumiaji wa mbinu ya kikongoo katika ubainishaji wa msamiati wa kamusi za Kiswahili ikiwa ni pamoja na namna ya kubainisha virai vitenzi kwa kutumia kongoo ya Kiswahili. Sehemu ya sita inabainisha mambo kadhaa ambayo yasipozingatiwa yanaweza kusababisha upungufu katika mbinu ya ubainishaji wa msamiati wa kamusi kikongoo.

Sehemu ya saba inadokezea utaratibu wa kuingiza virai vitenzi katika kamusi ya lugha moja na ya lugha mbili. Sehemu ya nane ni hitimisho la makala.

2.0 Msamiati wa Kikamusi

Kimapokeo, neno ndilo linalobeba maana na kamusi ni orodha ya maneno yaliyopangiliwa kialfabeti pamoja na maana ya/za kila neno. Pamoja na neno kuchukuliwa kama msingi wa maana katika kamusi, ziko pia maana nyingine kwenye kamusi ambazo hutokana na uambatano wa maneno. Hata hivyo maana hizo zimekuwa hazipewi umuhimu mkubwa zaidi ya kuingizwa tu kama nahau au istilahi (Sinclair 2006:26).

Kwa mujibu wa Sinclair (2006: 162), kufanywa kwa neno kuwa ndio msingi wa maana katika kamusi kunatokana na utamaduni wa kutenganisha kamusi na sintaksia. Katika utengano huo, sintaksia huhusishwa na kanuni za muundo wa lugha kwa kuzingatia uhusiano wa kikategoria huku kamusi ikihusishwa na maana ya neno mojamoya bila kutazama jinsi mahusiano ya kisintaksia yanavyoweza kuathiri maana ya neno. Kwa msingi huo, uchambuzi wa muundo wa lugha hauhusishi maana na ufafanuzi wa maana ya neno hautazami uhusiano wa maneno kama vile uhusiano wa kitenzi na yambwa. Kwa mujibu wa Sinclair (kameshatajwa) mgawanyo huo haupaswi kuwoko kwani hauakisi kabisa uhalisi wa lugha hasa katika mahusiano ya kimaana baina ya vipashio vya kisintaksia vya ngazi ya juu kuliko neno. Ushahidi wa hivi karibuni uliopatikana kwenye kongoo unaonyesha kwamba hakuna sababu yoyote na ni upotoshaji kutenganisha vipengele vya sintaksia na kamusi hasa wakati wa kushughulikia maana katika matini. Sinclair anadai kwamba vipengele hivyo vinaweza kushughulikiwa pamoja kwa kutumia kanuni mbili za ufafanuzi wa maana katika matini ambazo japo hufanya kazi pamoja kila kanuni huelemea upande wake.

2.1 Kanuni ya Uteuzi huru na Kanuni ya Unahau

Kwa mujibu wa Sinclair (2006:29) ubainishaji wa maana katika matini hauna budi kuhusisha viwango viwili vya ufafanuzi: kiwango cha neno mojamoya na kiwango cha uambatano wa maneno au virai. Viwango hivi viwili havina budi kushughulikiwa kwa kuzingatia kanuni mbili tofauti: kanuni ya uteuzi huru ambayo inahusu maana kwenye matini katika kiwango cha neno mojamoya na kanuni ya unahau inayohusu maana katika kiwango cha uambatano wa maneno au virai katika matini. Kimsingi, kanuni ya uteuzi huru inaoana na dhana kwamba neno ndio msingi wa maana katika lugha. Kulingana na kanuni hii, kila neno katika matini lina maana yake ambayo inaweza pia kupatikana katika kamusi. Aidha mahusiano ya maneno katika matini sio ya kimaana bali ni ya

kisarufi. Sinclair (kimeshatajwa) anadai kwamba kanuni ya uteuzi huru ambayo kwa kiasi kikubwa inaoana na ubainishaji wa maana kimapokeo ingetosheleza tu iwapo kweli kila neno katika lugha lingekuwa na maana moja tu kama vile ilivyo kwa istilahi. Lakini kwa vile hiyo sio hali halisi ya lugha. Sinclair (kimeshatajwa) anaendelea kudai kwamba ushahidi wa hivi karibuni wa kikongoo umethibitisha kwamba maana katika matini inahusiana zaidi na miambatano ya maneno au virai, hasa virai vitenzi, kuliko na neno mojamoya.

Kanuni ya unahau inaingiza maana zote katika matini zinazohusiana na miambatano ya maneno ikiwa ni pamoja na nahau, virai pamoja na misemo. Hivyo basi kwa mujibu wa kanuni hii na ile ya uteuzi huru, msingi wa maana unaweza kuwa neno au kirai. Utamaduni wa kutotambua virai kama msingi mmojawapo wa maana katika kamusi umesababisha upungufu katika kamusi nyingi, hasa kwa wanafunzi wa lugha na wafasiri, kwa kutoingiza virai mbalimbali na maana zake katika kamusi.

3.0 Vyanzo vya Msamiati wa Kamusi

Sinclair (1991) anabainisha vyanzo vitatu vya msamiati na maana zake katika kamusi ambavyo ni: (1) kamusi zilizokwishatungwa (2) ujuzi wa watumiaji wa lugha na (3) matumizi halisi ya lugha. Hata hivyo utumizi wa vyanzo hivi hutofautiana kulingana na mkabala unaotumika katika utunzi wa kamusi. Pearson (1996) anabainisha mikabala mikuu miwili ya ukusanyaji wa vidahizo vya kamusi: (1) mkabala wa kimapokeo unaotumia mbinu zilizoasisiwa na Johnson katika karne ya kumi na nane wakati akitunga kamusi yake ya kwanza ya Kiingereza ambayo ilichapishwa mwaka 1755, (2) mkabala wa kikongoo ambao hutumia mbinu za kikompyuta zilizoasisiwa na Mradi wa utafiti wa COBUILD katika Chuo Kikuu cha Birmingham katika miaka ya themanini.

3.1 Mkabala wa Kimapokeo

Mbinu za kimapokeo za ukusanyaji wa vidahizo vya kamusi huongozwa zaidi na uwezo na ridhaa ya mkusanyaji ambaye hufanya kazi hiyo kwa njia ya kupitia kwa makini maandiko mbalimbali, ikiwa ni pamoja na kamusi za zamani, na kukusanya msamiati ambao aghalabu huwa ni wa neno mojamoya. Kazi nzima hufanyika kwa mikono na kwa hiyo sio tu kwamba huchukua muda mrefu bali pia huchosha. Akielezea jinsi Johnson alivyokusanya msamiati pamoja na maelezo mengine kwa ajili ya kamusi yake ya Kiingereza, Pearson (1996:73-74) anasema kwamba Johnson alianza kwa kusoma maandiko mbalimbali ya maandishi wa Kiingereza ambao aliamini kwamba walitumia Kiingereza sanifu, huku akiweka alama kwa kalamu mwanzoni na mwishoni

mwa kila aya iliyokuwa na msamiati pamoja na mifano ya matumizi, kwa ajili ya kamusi yake. Kutokana na ugumu wa kazi, mkusanyaji wa msamiati hulazimika kuteua vidahizo vyake kutoka katika maandiko machache tu, aghalabu kamusi zilizoko.

Bila shaka mbinu hii ndiyo iliyotumika pia katika ukusanyaji wa vidahizo vya *Kamusi ya Kiswahili-Kingereza* (TUKI 2001) ambapo tunaeleza kwamba msamiati wake "...umetokana na Kiswahili sanifu na umekusanywa kutoka matini za msingi za Kiswahili sanifu, kama vile kamusi za Kiswahili zilizopo: TUKI (1981) *Kamusi ya Kiswahili Sanifu*, TUKI (2001) *Kamusi ya Kiswahili Sanifu* (toleo jipya), Johnson (1939) *Kamusi ya Kiswahili-Kingereza* na Feeley (1990) *Kamusi ya Kiswahili-Kingereza*. Maneno mengine yametokana na orodha ya msamiati wa fani mbalimbali ulioundwa au kusanifiwa na BAKITA, na kutumiwa kufundishia katika shule za msingi" (TUKI 2001:1). Sinclair (1991) anaeleza kwamba kutegemea sana kamusi zilizoko katika uteuzi wa vidahizo kuna hatari ya kushindwa kubaini kirahisi kama neno au maana inayoteuliwa bado inatumika au la, hali ambayo inaweza kusababisha uteuzi wa vidahizo na maana ambazo hazitamsaidia mtumiaji wa kamusi kutokana na kutotumika kabisa katika lugha. Tatizo hili limechochea kwa kiasi kikubwa kuzuka kwa mkabala mpya wa ukusanyaji wa vidahizo vya kamusi na maana zake kikongoo.

3.2 Mkabala wa Kikongoo

Chimbuko la mkabala wa kikongoo ambao pia hufahamika kama mkabala wa COBUILD (*Collins Birmingham University International Language Database*) ni mradi mkubwa wa utafiti ulioanzishwa katika Chuo Kikuu cha Birmingham katika miaka ya themanini kwa kugharimiwa na wachapishaji wa Collins. Lengo kuu la mradi lilikuwa ni kuunda hazina kubwa ya kongoo itakayowakilisha kwa utoshelevu lugha ya Kiingereza, hasa cha Uingereza na Marekani, kwa ajili ya kutumika kama chanzo cha data ya utafiti mbalimbali na ufundishaji wa lugha hiyo. Miongoni mwa matokeo makubwa ya mradi huo ni kuundwa kwa kongoo ya Kiingereza (*the Bank of English*) yenye zaidi ya maneno milioni 300, na kulichapishwa kamusi ya *Collins COBUILD English Dictionary* (1987 pamoja na matoleo mapya ya pili (1995), tatu (2001), nne (2003) na tano (2006). Kwa mujibu wa Sinclair (1987: xv), ambaye alikuwa mratibu wa mradi huo hadi alipostaafu mwaka 1997, Mradi ulitumia mbinu mpya kabisa za utunzi wa kamusi zilizotumia teknolojia ya kikompyuta ya hali ya juu. Katika mbinu hizo mpya, aina ya taarifa zilizoingizwa katika kamusi, ubora wake pamoja na uwasilishaji wake ni tofauti kabisa na zile za kimapokeo. Sinclair (1995: vii) anaeleza kwamba taarifa zote za kikamusi, ikiwa ni pamoja

na vidahizo, maana zake pamoja na mifano, huchotwa kwenye kongoo na wala sio kwenye vichwa vya watunga kamusi au vya wajuzi wa lugha. Vidahizo vinajumuisha neno mojamoya pamoja na virai vyenye hadhi ya vidahizo. Uteuzi wa vidahizo vya ngazi ya neno mojamoya hufanywa kwa kutumia programu yenye uwezo wa kuhesabu idadi ya urudio wa neno katika matini ambapo maneno yenye idadi kubwa ya urudio hupewa nafasi kubwa ya kuteuliwa. Maneno yenye idadi ndogo ya urudio (chini ya tano) au yasiyo na urudio kabisa pia huteuliwa baada ya kuchunguzwa umuhimu wake katika matumizi ya lugha.

Kwa upande wa uteuzi wa vidahizo vya ngazi ya virai, uteuzi hufanywa kwa msaada wa programu yenye uwezo wa kubaini miambatano mbalimbali ya maneno katika kongoo kwa msingi wa neno kuu katika uambatano au neno kifundo. Kwa mfano, ikiwa neno kifundo ni kitenzi cha Kiswahili 'cheza' programu inaweza kuonyesha ruwaza mbalimbali za utokeaji pamoja wa kitenzi hicho na maneno mengine katika kongoo. Kwa kawaida programu huonyesha maneno manne kulia na manne kushoto mwa neno kifundo. Hata hivyo inaweza kuonyesha maneno pungufu au zaidi kulingana na itakavyoelekezwa. Programu hii pia husaidia katika kupata mifano ya matumizi ya neno au kirai katika kongoo kwa kurejelea moja kwa moja matumizi ya neno kifundo katika kongoo. Sinclair (1995:ix) anasisitiza kwamba ukusanyaji wa msamiati wa kamusi kwa mbinu za kikongoo hauna budi kuzingatia matumizi ya kongoo inayowakilisha lugha halisi katika matumizi yake mbalimbali ya kimazungumzo na kimaandishi.

4.0 Uteuzi wa Msamiati katika Kamusi za Kiswahili

Historia ya utunzi wa kamusi za Kiswahili Sanifu inarudi nyuma hadi kwenye nusu ya kwanza ya karne ya 20 ambapo Frederick Johnson alitunga kamusi ya *A Standard Swahili-English Dictionary* (1939). Kwa mujibu wa Mdee (1997:89) kamusi hiyo ilitungwa kutokana na kudurusu na kuandika upya kamusi ya Madan (1903) ya Kiswahili cha lahaja ya Unguja Mjini. Johnson alitunga pia kamusi ya *A Standard English Swahili Dictionary* (1939). Baada ya hapo, kazi ya utunzi wa kamusi za Kiswahili sanifu, kwa kiasi kikubwa, imekuwa ikifanywa na Taasisi ya Uchunguzi wa Kiswahili ya Chuo Kikuu cha Dar es Salaam. Miongoni mwa kamusi ambazo zimeshatungwa na taasisi hiyo ni *Kamusi ya Kiswahili Sanifu* (1981); *Kamusi ya Kingereza - Kiswahili* (1996) na *Kamusi ya Kiswahili- Kiingereza* (2001).

Kwa jumla uteuzi wa vidahizo vya kamusi za Kiswahili hadi sasa ni wa kimapokeo kwa maana kwamba hautumii mbinu za kikompyuta. Watunga kamusi hukusanya vidahizo kutokana na vyanzo mbalimbali hasa kamusi zilizoko pamoja na maandiko mbalimbali. Uteuzi na uwasilishaji wa vidahizo

hutegemea kwa kiasi kikubwa ridhaa ya jopo la watunzi wa kamusi. Tukiangalia, kwa mfano, *Kamusi ya Kiswahili-Kingereza* (TUKI 2001:2) tunapata picha halisi ilivyo kuhusu vidahizo. Tunaelezwa kwamba “Vidahizo vingi katika kamusi hii ni neno mojawoja isipokuwa vichache tu ambavyo ni maneno ambatani yenye kuwakilisha dhana moja: **mama mkwe; mamantilie; mlango wa fahamu; pimajoto**”. Aidha “maneno ambatani yameingizwa kama kidahizo mfuto ambacho ni ... neno ambatani ambalo limeingizwa ndani ya maelezo ya kidahizo kikuu (ambacho ni umbo la msingi) na kutolewa maelezo ya maana na ya kisarufi. Iwapo neno ambatani lililoingizwa kama kidahizo mfuto halitumiki sana au halijapata kutumika lakini huenda likatumika siku moja hupatiwa maelezo ya kisarufi pekee.”

Kutokana na maelezo hayo ni dhahiri kwamba vidahizo vya ngazi ya kirai havipewi umuhimu katika uteuzi wa msamiati wa kamusi za Kiswahili zaidi ya kuhusishwa tu na maana mbalimbali za kidahizo kikuu. Bila shaka haya ni matokeo ya mwelekeo wa kimapokeo unaochukulia kwamba neno ndio msingi wa maana katika kamusi, mwelekeo ambao hautili manani maana zinazotokana na uambatano wa maneno au virai. Hata pale ambapo maneno ambatani yameingizwa katika kamusi kama sehemu ya maelezo ya kidahizo kikuu, bado kuna tatizo la ukosefu wa vigezo bayana vya uteuzi wa maneno ambatani yenye sifa ya kuwa vidahizo mfuto vinavyostahili na visivyostahili kutolewa maana katika kamusi. Kwa mfano, tunapoangalia ndani ya *Kamusi ya Kiswahili-Kingereza* (TUKI 2001:2) tunabaini kwamba baadhi ya maneno ambatani ambayo ni virai vitenzi, kama vile ‘*piga yowe*’; ‘*piga ramli*’; ‘*piga pambaja*’ ‘*piga pande*’, na ‘*piga picha*’ yameingizwa ndani ya maelezo ya vidahizo vikuu ambavyo ni nomino. Wakati huohuo, maneno mengine ambatani kama ‘*piga ngumi*’; ‘*piga chafya*’; ‘*piga kura*’; ‘*piga miayo*’; yameingizwa ndani ya maelezo ya kidahizo kikuu ambacho ni kitenzi. Mkanganyiko wa namna hii humkwaza mtumiaji wa kamusi katika kubaini wapi hasa katika kamusi linapatikana neno ambatani. Aidha, kwa utaratibu huo, sio rahisi kupata picha halisi ya maneno ambatani yaliyoingizwa kwenye kamusi chini ya maelezo ya kidahizo kikuu kimoja. Kwa mfano, si rahisi kufahamu idadi ya maneno ambatani yaliyoingizwa chini ya kidahizo kikuu ‘*piga*’ (katika kamusi hii ya TUKI maneno ambatani yaliyoingizwa chini ya kitenzi *piga* ni 12 tu) kutokana na mengine kuingizwa chini ya vidahizo vikuu vya kategoria ya nomino. Bila shaka matatizo ya aina hii ni changamoto inayosukuma kuelekea katika utumizi wa mbinu mpya katika utunzi wa kamusi za Kiswahili

5.0 Ubainishaji Msamiati wa Kamusi za Kiswahili Kikongoo

Mkabala wa kukusanya msamiati wa kamusi kwa kutumia mbinu za kikongoo ni mgeni katika utunzi wa kamusi za Kiswahili. Mkabala huo

umewahi kutumika katika uteuzi wa vidahizo ya kamusi ya *Swahili-Suomi-Swahili* (2002). Hata hivyo mbinu iliotumika katika uteuzi wa vidahizo vya kamusi hiyo ni tofauti na mbinu za COBUILD ambazo zimefafanuliwa hapo juu. Uteuzi huo umetumia mbinu ya kongoo iliyoalamishwa ilihali mbinu za COBUILD hutumia kongoo isiyoalamishwa. Mbinu hizi za kutumia kongoo isiyoalamishwa hazijawahi kutumika katika uteuzi wa vidahizo vya kamusi za Kiswahili. Hii inatokana na ukweli kwamba utekelezaji wa mbinu hizo hauna budi kutanguliwa na uundaji wa kongoo kubwa Kiswahili yenye kuwakilisha kwa utoshelevu lugha hiyo katika mazingira yake halisi ya matumizi mbalimbali ya kiuandishi na kimazungumzo.

5.1 Kongoo ya Kiswahili na Zana za Kikompyuta

Juhudi za awali za uundaji wa kongoo ya Kiswahili zilianza katika miaka ya 80 chini ya uratibu wa Taasisi ya Taaluma za Asia na Afrika ya Chuo Kikuu cha Helsinki, Finland. Hadi sasa juhudi hizo zimewezesha ukusanyaji wa kongoo ya Kiswahili yenye maneno milioni kumi na mbili na mia tano (<http://www.csc.fi/english/research/software/hcs>). Hii ni kongoo ndogo kwa ulinganisho ambayo imejengwa kwa kiasi kikubwa kwa Kiswahili cha magazeti ya Tanzania. Kwa kutambua hilo, pamoja na umuhimu wa kuwa na kongoo kubwa inayowakilisha kwa utoshelevu Kiswahili cha Afrika Mashariki, Taasisi ya Uchunguzi wa Kiswahili ya Chuo Kikuu cha Dar es Salaam (sasa hivi Taasisi ya Taaluma za Kiswahili) imeanzisha Mradi wa Kuandaa Mazingira ya Utumizi wa Kompyuta katika Utafiti na Ufundishaji wa Kiswahili. Miongoni mwa madhumuni ya mradi huo ulioanzishwa mwaka 2008 ni kuunda kongoo kubwa ya Kiswahili pamoja na kuwezesha upatikanaji na utumizi wa zana stahiki za kikompyuta kwa ajili ya utafiti na ufundishaji wa Kiswahili. Kazi ya uundaji wa kongoo tayari inaendelea kwa kukusanya na kuingiza kwenye kompyuta maandiko mbalimbali pamoja na lugha usemi ya Kiswahili. Sehemu ya kongoo hiyo imetumika katika makala haya kama chanzo cha data katika ubainishaji wa virai vitenzi vya kikamusi. Sehemu hiyo ya kongoo ina jumla ya maneno milioni mbili na mia tano yaliyokusanywa kutokana na maandiko mbalimbali ya kitaalamu na ya kawaida.

Kwa upande wa zana za kikompyuta, zipo ambazo ni za jumla yaani zinazoweza kutumika katika uchambuzi wa kongoo ya lugha yoyote, na zipo ambazo ni mahususi kwa ajili ya uchambuzi wa kongoo ya Kiswahili tu. Mfano wa zana za jumla ni kama vile programu ya Wordsmith (<http://www.lexically.net/wordsmith/>) na programu ya *Ant Conc*. (<http://ell.phil.tu-chemnitz.de/analysis/antConc.html>). Programu hizi zina zana ndogondogo zenye uwezo wa kufanya kazi mbalimbali za uchambuzi wa

kongoo. Miongoni mwa zana hizo, zipo za kupanga maneno yote ya kongoo katika orodha ya neno mojamoya na kuonyesha idadi ya urudio wa kila neno katika kongoo. Pia kuna zana za kuibua neno lolote katika kongoo pamoja na mistari lilikotumika huku neno hilo likiwa katikati ya kila mstari. Zana nyingine ni za kutambua ruwaza mbalimbali za utokeaji wa neno lolote katika kongoo pamoja na maneno mengine yaliyoko kulia na/au kushoto kwake pamoja na za kubainisha tungo za hadhi mbalimbali katika kongoo, na za kutambua mazingira mbalimbali ya matumizi ya kila neno katika kongoo. Kwa upande wa programu mahususi, tunazo programu za SALAMA (Swahili Language Manager) zilizobuniwa na kuundwa huko katika Chuo Kikuu cha Helsinki, Finland (<http://www.njas.helsinki.fi/pdf-files/vol8num2/hurskainen.pdf>). SALAMA ni kivunge cha programu zenye uwezo wa kufanya kazi mbalimbali zinazohusiana na kongoo ya Kiswahili, kama vile kukagua tahajia ya kila neno, kuunganisha maneno ambatani kwenye kongoo, kuingiza alama za kisarufi katika kila neno kwenye kongoo pamoja na kupanga maneno ya kongoo katika orodha ya neno mojamoya na kubainisha idadi ya urudio wa kila neno. Aidha programu za SALAMA zinaweza kuanzisha kisarufi na kietimolojia kila neno katika kongoo. Makala haya yametumia programu ya jumla ya *Anti-Conc* katika kubainisha virai vitenzi vya Kiswahili vyenye hadhi ya kikamusi.

5.2 Mbinu za Ubainishaji wa Virai Kikongoo

Msingi wa ubainishaji virai kikongoo ni kanuni ya unahau, iliyodokezwa hapo awali, kuhusu tabia ya maneno ya kuambatana pamoja katika lugha ya mazungumzo au ya maandishi na kuunda uambatano wa kimaana unaotawaliwa na neno kuu la uambatano, yani neno ambalo ndilo msingi wa uteuzi wa neno au maneno mengine katika uambatano. Uambatano wa kimaana wa neno kuu pamoja na neno au maneno mengine hujibainisha kwa kujirudiarudia katika matini na kujenga ruwaza za uambatano za neno kuu, ambazo ni neno au maneno yote yanayoambatana mara kwa mara na neno kuu katika kuwakilisha maana mojawapo ya neno hilo. Hivyo basi zoezi la ubainishaji wa virai kikongoo linahusu utambuaji wa ruwaza mbalimbali za uambatano wa kimaana za kila neno kuu linalohusika kwa kutumia programu yenye uwezo wa kufanya utambuzi huo. Hata hivyo katika kutambua ruwaza za uambatano za neno kuu husika, programu 'huokota' kila aina ya uambatano bila kujali kama ni wa kimaana au la. Hivyo mchakato mzima huhusisha upitiaji kwa makini wa data iliyotambuliwa na programu ili kubaini ruwaza za uambatano ambazo zinabeba maana mbalimbali za neno kuu. Hili hufanyika kwa kutumia umaizi wa lugha na utaalamu wa kiisimu wa mtafiti anayehusika. Katika sehemu inayofuata tunaonyesha mfano wa namna ya ubainishaji kikongoo wa ruwaza za

uambatano wa kimaana kwa kutumia baadhi ya vitenzi vya Kiswahili kama neno kuu la uambatano. Vitenzi vinavyohusika katika ubainishaji ni: 'piga' 'omba'; 'funga'; 'zaa' na 'twaa'.

5.2.1 Ubainishaji wa Ruwaza za Uambatano wa Kimaana za vitenzi

Katika kubainisha ruwaza za uambatano za kitenzi cha Kiswahili, tumezingatia kwamba kimuundo, neno kuu katika kirai cha Kiswahili hutokea kushoto ilihali kijalizo chake hutokea kulia. Kwa mfano, kitenzi elekezi hutokea kushoto mwa nomino ya yambwa. Hivyo ili kuweza kupata ruwaza mbalimbali za uambatano wa kimaana za kila kitenzi, inabidi programu ielekezwe kutambua katika kongoo uambatano mbalimbali wa kitenzi husika pamoja na neno moja la kulia kwake. Jedwali la 1 hapa chini linaonyesha sehemu ndogo ya ruwaza mbalimbali za uambatano za kitenzi 'piga' na neno la kwanza kulia kwake, kama zilivyoibuliwa na programu kutoka kwenye kongoo.

Jedwali la 1 Sampuli ya ruwaza za uambatano za kitenzi 'piga'

	PIGA	Kiambata cha kwanza kulia	Idadi ya urudijaji
1		Kura	45
2		Hatua	24
3		Vita	21
4		Kelele	19
5		Sindano	15
6		Kwa	15
7		Upesi	14
8		Marufuku	14
9		Chafya	13
10		Picha	11
11		Yai	10
12		Pasi	10
13		Mbio	10
14		Magoti	9
15		Chapa	9
16		Mswaki	6
17		Punyeto	5
18		Taipu	4
19		Sana	4
20		Debe	4
21		Na	3
22		Muziki	3
23		Kampeni	3
24		Jeki	3
25		Honi	3
26		Hadi	3

Katika sampuli hii, kama ambavyo tumedokeza hapo awali, baadhi ya ruwaza zinazoibuliwa na programu hazina hadhi ya kimaana kama ambavyo inajidhihirisha katika ruwaza namba 6, 7, 21 na 26. Aidha kila ruwaza ya uambatano inajirudiarudia ambapo urudio wa juu kabisa ni mara 45 na wa chini kabisa ni mara tatu. Baada ya programu kuibua ruwaza za uambatano za kila kitenzi na kuhaririwa kwa kuondosha zile ambazo hazikukidhi kimaana tulipata idadi ya ruwaza za kimaana kwa kila kitenzi kama ifuatavyo: *piga* 55; *omba* 9; *funga* 21; *zaa* 7 na *twa* 6. Ruwaza hizo ni kama zilivyoorodheshwa katika jedwali la 2 hadi la 6 hapa chini.

Jedwali la 2, Ruwaza za uambatano wa kimaana za kitenzi '*piga*'

	PIGA	Kiambata cha kwanza kulia
1		Bomba
2		bongo
4		Brashi
5		Breki
6		Chafya
7		Chapa
8		Chapuo
9		Chenga
10		Dana
11		Debe
12		Eksrei
13		Hatua
14		Hesabu
15		Hodi
16		Honi
17		Ishara
18		Jeki
19		Kampeni
20		Kelele
21		Kinubi
22		Kite
23		Kofi
24		Kope
25		Kura
26		Magoti
27		Maji
28		Makofi
29		Marufuku
30		Mateke
31		Mawe
32		Mbio
33		Mhuri

34		Miayo
35		Mishale
36		Mluzi
37		Mnada
38		Mswaki
39		Muziki
40		Pasi
41		Picha
42		Punyeto
43		Ramli
44		Rangi
45		Ripoti
46		Shabaha
47		Sindano
48		Stori
49		Suti
50		Taipu
51		Teke
52		Viboko
53		Vita
54		Yai
55		Yowe

Jedwali la 3, Ruwaza za uambatano wa kimaana za kitenzi 'omba'

	OMBA	Kiambata cha kwanza kulia
1		Daktari
2		Fedha
3		Ridhaa
4		Rushwa
5		Samahani
6		Ushuru
7		Kura
8		Mungu
9		Nafasi

Jedwali la 4, Ruwaza za kimaana za kitenzi uambatano wa 'funga'

	FUNGA	Kiambata cha kwanza kulia
1		Bao
2		Barabara
3		Choo
4		Dirisha
5		Gerezani
6		Jeraha
7		Kamba
8		Kazi
9		Kitambaa
10		Kuhara
11		Kukojoa
12		Mabao
13		Mahesabu
14		Mdomo
15		Mikanda
16		Mjadala
17		Ndoa
18		Pingu
19		Ramadhani
20		Timu
21		Virago
22		Warsha

Jedwali la 5 Ruwaza za uambatano wa kimaana za kitenzi 'zaa'

	ZAA	Kiambata cha kwanza kulia
1		Faida
2		Mapacha
3		Matunda
4		Mizozo
5		Chuki
6		Mtoto
7		Mbu

Jedwali la 6 Ruwaza za uambatano wa kimaana za kitenzi 'twaa'

	TWAA	Kiambata cha kwanza kulia
		Ardhi
		Jimbo
		Kombe
		Taji
		Ubingwa
		Ubunge

6.0 Upungufu wa Ubainishaji Msamiati wa Kamusi Kikongoo

Kimsingi ufanisi wa utumizi wa mbinu ya ukusanyaji wa vidahizo vya kamusi kikongoo hutegemea ubora wa kongoo kiukubwa na kiuwakilishi. Umuhimu wa ukubwa wa kongoo unatokana na ukweli kwamba kigezo kikubwa cha uteuzi wa msamiati ni idadi ya urudio wa neno au kirai katika kongoo. Endapo ukusanyaji wa msamiati wa kamusi utafanyika kwa kutumia kongoo ndogo, kutakuwa na upungufu wa kukosekana kwa msamiati wa kutosheleza. Kwa mfano, kutokana na udogo wa kongoo iliyotumika katika makala haya, baadhi ya virai vitenzi havikuweza kubainishwa. Mfano katika jedwali la 3 hapo juu, hatuna virai kama vile ‘omba radhi; omba msamaha; omba msaada; omba dua, na vingine vingi, chini ya kitenzi ‘omba’ japo virai hivyo hutumika sana katika Kiswahili.

Uwakilishi katika kongoo unahusu upatikanaji wa data ya lugha katika mazingira mbalimbali ya matumizi yake. Kipengele hiki kinahusu hasa suala la maana ya msamiati kimuktadha. Endapo msamiati wa kamusi utakusanywa kwa kutumia kongoo inayohusu mazingira ya matumizi ya aina moja tu, msamiati huo utakuwa na upungufu wa utoshelevu wa msamiati na maana yake kimuktadha. Kwa mfano, sio rahisi kupata maana ya kimichezo (mpira wa miguu) ya kirai ‘piga chenga’ kama kongoo inayohusika haina uwakilisho wa matumizi ya lugha katika eneo hilo.

Kwa upande mwingine, ukusanyaji wa msamiati wa kamusi kwa kutumia kongoo kubwa, kwa mfano kongoo ya maneno milioni mia tano, pamoja na kwamba utakuwa na utoshelevu kidata, bado unaweza kusababisha tatizo la kuwa na data kubwa mno. Hili linaweza kuwa ni tatizo kwa upande wa gharama za uchapishaji wa kamusi pamoja na soko la kamusi. Uchapishaji wa kamusi kubwa sana hugharimu pesa nyingi hali ambayo kwa upande mwingine huifanya bei ya kamusi kuwa kubwa na kusababisha ukosefu wa soko, hasa katika nchi maskini ambapo watumiaji wengi wa kamusi hawana uwezo mkubwa kifedha. Hata hivyo jambo muhimu la kuzingatia katika kuepuka tatizo hili ni kuteua data ya kamusi kwa kuzingatia malengo, gharama pamoja na soko la kamusi.

7.0 Uingizaji wa Virai katika Kamusi

Baada ya ubainishaji kikongoo wa ruwaza mbalimbali za uambatano wa kimaana, au virai, kwa kila neno husika, inakuwa rahisi kuamua ruwaza ambazo zitaingizwa katika kamusi kulingana na lengo la kamusi yenyewe. Uingizaji huo hauna budi kuzingatia kwamba kila ruwaza ya uambatano hubeba maana mojawapo ya neno kuu katika uambatano. Hivyo ruwaza ya uambatano haina budi kuingizwa kama kidahizo kidogo huru ndani ya kitomeo cha neno kuu

husika. Maana ya kila ruwaza ya uambatano itafafanuliwa kwa mifano kutoka kwenye kongoo. Kwa mfano, uingizaji wa kitenzi 'piga' pamoja na virai vyake, utanza kwa kuingiza kwanza kitenzi hicho na kufafanua maana yake. Halafu kila kirai kitaingizwa chote kama kidahizo huru kidogo kama sehemu ya kitomeo cha kitenzi hicho na kufafanuliwa maana yake kwa kuzingatia muktadha wa matumizi katika kongoo. Kwa utaratibu huo, katika kitomeo cha kitenzi 'piga' na virai vyake kama vilivyobainishwa katika makala haya, kutakuwa na kidahizo kikuu kimoja pamoja na vidahizo vidogo huru 55, kila kimoja kikiwa kimefafanuliwa kivyake.

Kwa upande wa kamusi ya lugha mbili, ufafanuzi wa maana ya kirai utahusu uingizaji wa kisawe cha kirai katika lugha lengwa, na ikibidi maelezo mengine ya matumizi yake katika kongoo. Kwa virai ambavyo kisawe cha neno kuu hakibadiliki, hivi vitaingizwa ndani ya maelezo ya neno kuu. Kwa mfano, kama neno kuu ni 'funga' na kisawe chake cha Kiingereza ni 'close' virai vyote ambavyo kisawe cha neno kuu ni 'close' vitaingizwa ndani ya maelezo ya kidahizo kikuu 'funga'. Kwa maana nyingine virai vya aina hiyo havitakuwa na hadhi ya vidahizo vidogo huru. Kwa utaratibu huu, kwa mfano, virai 'funga barabara', 'funga dirisha', 'funga warsha', vitaingizwa ndani ya maelezo ya kidahizo kikuu 'funga'. Kwa upande mwingine virai 'funga bao', 'funga choo', 'funga ndoa', na 'funga mikanda', vitaingizwa kama vidahizo vidogo huru pamoja na visawe vyake kama sehemu ya kitomeo cha neno kuu. Uingizaji wa virai kama vidahizo vidogo huru kama sehemu ya kitomeo cha neno kuu, utamnufaisha mtumiaji wa kamusi, hasa mfasiri na mjifunzaji lugha, anayetafuta visawe vya virai mbalimbali. Hii inatokana na ukweli kwamba virai vyote vya neno kuu vitakuwa sehemu ya kitomeo cha neno hilo. Aidha uingizaji wa virai kama vidahizo vidogo huru pamoja na maana au visawe vyake, utampatia mtumiaji maana au kisawe cha kila kirai kulingana na mahitaji yake. Kwa upande wa mtunga kamusi, aina hii ya uingizaji wa virai itamsaidia katika kufanya tathmini ya virai vya maneno mbalimbali katika kamusi kwa lengo la kubaini na kuongeza vile vinavyokosekana katika kamusi.

8.0 Hitimisho

Makala yalikusudiwa kuonyesha namna ambavyo ubainishaji wa virai vya kikamusi unaweza kufanywa kwa kutumia mbinu za kikompyuta, hususan za kikongoo. Katika kufanya hivyo, makala yamedokeza kanuni mbili za msingi ambazo hazina budi kuzingatiwa wakati wa uteuzi wa msamiati wa kamusi: kanuni hizo ni ile ya uteuzi huru na ya unahau. Aidha makala yamegusia kuhusu mikabala inayotumika katika kuteua msamiati wa kamusi, pamoja na mbinu za kila mkabala. Makala imebainisha kwamba baada ya kipindi kirefu cha uteuzi wa msamiati wa kamusi za Kiswahili kwa kutumia mbinu ya kimapokeo, juhudi

za utumiaji wa mbinu mpya za kikompyuta tayari zimeanza ambapo lengo ni kutumia kongoo kama chanzo cha msamiati wa kamusi ikiwa ni pamoja na msamiati wa ngazi ya virai. Kwa kutumia kongoo ya Kiswahili pamoja na programu ya *Ant-Con*, makala yamebainisha jinsi ambayo ruwaza mbalimbali za uambatano wa kimaana za vitenzi zinaweza kubainishwa kutoka kwenye kongoo kwa ajili ya kuingizwa kwenye kamusi. Kutokana na udogo wa kongoo iliyotumika katika makala haya, matokeo ya ubainishaji uliofanyika humu ni mfano tu wa zoezi zima la ubainishaji kikongoo wa virai vya Kiswahili vyenye hadhi ya kamusi.

Marejeo

- Anthony, L. (2005). *AntConc: Design and Development of a Freeware Corpus Analysis* (<http://ell.phil.tu-chemnitz.de/analysis/antConc.html>): COLLINS English Dictionary (Toleo la 3). Birmingham: HarperCollins Publishers.
- Hunston, S & G. Francis (1996). *Pattern Grammar: A corpus-driven approach to the lexical grammar of Englishi*. John Benjamins Publishing Company: Amsterdam/Philadelphia.
- Johnson, F. (1939) . *A Standard English- Swahili Dictionary*. Dar es Salaam, Nairobi: Oxford University Press.
- Johnson, F. (1939) . *A Standard Swahili- English Dictionary*. Dar es Salaam, Nairobi: Oxford University Press.
- Mdee, J. S. (1997). *Nadharia na Historia ya Leksikografia*. Dar es Salaam: Taasisi ya Uchunguzi wa Kiswahili, Chuo Kikuu cha Dare es Salaam.
- Pearson, J. (1998). *Terms in Context*. Amsterdam/Philadelphia: John Benjamins Publishing Company:
- Sinclair J. M. (2006). *Trust the Text: language, Corpus and Discourse*. London and New York: Routledge Taylor & Francis Group.
- Sinclair, J. M. (1991) *Corpus Concordance Collocation*. Oxford: Oxford University Press.
- Sinclair, J. M. 1987. "Grammar in the Dictionary" katika Sinclair (ed.), 104-115.
- Sinclair, J.M. (ed) 1987. *Looking Up: An Account of the COBUILD Project in Lexical Computing*. London: HarperCollins.

- Swahili-Suomi-Swahili sanakirja* (2002). Helsinki: Suomalaisen Kirjallisuuden Seura.
- Tognini-Bonelli, E. (1996) *Corpus Linguistics at Work*. John Benjamins Publishing Company: Amsterdam/Philadelphia.
- TUKI (2001) *Kamusi ya Kiswahili Kingereza*. Dar es Salaam: Taasisi ya Uchunguzi wa Kiswahili, Chuo Kikuu cha Dare es Salaam.