

ISIMU KOMPYUTA KATIKA UTAFITI NA UFUNDISHAJI WA LUGHA NA HALI ILIVYO KWA KISWAHILI

S. S. Sewangi

1.0 Utangulizi

Dhana ya isimu kompyuta ni tofauti na dhana nyingine za kiisimu ambazo hubainisha eneo la udadisi katika taaluma ya isimu, kwa mfano isimu fafanuzi na isimu jamii. Isimu kompyuta haibainishi eneo la udadisi katika taaluma ya isimu bali hubainisha eneo la uingizaji wa teknolojia ya kompyuta katika kusaidia udadisi wa kiisimu. Kwa mujibu wa Souter na Atwell (1993) lengo la wanaisimu kompyuta ni kutumia teknolojia ya kompyuta katika kudadisi data ya kiisimu ili kugundua tabia ya lugha na kuunda nadharia mpya za kiisimu. Kwa upande mwingine, lengo la wataalamu wa kompyuta katika isimu kompyuta ni kutumia data za kiisimu, aghalabu kutoka kwa wanaisimu kompyuta, katika kuunda programu na mbinu za kikompyuta kwa ajili ya kuiwezesha kompyuta kuelewa, kuhifadhi na kuchakata lugha ya binadamu ili kurahisisha utafiti na ufundishaji wa lugha pamoja na matumizi mengine ya lugha ya binadamu.

Mwanzoni mwa miaka ya 1950 uwanja wa isimu kompyuta ulitawaliwa zaidi na juhudi ya kuipatia kompyuta uwezo wa kutafsiri, juhudi hizo ambazo ziliendeshwa huko Marekani ziligharamiwa na Idara ya Ulinzi ya Marekani, na zilihusisha zaidi mabingwa wa kompyuta waliokuwa wamebobeza katika masuala ya kuchakata lugha kwa kutumia kompyuta.

(<http://humanities.uchicago.edu/faculty/goldsmith/CompLingCourse/History.htm>). Hata hivyo juhudi za wataalamu hao hazikuwa na mafanikio ya kuridhisha. Katika miaka ya 1970 hadi ya 1980, wataalamu wa isimu kompyuta walijishughulisha zaidi na uundaji wa programu na mbinu za kuiwezesha kompyuta kutambua lugha ya mazungumzo. Kutengenezwa kwa kompyuta zenye uwezo mkubwa zaidi wa kuhifadhi data ya lugha katika miaka ya 1990 ambako kulitokana na maendeleo makubwa ya teknolojia ya habari¹ kulisababisha mwelekeo mpya katika isimu kompyuta. Mkabala wa isimu kompyuta ulielekezwa zaidi katika kutumia lugha, hasa ya maandishi, iliyohifadhiwa kwenye kompyuta kama chanzo kikuu cha data ya utafiti wa kiisimu na hata ufundishaji wa lugha. Kwa mfano, kwa upande wa lugha ya Kiingereza, maandishi mbalimbali yenye maneno zaidi ya milioni mia moja

¹ Kwa lugha nyepesi, teknolojia ya habari inahusu taaluma ya nadharia na mbinu za ukusanyaji, uhifadhi, uchambuzi na usambazaji wa habari kwa kutumia teknolojia ya kompyuta

yameweza kukusanywa na kuhifadhiwa kwenye kompyuta kwa ajili ya shughuli mbalimbali za utafiti na ufundishaji wa lugha hiyo (taz. <http://www.hcu.ox.ac.uk/BNC/>). Kwa upande wa Kiswahili dhana nzima ya matumizi ya teknolojia ya kompyuta bado, kwa kiasi kikubwa, ni ngeni. Lengo la makala haya ni kutalii hali hiyo ilivyo na kubainisha uwezekano na ugumu wa uingizaji wa teknolojia ya kompyuta katika utafiti na ufundishaji wa Kiswahili.

2.0 Shabaha ya Isimu kompyuta

Kwa mujibu wa Hobbs² shabaha kuu ya isimu kompyuta ni kujenga mazingira ya kuwezesha kutumika kwa teknolojia ya kompyuta katika:

- 1) kuwezesha mawasiliano baina ya binadamu na mashine kwa kutumia mitambo maalumu.
- 2) kurahisisha uchambuzi wa sarufi ya lugha katika viwango vyake mbalimbali kwa kutumia programu maalumu za kompyuta.
- 3) kuhifadhi na kuchopoa taarifa mbalimbali za data ya lugha katika matini zilizohifadhiwa kwenye kompyuta.
- 4) kutafsiri kwa mashine, hasa matini sahili kama vile matangazo ya biashara na vijitabu vya mwongozo wa matumizi, kutoka lugha moja kwenda nyingine
- 5) Kudadisi na kugundua siri ya maarifa ya binadamu ya kutumia lugha.

Juhudi za kujenga mazingira ya kutumia teknolojia ya kompyuta katika utafiti wa kiisimu na wa matumizi mengine ya lugha ya binadamu zimechukua mikondo miwili tofauti: mkondo wa sayansi ya maarifa bandia na mkondo wa kongoo matini.

2.1 Mkondo wa Sayansi ya Maarifa Bandia.

Istilahi ya Maarifa Bandia ilibuniwa mwaka 1956 na mtaalamu John MacCarthy wa Taasisi ya Teknolojia ya Massachusetts ya huko Marekani.³ Maarifa bandia ni taaluma ya mbinu za kuijengea kompyuta mazingira ya kuiwezesha ifanye mambo afanyayo binadamu. Taaluma hii imegawanyika katika matawi mbalimbali, kama vile tawi la isimu kompyuta, tawi la uundaji wa programu za mifumo ya maarifa na ya kufikiri, na tawi la uundaji wa roboti⁴. Taaluma ya maarifa bandia katika tawi la isimu kompyuta, hujishughulisha na mbinu pamoja na programu za kuiwezesha kompyuta ielewe lugha ya binadamu, katika

² (taz. <http://www.lsadc.org/Hobbs.html>)

³ (taz. http://www.webopedia.com/TERM/a/artificial_intelligence.html).

⁴ (tz. <http://www.cs.toronto.edu/DCS/Research/artificial.html>)

maandishi na mazungumzo, na kuitumia katika mawasiliano mbalimbali kama afanyavyo binadamu. Juhudi za namna hiyo hufanywa na wataalamu wa sayansi ya kompyuta ambao pia wameboba katika masuala ya semantiki, lojiki na saikolojia. Wataalamu wa uwanja huu humithilisha ubongo wa binadamu na kompyuta ya kibiolojia wakiamini kwamba ugunduzi wa siri ya maarifa ya lugha ya binadamu utawasaidia katika kutengeneza programu zenye uwezo wa kutumia lugha kama binadamu. Mafanikio ya juhudi hizi yameonekana hasa katika tafsiri zinazofanywa kiotomati na mashine toka lugha moja hadi nyingine. Hata hivyo, bado mkabala huu unakabiliwa na changamoto kubwa ya kuiwezesha kompyuta itambue na kutumia lugha kama binadamu. Hii inatokana na ukweli kwamba maarifa ya lugha ya binadamu huhusisha pia vitu ambavyo viko nje ya lugha yenyewe, kama vile maarifa ya kuona, kusikia, na ya muktadha wa matumizi, vitu ambavyo sio rahisi kuvikamata kwa kutumia programu ya kompyuta. Aidha, pamoja na kwamba analojia ya ubongo na kompyuta ya kibiolojia ina mantiki, uwezo wa ubongo wa kuhifadhi kumbukumbu ni mkubwa sana kiasi kwamba hadi sasa haijawa rahisi kuunda mashine zenye uwezo huo.

2.2 Mkondo wa Kongoo matini

Kongoo matini au kopora ni lugha, ya maandishi au mazungumzo, iliyoteuliwa kwa kutumia vigezo mahususi ikakusanywa na kuhifadhiwa kikompyuta kwa ajili ya matumizi maalum (Pearson, 1998: 42-43). Mkondo wa kutumia kongoo matini katika isimu kompyuta ulianza mwanzoni mwa miaka ya 1960 ambapo kongoo ya kwanza ya kikompyuta iliundwa (Oostdijk *et al.* 1994). Mwanzoni kongoo ilitumika zaidi kama chanzo cha data kwa ajili ya kupima kiwango cha kurudiarudia kwa maneno au urefu wa sentensi katika aina mbalimbali za maandiko. Katika kipindi hicho cha mwanzo ziliundwa programu kwa ajili ya kubainisha orodha ya kiwango cha urudio wa maneno katika kongoo. Aidha ziliundwa programu kwa ajili ya kubainisha mazingira ya matumizi ya maneno na sentensi katika kongoo. Kwa kutumia programu za aina hiyo watafiti waliweza kunasa data mbalimbali walizozihitaji kutoka kwenye kongoo. Katika miaka ya 1970 juhudi zilianzishwa za kuunda programu kwa ajili ya kuingiza alama mbalimbali ili kurahisisha upatikanaji wa data za kiisimu kutoka kwenye kongoo, hasa za kategoria za kisarufi, katika kongoo. Lengo la mkondo wa kongoo matini katika isimu kompyuta ni kutumia data halisi ya lugha katika kuunda kanuni na mbinu zitakazoiwezesha kompyuta kufahamu lugha ya binadamu kama ilivyotumika katika kongoo matini. Uundaji wa programu za aina hii hutumia mbinu za kitakwimu au za kisarufi. Mbinu za kitakwimu huegemea katika nadharia ya uyumkini ambapo data yenye kiwango kikubwa cha kujirudia huchukuliwa kama yenye uelekeo wa usahihi. Kwa upande

mwingine, mbinu za kisarufi hutumia kanuni za kisarufi za lugha zilizofafanuliwa kwa kutumia data ya kongoo. Uwezo mkubwa wa kompyuta wa kuhifadhi matini umeupa uelekeo wa kongoo matini nafasi muhimu katika mkabala wa isimu kompyuta. Data kubwa ya kongoo matini, pamoja na kutumika katika kuunda kanuni za kisarufi kwa ajili ya programu za kompyuta za kuchambua lugha ya binadamu, pia imekuwa ikitumiwa na wataalamu wa maarifa bandia katika kutengeneza programu za kuipatia kompyuta uwezo wa kutambua na kutumia lugha ya binadamu.

2.2.1 Misingi ya mkondo wa kongoo matini

Misingi mkubwa wa mkondo wa kongoo matini ni upatikanaji wa kongoo kubwa katika muundo wa kikompyuta. Kongoo ya kikompyuta huundwa kwa kukusanya lugha ya maandiko au mazungumzo mbalimbali na kisha kuiingiza na kuihifadhi kwenye kompyuta. Ukusanyaji wa lugha ya kongoo hufanywa kwa kuzingatia lengo na matumizi ya kongoo inayokusanywa. Sinclair (1995), kwa kuzingatia kipengele cha lengo na matumizi ya kongoo, ameanisha kongoo za aina mbili: kongoo ya lugha ya kawaida na kongoo ya lugha maalumu. Kongoo ya lugha ya kawaida hukusanywa kwa lengo la kupata data ya lugha yenye kuwakilisha tabia ya jumla ya lugha. Aidha kongoo ya lugha ya kawaida hutumika sana katika uundaji wa kamusi za kawaida za lugha. Kongoo ya lugha maalum hukusanywa kwa lengo la kutumika kama chanzo cha data ya lugha yenye sifa mahususi, kwa mfano, data ya istilahi au ya simo. Mchakato mzima wa kuunda kongoo hutegemea kuwepo kwa zana za kuundia kongoo ya kikompyuta, kama vile vinaso sauti, skana na programu zake, pamoja na kompyuta zenye uwezo mkubwa wa kuchakata na kuhifadhi data ya lugha.

Matumizi ya kongoo katika utafiti wa lugha hutegemea uwezo wa kuchopoa data sahihi kutoka kwenye kongoo. Data huweza kuchopolewa kutoka ama kongoo ghafi au kongoo iliyochambuliwa. Kongoo ghafi ni ile ambayo haijafanyiwa uchambuzi na kuongezwa aina yoyote ya alama. Data ya kongoo ghafi aghalabu huhusu mazingira na takwimu za matumizi ya maneno pamoja katika maandiko. Kongoo iliyochambuliwa kwa kuongezwa alama za kategoria mbalimbali za kiisimu hurahisisha uchopoaji wa data ya lugha kwa vile alama zilizolingizwa kwenye kongoo hutumika kama msingi wa kuchopa kategoria mbalimbali za data. Kazi ya kuchambua kongoo ya kikompyuta hufanywa kiotomati kwa kutumia programu zilizoundwa mahususi kwa kazi niyo. Programu za kuchambua kongoo huundwa kwa kuzingatia aina ya data inayokusudiwa kuchopolewa. Kwa mfano, zipo programu za kuchambua kongoo kwa ajili ya uchopoaji wa kategoria za kisarufi za maneno, kategoria za kimofolojia, na kategoria za kisintaksia (taz. Sewangi, 2001: 21).

3.0 Data ya kongoo kama msingi wa ufundishaji na utafiti wa lugha

Lengo la mkabala wa isimu kompyuta ni kurahisisha ufundishaji na utafiti wa lugha. Maendeleo ya kasi katika teknolojia ya kompyuta yamewezesha kongoo ya kompyuta kuwa msingi wa data ya ufundishaji na utafiti wa lugha.

3.1 Ufundishaji wa lugha

Data ya kufundishia lugha inaweza kuwa ama razini au halisi. Data ya kirazini huhusisha mifano michache ambayo huchukuliwa kama kiwakilishi cha lugha halisi. Kwa upande mwingine, data halisi ya lugha hubainisha lugha kama inavyotumika katika mazingira halisi ya mawasiliano. Kongoo ya kikompyuta ni chanzo kikuu cha data halisi ya lugha. Mifano ya matumizi ya lugha iliyochopelewa kwenye kongoo humpatia mwanafunzi uzoefu wa lugha kama inavyotumika kwenye mawasiliano ya kila siku. Licha ya kuwa chanzo kikuu cha data ya lugha halisi, kongoo inaweza pia kutumika katika kutathmini mbinu zinazotumika kufundishia lugha hasa kwa lengo la kubaini kama mbinu hizo zinaakisi hali halisi ya lugha katika mawasiliano. Kwa mfano, zipo tafiti ambazo zimebaini kwamba ufundishaji wa lugha kwa mbinu ya kutumia mifano michache tu huweza kumpotosha mwanafunzi na kumkwaza katika kujifunza lugha hasa ya pili. Hivyo, kwa mujibu wa tafiti hizo, ni vyema kutayarisha vitabu vya kufundishia lugha kwa kutumia data halisi ya lugha kutoka kwenye kongoo.⁵

Kongoo iliyochambuliwa inaweza pia kutumika katika ufundishaji wa vipengele mbalimbali vya isimu ya lugha. Kwa mfano, Chuo Kikuu cha Lancaster, kimekuwa kikiendesha utafiti kuhusu namna ya kufundisha misingi ya uchambuzi wa kategoria za kisarufi kwa wanafunzi wa shahada ya kwanza kwa kutumia programu ya kuchambua kategoria za kisarufi (McEnery and Wilson, 1993). Programu inayoitwa *Cytor* husoma sentensi iliyoko kwenye kongoo iliyochambuliwa kwa alama za kategoria za kisarufi kisha huficha alama za kisarufi zilizopo na kuwaachia wanafunzi wachambue sentensi wenyewe. Pale wanafunzi wanapokwama, wanaweza kuitumia programu hiyo katika kutambua kategoria mbalimbali za kisarufi zilizoko. Kwa mujibu wa tathmini iliyofanywa na McEnery, Baker and Wilson (1995) kuhusu ubora wa njia hii ya kutumia programu, kwa jumla wanafunzi waliojifunza kwa kutumia programu walifanya vyema kuliko wale waliofundishwa na waalimu wa kawaida.

⁵ (taz. <http://bowland-files.lancs.ac.uk/monkey/ihe/linguistics/corpus4/4teach.htm>)

3.2 Utafiti wa lugha

Kwa upande wa utafiti wa lugha, kongoo matini hutumika kama msingi wa data ya utafiti. Kwa mujibu wa Souter na Atwell (1993:i-iii) matumizi ya kongoo ya kikompyuta katika utafiti hulenga hasa katika maeneo makuu mawili: eneo la kuhakiki na kuunda nadharia mpya za kiisimu na eneo la kuunda programu za kuchambua, kuelewa na kuzalisha lugha kwa kutumia mashine. Katika kuhakiki na kuunda nadharia mpya za kiisimu data ya kongoo hutumika kama ushahidi wa kufikia mahitimisho ya kinadharia. Huu ni mkabala tofauti na ule wa kimapokeo ambao hutumia mifano michache inayodaiwa kuwakilisha umilisi wa mzungumziaji wa lugha kama ushahidi wa mahitimisho ya kinadharia. Aidha, kwa upande wa uundaji wa programu za kuchakata lugha ya binadamu kwa kompyuta, data ya kongoo hutumika katika kuunda kanuni halisi za kiisimu ambazo hugeuzwa kikompyuta na kufanywa programu za kuchambulia lugha. Pamoja na kutumika kama chanzo kikuu cha data ya utafiti wa kiisimu, matumizi ya kongoo katika utafiti usio wa kiisimu yamekuwa yakiongezeka kwa kasi. Kwa mfano, hivi sasa kongoo hutumika kama msingi wa maneno na maana za kimatumizi katika eneo la uundaji wa kamusi za lugha ya kawaida na maalumu. Vilevile, data ya kongoo inaweza kutumika katika kubainisha mitindo ya matumizi ya lugha miongoni mwa waandishi mbalimbali wa kazi za fasihi.

Pamoja na uwezo mkubwa wa kusaidia katika utafiti wa lugha, mkabala wa isimu kompyuta katika utafiti wa lugha huhitaji maandalizi maalumu kabla ya kuanza kutumika. Vifuatavyo ni baadhi ya vipengele vya kuzingatiwa katika kuandaa mazingira kwa ajili ya kufanikisha matumizi ya kompyuta katika utafiti wa lugha:

- i. Uanishaji wa maeneo ya kipaumbele katika utafiti kulingana na tanzu mbalimbali za utafiti wa lugha, kama vile, fasihi, isimu, leksikografia, na teminografia.
- ii. Uwekaji wa malengo mahususi ya utafiti kulingana na maeneo ya kipaumbele katika kila tanzu ya utafiti.
- iii. Uundaji wa kongoo matini na programu husika kulingana na maeneo ya kipaumbele. Pale ambapo tayari programu zipo, basi maandalizi ya mafunzo ya kuzitumia programu hizo yapewe kipaumbele. Vilevile juhudi ziendelee za kuziboresha programu hizo kwa kuziongezea data mpya ya lugha
- iv. Ununuzi wa vifaa vinavyohitajika kwa utekelezaji wa utafiti ulioanishwa
- v. Uandaaji wa nguvu kazi inayosadifu matumizi yaliyolengwa, yaani wataalamu wa isimu kompyuta, wataalamu wa kuunda programu za kuchambua lugha na kutafsiri, na wajuzi wa kutumia kompyuta kulingana na programu husika.

4.0 Hali ilivyo kwa Kiswahili

Hadi sasa, matumizi ya kompyuta katika ufundishaji na utafiti wa lugha ya Kiswahili ama ni madogo sana au hayapo kabisa. Matumizi ya kongoo kama chanzo cha data ya kufundishia na utafiti bado ni mageni. Hivyo, maandalizi ya kadhaa yanahitajika kwa ajili ya kuingiza matumizi ya kompyuta katika kusaidia ufundishaji na utafiti wa Kiswahili. Hapa chini tunabainisha baadhi ya maandalizi muhimu yanayohitajika kwa kila upande.

4.1 Ufundishaji wa Kiswahili

Ufundishaji wa Kiswahili, katika viwango vyote, bado hufanywa kwa kutumia mbinu za kimapokeo ambapo mifano michache tu ya data ya lugha, inayochukuliwa kuwakilisha umilisi wa wazungumzaji wa Kiswahili, hutumika kama ushahidi wa kuthibitisha kanuni za tabia mbalimbali za lugha. Chuo Kikuu cha Dar es Salaam kimeanza kutoa mafunzo ya mbinu za kutumia kompyuta katika kusaidia ufundishaji wa Kiswahili. Kwa hiyo, ili kuanzisha matumizi hayo maandalizi hayana budi kufanywa katika maeneo yafuatayo:

- Uundaji wa kongoo matini kwa ajili ya data za lugha za kufundishia masomo mbalimbali
- Uundaji wa programu za kuchambua kongoo na kuchopoa data za kufundishia
- Uanzishaji wa hifadhi data ya masomo ya tanzu mbalimbali za lugha, kama vile masomo ya fasihi, isimu, leksikografia na teminografia ya Kiswahili
- Uundaji wa tovuti ya wanataaluma na wanafunzi wa masomo ya Kiswahili
- Utoaji mafunzo kwa wanataaluma na wanafunzi wa Kiswahili kuhusu matumizi ya programu za kufundishia kwa kompyuta

4.2 Utafiti wa Kiswahili

Matumizi ya kompyuta kusaidia utafiti wa Kiswahili bado sio makubwa. Hata hivyo kwa upande huu kuna juhudi kadhaa ambazo zimeshafanywa na zinaendelea kufanywa kwa lengo la kufanikisha matumizi ya kompyuta katika kusaidia utafiti wa lugha ya Kiswahili. Baadhi ya mambo yaliyokwishafanyika ni haya yafuatayo:

Kuanzia mwaka 1985, Taasisi ya Masomo ya Asia na Afrika ya Chuo kikuu cha Helsinki, Finland, ilianza kujenga mazingira ya kuwezesha matumizi

ya kompyuta katika kusaidia utafiti wa Kiswahili. Juhudi hizo ambazo zilianza chini ya mradi kabambe wa SALAMA⁶ hadi sasa zimewezesha kupatikana kwa:

- *Kikagua tahajia 2 ya Kiswahili (Orthografix 2 for Swahili)* Hiki ni kikagua tahajia kinachoweza kutumika katika toleo la hivi karibuni kabisa la programu ya *Microsoft Word*, kwa ajili ya Kiswahili sanifu⁷.
- *Kongoo Helsinki ya Kiswahili (Helsinki Corpus of Swahili)* Hii ni kongoo ya Kiswahili iliyochambuliwa yenye maneno kiasi cha takribani milioni 12. Kongoo hii hupatikana bila malipo kwa ajili ya utafiti wa kisayansi kutoka Kituo cha Kukokotoa Kisayansi (*Center for Scientific Computing*).
- *SWATWOL*. Hii ni seti ya programu za kuchambua mofolojia ya Kiswahili sanifu kwa kutumia data ya kongoo. Programu hizi zina uwezo wa kuainisha kategoria za kisarafu na za kimofolojia pamoja na kuondoa utata unaotokana na uchambuzi wa kimofolojia katika data ya kongoo
- *Programu ya Kutafsiri Kimashine Kiswahili kwenda Kiingereza*. Hizi ni programu za kutafsiri kwa kompyuta. Programu hizi zinajumuisha pia zile zilizotajwa hapo juu za *SWATWOL*. Kwa sasa programu hizo zina uwezo wa kutafsiri kiasi cha maneno 30,000 kwa dakika na juhudi bado zinaendelea za kuziboresha zaidi.

Kwa upande wa Chuo Kikuu cha Dar es Salaam, juhudi za kujenga mazingira ya kuwezesha matumizi ya kompyuta katika utafiti wa Kiswahili zimeanza. Taasisi ya Uchunguzi wa Kiswahili ina mradi wa kujenga kongoo ya Kiswahili kwa ajili ya matumizi mbalimbali. Hadi sasa mradi huo tayari umeshakusanya kongoo kutoka katika vitabu na vijitabu takribani 150 vya utaalumu wa afya, kilimo, na magonjwa ya mimea. Mradi huu unashirikiana kwa karibu na ule wa Chuo Kikuu cha Helsinki katika juhudi za kuunda na kutumia programu mbalimbali za kuchambua kongoo na kuchopoa data kwa ajili ya tafiti mbalimbali.

5.0 Matatizo na matarajio

Utumiaji kompyuta kusaidia utafiti na ufundishaji katika Kiswahili ni jambo ambalo lipo katika maandalizi lakini lenye kukabiliwa na matatizo kadha.

⁶ SALAMA ni finyazo iliyotokana na 'Swahili Language Manager' ambayo ni seti ya programu za kompyuta zinazoweza matumizi ya kompyuta katika kusaidia utafiti mbalimbali wa Kiswahili kwa kutumia kongoo matini (taz. <http://www.njas.helsinki.fi/salama/>)

⁷ www.lingsoft.fi

5.1 Matatizo

Juhudi za kuingiza matumizi ya kompyuta katika kusaidia utafiti na ufundishaji wa Kiswahili, hasa katika Afrika Mashariki zinakabiliwa na matatizo yafuatayo:

- Kutofahamika kwa mkabala huu mpya wa utafiti. Dhana ya utumiaji kompyuta katika kurahisisha utafiti na ufundishaji bado ni ngeni sana katika mazingira ya Afrika ya Mashariki. Hali hii imefanya baadhi ya watafiti na wakufunzi kutotilia maanani jambo hili. Hata hivyo tatizo hili ni la muda na litatoweka baada ya matunda ya mkabala huu kuanza kujidhihirisha.
- Ukosefu wa nguvu kazi inayosadifu mazingira ya matumizi ya kompyuta. Mkabala wa utafiti na ufundishaji kwa kompyuta hutumia mbinu zilizo tofauti kidogo na zile za kimapokeo. Kwa hali hiyo, maarifa mapya yanahitajika kabla ya kuanza matumizi hayo ya kompyuta. Ni wazi utoaji wa mafunzo hayo mapya utahitaji pesa na utaalamu ambavyo hatunavyo kwa sasa.
- Ukosefu wa fedha na vifaa ni mambo yanayokwaza matumizi ya kompyuta katika kusaidia kufundisha na kutafiti. Hii ni kwa sababu kunahitajika kiasi kikubwa cha fedha kwa ajili ya mafunzo, vifaa, programu, gharama za kujiunga katika mtandao wa Intaneti, na za kulipia nishati ya umeme. Ni dhahiri kwamba bajeti ndogo ya masuala ya utafiti ni kikwazo kikubwa katika uanzishwaji wa mkabala huu mpya wa utafiti. Kwa upande wa vifaa, mkabala huu wa utafiti huhitaji kompyuta zenye uwezo mkubwa wa kuchakata na kuhifadhi data ya lugha ambayo kwa kawaida ni kubwa. Kompyuta hizo ni tofauti na zile ndogo zilizozeleka kwa watumiaji wengi. Aidha mkabala huu huhitaji vifaa kama vile viyoyozzi ambavyo hufanya mazingira yanayofaa kwa ajili ya utunzaji wa kompyuta na data mbalimbali ili visiharibiwe na joto kali au unyevu mwingi.
- Ukosefu wa nishati ya umeme ya uhakika. Nishati ya umeme ni msingi wa uhai wa kompyuta kama vile maji yalivyo msingi wa uhai wa viumbe hai. Hali ya upatikanaji wa nishati hii bado haijatengemaa vilivyo katika takribani nchi zote za Afrika Mashariki. Hali hii ni kikwazo kikubwa katika matumizi ya kompyuta katika kusaidia utafiti na ufundishaji.

5.2 Matarajio

Pamoja na kwamba uingizaji wa matumizi ya kompyuta katika kusaidia utafiti na ufundishaji utahitaji kiasi kikubwa cha fedha na nguvu kazi, uzoefu unaonyesha kwamba ushirikiano baina ya nchi mbalimbali ndio ujia kuu ya kukabiliana na mahitaji haya makubwa. Kwa mfano, nchi za Ulaya zinashirikiana katika kuunda kongoo matini na programu za kuchambua lugha

zao. Uundaji wa kongoo matini unafanywa chini ya chombo kinachojulikana kama NERC (Network of European Research Corpora). Uundaji wa programu hufanywa katika vituo vya utafiti wa pamoja, kwa mfano kituo cha Xerox Research Centre Europe (XRCE)⁸

Matarajio ni kwamba utawekwa mkakati wa ushirikiano utakaoweza kujenga mazingira ya matumizi ya teknolojia hii mpya katika Afrika Mashariki.

Kwa upande wa gharama kubwa za programu za kompyuta, matarajio ni kwamba msisitizo utawekwa zaidi katika kutumia programu huria za *Linux* ambazo hupatikana bila malipo yoyote. Kwa sasa hivi Chuo Kikuu cha Dar es Salaam kinatekeleza mradi kabambe wa kutafsiri na kuingiza programu hizo katika lugha ya Kiswahili ili ziweze kutumiwa na watu wengi zaidi wa Afrika Mashariki⁹.

6.0 Hitimisho

Makala yametali mkabala wa isimu kompyuta kwa kuzingatia mazingira ya uelekeo wa kongoo matini na mustakabali wake katika kusaidia ufundishaji na utafiti wa lugha ya Kiswahili. Katika kufanya hivyo makala yamebainisha mambo muhimu katika ujenzi wa mazingira ya matumizi ya kompyuta katika utafiti na ufundishaji wa lugha. Aidha makala yamebainisha matatizo na matarajio ya uanzishaji wa matumizi ya kompyuta katika kusaidia ufundishaji na utafiti wa lugha ya Kiswahili katika Afrika ya mashariki.

MAREJEO

- Abrioux, D.A.M. (http://cade.athabasca.ca/vol14.1/7_abrioux.html) "Computer-Assisted Language Learning and Distance Education."
- Banbrook, G. (1996) *Language and Computers*. Edinburgh: Edinburgh University Press.
- Garside, R., Leech, G., and Sampson, G. 1987. (eds), *The computational analysis of English: A corpus-based approach*. London: Longman.
- Garside, R., Leech, G., and McEnry, T. (eds). (1997) *Corpus Annotation*. London & New York: Longman.
- Greenbaum, S., and Yobin, N. (1994) "Tagging the British ICE Corpus". In Oostdijk, N. and Haan, P. (eds), *Corpus-Based Research into Language, In Honour of Jan Aarts*. Amsterdam-Atlanta, GA.
- Hobbs, J. R. (<http://www.lsadc.org/Hobbs.html>), "Computers and Language"
- Hurskainen, A. (1992a) "A two-level computer formalism for the analysis of Bantu Morphology: An application to Swahili?" In *Nordic Journal of African Studies*. 1(1): 87-122.
- (199b) "Computer Archives of Swahili Language and Folklore – What is it?" In *Nordic Journal of African Studies*. 1(1): 123-127.

⁸ (taz. <http://www.xrce.xerox.com/research/mit/fst/>)

⁹ (taz. <http://www.kilinux.org/>)

- (1995) "Manual for the Computer Archives of Swahili" Manuscript.
- Leech, G. (1993) "Corpus Annotation Schemes." In *Literary and Linguistic Computing*, 8(4). Oxford: Oxford University Press. 275-281.
- (1997a) "Introducing corpus annotation". In Garside, R. *et al.* (eds.), 1-18.
- Leech, G., Garside, R., and Bryant, M. (1994) "The large-scale grammatical tagging of text: Experience with the British National Corpus". In Oostdijk, N. and Haan, P. (eds).
- Lynch, L. 1997. "Medical Terminology Management." In Wright, S. E and Budid, G. (eds.) 1997. *Handbook of Terminology Management, Vol.1, Basic Aspects of Terminology Management*. Amsterdam & Philadelphia: Benjamins Publishing Company.
- McEnery, T and Wilson, A. (1996) *Corpus Linguistics*. Edinburgh University Press.
- Sewangi, S. S. (2000) "Tapping the Neglected Resource in Kiswahili Terminology: Automatic Compilation of the Domain-Terms From Corpus." In *Nordic Journal of African Studies*. 9(2): 60-84.
2001. *Computer-Assisted Extraction of Terms in Specific Domains: The Case of Swahili*. Ph.D. Thesis, University of Helsinki.
- Sinclair, J. M. (1991) *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Souter, C. And E. Atwell (eds) (1993). *Corpus-based computational linguistics*. Amsterdam: Ropopi.