

# UP Diksiyunaryong Filipino: Sinonismong Walang Hangganan

Ramon G. Guillermo, Rhandley D. Cajote , at Aristeo Logronio

## ABSTRAK

Ang UP Diksiyunaryong Filipino (UPDF) (2001 at 2010) na pinamatnugutan ni Virgilio S. Almario ay isa sa pinakahuling ambag sa pagbubuo ng monolinggwal na diksyunaryo ng wikang Filipino. Nais ng papel na ito magbigay ng kritikal na perspektiba sa pagkakabuo ng UPDF sa pamamagitan ng kantitatibong analisis ng mga salitang binibigyan nito ng kahulugan at ng istruktura ng pagkakaugnay-ugnay (o kawalan ng pagkakaugnay-ugnay) ng mga entri. Hahantong ang analisis sa isang pangkalahatang pagpuna at obserbasyon sa mga suliranin hinggil sa pagbubuo ng korpus na dapat pinagbabatayan ng UPDF at ng anumang diksyunaryo.

Mga susing salita: Virgilio Almario, synonymy, language corpora, lexicography, self-referentiality

The UP Diksiyunaryong Filipino (UPDF) (2001 at 2010) which was edited by Virgilio S. Almario is one of the latest contributions towards developing a Filipino monolingual dictionary. This paper aims to give a critical perspective on the construction of the UPDF by means of a quantitative analysis of its definitions and the structure of cross-referencing (or lack thereof) of its entries. The analysis concludes with some general critical observations regarding problems of corpus development which should be the basis of the UPDF or any other dictionary.

## Introduksyon

Ang pinamatnugutan ni Virgilio S. Almario na dalawang edisyon (2001 at 2010) ng *UP Diksiyunaryong Filipino* (UPDF) ang huling pinakasubstansyal na ambag sa pagbubuo ng monolinggwal na diksyunaryo sa wikang Filipino. Ayon kay Lee (2010), ang dalawang edisyon lamang ng naunang monolinggwal na *Diksiyunaryo ng Wikang Filipino* (1989 at 1998) na inilathala ng Komisyon ng Wikang Filipino (KWF) ang maitatapat sa UPDF sa saklaw at ambisyon. Nabuo ang UPDF sa pamamagitan ng suporta ng Sentro ng Wikang Filipino (SWF) sa Unibersidad ng Pilipinas (naglalaman ito ng mga paunang salita ang mga dating pangulo ng UP na sina Emil Q. Javier at Francisco Nemenzo). Dahil inilathala ito ng isang kilala at prestihiyosong institusyon tulad ng UP, nagkaroon ng malawakang impluwensya at diseminasyon ang UPDF, hindi lamang sa UP, kundi sa mga paaralan, pablisar, brodkaster at mga opisinang pampamahalaan.

Layunin ng kasalukuyang pag-aaral ang magbigay ng pagsusuri ng UPDF mula sa aspetong kantitatibo, bagay na hindi pa nagagawa hanggang sa kasalukuyan.

## Ano ang Sinonimo?

Maaaring ihambing, bilang simula ng analisis, ang mga depinisyon ng *Oxford English Dictionary* (OED) at ng *UP Diksiyunaryo ng Wikang Filipino* (UPDF, ed. 2010) para sa “sinonimo.” Ayon sa OED, ang depinisyon ng “sinonimo” (*synonym*) ay ang sumusunod:

### **synonym**, n.

1.1 Strictly, a word having the same sense as another (in the same language); but more usually, either or any of two or more words (in the same language) having the same general sense, but possessing each of them meanings which are not shared by the other or others, or having different shades of meaning or implications appropriate to different contexts: e.g. serpent, snake; ship, vessel; compassion, fellow-feeling, sympathy; enormous, excessive, immense; glad, happy, joyful, joyous; to kill, slay, slaughter; to grieve, mourn, lament, sorrow. Const. for, of, †formerly to, with.

Ito naman ang depinisyon na makikita sa UPDF:

**síng·ka·hu·lu·gán** png [sing+kahulu-gan]: salita na may katulad na kahulugan sa ibang salita: equivalent, sinonimó, synonym.

**synonym** (sí·no·ním) png Lgw [Ing]: singkahulugan.

**synonymous** (si·nó·ni·mús) pnr [Ing] 1: may katulad na kahulugan 2: inihalintulad sa isa pa

Agad makikita rito ang matinkad na pagkakaiba ng dalawang depinisyon. Malinaw na para sa OED, ang mga sinonimo ay mga salita na matatagpuan sa loob ng *iisang wika lamang* (bagay na dalawang beses pa nga nitong idiniin). Hindi makikita ang ganitong restriksyon sa depinisyon ng UPDF. Dulot ng ganitong kondisyong itinakda ng OED, hindi sa gayon maituturing na mga sinonimo ang mga salitang: *bread* (Ingles), *Brot* (Aleman), *pain* (Pranses), samantalang sa depinisyon ng UPDF ay maaaring maituring na mga sinonimo ang tatlong nabanggit na salita.

Makikita ang ganitong pag-unawa ng UPDF sa “sinonimo” sa entri para sa “aso” (Fil./Tag.) na naglalaman ng mga “kasingkahulugan” nito sa iba’t ibang wika sa Pilipinas, kasama na ang Espanyol at Ingles:

**ÁSŌ:** [Ilk Mrw Seb Tag]: inaalagaan at karniborong mammal (Canis familiaris) na karaniwang may mahabang nguso, matalas na pang-amoy, at nakatatahol at umaalulong; inaalagaan din para sa pagtatrabaho o sa ilang uri ng isports: ásu, asù, áyam, banagán,

**dog, idô, irô, itu, tsitû, pérrō**

**ásu:** Zoo [Kap] áso

**asù:** Zoo [Mag] áso

**áyam:** Zoo [Akl Bik Tag War] áso Zoo ásong magandang palakihin at alagaan anumang hindi tiyak o halòhalò ang uri hal ásong mongrel [Ilk] larô [Iva] lákad

**banagán:** Zoo [ST] áso

**dog:** Zoo [Ing] áso

**idô:** Zoo [Bik Hil] áso

**irô:** Zoo [Seb] áso

**itu:** Zoo [Iba] áso

**tsitû:** Zoo [Iva] áso

**pérrō:** Zoo [Esp] áso

Tingnan din ang entri para sa “squirrel”:

SQUIRREL[Ing] iskuwírel→  
ískuwírel[Ing] squirrel hayop callosciurus mindanensis na may pino at matigas na balahibo at sarikulay ang katawan alalaksín,ardílya,bisíng,buót,láksoy, squirrel→  
alalaksín[Seb] ískuwírel  
ardílya[Esp] ardilla iskuwírel  
bisíng[Cuy] iskuwírel  
láksoy[Seb] iskuwírel

(Walang depinisyon sa buong UPDF ang “sarikulay.” At kung nangangahulugan ito ng “sari-saring kulay”, tila hindi rin angkop itong deskripsyon ng “squirrel.”)

Alinsunod sa prinsipyo ng “sinonismo” ng UPDF, makikita sa Larawan 1 ang dami ng samu’t saring salitang nanggaling sa mga wikang nakapaloob sa UPDF: Ingles (18,157 o h.k. 26% ng mga entri), Espanyol (10,643 o h.k. 15.6%), Sinaunang Tagalog (9,054 o h.k. 13.3%), Sebuwano (3,889), Ilokano (3,647), Bicolano (2,833), Hiligaynon (2,808), Waray (2,649). Mahigit limampung porsyento ang sinasakop na ng Ingles, Espanyol at Sinaunang Tagalog. Makikita naman sa Larawan 2 na ang mga domain na kinapapalooban ng mga marami sa mga ito ay sa botany (5,427), zoology (4,287), musika (1,240), anatomiya (1,224) atbp. Paano masasabi na ang lahat ng mga salitang ito ay “bumubuo” o “bahagi” ng isang totoong umiiral na wika? May sagot sa tanong na ito ang OED para sa wikang Ingles: Maituturing ang isang salita bilang bahagi ng wikang Ingles, at sa gayon, bilang salita na maaaring magamit bilang sinonimo ng iba pang mga salita sa loob ng wikang iyon kapag may mapapatunayang kahit iisang instansiya ng pagkagamit nito sa korpus ng wikang Ingles. May mga pagkakataon na kahit iisang beses lamang nagamit ang isang salita sa buong kasaysayan ng wikang Ingles ay isinasama na ito sa mga entri ng OED. Ang tawag sa mga ito ay hapax legomena na binibigyang depinisyon sa OED bilang:

**hapax legomenon** Pl. hapax legomena. Also simply hapax.  
A word or form of which only one instance is recorded in a literature or an author.

Sa katunaya’y maraming isinama sa mga entri ng OED na mga salitang iisang beses lamang lumitaw sa kasaysayan ng wikang Ingles sapagkat ginamit ni William Shakespeare, tulad ng *scamel*, *pulsidge*, *temperalitie*, *impeticos* atbp. (Willinsky 1994, 74-75). Dulot ng kaisa-isang pagkakataon ng paggamit ng isang napakasikat na awtor tulad ni Shakespeare ng mga hapax ay posibleng magkaroon ng mga mambabasa na hahanapin ang mga nakakabahalang salitang ito sa mga diksyunaryo. Hindi sila mabibigo ng OED. May posibilidad sa gayon ng pagsangguni o pag-look-up ng mga salitang ito ang mga nagbabasa kay Shakespeare. Ayon sa OED, ang depinisyon ng look-up ay ang sumusunod:

### look-up

2.2 The action of (or a facility for) looking something up in a dictionary, file, etc.; retrieval of information about items in an ordered collection.

Masasabi sa gayon na sa pinakamaluwag na pagkaunawa nito: ang isang salitang x ay maituturing na bahagi ng isang wikang L kapag mayroong kahit iisang instansya lamang ng paglitaw nito sa paggamit sa loob ng buong korpus ng L. Kapag may iisang mapapatunayang pagkakataon na ng paggamit ang isang salita ay magkakaroon ng posibilidad ng look-up para sa salitang iyon sa isang diksyunaryo katulad ng OED o UPDF. Ngunit dapat isaaalang-alang na napaka-espesyal na kaso ito na nag-aaplay lamang sa mga katulad ni Shakespeare at hindi angkop sa mga “ordinaryong” tagapagsalita lamang ng isang wika. Karaniwang daan-daan o libo-libong instansya ang kinakailangang paggamit ng isang salita sa isang korpus upang magkaroon ito ng pagkakataong maipaloob sa isang diksyunaryo.

Sa ganang ito, masasabing pambihira ang mambabasa ng diksyunaryo na binabasa itong parang isang aklat o nobela mula sa simula hanggang katapusan. Madalang din ang mambabasang nakakakuha ng aliw sa kaswal na pagbubuklat-buklat nito, tulad ng anekdota ni Almario sa simula ng kanyang “Paliwanag para sa Ikalawang Edisyon.” Aniya, “Naging libangan ko na ang gayong pagbuklat sa diksyunaryo kapag walang magawa” (v). Taliwas nito, karaniwang ginagamit ang diksyunaryo upang sumangguni ng mga kahulugan ng mga salitang naeengkwentro sa pagbabasa, pakikinig, pag-aaral at iba pang mga kahalintulad na aktibidad. Kapag may entring salita sa isang diksyunaryo ng isang linggwahe (na tunay na umiiral), na walang matukoy na kahit iisang malinaw na instansya lamang ng paglitaw sa buong korpus pangwika ay mawawalan ng probabilidad ng pag-look-up nito at hindi magkakaroon ng pagkakataong magamit ang entri na naglalaman ng salitang iyon. (Bagamat may masalimuot din na ugnayan ang dalas ng paggamit ng mga salita sa isang komunidad pangwika sa reyt ng pag-look-up sa mga ito. Natural lamang na ang pinakamadalas gamiting mga salita ay yaon ding karaniwang mababa ang reyt ng pag-look-up.) Ganito ang mangyayari kung halimbawa’y magkalat ng mga salitang Kapampangan sa isang diksyunaryo ng Bahasa Melayu. Walang Malay o Kapampangan na makikinabang mula dito. Hindi nakikita ang mga ito.

### ***Self-referentiality, Dead-end at Closed-Loop***

Masasabing walang malinaw na nakatala at mapapatunayang mga instansya ng pagkagamit sa mga karaniwang batis ng paggamit ng wikang Filipino ang maraming itinuturing na “sinonimo” mula sa samu’t saring wika ng Pilipinas na nakapaloob sa UPDF. Kung gayon, ano ang probabilidad ng look-up para sa mga ito? Pero bago ang lahat, dapat idiin na naiiba sa ganitong mga kaso ang sumusunod na halimbawa:

### **PINAKBÉT** Ilk

→

**pakbét** Ilk : **PINAKBÉT** ilk putahe ng iba’t ibang gulay na iginisa sa kamatis kaunting tubig at bagoong.

(Ang simbolong “→”ay tumutukoy sa paglipat mula sa isang entri sa diksyunaryo patungo sa iba.)

Malinaw na kahit matutukoy ang “pinakbet” o “pakbet” bilang salitang Ilokano ay masasabi namang marami na itong instansya ng pagkagamit sa wikang karaniwang sinasalita bilang “Filipino.” Dulot nito’y kahit wikang Ilokano ang dalawang entri na tumutukoy dito ay walang malaking suliranin na iniharap sa gumagamit ng diksyunaryo sapagkat, tulad ng marami pang salita galing sa ibang mga mayor na wika ng Pilipinas, maituturing na matagumpay nang nahiram at naging bahagi ng Filipino ang salitang “pinakbet.” (Ngunit dapat pansinin na labis na makitid ang depinisyon ng pinakbet sa UPDF. Maraming mga resipi nito na hindi “ginigisa sa kamatis” ang mga gulay ng pakbet.) Dapat talagang paramihin pa ang mga matagumpay na salitang hiram tulad nito. Ngunit ang paglalagay ba sa isang diksyunaryo ang tamang landas para sa ganitong layunin?

Ipinapakita sa Larawan 3 ang mga bilang, para sa pinamakataas na sampung wika, ng mga salitang iisang beses lamang lumitaw sa buong UPDF bilang head-word ng entri. Sa kabuuang h.k. 67,878 na entri ng UPDF, h.k. 9,198 ang mga ito o 13% ang may ganitong katangian (hindi kasama ang mga salitang “Filipino” (yaong mga salitang walang mga marker ng pinanggalingan sa diksyunaryo, Ingles, Espanyol at Tagalog). Ang pinakamarami ay nasa mga wikang Sinaunang Tagalog (1,691), Sebuwano (1,441), Ilokano (1,396), Bikolano (937), Waray (907). Sa praktikal na aspeto, maituturing ang mga entring ito, na walang ibang tumutukoy na entri, bilang mga potensyal na dead-end o blind-spot. Iisang beses lang lumilitaw ang mga salitang ito sa buong diksyunaryo bilang head-word at walang alinmang ibang pagtukoy o reperensya sa mga ito sa kahit alinmang ibang entri sa loob ng UPDF. Halimbawa, sa Talahanayan 1 ay nakalista ang mga salita sa wikang Ifugao para sa anatomiya at sa wikang Maranaw para sa heograpiya sa UPDF. Iisang beses lamang lumitaw ang mga ito bilang mga head-word sa buong UPDF. Malabong mangyari na maghahanap sa UPDF ang isang karaniwang tagagamit nito ng mga kahulugan para sa mga salitang Ifugao na “aluguug,” “halhip” o “tolto!” at gayundin ng mga salitang Maranaw na “balet,” “polangi,” o “talontonay.” Walang instansyang malinaw ng pagkagamit ng mga salitang ito sa umiiral at ginagamit na wikang Filipino. Maging ang dahilan sa likod ng pagpapaloob ng ilang termino galing mga wikang Europeo (maliban marahil sa pagkakasama ng mga ito sa ilang kilalang diksyunaryo sa wikang Ingles na sinangguni sa pagbubuo ng UPDF) ay tila hindi rin maipapaliwanag. Tulad halimbawa ng pagsama ng pariralang mula sa wikang Pranses na “j’y suis, j’y reste” (narito ako, dito ako titigil) o ng Aleman na “Hausen” (belúga). Ayon nga kay Lee, “there is no clear mechanism on how and why they have chosen these entry-terms to be part of the Filipino vocabulary” (394). Dahil sa kakulangan marahil sa mga espesyalista



sa wika, namumutiktik ang UPDF sa maling mga gabay sa pagbigkas ng mga salitang Europeo. Maging ang “juggernaut” na galing sa wikang Hindi na dumaan sa Ingles ay tinukoy bilang salitang Aleman dulot marahil ng maling pag-ugnay ng pagbigkas dito na “jágernót” sa Alemang “Jäger” (mangangaso) na ang tunay na pagbigkas ay “yay-guh” at wala sa UPDF mismo. Ang “Nazism” o “Nazi” ay hindi rin mga salitang maituturing na nanggaling nang direkta sa Aleman, tulad ng isinasaad sa UPDF, kundi pangunahing galing sa Ingles. Ang pinakaginagamit na termino sa Aleman ay “Nationalsozialismus” o “NS,” kung may “Nazism” man sa wikang Aleman ay nagiging “Nazismus” ito.

Ipinapakita rin sa Larawan 3 ang bilang, para sa pinamakataas na sampung wika, ng mga salitang dalawang beses lamang ang paglitaw sa buong UPDF. Sa unang pagkakataon bilang head-word ng entri, at sa pangalawang pagkakataon bilang sinonimo o sangkap ng depinisyong sa ibang entri (hindi kasama ang mga wikang banyaga, Filipino at Tagalog) (h.k. 1,060 na salita o 1.56% ng mga entri ang apektado). (Tingnan din ang Larawan 4.) Maaaring palasak sa Filipino ang salitang tumuturo sa mga salitang ito tulad ng kaso ng entri para sa salitang “aso” sa itaas na tumutukoy sa entring “tsitû.” Kung wala ang pagtukoy ng entring “aso” sa salitang Ivatan na tinuturing sa UPDF na “sinonimo” nito ay halos makakatiyak na walang magaganap na look-up para sa “tsitû” sa buong pag-iral nito sa diksyunaryo at magiging dead-end na lamang ito. Nagkakaroon sa gayon, kahit papaano, ng posibilidad ng look-up kung saan ang instansiya ng pagkagamit ng naturang salita ay ang mismong korpus na kinakatawan ng UPDF. Sa ganitong pangyayari, ang UPDF na mismo ang tumitindig bilang kaisa-isang akda sa “wikang Filipino” na nagsisilbing instansiya ng paglitaw ng naturang salita. Sa gayo’y nagkakaroon ang UPDF ng katangiang self-referential. Isang malinaw na halimbawa nito ang entri sa UPDF para sa salitang Ilokanong “ariwánas”:

**ARIWÁNAS** Ilk →

balátas Hil : pangkat ng liwanag na tumatawid sa kalawakan nakikita sa mga gabing walang buwan at unang natuklasan ni galileo ARIWÁNAS milky way silid

**MILKY WAY** Ing balátas

**SÍLID** War milky way

Ilokano [ariwánas]	→ Hiligaynon [balátas]	→ Ilokano [ariwánas]
	→ Waray [silid]	→ Ingles [Milky Way]
	→ Ingles [Milky Way]	→ Hiligaynon [balátas] → Ilokano [ariwánas]

Ang entri para sa “ariwánas” ay tumuturo sa isang salitang Hiligaynon na “balátas” na tumutukoy naman sa isa pang salitang Waray na “silid” at sa “sinonimong” Ingles ng mga ito na “Milky Way.” Kapag nanatili ang pagtutukuyan sa pagitan ng mga katagang Hiligaynon, Waray at Ilokano bilang nakasaradong sistema, napakahirap talagang maarok, para sa karaniwang tagapagsalita ng Filipino, kung ano ang binibigyang kahulugan. Sa pamamagitan lamang ng entri hinggil sa “Milky Way” makababagtas ang sumasangguni mula “balátas” patungo sa “ariwánas.” Hindi napapadali ng ganitong paraan ang gawain

ng pag-unawa hinggil sa kahulugan ng “Milky Way” para sa nagbabasa ng diksyunaryo. Ngunit dahil sa natatanging depinisyong ng UPDF ng “sinonimo” ay maaaring maipaloob ang lahat ng mga ito sa naturang diksyunaryo. (Gayumpama’y ang pariralang “pangkat ng liwanag” sa depinisyong ng “Milky Way” ay isang malinaw na maling pagkakasalin ng “band of light” mula marahil sa isang depinisyong Ingles na naturang termino.)

Maaari sa gayon tingnan ang isang posibleng asersyon sa leksikograpiya: Ang ispesikong korpus pangwika na sinasaklaw ng isang diksyunaryo (na maaaring sumaklaw sa lahat ng iba pang diksyunaryo hinggil sa parehong wika na nagawa na) ay hindi maaaring sumaklaw sa nilalaman ng diksyunaryo ding ito mismo (tingnan ang Larawan 5). Kung totoo ang kabaligtaran, maaaring maglaman ang isang diksyunaryo ng wikang “Filipino” ng entri tulad ng “++...<-+>” (lehitimong parirala ito galing sa wikang kompyuter na Brainfuck) na mabibigyang batayan ng instansiya ng pag-iral nito sa mismong diksyunaryo ding ito. Kung tutuusin ay pwedeng puro mga inimbentong salita ang nilalaman ng isang diksyunaryong “Filipino” (tulad ng mga talasalitaan ng “Maugnaying Filipino” (Guillermo 2002)) kung hindi nga susunod sa ganitong prinsipyo. Para sa UPDF, ang sarili na nito ang pangunahin nitong testigo hinggil sa pagkagamit ng mga salita nitong nilalaman sa wikang Filipino.

Ang problema ng self-referentiality ng UPDF ay lalong tumitingkad sa mga kaso kung saan ang mga salitang walang malinaw na nakatalang instansiya ng pagkagamit sa wikang Filipino ay tinutukoy ng isa pang salita sa ibang entri na wala ring malinaw na nakatalang instansiya ng pagkagamit sa wikang Filipino. Masasabing bumubuo ang dalawang salitang ito ng isang ganap na closed-loop. Makikita sa Larawan 6 ang mga wikang-magkakapares kung saan pinakamadalas naganap ang ganitong penomeno sa UPDF. Mga halimbawa:

**BALOKIKÍ**① ST : pako na nabaluktot ngunit hindi gaanong balokí→  
**balokí** ST : nabaluktot na pako **BALOKIKÍ** ②

Sinaunang Tagalog [balokikí] ↔ Sinaunang Tagalog [balokí]

**DAYÉNA**①Kan : katutubong awit para sa mga yumao var dayeng **eya éy**→  
**eya éy** Kap **DAYÉNA**②

Kankanaey [dayéna] ↔ Kapampangan [eya éy]

**PULALAKNÍT**①War →  
**Kabág** ST : **PULALAKNÍT**②pagbásag ng básong walang lamán o ang tunog ng nabásag

Waray [pulalaknít] ↔ Sinaunang Tagalog [kabág]

**PUSÔ**① Seb→

**bugnóy** Hil : bigas na isinaing sa loob ng hugis pusong dahon ng niyog **PUSÔ**② pritong itlog na pinagúlong sa arinang kulay kahel at isinasawsaw sa suka

Sebuwano [puso] ↔ Hiligaynon [bugnoy]

**TÁNGBAW**① Ilk →

**pandéng** Ilk : patayông kahoy na kinakabitan ng **TÁNGBAW**② at nakakabit sa angkod ng araro

Ilokano [tangbaw] ↔ Ilokano [pandeng]

Mga halimbawa ng matatawag na mga 2-node closed loop ang mga entri sa itaas. Tumutukoy ang Ilokano “tángbaw” sa katagang “pandéng” na Ilokano rin at vice versa. May magaganap kayang look-up para sa alinman sa dalawang salitang ito sa karaniwang gamit sa wikang Filipino? Ganito rin ang kaso ng ikalawang halimbawa ng Sebuwanong “pusô” na tumuturo sa Hiligaynon “bugnóy.” Ang Kapampangang “eya éy” ay tumuturo sa Kankanaey na “dayéna.” Ang head-word na “dayéna” ay tumuturo naman sa “eya éy.” Halos walang posibilidad ng look-up sa mga katagang nabanggit dahil walang salitang “Filipino” (sa karaniwang pakahulugan) na magbabagtas ng daan patungo sa mga ito. Sarado at hindi maaabot ng mga karaniwang tagasangguni ng UPDF ang mga salitang ito. (Maliban na lamang sa mga nagbabasa ng diksyunaryo nang dere-deretso tulad ng nobela o nagbubuklat-buklat dito kapag walang magawa.)

Mula sa mga 2-node closed loop sa itaas, maaaring makapagpalitaw ng 3-node closed loop. Halimbawa:

**MÁDMAD**① Igo : ritwal na pumupuri kay kabunyan at sa mga kaluluwa ng mga ninuno sa pamamagitan ng pag-aalay ng tapuy **ídn̄gaw ipútik**→

**ídn̄gaw** Kan **MÁDMAD**②

**ipútik** Kan **MÁDMAD**③

Mayroon ding mga 4-node closed loop. Halimbawa:

**BUDÓNG**① Kal : kasunduang pangkapayapaan na pinagtibay ng dalawang ili **pedin petyen**

**kalon**

**kalón** Tng **BUDÓNG**②

**pédin** Bon **BUDÓNG**③

**pétyen** Bon **BUDÓNG**④

Makikita naman sa halimbawa ng Kalingang “flandág” ang isang masalimuot na closed loop ng mga termino hinggil sa musika sa mga wikang Kalinga, Ifugao at Samal. Walang nakakalabas at nakakapasok sa saradong pag-uugnayan na ito ng mga termino. Walang posibilidad dito ng look-up para sa ordinaryong mambabasa.

**FLANDÁG** ①Kal →

**féyu** ① Kal : pipa na yari sa bukawe o tambo **feólenabéd FLANDÁG** ② **hupép kalulut kulatét oppípi sarúnay sawnáy**

**feyólenabéd** Kal **féyu** ②

**hupép** lfu **féyu** ③

**kulatét** Kal **féyu** ④

**oppípi** lfu **féyu** ⑤

**sawnáy** Sml **féyu**⑥

Kalinga [flandág]	→Kalinga [féyu]	→Kalinga [flandág]
→Kalinga [feyólenabéd]	→Kalinga [féyu]	
→Ifugao [hupép]	→Kalinga [féyu]	
→Ifugao [oppipi]	→Kalinga [féyu]	
→Samal [sawnáy]	→Kalinga [féyu]	

Nawawalang entri [sarúnay ]

Nawawalang entri [kalulut ]

Tulad ng makikita sa mga Larawan 3 at 4, hindi masasabing mga nakabukod at pambihirang halimbawa lamang ang mga ipinakita ditong mga dead-end at 2-node, 3-node at 4-node closed loop at mala-closed-loop na dulot ng kakaibang pagkaunawa sa UPDF ng pagka-“sinonimo” ng mga salita. Sintomas lamang ang mga ito ng mas malawak na problema ng UPDF sa pagbubuo ng korpus nito. Maaari pang paramihin nang paramihin ang mga halimbawa.

Bagama’t maituturing lamang bilang halimbawa ng mala-closed-loop, magandang panghuling halimbawa ng problema ng paggamit ng UPDF ng mga “sinonimo” ang entri nito para sa salitang “kantót.” Ang salita sa Ingles na kahalintulad nito na “fuck” ay nahirapang maipaloob sa OED dulot ng namamayaning moralistikong pananaw pangwika, kahit napakapasok ng gamit nito kung ikukumpara sa mga hapax legomena ni Shakespeare (Willinsky 138), 1972 na lamang ito unang naipaloob bilang entri sa OED. Magugulat ang mambabasang Pilipino sa kahulugan ng “kantót “ sa ikalawang edisyon ng UPDF:

**KANTÓT** ST salitang pangutya sa sinumang lumalakad nang naghahambog ST **kansóg** Kol **karát**

**kansóg** tunog ng tubig hábang hinahalò o inaalog ang banga o bote Cf bulbók kambóg **kantót**

**karát** pagkarát pakikipag karát [Kap Tag] ganap na ugnayang seksuwal ng dalawang tao Cf **fuck**  
hindót coitus iyót fuck hindútan kadô kanâ

**fuck** (fak) lng **karatín** o **makarat**

ST [kantót]	→ [kansóg]	→ ST [kantót]
	→ Kapampangan, Tagalog [karát]	→ Ingles [fuck]
		→ Ingles [fuck] → [karát] → Ingles [fuck]

Kapansin-pansin na ang unang ginamit na kahulugan nito ay sa Sinaunang Tagalog (ST) na galing pa sa entri sa *Vocabulario de la Lengua Tagala* (1860) nina Noceda at Sanlucar na “cantot” (may anyo pang ibinigay na “cacantotcantot”) na nangangahulugang “palabra de burla, contonearse” (mapangutyang salita, pagmamayabang sa paglakad). Mula rito ay itinuturo naman ng entri ang sumasanguni sa dalawa pang salita. Ang una ay ang salitang “kansóg” na may kinalaman sa tunog ng tubig na inaalog o tinitimpla. Ang seryeng “kantót” > “kansóg” > “kantót” ay may anyo bilang isang mala-*closed loop* na umiikot sa pakahulugan ng “kantót” bilang paghahambog habang naglalakad at paghahalo ng tubig. Halimbawang pangungusap ng “Filipino” ng UPDF: “Si Ver ay pakantot-kantot na naglalakad sa talontonay habang lumilikha ng mala-kantot na tunog ang polangi.” Talagang *it’s more fun in Filipinas*. Para makalabas sa loop na ito ay may isa pang salitang itinuro na “karát” (na ayon sa UPDF ay Kapampangan at Tagalog) na tumutukoy na sa wakas sa pinakapalask na depinisyong sekswal ng salitang “kantót” (ang “karát” ang simpleng isang salitang kahulugan na ibinigay sa unang edisyon ng UPDF). Hindi tumuturo pabalik sa “kantót” ang entring “karát,” sa halip ay bumubuo ito ng mala-closed loop sa salitang Ingles na “fuck”: “fuck” > “karát” > “fuck.” Halos hindi na makakarating ang mambabasa sa “kantót” mula sa “fuck” na siyang dapat sanang mangyari sa isang mas lohikal na pagkakabalangkas ng diksyunaryo. Hindi na kailangang idiin na mapanganib para sa isang diksyunaryo ang ganitong paglalaro sa mga “sinonimo” sapagkat maaaring magbunga ng maling okasyon ng paggamit ng sumasanguni ng ganitong sensitibong salita. Hindi malinaw ang patakaran ng UPDF sa pagkakasunud-sunod ng kahulugan. Kronolohiya ba o popularidad ng gamit ang batayan (Lee 391)?

## Konklusyon

Ang *raison d’être* ng bawat isang salita na piniling isama sa isang diksyunaryo bilang isang *information retrieval system* ay ang humigit-kumulang na mataas na posibilidad ng pagsanguni o pag-look-up dito ng isang nangangailangan ng impormasyon hinggil sa paggamit ng wika sa isang tiyak at partikular na pangwikang komunidad. Batay sa tinalakay na mga halimbawa, tila may mga kakulangan ang UPDF sa pagtugon sa mga pangangailangan ng dapat nitong pinagsisilbihang mga mambabasa at tagagamit. Lehitimo ang mga kwestyon hinggil sa *usability* o pagka-magagamit ng akdang ito.

Kung hindi maaaring tumayo bilang isang diksyunaryo para sa pangkalahatang gamit ang UPDF dahil sa problema ng pagkakabuo ng korpus ng mga nilalaman nitong entri, tila hindi rin ito maaaring tumayo bilang mas espesyalisadong batis ng kasaysayan ng wikang Filipino dahil sa hindi sistematikong pagkakaayos ng mga batis at reperensya, bagay na ibang-iba sa mga masinsinang aparatong nilalaman ng OED at ng *Deutsches Wörterbuch* ng magkapatid na Grimm. Kung hangad naman nitong maging kumparatibong sanggunian ng mga wika sa Pilipinas ay masasabing nagkukulang naman ito sa pagkakumprehensibo (Yap 1977).

Hindi nakasalalay ang UPDF sa prinsipyo ng pagtitiyak ng mga instansya ng paggamit ng mga salita nitong nilalaman sa isang malinaw at masasangguniang sistema ng mga korpus. Dulot nito, maliban pa sa ibang masasalimuot nitong problema na kabilang ang *self-referentiality* ng korpus nito, ay nagpapalitaw ito ng maramihang kaso ng mga entri na *dead-end*, *2-node*, *3-node* at *4-node closed-loop* at *mala-closed loop*. Labis nitong sinasagasaan ang anumang ideya ng anumang kabuuang “pangwika” dulot ng napakaluwag nitong pag-unawa sa “sinonimo” na sumasaklaw sa lahat ng mga salitang magkasing- o magkahawig ang kahulugan na walang kinikilalang tunay na umiiral na mga hangganang pangwika.

Hindi talaga maituturing ang UPDF, sang-ayon sa ambisyon nito, bilang monolingwal na diksyunaryo. Sa halip ay lumilitaw na diksyunaryo itong pira-piraso ng napakaraming wika. (Ang *Europa Wörterbuch* (1997) halimbawa, na naglalaman ng mga “magkasingkahulugan” na mga salita sa Aleman, Pranses, Espanyol at Italyano ay binansagang “apat na diksyunaryo” sa iisang bolyum.) Gayumpama’y hindi madaling malaman ang antas at hangganan ng pagkakabuo ng subset ng bawat wika sa loob ng kabuuan nito. Hindi rin malinaw ang mga hangganan ng korpus ng “pangwikang” penomeno na tinutukoy ng UPDF bilang “wikang Filipino.” Ang korpus ba nito ay ang lahat ng mga wikang sinasalita sa Pilipinas (kasama na ang Ingles at Espanyol)? Sino ang nakakapagsalita ng artipisyal na superwika na ito? Kung ang karamihan ng mga salita sa UPDF ay walang gumamit/gumagamit sa kasalukuyang populasyon na ipinapalagay na nagsasalita ng Filipino, ang layunin ba nito ay hindi maging sanggunian ng ginagamit na wika kundi upang magmungkahi at magpalaganap ng mga salitang dapat gamitin? Kung isa itong normatibong diksyunaryo ng isang wika na hindi pa umiiral at hindi pa sinasalita ng mga tao, sinu-sino ang pumili ng mga salita at gumagawa ng mga norm para sa nais pairaling wikang Filipino na ito? Ayon lamang ba ito sa panlasa ng isang indibidwal o isang pangkat ng tao? Kulang sa pagkalinaw at pagkahayag ang buong proseso ng pagkakabuo ng UPDF. Sa pangkalahatan, maituturing itong kahinaan ng isang diksyunaryong hindi nakabatay sa isang sistematiko, siyentipiko, at malinaw na proseso ng pagbubuo ng korpus.

Dapat salungguhitin nang buong diin na hindi dito tinutulan ang pagpapaloob ng paparami nang paparaming salitang-hiram mula sa iba’t ibang wika ng Pilipinas sa wikang Filipino. Kailangan talagang itaguyod ang pagpapayaman at pagpapaunlad ng wikang Pambansa batay sa mga umiiral na wika sa Pilipinas at iba pang mga wika tulad ng nakasaad sa



Artikulo 14, Seksyon 6 ng Saligang Batas ng Pilipinas. Gayumpaman, ang nakikitang mas wastong paraan ng paglulunsad ng ganitong proyekto ay mula sa laganap, masigla at buhay na paggamit ng wikang pambansa sa pagsusulat, pagsasalita, pagbabalita, pakikipagtalastasan ng mga mamamayan mula sa iba't ibang dako ng kapuluan, mula sa iba't ibang sektor at pangkat ng lipunang Pilipino. Hindi ito magagawa, at hindi dapat gawin, sa pamamagitan ng makapangyarihang pag-aangkin, pagdadampot, pagpipili, pagsasala at, maaari pa ngang mabansagang mapagsamantala, insensitibo at oportunistang pandarambong ng isang makapangyarihang "sentro" sa iba't ibang kultura at wika ng Pilipinas. Tutubo ito, kung tutubo man, sa buhay at demokratikong paggamit ng mamamayan sa wikang Filipino.

Dulot ng mga kahinaang ito, kinakaharap na ngayon ng mga leksikograpong Pilipino ang hamon ng pagbubuo ng mga monolingwal na diksyunaryo sa wikang Filipino na nakabatay sa korpus na sinadya at maingat ang pagkadisenyo. Nangangahulugan lamang ito na malinaw na dapat naitatakda ang mga batis tulad ng iba't ibang uri ng nakasulat at ibinigkas na mga teksto sa iba't ibang mga genre at register. Malinaw din dapat ang panahon na sinasaklaw ng korpus. Panghuli, mas mainam na pakinabangan na ang *digitization at automatic processing* upang lalong mapagaan, mapabilis at gawing mas tapat sa korpus ang pagbubuo ng mga bagong diksyunaryo.

*(Nagpapasalamat ang mga may-akda sa suportang pampinansyal para sa pananaliksik na ito ng Office of the Vice Chancellor for Research and Development (OVCRD) ng University of the Philippines (Diliman).)*

### Mga Sanggunian

Almario, Virgilio (punong editor). 2001. *UP Diksiyunaryong Filipino*. Pasig City: Sentro ng Wikang Filipino & Anvil Publishing Inc.

Almario, Virgilio S. (punong patnugot). 2010. *UP Diksiyunaryong Filipino*. Binagong edisyon. Diliman: UP Sentro ng Wikang Filipino.

*Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm auf CD-ROM und im Internet: <http://dwb.uni-trier.de/de/>*

*Europa Wörterbuch. Vier Wörterbücher in einem Band.* 1997. Geneve: LESA Printing & Production.

Guillermo, Ramon. 2002. Mga Suliranin sa Proyekto ng Maignaying Pilipino. Daluyan, Vol XI, No. 1. [MGA PAHINA?]

Lee, Aldrin P. 2010. The Filipino Monolingual Dictionaries and the Development of Filipino Lexicography. *Philippine Social Sciences Review*. Volume 62 No.2 July-December.

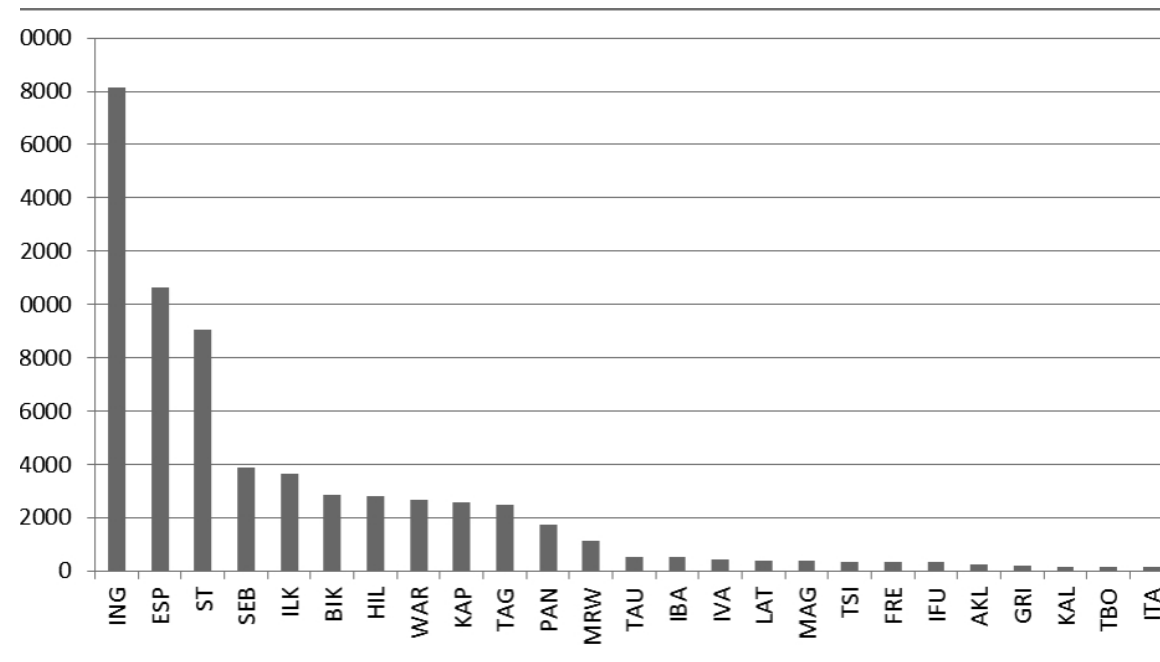
*Oxford English Dictionary Online.* <http://www.oed.com/>

Panganiban, Jose Villa. 1972. *Diksiyunaryo-Tesaurus Pilipino-Ingles*. Lungsod Quezon: Manlapaz Publishing Co.

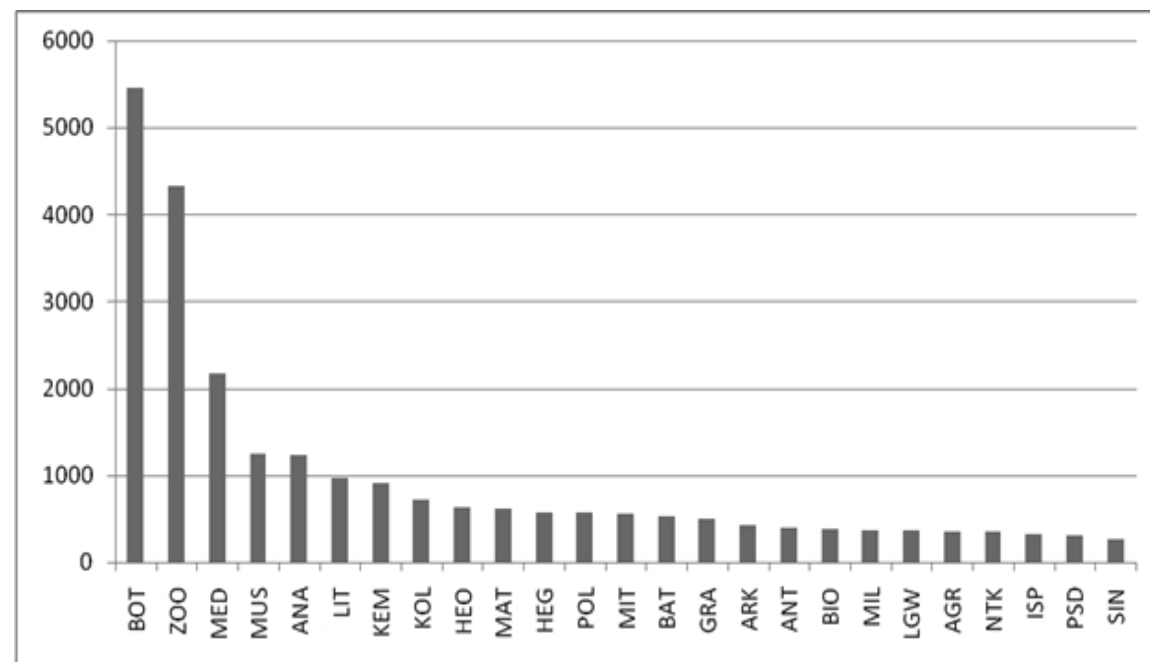
Willinsky, John. 1994. *Empire of Words: The Reign of the OED*. Princeton: Princeton University Press.

Yap, Fe Aldave. 1977. *A comparative study of Philippine lexicon*. Manila : Institute of National Language, Department of Education and Culture.

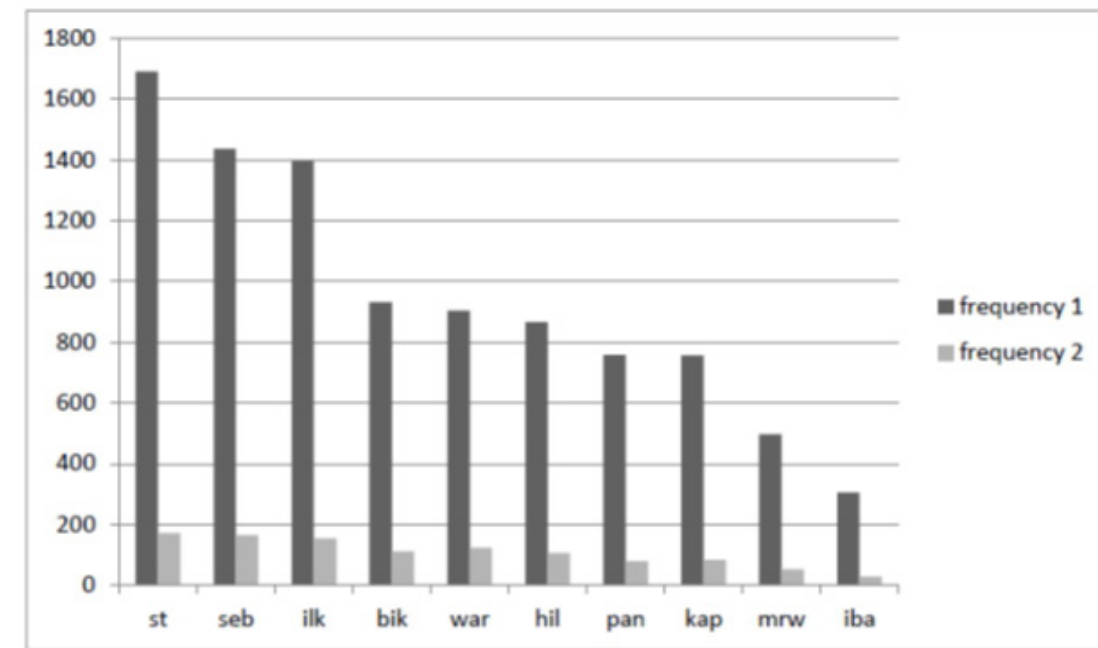
Larawan 1: Pinakamataas na 25 na wikang pinanggalingan ng mga entri ng UPDF (93 kabuuan)



Larawan 2: Pinakamataas na 25 na domain ng mga entri ng UPDF (51 kabuuan)



Larawan 3: Pinakamataas na sampung wika na may mga head-word entri na iisang banggit at/o dalawang banggit lamang sa buong UPDF (hindi kasama ang mga wikang banyaga, Filipino at Tagalog)

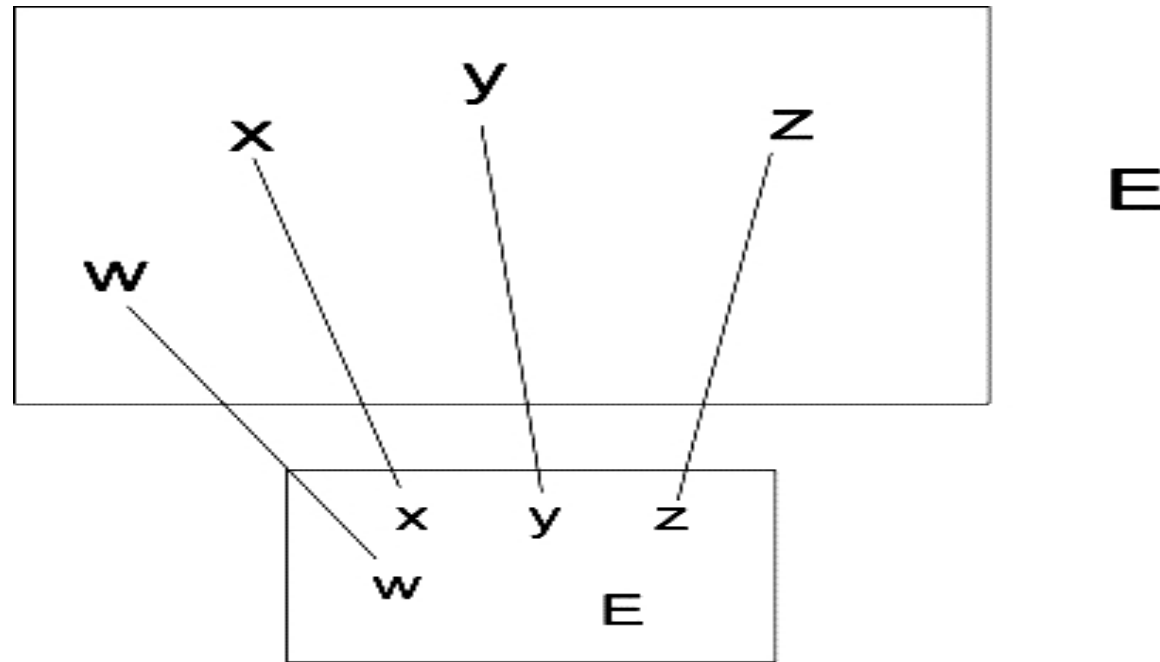


Larawan 4: Ang Y axis ay nagpapakita ng dami ng mga salitang head-word entri na iisang banggit at/o dalawang banggit lamang (hindi kasama ang mga wikang banyaga, Filipino at Tagalog), mula sa unang entring head-word sa X-axis hanggang sa dulong head-word ng UPDF. Gumamit ng computer program na may scan window na naka-set sa habang 3000 na salita. Umaabante ang scan window sa listahan ng mga head-word ng tigilimang salita. Pinakamababa ang graph sa bandang letrang D (y=283) at pinakamataas naman sa letrang I (y=1085).

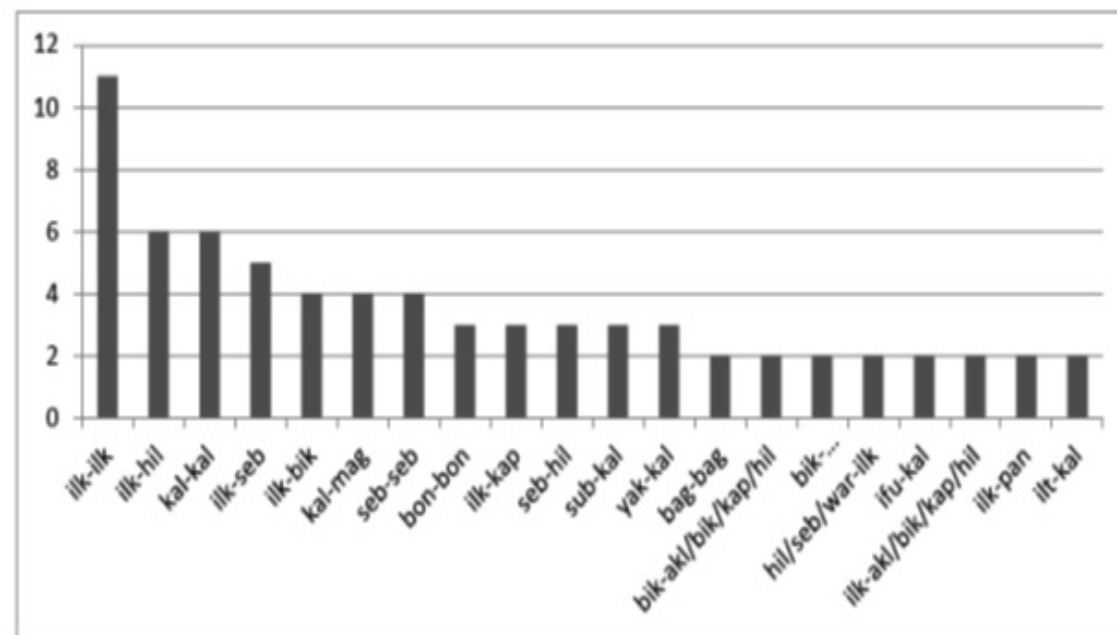




Larawan 5: Ang pagkakaugnay ng mga elementong nakapaloob sa pangwikang korpus (malaking parihaba) sa mga elementong nakapaloob sa diksyunaryo (maliit na parihaba). Ang “E” ay bahagi ng diksyunaryo pero hindi kasama sa mismong pangwikang korpus.



Larawan 6: Pinakamataas na 20 na mga pares ng wika ng mga head-word entri na dalawang banggit lamang sa buong UPDF (hindi kasama ang mga wikang banyaga, Filipino, Tagalog at Sinaunang Tagalog (ST))



Ifugao (anatomya)	Maranaw (heograpiya)
<b>ammáma</b> ana ifu hinlalakí o ang malaking daliri sa paa	<b>balíid</b> heo mrw libís
<b>apéng</b> ana ifu gilid ng mukha	<b>bowálan</b> heo mrw bálóng
<b>duntúg</b> ana ifu lulód	<b>bálet</b> heo mrw pulô
<b>húi</b> ana ifu paa at binti	<b>línik</b> heo mrw loók
<b>paypayáya</b> ana ifu singit	<b>lúwik</b> heo mrw loók
<b>putú</b> ana ifu sikmura	<b>polangí</b> heo mrw mag ilog
<b>tangód</b> ana ifu likod ng ulo	<b>pámpang</b> heo mrw talabís
<b>téngeng</b> ana ifu kasukasuan ng panga	<b>pángba</b> heo mrw pampáng
<b>alugúug</b> ana ifu leég	<b>péted</b> heo mrw buhangin
<b>amdodól</b> ana ifu hinlalato	<b>talontónay</b> heo mrw bangín
<b>gúngu</b> ana ifu túhod	<b>tádeng</b> heo mrw talabís
<b>haliphíp</b> ana ifu gulugód	<b>ílay</b> heo mrw loók
<b>híngul</b> ana ifu siko	
<b>lúlug</b> ana ifu tuhod	
<b>páupául</b> ana ifu balágat	
<b>píngal</b> píngas ana ifu kasukasuan ng balakang	
<b>tagláng</b> ana ifu pan tadyáng	
<b>talígtig</b> ana ifu gulugód	
<b>toltól</b> ana ifu bátok	
<b>tudú</b> ana ifu hintuturo	
<b>tábu</b> ana ifu tiyán pusón	
<b>áwa</b> ana ifu gilid ng baywang	

**Talahanayan 1:** Ang mga salita sa wikang Ifugao para sa anatomya at sa wikang Maranaw para sa heograpiya sa UPDF (mga dead-end na entri)

\*\*\*

Si **Ramon Guillermo** ay Assoc. Professor sa Depto. ng Filipino at Panitikan ng Pilipinas, Kolehiyo ng Arte at Literatura, UP Diliman. Naglulunsad siya ng mga proyekto sa digital philology at digital humanities.

Si **Rhandley Cajote** ay Assoc. Professor at tagpangulo ng Digital Signal Processing (DSP) laboratory ng Electrical and Electronics Engineering Institute sa UP Diliman. Ang Laboratoryo ng DSP ay isa sa mga institusyon na nagsasagawa ng pananaliksik sa speech recognition at synthesis ng wikang Filipino, natural language processing, koleksyon ng speech at text database collection para sa mga wika sa Pilipinas at iba pang aplikasyon ng signal processing.

Si **Aristeo Logronio** ay isang gradwadong mag-aaral sa Digital Signal Processing laboratory ng Electrical and Electronics Engineering Institute sa UP Diliman.