

ARTICLE

**Pagsusuri sa ortograpiya ng kambal-katinig
sa Filipino batay sa korpus:
Tuon sa reduplikasyon ng mga
hiram na salita at sa mga anyong may
<s(i)yon> / <s(i)ya>
(A corpus-based analysis of consonant clusters
in Filipino orthography:
On reduplication in borrowed terms and
on forms with <s(i)yon> / <s(i)ya>)**

Jem R. Javier

ABSTRAK

Kaakibat ng binubuonng *UP Monolingguwal na Diksiyonaryong Filipino (UP MDF)* ang pangangalap ng korpus ng wikang Filipino kapwa sa anyong pasalita at pasulat. Ang naturang korpus ang magsisilbing batayan sa pagpili ng mga salitang bubuo sa leksikograpikong akda, maging sa paglalarawan ng gramatika ng kontemporaryong gamit ng wikang pambansa. Pagtutuunan ng pansin ang pamamaraang ortograpiko ng mga pasulat na tekstong nakapaloob sa pasulat na korpus ng Filipino, partikular ang pagkakabit ng reduplicative affix C1(C2)V1- at ang pagbaybay ng mga salitang may <s(i)yon> at <s(i)ya>. Ihahambing ito sa mga pamantayang ipinatutupad ng Komisyon sa Wikang Filipino (KWF), na siyang pangunahing nagtatakda sa magiging anyo ng mga hiniram na salita patungong Filipino. Batay sa datos na nakalap, lumilitaw na bagaman may tiyak na tuntuning ipinatutupad ang KWF hinggil sa usapin ng pagbabaybay ng mga hiram na salita, makikita pa rin ang pagtaliwas dito kahit sa mga tekstong nailimbag na pambansa ang lawak ng sirkulasyon. Hindi maaaring isantabi ang mga kataliwasang ito sapagkat nilalayan ng *UP MDF* na ilarawan ang Filipino ayon sa mismong paggamit ng mga tagapagsalita nito sa kontemporaryong panahon, kung kaya, ang mga anyong may iba't ibang paraan ng pagbaybay ay itinuturing na mga *variant*. Ang mga variant na may mataas na bilang ang gagamitin sa diksiyonaryo. Magkagayunpaman, kinakailangan pa ring isaalang-alang ang mga tuntunin

ng KWF sa layuning makabuo ng batayan ang ortograpiyang Filipino, kung kaya inaasahan na ang diskurso hinggil sa paksang ito ay mananatili at magpapatuloy.

MGA SUSING-SALITA

leksikograpiyang Filipino, korpus ng wikang Filipino, ortograpiya,
hiram na salita, linggwistiks batay sa corpus

ABSTRACT

The *UP Monolingwal na Diksiyonaryong Filipino* [*UP Monolingual Filipino Dictionary*] (*UP MDF*) is currently compiling a Filipino language corpus in written and spoken forms. This corpus will be the basis of words that will be chosen as part of the lexicographic work, as well as the description of the grammatical structure of the contemporary use of the national language. This paper focuses on the orthographic systems of the published texts included in the written corpus of Filipino, particularly the forms that are derived using the reduplicative affix C1(C2)V1- and the spelling or respelling of forms with <s(i)yon> and <s(i)ya>. The results of the description of these orthographic processes are compared with the standardization rules of the Komisyon sa Wikang Filipino [Commission on the Filipino Language] (KWF), the primary state language agency that assigns the written forms of borrowed terms in Filipino. From the data collected, it appears that while the KWF has already come up with definitive rules with regard to the spelling of borrowed lexical items, there have been a number of instances that these rules have been disregarded as seen from the materials that were published for national circulation and were used for the Filipino language corpus. These forms that deviate from the KWF rules, however, cannot be dismissed since the goal of the *UP MDF* is to describe the current use of Filipino by its speakers, hence, the forms that are spelled differently are considered variants and the word with the higher frequency count becomes the primary entry in the dictionary. Nevertheless, there should be a consideration of the orthographic rules implemented by the KWF as a means of standardization such that discussions on the matter are expected to continue.

KEYWORDS

Filipino lexicography, Filipino language corpus, orthography,
lexical borrowing, corpus linguistics

Introduksiyon

Nais suriin ng pag-aaral na ito sa pamamagitan ng korpus ng wikang Filipino ang mga pamamaraan ng pagbaybay sa Filipino ng mga salitang hiniram at maging ng mga salitang naglalaman ng <s(i)yon> at <s(i)ya>, na nagdudulot madalas ng kalituhan sa mga manunulat at palimbagan. Likas man sa wika na iayon ang kayarian ng mga salitang hiniram sa ibang wika ayon sa sistemang ponolohikal nito, malimit pa rin na nagiging hamon para sa mga tagapagsalita, manunulat, at tagapagpatupad ng tuntunin hinggil sa paggamit ng wika ang pagbaybay sa mga hiram na anyong ito. Sa antas ng ponolohiya, madaling makita ang mga tuntunin sa pagsasaangkop ng mga salitang hiram sa pamamagitan ng pag-*transcribe* ng

mga ito upang makita ang mga pamamaraan sa paghulma ng mga ganitong anyo ng salitang hiram sa ponolohiya ng wikang nanghihiram. Halimbawa nito ang mga salitang Español at Ingles na pumasok na sa wikang Filipino, ang pambansang lingua franca ng Pilipinas. Sa paggamit ng kontemporaryong korpus ng Filipino mailalarawan ang kasalukuyang sistemang ortograpiko na siyang gagamiting batayan sa pagtukoy ng mga pangunahing *entry* at kaakibat nitong variant sa binubuong *UP Monolinggwal na Diksiyonaryong Filipino (UP MDF)*. Sa paglaot ng pag-aaral na ito, hihimayin ang mga salitang hiniram, partikular na ang mga nagtataglay ng kambal-katinig, kung paano binabaybay ang hinalaw na anyo sa pamamagitan ng pagkakabit ng *reduplicative affix*, maging ang mga salitang naglalaman ng <s(i)yon> at <s(i)ya> sa iba't ibang bahagi ng salita.

Ang proyektong *UP Monolinggwal na Diksiyonaryong Filipino* at korpus ng wikang Filipino

Opisyal na sinimulan noong 2014, ang *UP Monolinggwal na Diksiyonaryong Filipino (UP MDF)* ay proyektong naglalayong makabuo ng isang akdang leksikograpiko na nakabatay sa digital na korpus ng kasalukuyang gamit ng wikang Filipino, ang wikang pambansa, at bunga ng isang kritikal na pag-aaral ng *UP Diksiyonaryong Filipino* (Almario 2001 at 2010) (tingnan halimbawa ang isinulat nina Guillermo, Cajote, at Logronio (2015) hinggil sa kritikal na pagsusuri sa pagbibigay-kahulugan sa mga entry ng *UP Diksiyonaryong Filipino*). Ang kritikang ito ay mula sa mga nailimbag nang pag-aaral, panayam sa mga naging bahagi ng paglilimbag nito, at panayam sa mga gumagamit ng nasabing diksiyonaryo at binuonang gabay mula sa serye ng mga palihan ng mga mananaliksik ng proyekto upang makapagbuo ng bagong diksiyonaryo. Kasama sa gabay ang mga hakbang sa pagbuo ng isang diksiyonaryong angkop sa wika ng Pilipinas na may mga sariling katangian kung ihahambing sa ibang mga wika sa labas ng Pilipinas (Newell 1995). Ang digital na korpus ng kasalukuyang Filipino, kapwa sa anyong pasalita at pasulat, ang datos na ginamit sa pangangalap ng mga entry sa UP MDF. Inaasahang sa loob ng limang taon makalilikha ng masinop at sistematikong korpus ng wikang Filipino na binubuo ng 30 milyong salita na magiging daan sa paglilimbag ng bago at pinag-iging edisyon ng diksiyonaryo na papamagatang *UP Monolinggwal na Diksiyonaryong Filipino*.

Mayroong dalawang pangunahing layunin sa pagbubuo ng korpus ng wikang Filipino. Una, magsisilbi itong imbakan ng talasalitaang Filipino na siyang magiging sanggunian ng binubuonang diksiyonaryo. Sa pagtitiyak na ang mga tekstong kinalap sa anyong pasalita at pasulat ay nakapaloob sa kontesкто kung saan ang mga kalahok ay gumagamit ng wikang pambansa sa pagpapahayag o pakikipagtalastasan, maituturing na kinakatawan ng binubuonang korpus ng wikang Filipino ang kasalukuyang gamit nito na siyang nilalayong ilarawan ng binubuonang

monolingguwal na diksiyonaryo. Ikalawa, at kaugnay ng naunang layunin, inaasahang makapagbibigay ng sapat na linggwistikong datos ang korpus ng wikang Filipino upang makapagbigay-paliwanag sa mga kayariang gramatikal ng wikang pambansa, partikular na sa konteksto ng *language contact*, yayamang ang wikang Filipino ang siyang pambansang lingua franca. Kumakatawan din sa mga umuusbong na dibuho sa kayarian ng wika na nakabatay sa aktuwal nitong gamit sa kasalukuyan at sa iba't ibang sitwasyong ginagalawan ang korpus ng wikang Filipino.

Halaw sa International Corpus of English (ICE) na pinasimunan ni Sidney Greenbaum ang binubuong korpus ng wikang Filipino (tingnan ang akda ni Nelson 1996). Gayunpaman, iniangkop ang proseso ng pangangalap at pagsisinop ng korpus sa wikang Filipino sa sarili nitong konteksto. Higit na pinagtuunan ng pansin ang pasalitang korpus, sang-ayon na rin sa pinagbabatayang ICE, subalit tinakdaan pa rin ng malaking bahagdan ang pasulat na teksto upang mailahad ang kasalukuyang sistema ng ortograpiyang tinatalima at, bunga nito, pinalalaganap na paraan ng pagbaybay ng mga palimbagan (Nelson 1996). Samantala, ginamit namang batayan sa pagtukoy sa mga teksto ang pag-aaral nina Atkins, Clear, at Ostler (1991) sa pagbuo ng korpus ng wikang Filipino, at Lee (1992) para naman sa paggamit ng korpus bilang sanggunian ng bubuuing diksiyonaryo.

Ang pasalitang korpus ng Filipino ang bumubuo sa 60 bahagdan ng kabuuang korpus. Nahahati ito sa dalawang uri, dayalogo at monologo. Kumuha ang mga mananaliksik ng datos mula sa dayalogong pribado (hal., direktang pakikipag-usap, video call, skype) at publiko (hal., talakayan sa klase, talakayan sa broadcast, panayam sa broadcast, talumpating politikal, pag-uusap sa pampublikong lugar). Sa dalawa rin nahahati ang datos mula sa monologo, yaong walang gabay na nakasulat (hal., agarang komentaryo, panandaliang salita, talumpati sa mga demonstrasyon) at yaong may gabay na nakasulat (hal., pagbabalita, balitaan, mga palabas sa telebisyon).

Apatnapung (40) bahagdan naman ang itinakda para sa pasulat na teksto ng korpus ng Filipino. Nahahati sa dalawang pangunahing uri ang mga tekstong paghahangian ng pasulat na korpus: di-nakalimbag at nakalimbag. Manggagaling ang datos ng di-nakalimbag na teksto mula sa akademikong sulatin (hal., propesyonal na sulatin, sanaysay ng estudyante, bahaging sanaysay sa pagsusulit, blog) at pakikipagsulatan (hal., sulat, memo). Para naman sa nakalimbag na teksto, kumuha ang mga mananaliksik ng datos mula sa mga sumusunod: akademikong sulatin mula sa iba't ibang disiplina sa agham at teknolohiya, agham panlipunan, at araling pantao; di-akademikong sulatin kagaya ng mga tampok na lathalain; balita, kasama ang patungkol sa showbiz at palakasan; instruksiyonal na sulatin (hal., gabay, alituntunin, polyeto, technical vocational instruction); mapanghimok

na pagsulat gaya ng pangulong tudling; at malikhaing pagsulat (hal., nobela, kuwento, malikhaing sanaysay).

Binigyan ng karampatang pagsipi ang lahat ng sanggunian ng pasulat na teksto, at tiniyak na may kopya ng recording at/o video/audio link ang pangkat ng mga pasalitang teksto upang mapagtibay ang katotohanan at pinanggalingan ng binubuong korpus. Inaasahan na kaalinsabay ng pagkakalimbag ng bagong monolingguwal na diksyonaryo ang pagbubukas sa publiko ng korpus ng wikang Filipino upang maging sanggunian ng marami pang pananaliksik sa kasalukuyang gamit ng wikang pambansa.

Kinuha ang *frequency count of incident* at *frequency count of token per source* ng mga salita na magiging bahagi ng korpus. Ibig sabihin, kapwa inilahad ang aktuwal na bilang ng paglitaw ng isang partikular na salita sa korpus ng pasulat na teksto at ang bilang ng pinaghanguan kung saan lumitaw ang iba't ibang anyo ng mga salitang ito. Sa pamamagitan nito, matutukoy kung aling *variant* ng pagbaybay ng salita ang may mas mataas na bilang ng gamit kumpara sa variant na bagaman mas mababa ang frequency count of incident ay lumilitaw naman sa mas maraming sanggunian. Gagamitin ang pananda ng mga pasulat na teksto na siyang ginagamit ng *UP MDF* sa pag-aaral na ito:

Talahanayan 1: Mga uri ng pasulat na teksto na pinaghanguan ng datos ng korpus (cf. Nelson 1996)

	Uri ng Teksto	Pananda
Di nakalimbag	Akademikong sulatin Correspondence	W1A
Nakalimbag	Akademikong sulatin Di-akademikong sulating Pagbabalita Instruksiyonal na pagsulat Mapanghimok na pagsulat Malikhaing pagsulat	W2A W2B W2C W2D W2E W2F

Kahalagahan ng pasulat na korpus sa leksikograpiya

Sa unang bahagi ng pangangalap ng datos, inaasahang makakalap ng may 8,000,000 salita mula sa mga sanggunian na binanggit sa itaas. Habang isinusulat ang papel na ito, nakapangalap na ang proyekto ng 7,387,569 salita—katanggap-tanggap na ang bilang kung ihahambing sa nilalayong bilang. Mula sa bilang na ito, nakapangalap ang mga mananaliksik ng 2,359,614 salita mula sa mga tekstong nakalimbag. Higit ito sa kalahati ng kabuuang bahagi (40 bahagdan ng korpus ng Filipino) ng pasulat na korpus na kinakailangan.

Sa pag-aaral na ito, hinango ang datos mula sa pasulat na teksto sapagkat ito rin ang gagamiting batayan sa pagtukoy ng mga higit na gamiting baybay ng mga salita na sasangkap sa *UP MDF*. Samakatuwid, ang anyong may mas mataas na

bilang ang siyang ituturing na pangunahing entry samantalang ang may mas kaunting bilang ay mga variant ng salita. Hinango ang mga teksto sa mga materyal na lumitaw mula 1997 hanggang kasalukuyan. Minarapat na itakda bilang pinakamaagang taon ng pagkalimbag ang 1997, 10 taon mula nang isabatas ng Konstitusyon ng 1987 ang Filipino bilang wikang pambansa. Kinalap ang mga pasulat na teksto mula sa iba't ibang palimbagan at lunan, na pawang pambansa ang lawak ng sirkulasyon, at maaaring makuha ninuman at saan man sa Pilipinas.

Mga hiram na salitang may kambal-katinig

Sa kasalukuyan, mas malaya na ang panghihiram ng Filipino ng salita mula sa iba't ibang wika, lalo na sa Ingles. Ang mabilis na takbo ng teknolohiya at pagpasok ng mga bagong konsepto sa Pilipinas ay nagbubunsod ng mabilis ding panghihiram ng mga bagong salita upang tiyak ang pagtukoy sa mga ito. Upang makatugon sa mabilis na pagpasok ng mga hiram na salita sa Filipino, naglatag ng iba't ibang pamantayan ang Komisyon ng Wikang Filipino (KWF), maging ang iba't ibang institusyon halimbawa, ang Sentro ng Wikang Filipino ng Unibersidad ng Pilipinas Diliman, at mga palimbagan, alang-alang sa katapatan sa pagbaybay sa wikang Filipino (tingnan ang Zafra 2005 at 2008).

Sa inilimbag na *Revisyon ng alfabeto at patnubay sa ispelang* ng KWF, makikita ang pagiging bukas sa dalawang variant ng pagbaybay ng mga anyong nilalapan ng *reduplicative affix* na may kambal-katinig (2001). Sang-ayon sa gabay na ito, sa dalawang paraan maaaring sumailalim ang mga salitang sinisimulan ng kambal-katinig. Ang una ay ang pag-uulit lamang ng unang katinig at patinig, gaya ng mga hinalaw na anyong <paplantsahin> at <magpaplantsa> mula sa salitang-ugat na <plantsa> (2001, 37). Ang ikalawa naman ay ang pag-uulit ng buong kambal-katinig at ang kasamang patinig, gaya ng <plaplantsahin> at <magplaplantsa> (2001, 37). Mahalagang mapansin ang batayan ng tindig ng gabay na ito na nagsasaad na maaaring gamitin ang alinman sa dalawang paraang nabanggit, “batay...sa kinagawian ng nagsasalita o variant ng paggamit ng wika sa komunidad” (2001, 36).

Noong 2006, ibinaba ng Kagawaran ng Edukasyon ang Kautusang Pangkagawaran Blg. 42 s. 2006 na naglalahad na kasalukuyan noong pinag-aaralan ang *Revisyon ng alfabeto at patnubay sa ispelang* (2001). Kung kaya ipinag-uutos sa mga paaralan na pansamantalang huwag gamitin ang nasabing gabay; bagkus, sangguniin muna ang nakalakip na *Alfabeto at patnubay sa ispelang* sa pagsusulat ng mga kagamitang panturo at sa mga opisyal na pakikipagkomunikasyon (Kautusang Pangkagawaran Blg. 81, s. 1987). Sa inilakip nilang pansamantalang gabay, matingkad ang panuntunang kung may higit sa isang variant ng pagbaybay ang isang salita, lalo na ang hiniram, malayang makapipili sa dalawang pamamaraan:

kunin ang orihinal na anyo o muling baybayin ayon sa prinsipyong “...kung ano ang bigkas ay siyang sulat at kung ano ang sulat ay siyang basa” (1987, 6). Matapos isagawa ang pag-aaral, binago ang mga pamantayan at nabuo ang *KWF manual sa masinop na pagsulat* (Almario 2014). Gayunpaman, hindi nabanggit ang paksa hinggil sa pag-uulit ng mga salitang nagsisimula sa kambal-katinig.

Mga salitang may kambal-katinig na may reduplicative affix

Sa kabuuan, nakakalap ng 4,510 salitang sinisimulan ng kambal-katinig. Mula sa bilang na ito, 102 ang mga salitang kinakitaan ng nakakabit na reduplicative affix. Sa papel na ito pagtutuunan ng pansin ang mga kambal-katinig na umiiral sa unahan ng salita. Sa usapin ng phonotactics o pagtukoy ng maaari at di maaaring pagkakasunud-sunod ng mga tunog ng wika, mahalaga para sa tunog na pinag-aaralan ang posisyon sa unahan ng salita (Algeo 1978). Ang pag-iral ng kambal-katinig (o kumpol-katinig, dahil may anyong binubuo ng higit sa dalawang katinig) sa unahan ng salita kung gayon ay nagpapahiwatig na pinapayagan ito ng sistemang ponolohikal ng isang wika.

Gagamitin ang antas na ortograpiya, na nakasulat sa loob ng < >, sa pagsasaayos ng pasulat na korpus ng Filipino. Matapos isaayos ang datos ng inilimbag na teksto, nakakalap ng 44 na kambal-katinig: <bl, br, bw, by, ch, cl, cr, dr, dw, dy, fl, fr, gl, gr, gw, gy, kl, kn, kr, kw, ky, ph, pl, pr, ps, pw, py, sc, sh, sk, sl, sm, sn, sp, sq, sw, sy, th, tr, ts, tw, ty, wh, wr>. Mula sa 44 na kambal-katinig na ito, siyam ang kinakitaan ng pagsasailalim sa reduplication sa kinalap na korpus: <br, cr, dr, gr, kl, kw, pl, pr, tr>. Samantala, isinama rin ang mga katinig na bagama’t kumakatawan sa isa lamang tunog, ibinabaybay naman gamit ang dalawang titik: <ch> at <ts> na kapwa para sa *voiceless postalveolar affricate*, <sh> para sa *voiceless postalveolar fricative*, at <th> para sa *voiceless interdental fricative*.

Tatalakayin sa ibaba ang paraan ng pag-uulit ng mga nasabing kambal-katinig. Mayroong dalawang pangunahing klasipikasyon ng pag-uulit ng mga kambal-katinig: ang C1V1- at C1C2V1-.

Anyo 1: C1V1-

Inilalahad sa Anyo 1 na ang ginagamit na reduplicative affix sa mga salitang sinisimulan ng kambal-katinig ay binubuo ng unang katinig ng kambal-katinig at patinig ng unang pantig ng salitang-ugat. Mapapansing hindi nito isinasama ang ikalawang katinig na bahagi ng *word-initial consonant cluster*. Tingnan ang mga halimbawang ipinapakita sa Talahanayan 2. Gagamitin ang mga salitang “paghahalaw” at “hinalaw” sa antas na ortograpiko, upang tukuyin ang mga nilapiang anyo ng mga salitang kinalap.

Talahanayan 2: Paghahalaw gamit ang C1V1- ng mga salitang hiram sa Ingles

Kambal-katinig	Salitang-ugat	Hinalaw na anyo	Frequency count	Kabuuang bilang ng sanggunian	Uri ng mga sanggunian
br	bra	nagbabra	1	1	W2F (1)
cr	cramps	nacacrams	1	1	W2B (1)
dr	draw	dadraw	1	1	W2A (1)
pr	pray	nagpapray	1	1	W2B (1)
	promote	mapopromote	3	1	W2B (1)
	propaganda	pagpopropaganda	3	1	W2A (1)

Makikita sa Talahanayan 2 ang mga salitang hiram mula sa Ingles, at pinanatili ang orihinal na baybay. May anim na salita, na pawang nakalap sa nakalimbag na teksto, na nasa ilalim ng Anyo 1 o pagkakabit ng reduplicative affix na C1V1-. Sa pag-uulit ng mga salita, ang inuulit lamang ay ang unang katinig at ang patinig ng unang pantig, gaya ng <nagbabra> (mula sa salitang-ugat na <bra>), <mapopromote> (mula sa <promote>), at <pagpopropaganda> (mula sa <propaganda>). Nakakuha ng frequency count na tig-3 ang dalawang huling anyong nabanggit; gayunpaman, pawang nanggaling sa tig-iisang sanggunian lamang ang mga ito. Dalawa sa mga anyong inilatag sa Talahanayan 2 ay hindi sumailalim sa ibang pagbaybay, kung kaya ganoon din ang anyo ng reduplicative affix na ikinakabit dito—ang mga salitang <dadraw> (mula sa salitang-ugat na <draw>) na may reduplicative affix na ibinabaybay nang <da-> at <nagpapray> (mula sa <pray>) na may <pa->. Palalawigin ang pagtalakay ng ganitong katunayan sa Anyo 2 sa ibaba. Sa kabilang banda, nasa Talahanayan 3 ang mga naisa-Filipino na salitang Español na may C1V1- reduplicative affix.

Talahanayan 3: Paghahalaw gamit ang C1V1- ng mga naisa-Filipinong salita sa Español

Kambal-katinig	Salitang-ugat	Hinalaw na anyo	Frequency count	Kabuuang bilang ng sanggunian	Uri ng mga sanggunian
dr	droga	nagdodroga	3	2	W2A (1) W2F (1)
		pagdodroga	2	2	W2A (1) W2F (1)
gr	grupo	paggugrupo	7	3	W2A (3)
kl	klase	pagkaklase	1	1	W2A (1)
	klasipika	pagkaklasipika	1	1	W2A (1)
pl	plano	pagpaplanong	34	7	W1A (2) W2A (5)
		pagpaplano	29	14	W1A (1) W2A (9) W2B (1) W2C (1) W2D (1) W2F (1)

Talananayan 3: Paghahalaw gamit ang CIV1-
ng mga naisa-Filipinong salita sa wikang Español (Cont'n.)

Kambal- katinig	Salitang-ugat	Hinalaw na anyo	Frequency count	Kabuuang bilang ng sanggurian	Uri ng mga sanggurian
		nagpaplano	8	5	W2A (2) W2B (1) W2C (1) W2D (1)
		pinagpaplanuhan	2	2	W2B (2)
		nagpaplanong	1	1	W2B (1)
	plantsa	paplantsahin	1	1	W2F (1)
		ipaplantsa	1	1	W2F (1)
	platero	pagpaplatero	1	1	W2F (1)
pr	prepara	pagpepreparang	1	1	W2B (1)
	preserba	pagpepreserba	1	1	W1A (1)
	presyo	pagpepresyo	2	1	W2B (1)
	pribilehiyo	napipribilehiyo	1	1	W2A (1)
	prinsesa	piprinsesahin	1	1	W2A (1)
	prisinta	ipiprisinta	3	3	W2C (1), W2F (2)
	prito	pagpiprito	1	1	W2A (1)
	problema	pagpoproblema	1	1	W2F (1)
	proklama	nagpoproklama	1	1	W2A (1)
	propesiya	nagpopropesiya	1	1	W2A (1)
	proseso	pagpoproseso	13	4	W2A (4)
		pagpoprosesong	2	2	W2A (2)
		nagpoproseso	1	1	W2A (1)
	protekta	poprotektahan	1	1	W2F (1)
	protesta	nagpoprotesta	1	1	W2F (1)
	prusisyon	nagpuprusisyon	2	1	W2A (1)
		pagpuprusisyon	2	1	W2A (1)
tr	trabaho	nagtatrabaho	47	33	W1A (1), W2A (13), W2B (4), W2C (1), W2D (1), W2F (11)
		pagtatrabaho	28	18	W1A (1), W2A (10), W2B (1), W2C (2), W2D (1), W2F (3)
		pinagtatrabahuan	5	4	W2A (2), W2F (2)
		pinagtatrabaho	3	3	W2D (1), W2F (2)
		magtatrabaho	2	2	W2B (1), W2F (1)
		nagtatrabahong	2	2	W2A (1), W2D (1)
		pinagtatrabahuan	2	2	W2A (2)
		nakapagtatrabaho	1	1	W2B (1)
		pagtatrabahong	1	1	W2F (1)
		tatrabahuhin	1	1	W2F (1)
		tatrabahuin	1	1	W2F (1)
	trangaso	tatrangkasuhin	1	1	W2F (1)
	transforma	natattransforma	1	1	W2A (1)
	troso	pagtotroso	12	3	W2A (3)
		magtotroso	1	1	W2A (1)

Ang pagsasa-Filipino ng mga salitang Español ay nagaganap tuwing ang phonetics ng isang hiniram na salita, na kadalasan ay mula sa Ingles, ay binabago upang maging kahawig ng sa Español (Baklanova 2017). Isinama sa pag-aaral na ito hindi lamang ang mga salitang Español na hiniram sa Filipino kundi pati na rin ang mga salitang Ingles na binago upang magtunog-Español. Gaya ng makikita sa Talahanayan 3, nakahango ng 25 anyo ng naisa-Filipinong salita na wikang Español na umaayon sa Anyo 1. Sa 25 na salitang ito, apat ang may mataas na frequency count—higit sa 10. Ang mga naturang halaw na anyo ay mula sa mga salitang-ugat na <plano> (<pagpaplano(ng), nagpaplano(ng), pinagpaplanuhan>) na may 74 na frequency count mula sa kabuuang bilang na 29 sanggunian; <proseso> (<pagpoproseso(ng), ipoproseso, nagpoproseso>) na may 16 na frequency count na lumitaw sa pitong sanggunian; <trabaho> (<nagtatrabaho(ng), pagtatrabaho(ng), pinagtatrabahu(h)an, pinagtatrabaho, magtatrabaho, nakapagtatrabaho, tatrabahu(h)in>) na may pinakamataas na frequency count sa 93 na lumitaw sa 68 sanggunian; at <troso> (<pagtotroso, magtotroso>) na may frequency count na 13 na lumitaw sa apat na sanggunian. Bagaman isinama sa Talahanayan 3, mapapansin na ang paggamit ng “bagong” hiram na titik na <f> ang salitang <transforma> (kaya ang anyo ng nilapian ng reduplicative affix ay <natatransforma>). Nakapaloob pa rin ang anyong ito sa Anyo 1. Ang mga hinalaw na anyo ng mga salitang <presyo> at <prusisyon>, bagaman may frequency count na higit sa 1, ay lumilitaw sa tig-iisa lamang na sanggunian.

Nagpapakita naman ng mga salitang hiniram mula sa Ingles subalit sumailalim na sa respelling ayon sa ortograpiyang Filipino ang Talahanayan 4.

**Talahanayan 4: Paghahalaw gamit ang CIV1-
ng mga muling binaybay na salitang hiniram sa Ingles**

Kambal-katinig	Salitang-ugat	Hinalaw na anyo	Frequency count	Kabuuang bilang ng sanggunian	Uri ng mga sanggunian
dr	drayb	nagdadrayb	1	1	W2A (1)
	dribol	didribol	1	1	W2C (1)
	drowing	pagdodrowing	2	2	W2A (1), W2F (1)
gr	gradweyt	kagagradweyt	1	1	W2F (1)
pr	praktis	pagpapraktis	1	1	W2D (1)
		magpapraktis	1	1	W2F (1)
	predik	nagpepredik	1	1	W2A (1)
	produs	pagpoprodus	1	1	W2A (1)
	prodyus	ipoprodyus	3	1	W2A (1)
		pagpoprodyus	2	1	W2A (1)

Nagdaan man sa respelling ang mga hiram na salita mula sa Ingles gaya ng nakasaad sa Talahanayan 4, hindi naman kinakitaan ang mga ito ng mga hiram na titik na ipinakilala sa bagong alpabetong Filipino, gaya ng <nagdadrayb> na mula sa salitang-ugat na <drayb>, at hindi <drayv>. Gayundin, hindi pinananatili ang huling katinig sa kambal-katinig na umiiral sa dulo ng salita, gaya ng <nagpepredik> na mula sa salitang-ugat na <predik>, at hindi <predikt>. Ang pananatili naman ng katinig na <t> sa salitang <kagagradweyt> (mula sa <gradweyt>) ay dahil sa ang sinusundang titik na <y> ay kumakatawan sa tunog na [j] na isang malapatinig. Nakahango ng walong salitang-ugat na may kaakibat na nilapiang anyo sa pamamagitan ng pagkakabit ng reduplicative affix na C1V1-

Anyo 2: C1C2V1-

Ang Anyo 2 ay nagsasaad na isinasama sa pag-uulit ang dalawang katinig ng isang *consonant cluster* na nasa unahan ng salita. Tingnan ang mga halimbawang nasa Talahanayan 5.

Talahanayan 5: Paghahalaw gamit ang C1C2V1-
ng mga salitang hiram sa Ingles

Kambal-katinig	Salitang-ugat	Hinalaw na anyo	Frequency count	Kabuuang bilang ng sanggunian	Uri ng mga sanggunian
ch	chant	nagchachant	1	1	W1B (1)
sh	shabu	magshashabu	1	1	W2A (1)

Nakapagtala ng dalawang salitang-ugat na sumasailalim sa pag-uulit sa C1C2V1-na pattern: <nagchachant> (mula sa salitang-ugat na <chant>) at <magshashabu> (mula sa salitang-ugat na <shabu>). Mahalagang pansinin ang katangian ng dalawang salitang nakalap sa ilalim ng klasipikasyong ito—ang pagsisimula sa kambal-katinig na <ch> at <sh>. Sa antas ng ponolohiya, ang dalawang nabanggit ay itinuturing na *single segment* lamang—ang <ch> ay kumakatawan sa *voiceless affricate* at <sh> para sa *voiceless postalveolar fricative*. Gayunpaman, dahil sa ang tinatalakay sa papel na ito ay nauukol sa mga kambal-katinig sa antas na ortograpiya, minarapat na isama rin ang mga salitang nagsisimula sa dalawang nabanggit na bahagi, yayamang isinusulat ang mga ito sa pamamagitan ng dalawang titik.

Makikita naman sa Talahanayan 6 ang mga naisa-Filipino na salitang Español na kinakabit ng reduplicative affix sa Anyo 2.

**Talahanayan 6: Paghahalaw gamit ang C1C2V1-
ng mga naisa-Filipino na salitang Español**

Kambal-katinig	Salitang-ugat	Hinalaw na anyo	Frequency count	Kabuuang bilang ng sanggunian	Uri ng mga sanggunian
gr	grado	paggragrado	16	2	W1A (2)
		naggragrado	4	2	W1A (2)
kw	grupo	paggrugrupo	4	1	W2A (1)
	kwento	magkwekwentuhan	1	1	W1B (1)
		nagkwekwento	1	1	W2A (1)
pl	plano	pagplaplano	2	1	W2A (1)
tr	trabaho	nagtratrabaho	5	5	W1A (1), W2D (1), W2F (3)
		magtratrabaho	4	1	W2F (1)
		pagtratrabahuan	1	1	W1A (1)

Mayroong limang salitang-ugat na naisa-Filipino na naitala mula sa pasulat na korpus na mayroong ganitong uri ng pag-uulit, gaya ng <paggrugrupo> (mula sa salitang-ugat na <grupo>) at <pagplaplano> (mula sa <plano>). Dalawa sa mga salitang-ugat sa Talahanayan 6 sa itaas ay may mataas na frequency count na 10 o higit pa, dulot na rin ng mga hinalaw na anyo mula sa mga salitang ito: ang salitang-ugat na <grado> na ang kabuuang frequency count ng mga hinalaw na anyo ay 20, at <trabaho> na may 10. Samantala, lumitaw naman ang anyong <kwento>, na taliwas sa isinusulong ng KWF na wastong baybay <kuwento>, sang-ayon sa tuntunin nito hinggil sa pagsisingit ng <u> sa mga salitang may kambal-katinig kung saan ang ikalawang katinig ay ang malapatinig na [w] o [y]. Dahil dito, itinuturing ding salitang may kambal-katinig ang ganitong anyo ng salitang-ugat at nakapaloob sa pangkat ng mga salitang nilalapan ng Anyo 2 reduplicative affix. Ang mga hinalaw na anyong <paggrugrupo>, <pagplaplano>, at <magtratrabaho> ay nakalap sa tig-iisang sanggunian lamang kahit may frequency count na higit sa isa.

Paghahambing ng mga anyong kapwa may C1V1- at C1C2V1-

Mapaghahambing ang ilang mga salitang sumailalim sa reduplicative affix sa dalawang anyo na tinalakay sa itaas. May magkahawig o iisang salitang-ugat subalit kinakitaan ng mahigit sa isang paraan ng pagkakabit ng reduplicative affix ang mga ito. Sa pagtukoy sa kung alin ang pangunahing variant, dalawang aspekto ang titingnan: ang frequency count at ang kabuuang bilang ng sanggunian kung saan lumitaw ang partikular na anyo ng mga salitang ginamit.

May tatlong naitalang salita na pumapaloob sa kalagayang binanggit, gaya ng isinasaad sa Talahanayan 7.

**Talahanayan 7: Paghahalaw gamit ang C1C2V1-
ng mga naisa-Filipino na salitang Español**

Salitang-ugat	Anyo	Hinalaw na anyo (Frequency count)	Kabuuang frequency count	Kabuuang bilang ng sanggunian		
draw	1 (E)	dadraw	1	1		
	1 (E-R)	pagdodrawing	2	2		
plano	1 (H)	pagpaplanong	74	29		
		pagpaplano				
		nagpaplano				
		pinagpaplanuhan				
		nagpaplanong				
trabaho	2 (H)	pagplaplapano	2	1		
	1 (H)	nagtatrabaho	93	68		
		pagtatrabaho				
		pinagtatrabahuhan				
		pinagtatrabaho				
		magtatrabaho				
		nagtatrabahong				
		pinagtatrabahuan				
		nakapagtatrabaho				
		pagtatrabahong				
		tatrabahuhin				
		tatrabahuin				
		nagtratrabaho			10	7
		magtratrabaho				
pagtratrabahuan						

42

Sa ilalim ng hanay na pinangalanang Anyo sa Talahanayan 7, ang E ay kumakatawan sa mga salitang hiniram mula sa Ingles at pinanatili ang orihinal na anyo nito sa nasabing wika. Ang H naman ay tumutukoy sa mga salitang Español na naisa-Filipino subalit sumailalim sa respelling sang-ayon sa ponolohiya ng huli. Samantala, ang panandang E-R naman ay kumakatawan sa mga salitang hiniram sa Ingles subalit ang anyo ay muling binaybay sang-ayon sa ortograpiyang Filipino.

Sa tatlong salitang may iba't ibang uri na inilatag sa Talahanayan 7, dalawa ang kapansin-pansing may malaking pagkakaiba sa frequency count ng mga variant: ang <plano> at <trabaho>. Ang unang variant <plano> na kinakabitan ng reduplicative affix na C1V1- ay may kabuuang frequency count na 74 at kinakitaan ng limang pagkakataon ng paghalaw: <pagpaplanong, pagpaplano, nagpaplano, pinagpaplanuhan, nagpaplanong>. Samantala, ang isa pang variant na kinakabitan ng reduplicative affix na C1C2V1- ay may frequency count na dalawa at kinakitaan lamang ng isang halimbawa, <pagplaplapano>. Malaki rin ang pagkakaiba ng kabuuang bilang ng sangguninan ng dalawang variant, 29 para sa C1V1- samantalang isa lamang para sa C1C2V1-. Malaki rin ang kaibahan ng dalawang variant ng <trabaho>, tulad ng makikita sa Talahanayan 7. Bagaman kapwa nasa anyong naisa-Filipino; gayunpaman, ang variant ng <trabaho> kung saan ang reduplicative affix na ikinakabit ay nasa anyong C1V1- ay may kabuuang

frequency count na 93 at may 11 halimbawa ng hinalaw na anyo: <nagtatrabaho, pagtatrabaho, pinagtatrabahuan, pinagtatrabaho, magtatrabaho, nagtatrabahong, pinagtatrabahuan, nakapagtatrabaho, pagtatrabahong, tatrabahuhin, tatrabahuin>. Ang mga anyong ito ay nakuha sa kabuuang bilang na 68 sanggunian. Ang ikalawang variant naman, yaong kinakabitan ng reduplicative affix na C1C2V1-, ay may kabuuang frequency count na 10 at may tatlong halaw na anyo: <nagtratrabaho, magtratrabaho, pagtratrabahuan>. Ang mga ito ay lumitaw sa kabuuang bilang na pitong sanggunian.

May dalawang variant ang lumitaw sa salitang <draw> na hiniram sa Ingles: (a) anyong pinanatili ang orihinal na baybay at kinakabitan ng reduplicative affix na C1V1-; at (b) anyong muling binaybay at may reduplicative affix na C1V1-, katulad ng naunang variant. Sa dalawang variant na ito, higit ang paglitaw ng ikalawang variant sa frequency count na dalawa na lumitaw sa dalawang sangguian. Samantala, ang unang variant ay may frequency count na isa. Bagaman masasabing hindi gaanong mahalaga ang kaibahan kung pag-uusapan ang estadistika, mahalaga pa ring itala ang ganitong anyo bilang sanggunian.

Mga salitang may <s(i)yon> at <s(i)ya>

Bagaman tinitingnan bilang halimbawa ng pantig na may kambal-katinig sa antas ortograpiko, mahihinuha na ang kumpletong anyo ng <syo> at <sya> sa maraming salita sa Filipino ay <siyo> at <siya>. Mababakas sa kasaysayan ng mga anyong nabanggit ang pagiging bunga ng pagdaragdag ng malapatinig na <y> sa pagitan ng orihinal nitong anyo na kambal-patinig na <io> at <ia>, yayamang ang mga kambal-patinig na ito ay hango sa mga hiram na salitang Español, habang ang ilan naman ay naisa-Filipino na anyo.

Sa respelling ng mga salitang may kambal-patinig na <io> at <ia>, bumuo ng gabay ang noon ay Surian sa Wikang Pambansa (KWF ngayon) sa pamamagitan ng *Mga tuntunin ng ortograpiyang Pilipino/Patnubay sa pagwawasto ng mga aklat babasahin* (1977). Ayon sa gabay na ito, nananatili ang kambal-patinig subalit sinisingitan ng malapatinig na <y> sa mga salitang ang pag-iral ng kambal-patinig ay nasa unang pantig, gaya ng <siyá>. Nawawala naman ang unang patinig ng kambal-patinig kapag nasa ikalawa o iba pang pantig ng salita, gaya ng <tilapyà>.

Mula noong 2009, bahagyang binago ng KWF ang mga gabay sa ortograpiya sa pamamagitan ng pagbibigay ng karagdagang mga tuntunin sa pagbaybay ng mga salitang may kambal-patinig, hanggang sa ilathala noong 2014 ang mga pagbabagong ito. Bilang tugon sa *Revisyon* at gabay sa pagbaybay na nakapaloob sa Kautusang Pangkagawaran Blg. 42 s. 2006 na nagbibigay-kalayaan sa manunulat na pumili ng variant ng pagbaybay sa Filipino, isinasaad sa inilathalang *Manwal* noong 2014 ang ilang tuntunin sa layuning buwagin ang diumano’y malaya-basta

pare-pareho na lapit (2001). Bagaman pinananatili ang nakasanayan na pagtingin sa mga kambal-patinig, tumukoy ang KWF ng apat na kataliwasan sa tuntuning ito:

- (1) kapag ang kambal-patinig ay sumusunod sa katinig sa unang pantig ng salita (hal., <tiya> na mula sa <tia> at <kiyosko> na mula sa <kiosco>);
- (2) kapag ang kambal-patinig ay sumusunod sa dalawa o mahigit pang kumpol-katinig (consonant cluster) sa loob ng salita (hal., <ostiya> na mula sa <hostia> at <leksiyon> na mula sa <leccion>);
- (3) kapag ang kambal-patinig ay sumusunod sa tunog na [h] (hal., <estrategiya> na mula sa <estrategia> at <kolehiyo> na mula sa <colegio>); at
- (4) kapag ang kambal-patinig ay nasa dulo ng salita at may diin ang bigkas sa unang patinig ang orihinal (hal., <ekonomiya>, na mula sa salitang Espanyol na <economía>). (2014, 25)

Tinukoy sa pag-aaral na ito sa pamamagitan ng pasulat na korpus kung nasusunod ang mga gabay sa pagsusulat ng mga salitang may <s(i)ya> at <s(i)yon> at paghahambingin ang dalas ng paglitaw ng mga anyo sa pamamagitan ng pagsusuri sa bilang ng mga ito, maging ang bilang ng sanggunian kung saan lumitaw ang mga anyong sinusuri. Ang mga anyong may <s(i)ya> at <s(i)yon> na susuriin ay yaon lamang matatagpuan sa hulihan ng mga salita, yayamang partikular sa <s(i)yo>, ibang usapin pa ang paggamit ng <o> vs. <u> kung nasa loob ng salita. Ang kahihinatnan ng pagsusuri sa pasulat na korpus ang siyang magiging gabay sa pagtukoy ng mas gamiting variant ng mga nasabing salita sa binubuong monolingguwal na diksiyonaryong Filipino.

Mga anyong may <s(i)yon>

Mula sa pasulat na korpus ng Filipino, nakahango ng 670 salitang naglalaman ng <s(i)yon> na may hindi bababa sa frequency count na dalawa. Nilinis ang datos sa pamamagitan ng pagsasama-sama ng mga anyong may iisang salitang-ugat lamang, kung kaya ibinibilang na isa ang anyong <aksiyon> para sa mga variant na <aksiyon>, <umaksiyon>, at <aksiyong> at ang anyong <lisensiya> para sa mga variant na <lisensiya>, <lisensiyadong>, at <lisensiyang> (sa kaso ng mga salitang naglalaman ng <s(i)ya>). Pagkatapos linisin ang datos na nakalap, mula 670 ay naging 394 ang mga anyong pinag-aralan at 340 ang mga salitang naglalaman ng <syon> habang 17 naman ang may <siyon>. Nakahango naman ng 37 pares ng mga salitang kapwa naglalaman ng <siyon> at <syon>.

Makikita sa Talahanayan 8 ang mga pares ng salita kung saan ang anyong may <syon> ay nakalalamang sa frequency count kaysa anyong may <siyon>.

Talahanayan 8: Pares ng salitang may <s(i)yon>, kung saan mas mataas ang frequency ng may <syon>

Salita	Hinalaw na anyo (Frequency count)	Uri ng mga sanggunian	Kabuuang frequency count	Kabuuang bilang ng sanggunian
1	adiksiyon	adiksiyon (1)	W2F (1)	1
	adiksiyon	adiksiyon (4)	W2A (1)	1
2	aksiyon	aksiyon (24)	W2A (9), W2C (1), W2F (1)	14
	aksiyon	umaksiyon (2) aksiyong (3) aksiyon (20)	W2A (1) W2A (2) W1A (3), W2A (10), W2C (4), W2E (1)	25
3	deskripsiyon	deskripsiyon (8)	W1A (1), W2A (5), W2F (1)	9
	deskripsiyon	deskripsiyong (4) deskripsyon (7)	W2A (2) W2A (5), W2F (1)	8
	deskripsiyon	deskripsiyong (7)	W2A (2)	
4	dimensiyon	dimensiyon (13) dimensiyong (3)	W2A (8) W2A (3)	11
	dimensiyon	dimensiyon (19) dimensiyong (1)	W2A (6) W2A (1)	7
5	instruksiyon	instruksiyon (9)	W2A (7)	7
	instruksiyon	instruksiyon (19) instruksiyong (5)	W1A (2), W2A (4), W2F (1) W1A (1), W2A (1)	9
6	interaksiyon	interaksiyon (13) interaksiyong (1)	W2A (10) W2A (1)	11
	interaksiyon	interaksiyon (20) interaksiyong (3)	W1A (2), W2A (7) W2A (2), W2C (1)	12
7	introduksiyon	introduksiyon (8)	W2A (4), W2F (2)	6
	introduksiyon	introduksiyon (24) introdaksyon (2)	W1A (2), W2A (9), W2F (4) W2A (1)	16
8	komersiyo	komersiyo (5)	W1A (1), W2A (2), W2D (1)	6
	komersyo	komersiyong (1) pangongomersiyo (2) komersyo (11)	W2F (1) W2F (1) W2A (4)	4

Talahanayan 8: Pares ng salitang may <s(i)yon>, kung saan mas mataas ang frequency ng may <syon> (Cont'n.)

	Salita	Hinalaw na anyo (Frequency count)	Uri ng mga sanggurian	Kabuuang frequency count	Kabuuang bilang ng sanggurian
9	komprehensiyon	komprehensiyon (2)	W2A (1)	2	1
	komprehensyon	komprehensyon (8)	W2A (2)	8	2
10	kontradiksiyon	kontradiksiyon (5)	W2A (4), W2F (1)	5	5
	kontradiksyon	kontradiksyon (15)	W2A (10), W2C (1), W2F (1)	18	15
		kontradiksyong (3)	W1A (1), W2A (2)		
11	korupsiyon	korupsiyon (5)	W2A (3), W2F (1)	5	4
	korupsyon	korupsyon (11)	W1A (2), W2A (1), W2C (1)	11	4
12	pensiyon	pensiyon (2)	W2A (1)	2	1
	pensyon	pensyon (4)	W2A (1), W2F (1)	4	2
13	persepsiyon	persepsiyon (4)	W2A (2)	4	2
	persepsyon	persepsyon (17)	W1A (1), W2A (4)	18	6
		persepsyong (1)	W2A (1)		
14	proteksiyon	proteksiyon (20)	W2A (12), W2F (1)	21	14
	proteksyon	proteksiyong (1)	W2A (1)	27	12
		proteksyon (25)	W1A (2), W2A (3), W2B (2), W2D (2), W2F (1)		
		proteksyong (2)	W2A (1), W2F (1)		
15	seleksiyon	seleksiyon (2)	W2A (1)	3	2
	seleksyon	seleksiyong (1)	W2A (1)	5	3
		seleksyon (4)	W2A (2)		
		seleksyong (1)	W2F (1)		
16	intensiyon	intensiyon (6)	W2A (4), W2F (1)	10	8
	intensyon	intensiyong (4)	W2A (3)	10	10
		intensyon (7)	W1A (2), W2A (3), W2B (1), W2F (1)		
		intensyong (3)	W2A (3)		

Sa 16 na pares ng mga anyong nakalap at inilatag sa Talahanayan 8, limang pares ang kinakitaan ng malaking kalamangan ng frequency count (10 o higit pa) sa <syon> kung ihahambing sa mga salitang may <siyon>: <aksyon> vs. <aksiyon> na may frequency count difference na 22; <instruksyon> vs. <instruksiyon> sa 15; <introduksyon> vs. <introduksiyon> sa 18; <kontradiksiyon> vs. <kontradiksiyon> sa 13; at <persepsyon> vs. <persepsiyon> sa 14. Samantala, ang <intensyon> at <intensiyon> ay may magkaparehong frequency count na 10, gaya ng makikita sa itaas. Gayunpaman, minarapat na isama ang pares na ito sa Talahanayan 8 sapagkat kung titingnan ang kabuuang bilang ng sanggunian, maging ang kaibahan ng mga pinagkunan ng mga salita, higit na nakalalamang ang anyong <intensyon>.

Ipinapakita naman ng Talahanayan 9 ang mga salitang may <siyon> na may mas mataas na frequency count kaysa katambal nitong salitang may <syon>.

Talahanayan 9: Pares ng salitang may <s(i)yon>, kung saan mas mataas ang frequency ng may <siyon>

Salita	Hinalaw na anyo (Frequency count)	Uri ng mga sanggunian	Kabuuang frequency count	Kabuuang bilang ng sanggunian
1 asersiyon	asersiyon (3)	W2A (2)	3	2
	asersiyon (2)	W2A (2)	2	2
2 atensiyon	atensiyon (26)	W2A (8), W2C (1), W2F (6)	28	17
	atensiyong (2)	W2A (1) W2F (1)		
atensyon	atensyon (22)	W1A (2), W2A (8), W2B (2), W2C (1), W2D (1), W2E (1)	22	15
3 atraksiyon	atraksiyon (3)	W2A (2), W2F (1)	5	5
	atraksiyong (2)	W2A (2)		
atraksyon	atraksyon (3)	W2A (2)	3	2
4 bersiyon	bersiyon (36)	W1A (1), W2A (9), W2F (4)	60	20
	bersiyong (24)	W1A (1), W2A (3), W2F (2)		
bersyon	bersyon (32)	W1A (5), W2A (9), W2B (1), W2C (1), W2F (2)	42	21
	bersyong (10)	W1A (1), W2A (2)		

Talahanayan 9: Pares ng salitang may <s(i)yon>, kung saan mas mataas ang frequency ng may <siyon> (Cont'n.)

	Salita	Hinalaw na anyo (Frequency count)	Uri ng mga sanggunian	Kabuuang frequency count	Kabuuang bilang ng sanggunian
5	eksepsiyon	eksepsiyon (5)	W2A (2)	5	2
	eksepsyon	eksepsyon (2)	W2A (2)	2	2
6	ekstensiyon	ekstensiyon (6)	W1A (1), W2A (3), W2F (1)	7	6
	ekstensyon	ekstensiyong (1) ekstensyon (3)	W2A (1) W1A (1), W2A (2)	3	3
7	eleksiyon	eleksiyon (25)	W1A (2), W2A (6), W2E (1), W2F (2)	25	11
	eleksyon	eleksyon (15)	W1A (5), W2A (1), W2C (1), W2F (1)	17	10
		eleksyong (2)	W2A (1), W2F (1)		
8	funsiyon	funsiyon (5)	W2A (4)	5	4
	funksyon	funksyon (1)	W2A (1)	2	1
9	imbensiyon	imbensiyon (6) imbensiyong (6)	W2A (4) W2A (2)	12	6
	imbensyon	imbensyon (5)	W2A (1), W2A (4)	5	5
10	impeksiyon	impeksiyon (28) impeksiyong (1) naimpeksiyon (8)	W2A (7) W2F (1) W2A (2)	37	10
	impeksyon	impeksyon (4)	W2A (2), W2D (1)	4	3
11	interseksiyon	interseksiyon (11) interseksiyong (2)	W2A (6) W2A (2)	13	8
	interseksyon	interseksyon (3)	W1A (1), W2A (1)	3	2
12	koleksiyon	koleksiyon (22)	W1A (1), W2A (5), W2F (4)	30	14
		koleksiyong (8)	W2A (1), W2F (3)		
	koleksyon	koleksyon (19)	W1A (1), W2A (5), W2B (1), W2C (1), W2F (3)	23	15
		koleksyong (4)	W1A (1), W2A (2), W2F (1)		

Talahanayan 9: Pares ng salitang may <s(i)yon>, kung saan mas mataas ang frequency ng may <siyon> (Cont'n.)

	Salita	Hinalaw na anyo (Frequency count)	Uri ng mga sanggunian	Kabuuang frequency count	Kabuuang bilang ng sanggunian
13	koneksiyon	koneksiyon (15)	W2A (4), W2E (1), W2F (3)	17	10
	koneksiyon	koneksiyong (2) koneksyon (10)	W2A (2) W2A (5), W2B (1), W2F (1)	11	8
		koneksyong (1)	W2A (1)		
14	konstruksiyon	konstruksiyon (24)	W1A (1), W2A (7), W2F (1)	26	11
	konstruksyon	konstruksiyong (2) konstruksyon (17) konstruksyong (5)	W2A (2) W2A (9) W1A (2), W2A (2)	22	14
15	kumbensiyon	kumbensiyon (6)	W2A (3), W2F (2)	15	11
		kumbensiyong (2)	W2F (1), W2A (1)		
	kumbensyon	kumbensiyonal (7) kumbensyon (3)	W2A (4) W2A (2), W2F (1)	3	3
16	leksiyon	leksiyon (118)	W1A (2), W2A (12), W2F (2)	124	21
		leksiyong (6)	W2A (4) W2F (1)		
	leksyon	leksyon (6) leksyong (1)	W1A (1), W2A (3) W2A (1)	7	5
17	produksiyon	produksiyon (67)	W1A (1), W2A (23), W2F (4)	70	30
	produksyon	produksiyong (3) produksyon (53)	W2A (2) W1A (4), W2A (17), W2C (3)	62	31
		produksyong (9)	W2A (6), W2C (1)		
18	reaksiyon	reaksiyon (29)	W2A (17), W2F (4)	47	30
	reaksyon	reaksiyong (18) reaksyon (40)	W2A (9) W1A (2), W2A (11), W2C (3), W2D (1), W2F (2)	46	21
		reaksyonaryong (5) reaksyong (1)	W2A (1) W2B (1)		

Talahanayan 9: Pares ng salitang may <s(i)yon>, kung saan mas mataas ang frequency ng may <siyon> (Cont'n.)

	Salita	Hinalaw na anyo (Frequency count)	Uri ng mga sanggunian	Kabuuang frequency count	Kabuuang bilang ng sanggunian
19	restriksiyon	restriksiyon (3)	W1A (1), W2A (1)	3	2
	restriksiyon	restriksiyon (2)	W2A (2)	2	2
20	seksiyon	seksiyon (44)	W1A (1), W2A (7), W2F (1)	74	20
		pagseseksiyon (3) seksiyong (27)	W2A (1) W2A (9), W2F (1)		
	seksyon	seksyon (49)	W1A (2), W2A (10), W2C (1), W2F (1)	57	21
		seksyong (8)	W2A (6), W2F (1)		
21	transaksiyon	transaksiyon (5)	W2A (4), W2D (1)	6	6
		transaksiyong (1)	W2A (1)		
	transaksyon	transaksyon (2)	W1A (1), W2A (1)	6	4
		transaksyong (4)	W2A (2)		

Sa dalawang pares ng mga anyong nakalap gaya ng nakasaad sa Talahanayan 9, anim na pares ang kinakitaan ng malaking kalamangan ng frequency count (10 o higit pa) sa <siyon> kung ihahambing sa mga anyong may <syon>: <bersiyon> vs. <bersyon> na may frequency count difference na 18; <impeksiyon> vs. <impeksyon> sa 33; <interseksiyon> vs. <interseksyon> sa 10; <kumbensiyon> vs. <kumbensyon> sa 12; <leksiyon> vs. <leksyon> sa malaking frequency count difference na 117; at <seksiyon> vs. <seksyon> sa 17. Samantala, nakahango ng isang pares ng salitang naglalaman ng <s(i)yon> na may magkaparehong frequency count: <transaks(i)yon> (na may frequency count na tig-6). Gayunpaman, dahil sa mas maraming sanggunian ang pinag-iiralan ng variant na may <siyon>, isinama ang nasabing pares sa Talahanayan 9.

Mga anyong may <s(i)ya>

Para naman sa mga salitang may <s(i)ya>, nakapagtala mula sa korpus ng 260 salita na may pinakamaliit na frequency count na dalawa. Matapos linisin ang datos sa pamamagitan ng prosesong inilarawan sa itaas, nakahango ng 98 salita. Mula sa bilang na ito, 48 ang naglalaman ng <sya> habang 13 naman ang may <siya>. Samantala, nakatukoy naman ng 27 pares ng mga salitang naglalaman ng parehong <siya> at <sya>.

Ang Talahanayan 10 ay nagpapakita ng mga salitang may <sya> na mas lamang sa frequency count kaysa mga salitang may <siya>.

Talahanayan 10: Pares ng salitang may <s(i)y>, kung saan mas mataas ang frequency ng may <sya>

	Salita	Hinalaw na anyo (Frequency count)	Uri ng mga sanggunian	Kabuuang frequency count	Kabuuang bilang ng sanggunian
1	ebidensiya	ebidensiya (11)	W1A (1), W2A (4), W2C (1), W2E (1) W2F (1)	15	11
		ebidensiyang (4)	W2A (2) W2E (1)		
	ebidensya	ebidensya (22) ebidensyang (1)	W1A (1), W2A (5), W2F (1) W2A (1)	23	8
2	ganansiya	ganansiya (2)	W2A (1) W2F (1)	2	2
	ganansya	ganansya (3)	W1A (1), W2A (2)		
3	hustisiya hustisya	hustisiya (2) hustisya (22)	W1A (1) W2A (5), W2C (4), W2E (1), W2F (3)	26	15
		hustisyang (4)	W2A (1), W2C (1)		
	komersiyal	komersiyal (24)	W2A (8), W2F (4)	25	13
	komersyal	komersyalisasyon (1) komersyal (22)	W2A (1) W1A (2), W2A (10), W2B (1)	28	16
		komersyalisasyon (6)	W1A (1), W2A (1), W2D (1)		
5	konsensiya konsensya	konsensiya (4) konsensya (5)	W2F (2) W2A (2), W2B (1)	4 5	2 3
	6	kontrobersiya	kontrobersiya (10)	W1A (1), W2A (2), W2C (1), W2E (1), W2F (1)	14
kontrobersiyal (4)			W1A (2), W2C (2)		
kontrobersya		kontrobersya (7) kontrobersyal (10)	W1A (1), W2A (3) W2A (5), W2B (1), W2C (3)	17	13

Talahanayan 10: Pares ng salitang may <s(i)y>, kung saan mas mataas ang frequency ng may <sya> (Cont'n.)

	Salita	Hinalaw na anyo (Frequency count)	Uri ng mga sanggunian	Kabuuang frequency count	Kabuuang bilang ng sanggunian
7	pasensiya	pasensiya (12)	W1A (2), W2A (2), W2F (5)	14	11
		mapagpasensiya (2)	W1A (1), W2A (1)		
	pasensya	pasensya (26)	W1A (1), W1B (1), W2A (3), W2B (2), W2C (1), W2F (4)	26	12
8	potensiyal	potensiyal (9)	W2A (5), W2F (2)	9	7
	potensya	potensyal (17)	W1A (5), W2A (4), W2B (1), W2C (1)	17	11
9	pransiya	pransiya (3)	W2A (1)	3	1
	pransya	pransya (7)	W2A (3), W2F (1)	7	4
10	probinsiya	probinsiya (36)	W1A (1), W2A (10), W2C (1), W2F (8)	36	20
	probinsya	probinsya (44)	W1A (2), W2A (13), W2B (2), W2C (1), W2F (3)	44	21
11	tersiyarya	tersiyaryong (1)	W2A (1)	1	1
	tersyarya	tersyarya (20)	W1A (1), W2A (2)	20	3
12	intelihensiya	intelihensiya (2)	W2A (1)	2	1
	intelihensya	intelihensya (2)	W1A (1), W1A (2)	2	3

Sang-ayon sa Talahanayan 10, sa 12 pares na nakalap, tatlo ang kinakitaan ng malaking kalamangan ng frequency count (10 o higit pa) sa mga anyong may <sya> kung ihahambing sa may <siya>: <hustisya> vs. <hustisiya> na may frequency count difference na 24; <pasensya> vs. <pasensiya> sa 12; at <tersyarya> vs. <tersiyarya> sa 19. Ang variant na <ebidensya>, bagaman nakalalamang sa <ebidensiya> sa usapin ng kabuuang frequency count ay lumilitaw sa mas kaunting bilang ng sanggunian (walo, kung ihahambing sa 11 ng <ebidensiya>). Samantala, ang pares ng salitang <intelihensiya> vs.

<intelihensya> ay may tig-dalawang frequency count. Gayunman, inilagay ang pares na ito sa Talahanayan 10 sapagkat ang variant na <intelihensya> ay lumilitaw sa higit na bilang ng sanggunian.

Para naman sa mga pares ng salita kung saan nakalalamang sa frequency count ang may <siya> kaysa <sya>, tingnan ang Talahanayan 11.

Talahanayan 11: Pares ng salitang may <s(i)ya>, kung saan mas mataas ang frequency ng may <siya>

	Salita	Hinalaw na anyo (Frequency count)	Uri ng mga sanggunian	Kabuuang frequency count	Kabuuang bilang ng sanggunian
1	ahensiya	ahensiya (31)	W1A (3), W2A (13), W2C (1), W2D (1), W2E (1), W2F (2)	31	21
	ahensya	ahensya (35)	W1A (3), W2A (7), W2C (1)	35	11
2	alkansiya	alkansiya (3)	W2D (1)	3	1
	alkansya	alkansya (2)	W2F (2)	2	2
3	diperensiya	diperensiya (4)	W2A (1), W2C (1), W2F (1)	4	3
	diperensya	diperensya (1)	W2A (1)	1	1
4	distansiya	distansiya (23)	W2A (11), W2F (1)	23	12
	distansya	distansya (3)	W1A (1), W2A (1), W2F (1)	5	5
		dumistansya (2)	W1A (1), W2A (1)		
5	esensiya	esensiya (22)	W1A (1), W2A (9), W2C (1), W2F (3)	22	14
	esensya	esensya (9)	W1A (1), W2A (2), W2B (1), W2F (1)	9	5
6	impluwensiya	impluwensiya (98)	W1A (5), W2A (22), W2B (1), W2F (7)	98	35
	impluwensya	impluwensya (81)	W1A (3), W2A (22), W2B (1), W2C (2), W2F (2)	81	30

Talahanayan 11: Pares ng salitang may <s(i)ya>, kung saan mas mataas ang frequency ng may <siya> (Cont'n.)

	Salita	Hinalaw na anyo (Frequency count)	Uri ng mga sanggurian	Kabuuang frequency count	Kabuuang bilang ng sanggurian
7	importansiya	importansiya (9)	W1A (1), W2A (3), W2C (1)	9	5
	importansya	importansya (2)	W1A (1), W2A (1)	2	2
8	kumperensiya	kumperensiya (19)	W1A (2), W2A (5), W2C (1)	19	8
	kumperensya	kumperensya (16)	W2A (5), W2C (2), W2F (2)	16	9
9	lisensiya	lisensiya (11)	W2A (2), W2C (1), W2F (1)	11	4
	lisensya	lisensya (6)	W2A (3), W2C (1)	6	4
10	pinansiya	pinansiya (43)	W2A (6), W2B (1), W2D (1), W2F (2)	43	10
	pinansya	pinansya (15)	W1A (2), W2A (5), W2B (1), W2D (1)	15	9
11	presensiya	presensiya (7)	W2A (4), W2F (2)	7	6
	presensya	presensya (3)	W1A (1), W2B (1)	3	2
12	resistensiya	resistensiya (18)	W2A (2)	18	2
	resistensya	resistensya (5)	W2A (2), W2B (1)	5	3
13	siyensiya	siyensiya (14)	W2A (9), W2F (1)	14	10
	siyensya	siyensya (6)	W2A (3), W2C (1), W2F (1)	6	5
14	sustansiya	sustansiya (23)	W1A (1), W2A (2), W2D (1), W2F (3)	23	8
	sustansya	sustansya (9)	W2B (1), W2D (1), W2F (3)	9	5
15	tendensiya	tendensiya (15)	W2A (4), W2C (1)	15	5
	tendensya	tendensya (5)	W2A (3)	5	3

Labinlimang (15) pares ang may anyong may <siya> na nakalalamang pagdating sa frequency count kung ihahambing sa may <sya>, gaya ng makikita sa Talahanayan 11 sa itaas. Pito sa mga ito ay may mataas na frequency count difference (10 o higit pa), gaya ng mga sumusunod: <distansiya> vs. <distansya> na may frequency count difference na 19; <impluwensiya> vs. <impluwensya> sa 15; <kumperensiya> vs. <kumperensya> sa 18; <pinansiya> vs. <pinansya> sa 26; <resistensiya> vs. <resistensya> sa 12; <sustansiya> vs. <sustansya> sa 13; at <tendensiya> vs. <tendensya> sa 12.

Ang <s(i)yon> / <s(i)ya> sa korpus vs. Manwal ng KWF

Litaw ang malaking pagkakaiba sa bilang ng mga salitang naglalaman ng <syon> (340 salita) at <siyon> (17 salita) sa nakalap na datos na 394 na salita. Gayunpaman, kapansin-pansin na ang baybay (partikular na sa bahagi ng <siyon>) ng lahat ng 17 salitang ito ay nakaayon sa pamantayang ipinasusunod ng KWF. Sa 340 salitang nagtataglay ng <syon>, 12 ang dapat ay nasa anyong <siyon>: <detensyon> (na may frequency count na 2), <insureksyong> (2), <kapsyon> (2), <konskripsyon> (2), <konsumpsyon> (3), <opsyon> (4), <prebensyon> (4), <prediksiyon> (7), <reduksyon> (2), <repleksyon(g)> (25), <seduksyon> (2), at <subersyon> (3). Gayundin, kung sisipating muli ang talahanayan sa itaas, ang lahat ng 21 pares kung saan nakalalamang sa frequency count ang salitang may <siyon> kaysa katambal nitong may <syon> ay nakaayon din sa pamantayan ng KWF. Sa kabilang banda, ang lahat ng 16 na pares ng salita kung saan nakalalamang sa frequency count ang may <syon> ay tumataliwas sa pamantayang ipinatutupad ng KWF.

Nasa parehong kalagayan makikita ang datos na nakatuon sa mga salitang may <s(i)ya>. Malaki ang kalamangan ng mga salitang may <sya> (48 salita) kung ihahambing sa may <siya> (13 salita), mula sa 98 anyong nakalap mula sa teksto. Gayunpaman, kagaya ng datos hinggil sa <siyon> na itinanghal sa itaas, lahat ng 13 salitang may <siya> ay nakaayon sa tuntunin ng KWF sa pagbaybay. Samantala, sa 48 salitang nagtataglay ng <sya>, 9 ang dapat ay nasa anyong <siya>, kung pagbabatayan ang KWF: <ambulansya> (na may frequency count na 6), <eksistensyal> (3), <independensya> (4), <instansya> (12), <preperensya> (18), <putsya> (2), <sirkumstansya> (4), <substansya> (4), at <worksyp> (4). Kung susuriin naman ang 15 pares ng salitang kapwa nagtataglay ng <siya> at <sya> kung saan nakalalamang ang may <siya>, iisa lamang ang umaayon sa pamantayan ng KWF, ang <hustisiya> (na may frequency count na 28) vs. <hustisiya> (na may 2).

Maliban sa bilang ng mga token, isinasalang-alang din sa pagsusuri ng datos mula sa korpus ng pasulat na teksto ang bilang ng sanggunian na pinanggagalingan ng mga variant, lalo na sa mga anyong may parehong variant (kapwa may <i> at

wala), sa pagtukoy ng kung aling variant ang mas malaganap ang gamit sa kontemporaryong Filipino. Gayunpaman, kinakitaan din naman ng anyo sang-ayon sa mga talahanayang nakalatag sa itaas na kung aling anyo ang nakakuha ng mas mataas na frequency count, ay siya ring matatagpuan sa mas maraming bilang ng sanggunian ng pasulat na teksto. Maging ang pagkaiba ng sangguniang teksto ay makikitang mas mataas ang bilang kung kaya hindi pa ito itinuturing na suliranin sa pagtukoy ng pangunahing entry sa binubuong diksiyonaryo.

Konklusyon

Sa pagtatapos ng proyektong *UP MDF*, inaasahang makabuo ng isang diksiyonaryong naglalaman ng 20,000 entry na hinango mula sa 8,000,000 salita ng korpus ng wikang Filipino. Sa paunang pagsipat sa kasalukuyang nakalap na korpus ng pasalita at pasulat na teksto, ang pinakamaliit na frequency count na kailangan upang ang isang salita ay maisaalang-alang bilang entry ng binubuong diksiyonaryo ay apat. Kung titingnan ang datos na itinanghal sa itaas, marami-rami ang mga salitang malaki ang pagkakataong mapabilang sa diksiyonaryo, kalakip ang mga hinalaw na anyo at ang kahulugang kaakibat ng mga ito. Kung gayon, mahalagang salik ang paglutas sa mga usaping may kinalaman sa ortograpiya nang sa gayon ay ganap na mailarawan ang lexicon at gramatika na kasalukuyang gamit ng wikang pambansa.

Sa paglutas sa mga usaping ito, marapat na isipin kung ano ang layunin na nais mapagtagumpayan ng binubuong diksiyonaryo. Ayon kay Sidney Landau, ang diksiyonaryo ay paglalarawan ng isang wika sa isang partikular na panahon sa kasaysayan ng pag-unlad nito (1984). Sinususugan ito ni Leonard Newell sa pagsasabing ang diksiyonaryo ay sumasalamin sa buhay at nagbabagong wika na ginagamit sa pasalita at pasulat na teksto (1995). Kung gayon, tungkulin ng tagakalap ng wika na suriin at sinupin ang ganitong pamamaraan ng paglalarawan sa wika at hindi kung ano ang dapat na maging katangian nito. Ibig sabihin, ang mga entry na bumubuo sa leksikograpikong akda ay sumasalamin sa kung ano ang gamit-gamit ng mga nagsasalita sa isang partikular na panahon, sa kaso ng *UP MDF*, ang kasalukuyan. Ang diksiyonaryo (at alinmang akdang panlingguwistika sa puntong ito) ay hindi nililikha upang gawing pamantayan o maykapangyarihan sa isang partikular na wika, bagaman ito ang kadalasang pagtuturing sa diksiyonaryo. Sa ganang ito, ang ganitong uri ng akda ay paglalarawan ng wika batay sa kung paano ito ginagamit ng mga tagapagsalita na siyang tunay na nagtatakda ng hugis at anyo nito. Kung gayon, marapat lamang na ilangkap ang lahat ng variant sa pagbaybay ng mga salita, lalo na ang may kinalaman sa hiniram na anyo at kung paano iniaakma ang pagbaybay sa mga ito na mahihinuha naman sa sistematiko at masinop na pangangalap ng korpus ng inilimbag na pasulat na teksto.

Mahalagang isaalang-alang ang mga inihahaing *prescriptive* na alituntunin na ipinatutupad ng mga ahensiyang pangwika, upang mapanatili ang pagkakapareho hindi lamang sa pagbaybay, kundi maging sa paggamit ng wika sa kabuuan. Ang mga alituntuning ito ang nagsisilbing gabay sa pagtuturo ng wika sa kabataan at mga taong naghahangad na matutunan ang wika, bagay na hindi lubusang tinutugunan ng *descriptive* na lapit na pag-aaral kagaya ng lenteng ginamit sa papel na ito. Hindi nagtatalaban ang mga alituntuning descriptive at prescriptive sa pagsasalarawan ng mga aspekto ng wika; bagkus, ang mga descriptive na alituntunin ay nagpapahayag ng kasalukuyang umiiral na anyo lamang samantalang ang tiyak na alituntuning prescriptive ay may kinalaman sa kaangkupan ng gamit ng iba't ibang variant ng mga salita.

Tinatanaw na mananatili at magpapatuloy pa ang talakayan hinggil sa ortograpiya ng wikang pambansa, partikular na ang mga bagong salitang ipinapakilala ng nagbabagong panahon at pamalagiang pakikipag-ugnayan sa labas ng Pilipinas. Bunga nito, hindi dapat isantabi ang tuntuning isinusulong ng mga ahensiyang pangwika, gaya ng KWF na naglalayong gawing estandardisado ang sistema ng ortograpiya ng pambansang lingua franca upang higit na mapatingkad ang kakayanan nitong maisalin ang pambansang wika sa susunod na salinlahi sa pamamagitan ng pedagohiya, palimbagan, at pambansang diskurso.

Pasasalamat

Pasasalamat sa suportang pinansiyal ng Opisina ng Bise-Tsanselor sa Pananaliksik at Pagpapaunlad (OVCRD), Unibersidad ng Pilipinas Diliman, upang maisakatuparan ang proyektong *UP Monolinggwal na Diksiyonaryong Filipino*, korpus ng wikang Filipino, at ang pananaliksik na ito.

Sanggunian

- Algeo, John. 1978. "What consonant clusters are possible?" *Word* 29, no. 3: 206-224. DOI: 10.1080/00437956.1978.11435661
- Almario, Virgilio S., ed. 2001. *UP Diksiyonaryong Filipino*. Lunsod ng Pasig: Sentro ng Wikang Filipino at Anvil Publishing, Inc.
- Almario, Virgilio S., ed. 2010. *UP Diksiyonaryong Filipino* (binagong edisyon). Lunsod ng Quezon: UP Sentro ng Wikang Filipino.
- Almario, Virgilio S., ed. 2014. *KWF Manwal sa masinop na pagsulat*. Lunsod ng Quezon: Komisyon sa Wikang Filipino.
- Atkins, Sue, Jeremy Clear, at Nicholas Ostler. 1992. "Corpus design criteria." *Literary and Linguistic Computing* 7, no. 1: 1-16.
- Baklanova, Ekaterina. 2017. "Types of borrowings in Tagalog/Filipino." *Kritika Kultura* 28: 35-54.

- Guillermo, Ramon G., Rhandley D. Cajote, at Aristeo Logronio. 2015. "UP Diksiyunaryong Filipino: Sinonismong walang hangganan." *Daluyan*, no. 2: 54-71.
- Kagawaran ng Edukasyon. 2006. *Kautusang Pangkagawaran Blg. 42 s. 2006: Pag-review ng "2001 Revisyon ng alfabeto at patnubay sa ispelng ng wikang Filipino"*. Maynila: Komisyon sa Wikang Filipino. 1987. *Alpabeto at patnubay sa ispelng*. Maynila: Komisyon sa Wikang Filipino.
- Komisyon sa Wikang Filipino. 2001. *Revisyon ng alpabeto at patnubay sa ispelng ng Wikang Filipino*. Maynila: Komisyon sa Wikang Filipino.
- Landau, Sidney I. 1984. *Dictionaries: The art and craft of lexicography*. New York: The Scribner Press, Charles Scribner's Sons.
- Lee, Sangsup. 1992. "The Yonsei corpora of Korean and Lexicographic projects." EURALEX '92 Proceedings I-II: Papers submitted to the 5th EURALEX International Conference on Lexicography in Tampere, Finland, *Studia Translatologica*, Ser. A, Vol. 2, Tampere.
- Nelson, Gerald. 1996. "The design of the corpus." *Nasa Comparing English worldwide: The international corpus of English*, Sidney Greenbaum, ed., 27-35. Oxford: Oxford University Press.
- Newell, Leonard. 1995. *Handbook on lexicography for Philippine and other languages: With illustrations form Batad Ifugao dictionary*. Manila: Linguistic Society of the Philippines.
- Surian ng Wikang Pambansa. 1977. *Mga tuntunin ng ortograpiyang Pilipino/Patnubay sa pagwawasto ng mga aklat babasahin*. Manila: Government Printing Office.
- Zafra, Galileo, et al., eds. 2005. *Gabay sa editing sa wikang Filipino (Tuon sa pagbaybay)*. Sentro ng Wikang Filipino, Unibersidad ng Pilipinas Diliman.
- Zafra, Galileo, et al., eds. 2008. *Gabay sa ispelng*. Sentro ng Wikang Filipino, Unibersidad ng Pilipinas Diliman.

TALA TUNGKOL SA MAY-AKDA

Si Jem R. Javier ay katuwang na propesor sa Departamento ng Linggwistiks, UP Diliman. Siya ay kasalukuyang mag-aaral sa PhD Linggwistiks sa nasabing institusyon. Siya ay naging bahagi ng iba't ibang pananaliksik sa wika at kultura ng Pilipinas, gaya ng pagbuo ng UP Monolingwal na Diksiyonaryong Filipino. Maaaring makipag-ugnayan sa kanya sa pamamagitan ng email address: jem.javier@up.edu.ph

ABOUT THE AUTHOR

Jem R. Javier is Assistant Professor at the Department of Linguistics, University of the Philippines, Diliman, where he is also pursuing a PhD in linguistics. He is engaged in various research endeavors on Philippine languages and cultures, such as the UP Monolingwal na Diksiyonaryong Filipino project. He may be reached at: jem.javier@up.edu.ph