



Pragmatikong Analisis sa AI Image-to-Text Generation: Ang Kaso ng ChatGPT, Grok, at Gemini

Pragmatic Analysis of AI Image-to-Text Generation: The Case of ChatGPT, Grok, and Gemini

Ariel Ursolon Bosque 
University of the Philippines Diliman
aubosque@up.edu.ph

DOI: <https://doi.org/10.70922/x2sdfx87>

[Submit to this journal](#) 

[Other related articles](#) 

[Announcements](#)

[Other PUP Journals](#) 

Article History:

Date received: July 20, 2025

Date accepted: January 7, 2026

Other citation formats:

How to Cite this paper:

Bosque, A. U. (2026). Pragmatikong Analisis sa AI Image-to-Text Generation: Ang Kaso ng ChatGPT, Grok, at Gemini. *Social Sciences and Development Review*, 18(1), pp. 277-298.

<https://doi.org/10.70922/x2sdfx87>.

Please contact the publisher for any further use of this work at
socialsciencesreview@pup.edu.ph.

Pragmatikong Analisis sa AI Image-to-Text Generation: Ang Kaso ng ChatGPT, Grok, at Gemini

Pragmatic Analysis of AI Image-to-Text Generation: The Case of ChatGPT, Grok, and Gemini

Ariel Ursolon Bosque 

University of the Philippines Diliman

aubosque@up.edu.ph

ABSTRAK

Layunin ng pag-aaral ang pagsasagawa ng komparatibong pagsusuri sa *ChatGPT*, *Grok*, at *Gemini* upang mataya ang mga kapasidad nito sa *image-to-text generation*. Gumamit ang mananaliksik ng tatlong *AI-generated image prompts* mula sa *Canva*, at isinalang sa mga nabanggit na *generative AI (GenAI)*. Sa tulong ng isang *text prompt*, nabigyang-panuto ang *GenAI* na makabuo ng labingwalong pangungusap na makapaglalarawan sa *image prompts*. Sinuri sa pag-aaral ang kaakmaang kontekstuwal at pragmatikong katugmaan, at mga halusinasyon at pragmatikong kamalian ng mga nabuong pangungusap sa tulong ng pragmatikong pagsusuri. Natuklasan sa pagsusuri ang paggamit ng angkop na referents, palatandaang modal, mga deskriptibong detalye upang angkop at pragmatikong mailarawan ang mga larawan. Sa likod nito, may lumutang pa ring ilang lisyang estruktural at kalabuang leksikal. Bagaman nagpakita ang mga awtput ng lohika sa tulong ng rekognisyon ng mga padron ng mga salita, masasalamatin pa rin dito ang labis na pagkatig ng mga modelo sa probabilitikong asosasyon kaysa totoong semantikong pag-unawa. Masusuri naman ang limitasyon ng *GenAI* sa mga nalikha nitong deskripsiyong nagtataglay ng halusinasyon, labis na inferens, at sintaktikong kamalian, lalo sa mga pagkakataong wala o di-sapat ang biswal na ebidensiya sa mga ipinahihiwatig na ideya. Sa kabuoan, nangibabaw ang *Gemini* sa pagbibigay ng tumpak at makabuluhang deskripsiyong nagtataglay ng mataas na kontekstuwal at pragmatikong katumpakan, sunod naman ang *ChatGPT* na may mga nabuong pangungusap na makahulugang lisyang. Samantala, *Grok* naman ang nagtataglay ng pinakamaraming awtput na halusinado at may pragmatikong kamalian.

Mga Susing Salita: *generative artificial intelligence, inferentiality, I2T generation, pragmatikong pagsusuri; referentiality*

ABSTRACT

The study aims to conduct a comparative analysis of *ChatGPT*, *Grok*, and *Gemini* to evaluate their capacities in *image-to-text generation*. The researcher utilized three *AI-generated image prompts* from *Canva*, which were processed through the mentioned generative AI (*GenAI*) models. With the aid of a standardized text *prompt*, each model was instructed to produce eighteen descriptive sentences corresponding to the *prompts*. The study examined the contextual alignment, pragmatic appropriateness, as well as hallucinations and pragmatic errors present in the generated sentences through pragmatic analysis. The analysis revealed the use of appropriate referents, modal markers, and descriptive details that contributed to contextually and pragmatically accurate descriptions of the images. Despite this, several structural inconsistencies and lexical ambiguities surfaced. While the outputs demonstrated logical coherence through the recognition of lexical patterns, the findings indicate that the models heavily rely on probabilistic associations rather than genuine semantic understanding. The limitations of *GenAI* became evident in descriptions that contained hallucinations, excessive inference, and syntactic errors, particularly in cases where visual evidence was absent or insufficient to support the inferred *content*. Overall, *Gemini* outperformed the other models in producing accurate and meaningful descriptions with high contextual and pragmatic precision, followed by *ChatGPT*, whose outputs showed some structural flaws, *Grok* produced the highest number of hallucinated outputs and pragmatic errors.

Keywords: *generative artificial intelligence, inferentiality, I2T generation, pragmatic analysis, referentiality*

INTRODUKSIYON

Kasama ang pagpoproseso ng iba't ibang risors at kakayahan nitong komprehensibong makapagsuri ng *multimodal* na datos sa mga kapasidad na patuloy pang ineensayo sa *Generative Artificial Intelligence (GenAI)* sa kasalukuyan, tulad ng mga ipinamamalas ng *ChatGPT*, *Grok*, at *Gemini*. Bentaha ng *Grok* ang integrasyon nito sa X (Twitter), samantalang nangingibabaw naman sa *Gemini* ang *multimodalidad* ng modelo at pagbibigay ng pleksibleng dulong sa iba't ibang tipo ng datos. Malawakan namang nagagamit ang *ChatGPT* sa pagbibigay-solusyon sa samot-saring suliranin, kaya mayroon itong malakas na presensiya bilang

isang conversational AI (Souza at Weigang, 2025). Sa likod ng mga nabanggit na kapasidad, matibay ang argumento nina Evkaya at de Carvalho (2024) na mahalagang dumaan sa pagkikritika at pagsusuri ng tao ang anumang datos na hango sa AI-powered software. Gaya rin ng kadalasang disclaimer na nakalakip sa automated *content* ng large language models na produkto ng ini-input na *prompts*, ipinaaalala nilang may potensiyal na makalikha ang *GenAI* ng mga hallucinated, biased, at okasyonal na maling nilalaman (Evkaya & de Carvalho 2024). Pinahahalagahan ng ganitong sitwasyon ang presensiya ng masuri at nakasentro-sa-taong (human-centered) paggamit nito, kasama ang kritikal na pagtanggap at pagpili ng impormasyong magmumula sa *AI-generated responses*.

Kung sa usapin ng pagpoproseso ng *GenAI* ng *multimodal* na risors, mas karaniwang nagagamit ang text-to-image generation (T2IGen) (i.e., pagbuo ng *GenAI* ng mga larawan batay sa mga pasulat na deskripsiyon). Bumabase ang *GenAI* sa mga ini-input na *prompt* ng user. Isinasalin nito ang mga salita patungong larawan sa pamamagitan ng mga padrong natututuhan at ineensayo mula sa milyon-milyong halimbawang larawan at teksto.

Kadalasan, walang kasiguraduhan ang kalalabasan ng T2IGen, halimbawa, makakakita ang user sa *AI-generated* pictures ng mga: anatomikong kamalian o kakulangan; di-natural na elemento at detalye; di-mabasang teksto; imposibleng bagay o tanawin, at iba pang kakaibang bagay na masisipat naman kung susuriing mabuti ang larawan. Pinatutunayan nitong mahirap na ganap na makontrol ang uri ng larawang malilikha nito kahit na may ‘*prompting*’ o ‘*prompt engineering*’ (Oppenlaender et al., 2024). Itinuturing ang *prompting* bilang iteratibo at interaktibong diyologo sa pagitan ng tao-bilang-user at AI (Oppenlaender, 2022) upang mabuo ang ipinasasadyang hitsura ng mga nais likhaing larawan ng *GenAI*. Gayumpaman, hindi rin maisasantabi ang naobserbahang pagsisikap ng *content developers* upang maitugma, magaya, at maisaayos ang mga *AI-generated content*.

Kaugnay ng mga limitasyong ito sa kontrol ng awtput, lumalawak din ang mga pananaliksik sa iba’t ibang aspekto ng T2IGen upang higit na maunawaan, mapaunlad, at masukat ang kakayahan ng mga sistemang ito. May mga pag-aaral na nakatuon sa: pagtataya ng T2IGen (Chen et al., 2023); pagpapaunlad, pag-aarkitektura, at pagsasanay ng mga T2I na modelo (Wang & Zhang, 2025); at *prompt engineering* at automizer ng mga algoritmo (He et al. 2024). Masusuri rito ang pagtalakay ng mga iskolar sa antas-sistema at antas-user na mga hamon at oportunidad ng T2IGen. Nananatiling aktibo ang pananaliksik sa image-to-text (I2T), bagaman madalas na nakapaloob na lamang ang mga gawaing I2T, gaya ng pagka-caption ng mga larawan, paglalagay ng alternative text, o pagsagot sa mga biswal na tanong, sa mas malalawak na sistema (e.g., *multimodal* chatbots o visual reasoning models).

Sa Pilipinas, interdisiplinaryo ang dulog sa pag-aaral ng punsiyon at gampanin ng mga kagamitan at platapormang AI. Mas marami ring saliksik ang nakasulat sa konteksto ng

wikang Ingles, gaya ng pag-aaral sa kakayahan ng AI sa pagsapol ng esensiya ng wika (Estrella, 2023), at kontrastibong pagsusuri sa mga naisulat ng tao at AI (Briana, 2024). May ilan mang nakasulat sa konteksto ng wikang Filipino, ngunit nakasulat pa rin sa wikang banyaga o mas nakatutok sa teknikalidad ng AI, tulad ng pagsusuri sa speech recognition sa pasalitang Filipino (Corpus at Villanueva, 2024); mga kaakibat na usapin sa pagsasalin at AI (Endriga at Rosario, 2023; Liwanag et al., 2024); at kagamitang ginagamit sa pagwawasto ng gramatika (Montalan et al., 2025). Kabilang naman sa mga pag-aaral na nakasulat at nakatuon sa konteksto ng wikang Filipino ang mga paksang may kinalaman sa pagsasalin gamit ang mga kagamitang AI (Andrada, 2021; Bosque, 2024) at integrasyon sa pagtuturo at edukasyon. Sa likod ng mga inisyatiba, masusuring hindi pa napagtutuunang-pansin ang pananaliksik sa training capacity ng AI para sa I2T Generation.

Kapansin-pansin ding hindi pa nasisiyasat ang pragmatikong interpretasyon ng I2T *AI-generated content*. Kritikal na komponent ang mga nabanggit sa pag-unawa sa pagbuo ng kahulugan ng mga *AI-generated content*. Pinahihintulutan ng kontekstuwal na inferens ang pagtukoy at pagpoproseso ng AI nang higit pa sa literal na rekognisyon, kasama ang pagbibigay-kahulugan sa resepsiyon sa emosyon, intensiyon, o anumang kultural na reperensiyang lumilitaw sa image *prompt* at hindi bunga ng anumang halusinasyon. Mahalaga ito para sa pagbuo ng malagawa-ng-tao, tumpak, at angkop na deskripsiyon.

Sa kasong ito, hindi maaaring mawala ang pagsapol sa mga kultural, kontekstuwal, at pragmatikong kaibahan at sensitibidad upang wastong matukoy ang pakay na mensahe (Naveen at Trojovsky, 2024). Mahalagang maunawaan ang mga pragmatikong estratehiyang ginagamit ng *GenAI* upang malinang ang pagpapaliwanag, lingguwistikong kakayahan, at pagpapaunlad ng modelo. Nagbibigay ng dagdag na paliwanag sa kapasidad ng *GenAI* ang pragmatikong pagsusuri lalo na sa *multimodal* na gawain tulad ng *image-to-text generation* (I2TGen). Mababatid ditong bagaman may patuloy pang pag-unlad sa *GenAI*, kakaunti pa lamang ang mga paniniyasat sa pagkakaiba-iba ng mga modelo sa usapin ng mga pragmatikong estratehiya sa paglikha ng mga tekstong nakabatay sa mga larawan.

LAYUNIN NG PAG-AARAL

Mula sa mga nabanggit, pinakalayunin ng pag-aaral ang pagsasagawa ng komparatibong pagsusuri upang mataya ang kapasidad ng *generative artificial intelligence platforms* para sa *image-to-text generation*. Pinatitingkad nito ang pagkokompara sa kakayahan ng *ChatGPT*, *Grok*, at *Gemini* na makapagsagawa ng pragmatikong interpretasyon sa *image-to-text generation*. Dahil dito, sinuri sa pag-aaral ang kaakmaang kontekstuwal at

pragmatikong katugmaan, at mga halusinasyon at pragmatikong kamalian ng mga nabuong pangungusap.

METODOLOHIYA

Disenyo ng Pag-aaral

Kuwalitatibo ang kalikasan ng pag-aaral dahil sa layon nitong maimbestigahan ang kapasidad ng *ChatGPT*, *Grok*, at *Gemini* sa *image-to-text generation* o pagbuo ng mga nakapangungusap na deskripsiyon mula sa *image prompts*. Inilapat ang kuwalitatibong dulong upang masuri sa pag-aaral ang mga pragmatikong padrong inilalapat ng mga tampok na generative AI.

Mga Materyal at Proseso ng Pag-aaral

Ginamit ng mananaliksik ang libreng AI-powered text-to-image app ng *Canva* upang makalikha ng *image prompts*. Narito ang *prompts* na in-input para sa pag-generate ng mga larawan:

- a. *An ink drawing of a child reading a book under a tree;*
- b. *An ink drawing of a dog chasing a chicken; at*
- c. *An ink drawing of children picking mangoes from a tree.*

Orihinal na nakasulat sa wikang Filipino ang text *prompts* na inilagay sa *CanvaAI* ngunit hindi nito masapol ang inaasahang larawan kaya isinalin sa wikang Ingles. Sa artikulo ni Villemaire (2025), binanggit niyang walang anumang pag-aangkin sa karapatang-ari ang *Canva* sa mga larawang malilikha sa *CanvaAI*. Gayumpaman, hindi rin nangangahulugang awtomatikong maituturing na may-ari ng mga larawan ang user, hindi rin kikilalanin ang anumang ganap na akwisisyon ng eksklusibong karapatang-ari sa mga ito. Ipinapaalala rin ng kompanyang mahalaga ang pagbabanggit na hindi orihinal na likha ng tao ang mga larawan, sa ganitong gana, ituring na rin ito bilang deklarasyon ng mananaliksik sa paggamit ng *AI-generated* images sa kasalukuyang pag-aaral bilang *image prompts*. Narito ang mga larawang ginamit:

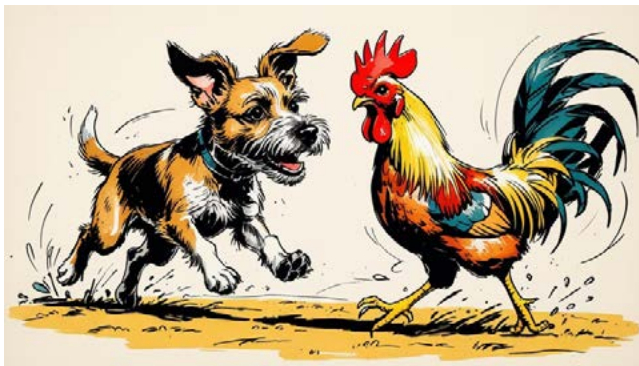
Pigura 1

Larawan ng batang nagbabasa ng aklat sa ilalim ng puno



Pigura 2

Larawan ng manok na hinahabol ng aso



Pigura 3.

Larawan ng mga batang masayang namimitas ng mga prutas



Napili ng mananaliksik ang *ChatGPT*, *Grok*, at *Gemini* dahil sa matatag nitong *multimodal* na kakayahan, malinaw na pundasyong arkitektural, at aksesibilidad. Nagagawa ng mga nabanggit na masuportahan ang I2TGen sa pamamagitan ng pagpoproseso ng *image prompts* at produksiyon ng deskriptibong tekstuwal na awtput na mahalaga sa pagsusuri ng pragmatikong interpretasyon. Kinakatawan ng bawat modelo ang naiibang arkitektural na dulong. Kilala ang *ChatGPT* sa konstruksiyon nito ng tekstuwal na datos at malinaw na pagpapangungusap. Nakapagbibigay naman ang *Grok* ng malikhaing awtput na naiimpluwensiyahan ng singkronikong datos; at matingkad naman sa *Gemini* ang estrukturalisadong tugon sa tulong ng integrasyon nito sa iba't ibang web-based knowledge system.

Ginamit sa pag-aaral ang mga sumusunod na bersiyon: *OpenAI* o 4-mini (Abril 2025 ang huling update) para sa *ChatGPT*, *Grok* 3 (Pebrero 2025) para sa *Grok*, at *Gemini* 2.5 (Marso 2025) para sa *Gemini*. Nang makaiwas sa personalisadong bias at iba pang suliraning dulot ng taglay na memorya ng isang account na maiuuri bilang matinding epekto ng teknolohikong kustomisasyon (Warr, 2024), lumikha ang mananaliksik ng panibagong accounts na para lamang sa layunin at paggamit sa kasalukuyang pag-aaral. Ginamit naman ang *prompt* na: “Bumuo ng tigdalawang pangungusap sa Filipino batay sa mga larawang ilalakip sa *prompt*,” para sa I2TGen.

Pagsusuri sa mga Datos

Sininop ng mananaliksik ang labingwalong *AI-generated* na mga pangungusap at nilagyan ng leyenda upang mas maging sistematiko at malinaw ang pag-uulat ng resulta. Sa pagsusuri ng datos, inilapat ang pragmatikong analisis ng mga *AI-generated content*, saklaw nito ang lohika ng mga modelo sa pagbuo ng mga pangungusap. Ayon kay Duffy (2008), ginagamit sa pragmatikong analisis ang set ng mga lingguwistiko at lohikong kagamitan upang matukoy ang ganap na lawak ng mga inferens na mabubuo sa teksto. Inilapat ito sa pag-aaral upang masuri kung paano nakabubuo ng kahulugan ang mga pangungusap batay sa biswal nitong konteksto, ipinahihiwatig nitong pakay, at koherens na ipinamamalas ng diskurso. Mula sa mga estratehiyang pragmatiko, tumutok ang pagsusuri sa *reference tracking*, modalidad, deixis, inferens, at kaangkupang kontekstuwal, may ilang bahagi rin ng pagsusuri ang umugat sa lingguwistiko-estruktural na katangian ng mga pangungusap. Nakatulong ang mga nabanggit upang magkaroon ng sistematikong ebalwasyon kung angkop at nasusuportahan ng mga available na biswal na ebidensiya ang mga *AI-generated* na deskripsiyon.

Sa pag-uulat, gumamit ang mananaliksik ng koda upang tukuyin ang pagkakasunod-sunod ng mga larawan para sa pagkakahanay ng mga tema:

Talahanayan 1

Pagkokoda

Leyenda	Mga Koda	Kategorya
Mga Larawan	1	Batang nagbabasa ng libro sa ilalim ng puno
	2	Manok na hinahabol ng aso
	3	Batang namimitas ng prutas
Kagamitang GenAI	a	ChatGPT
	b	Grok
	c	Gemini
Mga Pangungusap	x	Pangungusap 1
	y	Pangungusap 2

Halimbawa, kung tinutukoy sa talakayan ang ikalawang pangungusap na mula sa unang larawang na-generate ng *Grok*, mabibigyan ito ng kowd na gaya nito: 1By.

RESULTA AT TALAKAYAN

Ugnayan ng Kalagayang Lingguwistiko-Estruktural at Kalinawang Pragmatiko

Natuklasan ang mga sumusunod mula sa pagtatalaban ng mga larawan at tekstuwal na deskripsiyong nabuo mula sa I2TGen:

(1bx) Ang bata ay payapaang [*sic*] nagbabasa ng aklat sa ilalim ng puno, na puno ng sariwang dahon.

Maliban sa maling baybay sa salitang ‘payapa-ng,’ taglay rin ng (1bx) ang *right-branching construction* (RBC) na karaniwang makikita sa wikang Tagalog na nagpapamalas ng head-initial na estruktura ng sugnay (Reid at Liao, 2004; Nagaya, 2007), bagaman may ilang kaso rin ng mixed-headedness lalo na sa mga komplikadong konstruksiyon ng pangungusap. Nailagay ang ‘na punô ng sariwang dahon’ (akin ang pagtutuldik, walang tuldikang na-generate ang *GenAI*) upang ilarawan ang salitang ‘puno’ bilang nominal head nito, kaakibat nito ang sintaktikong ambigwiti. Kung higit pang palilinawin ang pangungusap, maimumungkahi ang pagbalangkas dito bilang ‘pûnong punô ng sariwang dahon’ (ginamit lamang ang pagtutuldik para sa pabigkas na pagsusuri sa heteronym na ‘puno’) gayumpaman, sa paglalapat ng pang-angkop na ‘-ng,’ maaaring magkaroon ng kalituhang ortograpiko at semantikong nakakaapekto sa sintaks nito, sa halip na ‘punong puno’ (i.e., ‘a tree full of...’) maaari itong basahin bilang

‘punong(-)puno’ (i.e., ‘full of...’). Alternatibo rin dito ang paggamit ng ‘hitik’ bilang deskriptor, gaya ng: ‘púnong hitik sa sariwang dahon.’

(1cx) Ang batang lalaki ay masayang nagbabasa ng libro sa ilalim ng malaking puno.

Direktang inilalarawan ng (1cx) ang aktuwal na larawan. Wala itong anumang pagdaragdag o pagbabawas ng impormasyong lantad sa biswal. Bagaman may emosyong inilatag dito, hindi naman ito maituturing bilang halusinasyon pagkat mababakas naman ito, tao man ang magbigay ng deskripsiyon.

Malinaw na pamantayan para sa pagtukoy ng halusinasyon ang mga ipinakikilalang entidad, intensiyon, emosyon, o abstraktong relasyong hindi direktang naoobserbahan o maipagpapalagay mula sa image *prompt*. Nakatutugon ang ganitong operasyonalisasyon sa isinasaad ng ilang kontemporaneong saliksik ukol sa *multimodal* AI, matingkad sa mga pag-aaralang pagpapakahulugan sa halusinasyon bilang isang *content* na hindi nasusuportahan ng input data o visual grounding (Ji et al., 2024; Lim & Shim, 2024). Dahil dito, maikakategorya ang anumang pahayag na lampas sa ebidensiyal na saklaw ng isang larawan bilang halusinasyon.

Malinaw sa (1bx) at (1cx) ang mga elemento gaya ng presensiya ng batang lalaking nagbabasa sa ilalim ng punong mabunga at madahon. Tumutugma ang paggamit ng mga pariralang pandiwang ‘masayang nagbabasa’ at ‘payapaang [sic] nagbabasa,’ at pariralang pangngalang ‘malaking puno’ bilang tekstuwal na representasyon ng larawan. Angkop ang deskripsiyon ng mga tampok na pangungusap sa image *prompt* (IP) na ginamit, pagkat lantad dito ang mga pangunahing elementong makikita rin sa larawan, tulad ng bata, aklat, at malaking puno.

Dagdag pa, kapuna-puna rin ang akto ng pagbabasa at kapansin-pansin ang paggamit ng tatlong pangungusap ng pandiwang imperpektibong ‘nagbabasa’ na nagpapahiwatig na may nangyayaring kilos sa temporal na pagbalangkas ng ‘kasalukuyang panahon’ ng larawan. Masusuri rin sa paggamit ng pariralang pangngalang ‘batang lalaki’ ang pagtitiyak sa karakteristiko ng paksa bilang ‘bata’ at ‘lalaki,’ lumilitaw rito ang kamalayan ng *ChatGPT* sa paglalarawan ng aktor. Mababatid din sa paggamit nito ng mga pang-abay na pamaraang ‘masaya’ at ‘payapaa-ng’ [sic] ang pagpapalawig pa sa mental na depiksiyong naluluguran at walang gumagambala sa pagbabasa ng bata. Maituturing namang spatial deixis ang paggamit ng lokatibong pariralang pang-abay na ‘sa ilalim ng puno’ upang ipahiwatig na may lugar na sangkot sa paglalarawan. Nakatutulong din ang deskripsiyon sa malaking puno at sa mga mayayabong at sariwang dahon nito upang higit pang mailatag ang angkop na konteksto.

(2ax) Masiglang hinahabol ng aso ang tandang sa gitna ng bakuran.

(2cx) Ang aso ay tila hinahabol ang tandang sa bakuran.

Bagaman hindi naman kailangang banggitin sa set ng mga pangungusap sa Tema 2 ang ispesipikong katawagan, prominente sa mayorya ng mga pangungusap sa alinmang kategorya ang paggamit ng ‘tandang’ sa halip na ang katapat nitong *gendered co-hyponym* na ‘inahin’ o ang mas generic na katawagang ‘manok.’ *Hyponym* ang terminong ginagamit upang para sa mga tiyak na salitang maaaring mapabilang sa mas malalawak na kategorya. Tinatawag ito bilang hypernym (Nordquist, 2024). Maituturing ang kaso ng (2ax) at (2cx) bilang mga halimbawa ng *lexical markedness* na nangyayari sa *hyponyms*, gaya ng ‘tandang’ at ‘inahin.’ Maaaring nag-uugat ang naturang *markedness* mula sa *gender information* (Mignot, 2012) at *semantic specificity* o *word-to-referent fit* (Bergelson at Aslin, 2017). Kaugnay nito, sa T2lGen *prompt* na inilagay sa CanvaAI, ‘rooster’ naman talaga ang inisyal na ginamit.

(2cy) Mukhang naglalaro ang dalawang hayop.

Kapansin-pansin naman ang paggamit ng hypernym na ‘dalawang hayop’ (2cy) upang katawanin ang ‘aso’ at ‘tandang.’ Tumutukoy ang hypernym sa isang pangkalahatang salitang sumasaklaw sa mga kahulugan ng mga mas tiyak na salita (Nordquist, 2019). Sa paggamit ng hypernym, nakaiwas ang pahayag sa paggigiit ng anumang inilalapat na hayrarkiya o ebalwatibong bias sa pagitan ng dalawang hayop, samakatwid nababalangkas ang depiksiyon sa mas neutral at pangkalahatang paraan. Masusuri rin sa huling dalawang pangungusap ang paggamit ng epistemikong modal na ‘tila’ at ‘mukha-ng’ na bumabalangkas sa pahayag bilang di-tiyak. Palatandaan ito ng pagiging tentatibo ng aksiyon at paggamit ng *Gemini* ng pag-aalinlangan sa interpretasyon nito. Pagkat wala namang imbento o walang kaugnayang elementong makikita sa larawan, hindi ito maituturing bilang isang halusinasyon. Sa halip, sinasalamain ng paggamit ng mga epistemikong modal na ‘mukha-ng’ ang modalisadong interpretasyong nananatiling nakasalig sa biswal na input.

Umaayon ang deskripsiyon ng mga nabanggit na pangungusap sa eksenang ipinamamalas ng larawan batay sa mga nabuong pangungusap ng *GenAI*. Pinatitingkad ng mga bersiyon nito ang maaaring pagbibigay-interpretasyon sa iisang biswal na input sa magkakaibang paraan. Ipinakikita ng mga pagkakaibang ito ang suhetibong kalikasan ng I2TGen, salik dito ang pagkakaroon ng iba’t ibang kontekstong mapagmumulan ng kakaiba ngunit rasonableng deskripsiyon.

(3ax) Tatlong bata ang masayang pumipitas ng hinog na prutas mula sa punong may masaganang bunga.

May ginagampanang pragmatikong punsiyon ang ‘tatlong bata’ sa (3ax), partikular sa *referent tracking*. Sangkot sa *referent tracking*, gaya ng binibigyang-diin ni Mojapelo (2023), ang pagpapanatili sa kaugnayan ng mga reperensiya sa diskurso. Bagaman parating

naidaramay ang mga pamilyar na referents (o ang tinutukoy sa pangungusap), ipinakikita rin nito kung paanong naipakikilala at nasusubaybayan ang panibagong referents, gaya ng ‘tatlong bata’ sa (3ax). Sa kaso nito, ipinababatid nitong may tatlong indibidwal na gumaganap na nagpapakongkreto sa larawang nabubuo. Sa kabilang banda, mas angkop naman ang paggamit ng pandiwang ‘namimitas’ sa halip na ‘pumipitas’ sa pangungusap. Parehas namang tama ang dalawang pandiwa sa paglalarawan ng aksiyon sa larawan.

(3cx) Ang mga bata ay tuwang-tuwa na namimitas ng mga prutas mula sa puno.

(3cy) Sila ay nagtutulungan upang maabot ang mga hinog na bunga.

Mapapansin sa (3ax), (3cx), at (3cy) ang paggamit ng mga salitang ‘prutas’ at ‘bunga’ bilang hypernyms upang katawanin ang makikita sa larawan. Gayumpaman, ‘mangga’ ang tiyak na prutas na inilahad sa *prompt*, ngunit malinaw na ‘limon’ ang makikita sa *AI-generated* na larawan. Napapahihina nito ang reperensiyal na katumpakan ng mga pahayag. Sinasalamain ng ganitong kamalian ang text-to-image hallucination, nangyayari ito tuwing hindi naipapakita ng *AI-generated* na mga larawan ang makatotohanang impormasyon (Lim & Shim, 2024). Maliban pa ryan, masusuri rito ang suliraning dinaranas ng *vision-language models* upang mailinya ang mga entidad (Alonso et al., 2025), o tinatawag bilang ‘*systemic failure of alignment*’ sa pagitan ng modalidad, o sa kaso ng pag-aaral mula sa impormasyong ‘mangga’ ang kahingian ng *prompt* patungong paglikha ng larawang ‘limon’ ang depiksiyon.

Nakaaapekto ang ganitong *prompt* engineering sa pagbuo ng *AI-generated multimodal content*, dahil esensiyal sa pagbuo ang malinaw ring panuto. Sa kaso ng pag-aaral, nakompromiso ang kalidad ng *AI-generated* na larawan dahil sa hindi sapat na katiyakan sa *prompt*, kahit na may malinaw na pagbanggit sa salitang ‘mangga.’

(3ay) Kitang-kita ang tuwa sa kanilang mga mukha habang nagkakatuwaan sila sa ilalim ng puno.

(3by) Ang mga bata ay nagpapakita ng pagtutulungan at kaligayahan sa pagsasama-sama.

Sa kabilang banda, hindi naman direktang nabanggit sa (3ay) at (3by) ang pinipitas ng mga bata. Sa halip, mas nakatuon ito sa emosyonal na kalagayan (tuwa) at interaksyon (pagtutulungan) ng mga aktor sa larawan. Parte ang ganitong pagbabago sa binabanggit nina Alikhani et al. (2023) na image-text coherence o *multimodal* coherence. Higit na nabigyang-prioridad sa mga pahayag ang apektibo o interpersonal na kahulugan kaysa referents.

Nagbibigay naman ng kolektibong perspektiba ang paggamit ng pariralang nominal na ‘ang mga bata’ sa (3by) at (3cx), samantalang reperensiyang pronominal naman ang umiiral

sa (3ay) at (3cy). Gumaganap na anaporikong panghalip ang ‘sila’ at ‘kanila’ na tumutukoy sa mga naunang nominal, patunay rin ito sa kakayahan ng *GenAI* na makabuo ng lohikong pagrereperensiya sa mga konstituyent ng pangungusap, bagaman malinaw na ibinabatay ito sa pattern recognition (e.g., word embeddings) sa halip na aktuwal na pag-unawa.

Sa kabuoan, binigyang-pansin sa mga nabuong deskripsiyon ang presensiya ng mga bata, ang pagpapahayag ng kanilang kasiyahan, at ang kanilang pamimitas ng prutas, nakatutugon ang paggamit ng mga pariralang naglalarawan upang magkaroon ng biswal na depiksiyon. Masusuring pinagtibay ang mga deskripsiyon ng mga naoobserbahang elemento sa mga larawan, kasama ang mga paksa (batang lalaki, aso, tandang, mga bata), at kanilang mga ikinikilos (pagbabasa, paghahabol, at pamimitas ng prutas), at ang lugar (sa ilalim ng puno, at sa bakuran).

Ipinakikita ng unang bahagi ng pagsusuri ang kakayahan ng lingguwistikong seleksiyon ng *GenAI* na mahugis ang kalinawan at detalye ng mga deskripsiyon batay sa IPs habang pinatitingkad nito ang ugnayan sa pagitan ng pragmatiks, semantiks, at mga palatandaang kontekstuwal. Binibigyang-pansin nito ang kompleksidad ng *AI-generated* na wikang nagbubunsod lamang ng mga padron at probabilistikong ugnayan sa halip na totoong komprehensiyon sa larawan. Maituturing ito bilang kritika sa kasalukuyang repleksiyon ng tao sa kakayahan at limitasyon ng mga modelo sa I2TGen at *multimodal content creation*.

Mga Halusinasyon at Pragmatikong Kahinaan at Kamalian

Sa likod ng pagbuo ng mga *GenAI* ng mga deskripsiyong may kontekstuwal na kaakmaan at pragmatikong katugmaan, nakabuo rin ang *ChatGPT*, *Grok*, at *Gemini* ng mga pangungusap na nagtataglay ng paghahalusinasyon at kamaliang pragmatiko.

(1ay) Tila ba natututo siya ng maraming kaalaman habang nakaupo sa damuhan.

(1cy) Tila siya ay lubos na nasisiyahan sa kanyang binabasa at sa sariwang hangin sa labas.

Hindi tulad ng mga pangungusap na (2cx) at (2cy) na mas direktang umaayon sa nakikitang detalye, nagtataglay ang (1ay) at (1cy) ng impormasyong mahirap matiyak. Palatandaan ang paggamit ng tentatibo at epistemikong pariralang ebidensiyal na ‘tila ba’ ng panghihinuha o pagpapalagay na gumagamit ng suhetibong inferens at hindi malinaw na makikita sa nilalaman ng image *prompt* (IP). Bagaman hindi tulad ng paggamit ng ‘daw/raw’ na reportative evidential at ‘kayâ’ na speculative evidential (Almazán, 2019), nagsisilbing *lexical marker* ng *inferential evidentiality* ang ‘tila (ba)’ sa mga pangungusap. Ipinahihiwatig nito ang interpretasyong nakabatay sa hindi direktang ebidensiya, gaya ng panlabas na anyo o kilos.

Dahil sa pagiging abstrakto at pag-uuri nito bilang isang mental na estado, hindi matitiyak sa empirikal na paraan kung maraming kaalaman ba ang natututuhan ng aktor at kung lubos nga ba siyang nasisiyahan habang nakaupo sa damuhan.

(1by) Ang bata ay sumasalamin sa halaga ng pag-aaral sa gitna ng kalikasan.

Dahil din sa kakulangan sa simbolikong representasyon at malinaw na kontekstong maaaring makapagpatibay sa interpretasyon ng (1by), maituturing din ito bilang isang halusinasyon. Isang interpretasyong hindi sinusuportahan ng IP ang ideya ng pagsalamin ng bata sa ‘halaga ng pag-aral’ at ang kalugurang dala ng sariwang hangin. Katulad ito ng obserbasyon nina Ji et al. (2024) ukol sa malalaking *pre-trained language models* (PLMs) na ginagamit para sa pagsasagawa ng *natural language generation* (NLG) tasks. Anila, makapangyarihan ang PLMs sa pagbibigay ng generalizability (i.e., kakayahan ng modelong makaganap sa mga di-nakikitang datos o mga gawaing higit pa sa nalalaman nitong training set) at coverage (i.e., hanggahan ng training data ng modelo). Gayumpaman, natuklasan nina Longpre et al. (2021, sinipi sa Ji et al., 2024) na binibigyang-prioridad ng mga modelong ito ang parametrikong kaalaman kaysa sa mga impormasyong ini-input. Kadalasan itong nagreresulta sa halusinasyong nagbubunga ng labis-labis na impormasyon sa awtput nito. Kaugnay ito sa ipinakita ng mga nabanggit na interpretasyon ng kognitibo at emosyonal na estado at mga kondisyong pangkapaligirang higit pa sa mamamalas. Samakatwid, maituturing ito bilang isang ispekulasyon dahil hindi sapat ang ebidensiyang biswal.

(2ay) Ang tandang naman ay mabilis na tumatakbo upang makaiwas sa aso.

Ipinakikita ng pariralang pang-abay na ‘upang makaiwas sa aso’ sa (2ay) ang layunin ng kilos ng sa tandang¹ na hindi naman direktang maoobserbahan sa image *prompt* (bagaman nakasandig sa mga palatandaang biswal at kaalamang nakaimbak ang AI), kaya maituturing ito bilang pragmatikong inferens, partikular sa inferensiyal na atribusyon ng intensiyon. Sumasang-ayon ito sa binabanggit sa ostensive-inferential communication nina Sperber at Wilson (2004) na nabubuo ang kahulugang batay sa inferensiyal na ugnayan sa pagitan ng ipinahahayag at ng konteksto. Dagdag pa, masusuri sa interpretasyong nailapat ng AI sa larawan ang teleological stance o pagpapalagay na may partikular na layunin ang aksiyon (Gergely & Csibra, 2013). Nabigyang-kahulugan ng *ChatGPT* ang di-umanong ‘pagtakbo’ ng tandang bilang isang inferens na iiwas ito sa aso.

(2bx) Ang aso ay masayang tumatakbo palapit sa manok sa bukirin, na puno ng sigla at kulay.

1 Ginamit na lamang ang ‘tandang’ upang maging konsistent sa mga nabuong deskripsiyon.

Labis na pragmatikong interpretasyon at bahagyang pragmatikong halusinasyon naman ang paglalagay ng pariralang ‘puno ng sigla at kulay’ sa (2bx) dahil sa suhetibong paglalarawan gamit ang piguratibong wika. Hindi malinaw ang referent ng parirala kaya maituturing din ito bilang kalabuang reperensiyal dahil sa malabong pagdirikit ng head at referent. Hindi malinaw aling konstityuwent sa pangungusap ang inilalarawan ng parirala, pagkat kung:

- a. ang aso ang nais ilarawan, maaaring ebidensiyal na masigla ito, hindi naman puno ng kulay;
- b. kung manok naman, hindi agarang mababatid ang kasiglahan nito;
- c. kung bukirin naman, maaaring ebidensiyal na puno ito ng kulay, walang kuwalifayer na makapagtatakda ng kasiglahan nito.

Ipinakikita ng ganitong kalabuan ang referential indeterminacy o di-tiyak na pagrereperensiya sa mga bagay-bagay (Assadian, 2021). Sa madaling sabi, umiiral ito bilang isang kondisyon sa oras na nakokompromiso ang koherens ng diskurso dahil sa malabong reference tracking.

(2by) Ang mga hayop ay nagpapakita ng kaibigan na maaaring magkaroon sa kalikasan.

Maoobserbahan naman sa (2by) ang malinaw na kaso ng halusinasyon. Ipinakikilala ng pangungusap ang abstraktong konsepto ng ‘pagkakaibigan’ (mamamalas sa paggamit ng salitang ‘kaibigan’) na walang direktang paglalatatag ng biswalidad sa pagitan ng mga karakter sa larawan. Taglay ng deskripsiyon ang pagsalamin nito bilang isang inferential leap o pagpapatibay sa anumang higit pa sa ipinapakitang kahulugan ng aytem (Hinzen, 2001) na ikinakabit sa anumang panlipunang kahulugang hindi masuportahan ng biswal na ebidensiya. Bunsod din ito ng kalabuan sa paglalarawang bitbit ng pangungusap.

Mas angkop kung irereestruktura ang pangungusap bilang: “Ang mga hayop ay nagpapakita ng pagkakaibigang maaaring makita sa kalikasan.” Masusuring mula sa pangkaraniwang pangngalang ‘kaibigan’ naisalin ito bilang pangngalang abstraktong ‘pagkakaibigan-g’ na higit na nagpapakita ng pragmatikong katuturan sa pangungusap. Ipinakikita nito ang madalas na paglampas ng AI sa hanggahan ng deskriptibong katugmaan nito sa pamamagitan ng pag-uugnay sa mga tematiko o simbolikong nilalamang hindi naman nakaangkla sa konteksto.

Halusinasyon ang atribusyon sa layunin ng tandang na ‘iwasan’ ang asong hindi naman direktang makikita sa larawan. Gayundin ang halusinasyong ‘masaya-ng’ paggalaw ng aso at di-beripikadong sitwasyon ng lugar na ‘puno ng sigla at kulay,’ kasama na rin ang pabrikasyon sa nosyong relasyonal ng ‘pagkakaibigan’ sa pagitan ng mga hayop na walang

sapat na biswal na ebidensiya. Nagpapakita ang mga dagdag na deksripsiyon ng higit pa sa aktuwal na depiksiyon ng paghabol ng aso sa manok sa bakuran.

(3bx) Ang mga bata ay naghihikayat sa isa’t isa habang kinuha ang mga dilaw na prutas mula sa puno.

Masusuri naman sa pariralang ‘naghihikayat sa isa’t isa’ sa (3bx) ang social interaction inference na hindi naman nagpapakita ng anumang biswal na pagpapatotoo sa IP. Tumutukoy ito sa malinaw na paggamit ng tao ng mataas na antas ng inferens upang matukoy at maunawaan ang maraming aspekto ng panlipunang interaksyon (Malik at Isik, 2023). Dahil sa kakulangan ng mga biswal na palatandaan upang katawanin ang akto ng panghihikayat, maituturing ang kasong ito bilang halusinasyon, o isang interpretatibong pagmamalabis. Bagaman nakasalalay ang *GenAI* sa mga pamilyar na naratibong padron o asosasyon sa paggawa ng mga deskripsiyon, sinasalamin ng sitwasyong ito ang imbentong ugaling panlipunang ginagawa ng totoong tao.

Ipinakikita ng mga nabanggit sa ikalawang categoryang bagaman may kakayahan ang *GenAI* na maisakatuparan ang I2TGen, hindi pa rin maiiwasan ang pagbuo ng halusinadong nilalaman sa mga nalilikha nito. Manipestasyon ito ng limitado pang kakayahan ng *artificial intelligence* na tanging tao lamang ang may natural na kakanyahan.

Dagdag pa, may masusuring varyasyon sa tatlong modelo sa usapin ng pragmatikong katugmaan at kaakmaang kontekstuwal. Konsistent naman ang akto ng pamimitas ng mga dilaw na prutas sa image *prompt*, ngunit hindi ganap na makikita sa larawan ang espesipikong detalyeng nais iparating ng “ang mga bata ay naghihikayat.” Ipinakikita ng mga natuklasang tugma naman ang representasyon ng mga pangunahing paksa at aksiyon, gayumpaman, maituturing ang mga dagdag na detalye gaya ng emosyon, intensiyon, at mga di-maobserbahang konteksto bilang halusinasyon pagkat hindi ito direktang sinusuportahan ng anumang empirikal na ebidensiya sa mga larawan.

Talahanayan 2

Kategorisasyon ng mga Pangungusap

GenAI Tool	May kaakmang kontekstuwal at pragmatikong katugmaan	May halusinasyon at pragmatikong kamalian
ChatGPT	4	2
Grok	2	4
Gemini	5	1
Kabuoan	11	7

Masusuma sa Talahanayan 2 ang bilang ng mga pangungusap na maikakategorya bilang nagtataglay ng kaakmaang kontekstuwal at pragmatikong katugmaan, at halusinasyon at pragmatikong kamalian. Lima sa anim na pangungusap ng *Gemini* ang nagpakita ng kawastuhang kontekstuwal at pragmatiko, habang isa lamang ang halusinado. Sinundan ito ng *ChatGPT* na may apat na wastong pangungusap at dalawang nagtataglay ng kamalian. Pinakamababa naman ang *Grok* sa unang kategorya (2) at pinakamataas sa ikalawa (4).

KONGKLUSYON AT REKOMENDASYON

Dahil sa naoobserbahang mabilis at di-mapipigilang ebolusyon ng Generative Artificial intelligence (*GenAI*), may agarang pangangailangan din upang kritikal na mataya ang kakayahan ng mga sistema upang makabuo at makalikha ng kahulugan mula sa samot-sari at magkakaibang modalidad. Lalo itong totoo sa mga gawaing mangangailangan ng pagsasalin ng biswal na datos patungo sa magkakaugnay at angkop-sa-kontekstong lingguwistikong awtput o *image-to-text generation* (I2TGen). Natuklasan sa pag-aaral na limitado at di-gaanong napagtutuunan ng pansin ang naturang paksa at mas nabibigyan ng atensiyon ang *text-to-image generation* (T2IGen).

Nakaaambag ang kasalukuyang pag-aaral sa patuloy pang lumalawak at lumalaking literatura ukol sa *multimodal* AI systems, dahil na rin sa demand nito. Partikular itong nakapokus sa nagiging gampanin ng wika sa pag-eenkord ng mga elemento gaya ng mga ahente, aksiyon, ugnayang pang-espasyo, emosyon, at mga konseptong abstraktong hinahango sa *image prompts*.

Pinatitingkad din ng pag-aaral ang katotohanang hindi sapat na may kakayahan lamang ang generative AI na makalikha ng mga tekstong may wastong gramatika, bagkus, mahalagang taglayin din ng mga deskripsiyon ang pragmatikong kaayusan at kalinawan, lohikong inferens, at sensibilidad sa kulturang pinagmumulan, pinagbabatayan, o pinaghahanguan ng *image prompts*, at sa anumang modipikasyon o *prompt engineering* na gagawin ng user. Sa pagsusuma, nagpakita ang *ChatGPT*, *Grok*, at *Gemini* ng kakayahang maka-generate ng mga pangungusap na may maayos na estruktura at pagsalamin sa mga mahahalagang komponent ng *image prompts*, bagaman nagtataglay ng lisyang ilan sa mga ito. Sa kabila nito, makahulugan ang pagkakaiba-iba ng mga awtput kung ibabatay sa pragmatikong katumpakan at inferensiyal na katiyakan.

Mula sa mga obserbasyon, lumitaw na *Gemini* ang may pinakasensitibo sa konteksto at pinakamalinaw na deskripsiyon sa tatlo. Taglay rin ng mga nabuo nitong pangungusap ang tumpak na referent identification. Balanse naman ang resulta ng pagsusuri sa mga nabuong pangungusap ng *ChatGPT*, bagaman may ilang deskripsiyong nagpamalas ng bahagyang pagmamalabis sa interpretasyon.

Sa kabilang banda, naging lantad naman sa mga pangungusap na na-generate ng *Grok* ang halusinasyon. Nakita sa mga deskripsiyon ang mga lohikong inferens ngunit kapos sa batayang biswal at sobra naman sa pagkatig sa abstrakto o emosyonal na wikang di-mamamalas sa larawan. Naging madalas ang paggamit ng *Grok* ng mga pariralang metaporiko at piguratibo bagaman hindi naman ebidensiyal ang mga temang ito. Nailalantad ng ganitong sitwasyon ang tendensiyang magmalabis sa mga naratibo kaysa tumuon sa mga deskripsiyong tuwirang maoobserbahan sa mga *prompt*.

Binibigyang-diin ng mga resulta ang kapwa mga potensiyal at limitasyon ng kasalukuyang *GenAI* systems na kaugnay sa gawaing *image-to-text generation*. Naipamalas man ng mga plataporma ang katatasan, batayang kalinawang reperensiyal, at kakayahang maenkwid ang mga elemento (e.g., aso, bata, puno, tandang, at iba pang pangkaraniwang biswal na palatandaan sa mga larawan), lantad na kinapos ang tatlo sa pag-generate ng mga deskripsiyong nagtataglay ng pragmatikong pag-iingat at katapatang kontekstuwal. Mailalapit ang mga limitasyong ito sa mga salik tulad ng biases sa training data, kawalan ng makatotohanang batayang perseptuwal na bentaha ng tao kompara sa kakayahan ng LLMs, at mga hadlang na bunga ng inferensiyal na pagrarasong isinasakatuparan ng mga kasalukuyang arkitekturang pang-AI.

Mula rito, may ilang deduksiyon at pangangatwirang nagmula sa mga natuklasang maaari pang bigyang-pansin. Una, hindi pa sapat ang kakayahan ng *GenAI* system upang pragmatikong makapagrason. Kadalasan, sumasangguni lamang ito sa antas-paimbabaw na korelasyon sa pagitan ng mga larawan at padrong pangwikang nakaimbak sa database nito. Ikalawa, nagmumula ang halusinasyon sa tendensiya ng mga modelong pilit at puwersahang mapunan ang mga inferensiyal na puwang ng mga maituturing na generic na kaisipan o lubhang mabulaklak na wika. Ikatlo, nananatiling isyu ang referential indeterminacy, lalo na kung sangkot sa pangungusap ang maraming ahente o malabong mga konteksto. Samakatwid, mayroon pang pangangailangan upang tuluyang mapaunlad ang mga balangkas na gagamitin para sa ebalwasyon ng mga platapormang binubuo hindi lamang ng mga sintaktiko at semantikong pamantayan, bagkus maging ng pragmatiko at kultural na katumpakan.

Bagaman nagpapamalas ng kontemporaneong kapasidad ang mga platapormang *GenAI* gaya ng *ChatGPT*, *Grok*, at *Gemini* sa *multimodal* text generation, naisiwalat pa rin ng mga nabanggit na pragmatikong limitasyon ang hangganan ng kasalukuyang nauunawaan o ‘naieensayo’ sa pag-unawa ng AI. Sa ganitong gana, lantad na magkakaiba ang kakayahan ng mga modelo sa pagbibigay ng kahulugan sa mga biswal na konteksto. Pinatitingkad ng mga nasuring pagkakaiba ang pangangailangang mataya hindi lamang ang katatasan ng tekstong *AI-generated*, bagkus maging ang mga katumpakan at kaugnayan nito sa biswal na input, partikular sa mga gawaing nangangailangan ng angkop na interpretasyon ng mga larawan.

Hindi lamang dapat limitahan ang mga pag-aaral ukol sa T2IGen sa mga pagpapahusay ng mga modelo at protokol sa pagsasanay nito, bagkus maging sa paniniyasat sa mga interdisiplinaryong solusyon. Habang kinakailangan ang pagtalakay sa debelopment ng large language models (LLMs), gaya ng mabisang mekanismo sa pagtutok ng atensiyon sa mahahalagang impormasyong bitbit ng *prompt*, paggamit ng malinaw na teknik sa paghahanay, at pagsasagawa ng mga estratehiyang makapagtatahi sa mga *multimodal* na elemento, hindi sapat na nagsasarili ang AI sa pagsasakatuparan nito. Nangangailangan ng holistiko at interdisiplinaryong interbensiyon ang pagharap sa bantang hatid ng halusinasyon, kalabuang reperensiyal, kultural na misinterpretasyon, at lisyang inferensiyal sa mga awtput ng *GenAI*.

Sa usaping lingguwistiko, maaaring mailapat sa mga susunod pang saliksik ang iba pang prominenteng pananaw at kaisipang nagmumula sa semantiks at pragmatiks upang matulungan ang mga modelong higit na maunawaan at makagamit ng angkop sa kontekstong pangwika. Mahalagang maisanib ang mga teorya at konseptong may kaugnayan sa pandiskursong koherens, *referent tracking*, at evidentiality sa mga tiyak na layunin ng bawat modelo at mga balangkas na ginagamit sa pagtataya at pagtatasa ng mga datos upang mawaksi ang paglikha ng mga ligaw at mapanlinlang na awtput. Maaari ring palawakin pa ang dataset ng kasalukuyang pag-aaral sa pamamagitan ng paglalahok ng pragmatic markers, speaker/writer intention, at implicatures, nang maensayo ang *GenAI* Models sa pag-generate ng mga tugong mas sensitibo sa konteksto.

May mariing pangangailangan din upang makadebelop ng auditing tools na makatutulong sa users at developers upang makatukoy ng halusinasyon, lalo na sa mga *AI-generated content* na nagtataglay ng mga makahulugang sosyal, politikal, o medikal na implikasyong ipinamamahagi sa publiko—malay man o hindi ang madla. Magagamit ang mga kagamitang ito upang mamarkahan ang mga posibleng pagkakamali at makapagmungkahi ng sagot na mapagtitibay ng mga patotoo. Dagdag pa, kinakailangan ding lubog sa datos at angkop na pagdidisenyo ang pag-aaral ng paksa sa lente ng kros-kultural na pragmatiks. Bilang dominante ang pagsasanay ng mga modelo sa Kanluraning dataset, maaari itong magbunga ng misinterpretasyon o misrepresentasyon sa mga gawî, simbolo, o norm sa mga kontekstong di-Kanluranin.

Mahalagang masaliksik ang nababagay na lokalisasyon at adaptadong pagpipino upang nakaugnay sa lokal na kultura ang korpus at materyal na gagamitin sa pag-aaral. Kailangang maisaalang-alang din ang inklusyon at pagsasangkot ng mga multilingguwal at multikultural na eksperto sa anotasyon at ebalwasyon ng magkakarugtong na proseso. Sa pamamagitan nito, mababawasan ang banta ng kultural na pagbubura, pag-iistiryotipo, at inferensiyal na pagmamalabis. Sa pagsusuma, higit pa sa teknikal na optimisasyon

at apropyadong paggamit ang responsableng pagtataguyod at pagsusulong sa *GenAI*. Mahalagang mailinya ang mga naturang pagpapaunlad sa arkitektura nang may malalimang pag-unawa sa wika at kulturang taglay at masasalamín sa mga mapagkakatiwalaang awtput nito.

MGA REPERENSIYA

- Alikhani, M., Khalid, B., & Stone, M. (2023). Image-text coherence and its implications for multimodal AI. *Frontiers in Artificial Intelligence*, 6, 1-13.
<https://doi.org/10.3389/frai.2023.1048874>
- Almazán, J. (2019). *Evidentiality in Tagalog*. [Doktral na Disertasyon, Universidad Autónoma de Madrid]. AUM Repositorio.
- Alonso, I., Salabierra, A., Azkune, G., Barnes, J., & de Lacalle, O. (2025). *Vision-language models struggle to align entities across modalities*. ArXiv.
<http://dx.doi.org/10.48550/arXiv.2503.03854>
- Andrada, M. (2021). Subersibong potensiyal ng makina ng pagsasalin: Google translate at tula ni Carlos Bulosan. *Malay Journal*, 33(2), 29-45.
<https://doi.org/10.59588/2243-7851.1071>
- Assadian, B. (2021). Deflationary reference and referential indeterminacy. Sa M. Mojtahedi, S. Rahman, & M. Zarepour (mga Pat.), *Mathematics, Logic, and their Philosophies. Logic, Epistemology, and the Unity of Science*, 49 (pp. 365-377). Springer, Cham.
https://doi.org/10.1007/978-3-030-53654-1_13
- Bergelson, E., & Aslin, R. (2017). Semantic specificity in one-year-old's word comprehension. *Language Learning and Development: The Official Journal of the Society for Language Development*, 13(4), 481-501. <https://doi.org/10.1080/15475441.2017.1324308>
- Bosque, A. (2024). Analisis sa lexical choice ng *ChatGPT*-generated na pagsasalin. *Malay Journal*, 37(1): 94-106. <https://doi.org/10.59588/2243-7851.1008>
- Briana, J. (2024). Is *ChatGPT*-produced text authentic? a contrastive analysis of cohesive markers in human and *AI-generated* text. *Journal of English and Applied Linguistics*, 3(2): 38-54. <https://doi.org/10.59588/2961-3094.1120>
- Chen W., Hu H., Li Y., et al. (2023). *Subject-driven text-to-image generation via apprenticeship learning*. arXiv. <https://doi.org/10.48550/arXiv.2304.00186>

- Corpus, S. and Villanueva, A. (2024). Speech emotion recognition in Filipino spoken language using deep learning. Sa *2024 IEEE International Conference on Artificial Intelligence in Engineering and Technology*, 430-435.
<https://doi.org/10.1109/IICAET62352.2024.10730717>
- Csibra, G., & Gergely, G. (2013). Teleological understanding of actions. Sa M. Banaji & S. Gelman (mga Pat.), *Navigating the Social World: What Infants, Children and Other Species Can Teach Us* (pp. 38-43). Oxford Academic.
<https://doi.org/10.1093/acprof:oso/9780199890712.003.0008>
- Duffy, G. (2008). Pragmatic analysis. Sa A. Klotz, & D. Prakash (mga Pat). *Qualitative Methods in International Relations* (pp. 168-186). London, Palgrave Macmillan.
- Endriga, D., & Rosario F. (2023). Gender bias in machine translation: the case of Filipino-English translation in Google Translate. Sa R. Moratto & M. Bacolod (mga Pat). *Translation Studies in the Philippines: Navigating a Multilingual Archipelago* (pp. 83-99). Routledge.
- Estrella, J. (2023). Does *artificial intelligence* genuinely capture the essence of language? *UP Working Papers in Linguistics*, 2(1), 120-124.
<https://linguistics.upd.edu.ph/wp-content/uploads/2023/10/21-Artificial-Intelligence.pdf>
- Evkaya, O. & de Carvalho, M (2024). Decoding AI: The story of data analysis in *ChatGPT*. arXiv.
<https://doi.org/10.48550/arXiv.2304.00186>
- He Y., Robey A., Murata N., et al. (2024). *Automated black-box prompt engineering for personalized text-to-image generation*. arXiv.
<https://doi.org/10.48550/arXiv.2403.19103>
- Hinzen, W. (2001). The pragmatics of inferential content. *Synthese*, 128(1-2), 157-181.
<http://dx.doi.org/10.1023/A:1010362521497>
- Ji, Z., Li, N., Frieske, R. et al. (2024). *Survey of hallucination in natural language generation*. ArXiv. <https://doi.org/10.48550/arXiv.2202.03629>
- Lim, Y., & Shim, H. (2024). *Addressing image hallucination in text-to-image generation through factual image retrieval*. ArXiv. <https://doi.org/10.48550/arXiv.2407.10683>

- Liwanag L., Liwanag G., & Liwanag L.. (2024). AI in anthem: A comparative analysis of the English and Filipino *ChatGPT* 4 translations from the existing translations of the Philippine National Anthem. *Recoletos Multidisciplinary Research Journal*, 12(2), 91-102. <https://doi.org/10.32871/rmrj2412.02.07>
- Longpre, S., Perisetla, K. Chen, A., Ramesh, N., DuBois, C., Singh, S. (2021). *Entity-based knowledge conflicts in question answering*. ArXiv. <https://doi.org/10.48550/arXiv.2109.05052>
- Malik, M., & Isik, L. (2023). Relational visual representations underlie human social interaction recognition. *Nature Communications*, 14(7317), 1–11. <https://doi.org/10.1038/s41467-023-43156-8>
- Mignot, E. (2012). The conceptualization of natural gender in English. *Anglophonia: French Journal of English Studies*, 16(32), 39-61. <https://doi.org/10.4000/anglophonia.140>
- Mojapelo, M. (2023). Aspects of *referent tracking* in Northern Sotho. Sa B. Achiri-Taboh (Pat.), *The Bantu Noun Phrase: Issues and Perspectives* (pp. 210-226). London: Routledge. <http://dx.doi.org/10.4324/9781003254188-13>
- Montalan, J., Layacan, J., Africa, D., et al. (2025). *Batayan: A Filipino NLP benchmark for evaluating large language models* arXiv. <https://arxiv.org/abs/2502.14911>
- Nagaya, N. (2007). Information structure and constituent order in Tagalog. *Language and Linguistics*, 8(1), 343-372.
- Nordquist, R. (2019, Hulyo 3). Definitions and examples of hypernyms in English. *ThoughtCo*. <https://www.thoughtco.com/hypernym-words-term-1690943>
- Nordquist, R. (2024, Abril 30). What are *Hyponyms* in English? *ThoughtCo*. <https://www.thoughtco.com/Hyponym-words-term-1690946>
- Oppenlaender, J. (2022). The creativity of text-to-image generation. In *Proceedings of the 25th International Academic Mindtrek Conference (Academic Mindtrek '22)* (p. 11). ACM. <https://doi.org/10.1145/3569219.3569352>
- Oppenlaender, J., Linder, R., & Silvennoinen, J. (2024). *Prompting AI art: An investigation into the creative skill of prompt engineering*. *International Journal of Human-Computer Interaction*, 1–23. <https://doi.org/10.1080/10447318.2024.2431761>
- Reid, L., & Liao, H. (2004). A brief syntactic typology of Philippine languages. *Language and Linguistics*, 5(2), 433-490. <http://hdl.handle.net/10125/32991>

- Souza, M. E., & Weigang, L. (2025). *Grok, Gemini, ChatGPT, and Deepseek: Comparative and applications in conversational artificial intelligence*. *Inteligencia Artificial*, 2(1), 1–7. <http://dx.doi.org/10.5281/zenodo.14885243>
- Sperber, D., & Wilson, D. (2004). Relevance theory. Sa L. Horn & G. Ward (mga Pat.), *Blackwell's Handbook of Pragmatics* (pp. 607-632). Blackwell.
- Trojovsky, P., & Naveen, P. (2024). Overview and challenges of machine translation for contextually appropriate translations. *iScience*, 27(10), 1–25. <https://doi.org/10.1016/j.isci.2024.110878>
- Villemaire, S. (2025). *Canva*. <https://www.Canva.com/learn/how-to-convert-text-images-ai-magic/>.
- Wang, Y., & Zhang, G. (2025). Lightweight text-to-image generation model based on contrastive language-image pre-training embeddings and conditional variational autoencoders. *Electronics*, 14(11), 1–31. <https://doi.org/10.3390/electronics14112185>
- Warr, M. (2024). Beat bias? Personalization, bias, and generative AI. Sa J. Cohen & G. Solano (Eds.), *Proceedings of Society for Information Technology and Teacher Education International Conference* (pp. 1481–1488).

TUNGKOL SA MAY-AKDA

Si Ariel Ursolon Bosque ay isang lisensiyadong propesyonal na guro at kasalukuyang katuwang na propesor sa Departamento ng Filipino at Panitikan ng Pilipinas, Kolehiyo ng Arte at Literatura, Unibersidad ng Pilipinas Diliman. Nagtapos siya ng programang Batsilyer sa Pansekondaryang Edukasyon, Medyor sa Filipino, magna cum laude, sa Rizal Technological University, at programang Master ng Sining sa Araling Pilipino (Wika at Antropolohiya) sa UP Diliman. Sa ngayon, tinatapos niya ang programang Doktor sa Pilosopiya ng Filipino na may konsentrasyon sa Istrukturang Wikang Filipino sa kaparehong unibersidad. Gumaganap din siya bilang Kawaksing Kasapi ng National Research Council of the Philippines sa ilalim ng Dibisyon VIII, Agham Panlipunan. Kabilang sa kaniyang interes sa pananaliksik ang Araling Wika, Gramatika, Sintaks at Semantiks, Mga Teknolohiyang Pangwika, Filipino bilang Ikalawang Wika, Pagsasalin, at Plain Language.