Case marking determiners and the role of register in Philippine NP-interpretation

Daniel Kaufman

Queens College, CUNY & Endangered Language Alliance

dkaufman@qc.cuny.edu

Abstract

Almost all descriptions of Philippine languages note a connection between certain types of arguments and (in)definiteness, the clearest of which is that actor voice objects tend to receive an indefinite interpretation while the syntactically privileged argument (alternatively analyzed as nominative, absolutive and pivot) receives a definite interpretation. However, this is where the consensus ends. In some analyses of Tagalog, the primary focus of the present chapter, these interpretive differences have been attributed to the case markers that introduce arguments. Other analyses attempt to derive these differences solely through syntactic position, with lower verb phrase internal positions typically receiving an indefinite interpretation and higher positions receiving a definite one. Here, I show that it is impossible to deny the semantic contribution of the nominative case marker, in particular, which functions as a definite determiner in both sentential and non-sentential contexts. I focus on a set of exceptions in which this determiner is followed by the numeral 'one' (e.g., *ang isang guro*) and show that this construction is most likely due to contact with European languages and is highly restricted both semantically and syntactically.

1.0   Ricardo Nolasco: an appreciation

It is an honor to offer this paper to my former professor and lifelong mentor, Ricardo Nolasco, who profoundly shaped my thinking about Philippine languages and linguistics during my formative years and whose work continues to inspire me. I would like to first consider Nolasco's approach to Philippine linguistics, which, in many ways, stands in subtle opposition to what can be considered the general western approach. For the last several decades, research in Philippine linguistics has been dominated by largely American theoretical agendas. As a result, an enormous amount of ink has been spilled on constraints that are mystifying from an English perspective but so fundamental to Philippine grammar that they have been perhaps too obvious to evoke the same interest among Filipino scholars, a prime example of this being the well-trodden ungrammaticalities shown in (1).

(1)a.   **Ano ang kumain ang bata?*   b.   **Sino ang kinain ang isda?*
    anu   aŋ=k\<um>aːʔin   aŋ=baːtaʔ   siːnu aŋ=k\<in>aːʔin-∅   aŋ=isdaʔ
    what   NOM=\<AV.INIT>eat NOM=child   who NOM=\<INIT>eat-PV NOM=fish
    (For 'What the child eat?')   (For 'Who ate the fish?')

These are neither the types of errors that children make nor do they represent an area of variation in Philippine languages. They are (as I have argued elsewhere) a result of the basic copular style of predication employed by Philippine languages, an insight that was already suggested by

Cecilio Lopez and other early Filipino linguists. On the other hand, the types of questions that have been of more immediate interest to Filipino linguists have been in areas where there does exist cross-linguistic variation and even optionality. One such area where Nolasco has focused his energies are the factors involved in voice selection in declarative main clauses. The subtle, multifactorial and discourse sensitive nature of this process has not made it highly amenable to current generative frameworks and it has thus received far less attention from foreign linguists. Approaching the voice system from a speaker-centric perspective, Nolasco has made ample use of vernacular literature. As an undergraduate, I would often find him in his office collecting examples from literary magazines such as Bannawag, Liwayway, and Hiligaynon to demonstrate one or another factor in voice selection. Nolasco's approach has differed in other important ways from his Philippinist counterparts in America and elsewhere in that his speaker-centrism extends to advocating for the maintenance of regional languages in all domains of life. For the last two decades he has dedicated himself to training school teachers to better understand the structure of their own languages so that they can explain grammatical principles in a scientific manner to their students. Nolasco's linguistics has been humanitarian in the broadest sense - based on the real speech and writing of native speakers, focused on subtle points on which Philippine languages differ but which are sensible to average speakers, and part and parcel of social and political advocacy to protect the fantastic linguistic diversity which has defined the Philippines for milennia.

It is thus firmly in the Nolasconian style which I present the current paper on the different strategies for expressing definiteness across Philippine languages. Most importantly, I argue that it is only when we take a critical eye to the data and take discouse context more seriously that we can rectify errors of judgment and analysis in the theoretical literature.

2.0   Introduction

Many approaches to Philippine-type alignment systems rely heavily on the way certain arguments naturally obtain a definite or indefinite interpretation (all too often conflated with the specific vs. non-specific distinction). It has long been noted that the *ang*-phrase or "pivot" in Tagalog and its equivalents in other Philippine languages receive a definite interpretation and that the actor voice object receives an indefinite interpretation. At the same time, genitive agents may obtain either a definite or indefinite reading. The basic paradigm is shown in (2) with its various interpretive possibilities.[1]

---

[1] Glossing abbreviations: AV - actor voice; CV - conveyance voice; GEN - genitive case; IMPRF - imperfective; INIT - initiated action; LV - locative voice; NOM - nominative/pivot/absolutive case; PV - patient voice. Note that on this analysis the initiated action morpheme alone gives rise to a perfective interpretation; initiated action in combination with the imperfective yields a progressive interpretation and the imperfective alone gives rise to a prospective aspect interpretation. In the tier containing segmented underlying forms in the examples, I do not indicate the initial glottal stop in words that are written with an initial vowel, as this glottal stop is predictable.

(2)a. *Hinabol ng aso ang pusa.*   b. *Humabol ng pusa ang aso.*

    h⟨in⟩aːbul-∅   naŋ=aːsu aŋ=puːsaʔ   h⟨um⟩aːbul   naŋ=puːsaʔ aŋ=aːsu

    <INIT>chase-PV GEN=dog NOM=cat   <AV.INIT>chase GEN=cat   NOM=dog

    '(The/A) dog chased (the/*a) cat.'   '(The/*A) dog chased (a/*the) cat.'

Yet the literature is rife with disagreements over some of the basic facts. For instance, Sabbagh (2016), Paul et al. (2015) and Collins (2019) argue that the restrictions on definiteness and indefiniteness are defeasible and therefore merely represent tendencies while the traditional view holds these to be rigid principles.

    This question of NP interpretation in Tagalog has often been approached from the perspective of the markers that ubiquitously precede all types of arguments. Specifically, are markers such as *ang* and *ng* pure determiners indicating referentiality, pure case markers, portmanteau case-marking determiners, or something else, a question which has been answered in every possible way (cf. see Kaufman, 2017 and Hsieh, 2020 ch.3 for overviews).

    I will argue that they are portmanteau case-marking determiners and that attempts to derive their interactions with definiteness purely through syntactic position are doomed to fail. I will not examine definite/specific AV objects here, but rather focus solely on the phenomenon of indefinite nominatives/pivots. As a preliminary, I give the following informal definitions of the key concepts (see von Heusinger, 2019 for further elaboration)[2]:

    EXISTENTIAL INDEFINITE: A referent that is neither identifiable to the speaker nor to the hearer (e.g. 'If you find <u>a taxi</u>, let me know.')

    SPECIFIC INDEFINITE: A referent that is identifiable to the speaker but not the hearer (e.g. 'So I met <u>an interesting linguist</u> last night.')

    DEFINITE: A referent uniquely identifiable to the speaker and presented as uniquely identifiable to the hearer (e.g. '<u>The Queen </u>is arriving.')

The typical mapping between grammatical relations and NP-interpretations in Philippine languages is summarized in (3).

(3)   PHILIPPINE-TYPE SYNTAX-PRAGMATICS MAPPING

    a. The nominative marked argument receives a definite interpretation.

    b. Non-oblique, non-nominative undergoers receive an indefinite interpretation.

    c. Genitive agents of non-actor voice predicates have an unrestricted interpretation (existential indefinite, specific indefinite or definite).

---

[2] Specific definites are also referred to as referential definites while existential indefinites are also referred to as non-specific or non-referential indefinites in the literature.

The referentiality[3] of the nominative phrase/pivot posited in (3a) has been noted by nearly all descriptions and typological overviews of Tagalog and many related Philippine languages (cf. Bloomfield, 1917; Cena, 1977; McFarland, 1978; Schachter and Otanes, 1982; Schachter, 1977, 1976; Himmelmann, 1991, 1997, 2005; Kroeger, 1993; Foley, 1998; Nagaya, 2011; Reid, 2000, 2002; Reid and Liao 2004; Nolasco, 2003; Kaufman, 2024, 2018), although precise descriptions can vary widely.[4]

While there is widespread agreement on the definiteness of 'plain' pivot phrases, apparent exceptions to the definite interpretation are found when the pivot is modified by the numeral 'one' or weak quantifiers. We will focus here on pivot phrases containing the cardinal modifier 'one', labelling this the NOM+one construction. With only few exceptions, the NOM+one construction had largely escaped scrutiny in the 27 years between Adams and Manaster-Ramer (1988) and Paul et al. (2015). One exception, Himmelmann (2005:368), notes: "It is also common to make indefinite patients and themes the subject if they are going to be major participants in the ongoing discourse, in particular if they are animate", offering (4) as an example.

(4)   *Doon ay nakita nila ang isang malaking higante.*
        duʔun   ay na-ki:taʔ=nila   aŋ=isa=ŋ   ma-laki=ŋ   higante
        DIST.LOC TOP INIT.STA-see=3p.GEN   NOM=one=LNK ADJ-big=LNK giant
        'There they saw a great giant . . . [the giant is going to be the main protagonist of the ensuing episode]'

Himmelmann thus sees examples like (4) as instantiating an exceptional mapping of an indefinite argument to pivot, countenanced by the argument having higher level discourse prominence. I believe the keys to understanding apparent exceptions to the pattern in (3) can be reduced to two factors: (i) competition between alternative expressions and (ii) a more nuanced view of sociolinguistic register and speech style. With this, we can restore the traditional analysis of Foley (1988), Himmelmann (1998, 2005), Reid (2000, 2002) and Kroeger (1988) that attributes a determiner function to Philippine style case markers. In the following, I first argue that an alternative view which takes NP interpretation to be determined purely by syntactic position without any contribution from the case markers is untenable for Tagalog.

3.0   The positional analysis

Collins (2019) offers a detailed argument that the case marker *ang* cannot mark definiteness because *ang*-marked phrases receive an indefinite interpretation when combined with certain

---

[3] Some authors claim that the Tagalog pivot bears a definite interpretation while others have made the weaker claim that the pivot is only required to be specific. I use 'referentiality' here to cover either of these possibilities, both of which share the property of being speaker-identifiable.
[4] Nagaya (2011) claims that the *ang* phrase itself is losing referentiality and is being replaced by demonstrative headed phrases in Tagalog (i.e. *yung*) and other Philippine languages.

types of quantificational expressions such as *isa* 'one'. Without delving into the mechanics of his analysis, which are irrelevant here, I focus on the empirical claims, arguing that the interpretive contributions of the case markers cannot be so easily discounted.[5]

The claim that case markers do not determine referentiality in Tagalog is largely based on data such as (5), first discussed by Adams and Manaster-Ramer (1988), which ostensibly shows that the *ang* phrase may freely obtain an indefinite interpretation.

(5)   *Dumating ang isang lalaki.*
        d⟨um⟩atiŋ   aŋ=isa=ŋ   lala:ki
        ⟨AV.INIT⟩arrive   NOM=one=LNK man
        'A man arrived.'

However, in competition with the NOM+one construction, we must at least include the existential construction in (6).

(6)   *May dumating na lalaki.*
        may d⟨um⟩atiŋ   na   lala:ki
        EXT ⟨AV.INIT⟩arrive   LNK   man
        'A man arrived.'

Any Tagalog speaker will likely say that both (5) and (6) are perfectly grammatical and that they "mean the same thing", and here our diffculties begin. Taking (5) at face value, Paul et al. (2015) and Collins (2019) treat *ang* and *ng* as pure case markers whose relation to definiteness and specificity is only indirect. The definite reading of the *ang* phrase is derived by covert movement to a pragmatically prominent position in the A' layer (ala Richards, 2000; Pearson, 2005 inter alia). When that movement is blocked or the definite reading is otherwise countered by an overt quantifier, an indefinite reading is freely available. The other side of the coin is that when an argument remains in its base position within VP (or vP) then it receives an (existential) indefinite interpretation (as in Diesing, 1992). It is uncontroversial that a bare NP object of an AV predicate receives an existential indefinite interpretation not only in Tagalog but also in the majority of Philippine languages. On this basis, Collins (2019) similarly proposes that transitive verbs in Tagalog are inherently quantificational in a fashion similar to object-incorporating verbs in languages such as Greenlandic (Van Geenhoven, 1998). The empirical lynchpin in this syntactic analysis of definiteness is the putative indefiniteness of weakly quantified *ang* phrases, which appear to cancel out any definiteness effects. Collins provides evidence for this in the form of examples such as (7).

---

[5] In short, Collins (2019) employs type-shifting operations triggered by type mismatch in particular syntactic configuration in order to derive the indefinite reading of VP-internal objects and the definite readings of pivots, which are VP-external.

(7) Context: The teacher is running a seminar in which six students signed up:

a.   I-p⟨in⟩asa   ng   guro   ang   mag–aaral.
           PV-⟨INIT⟩.pass GEN   teacher NOM   student
           'The teacher passed the student.'
           Consultant response: Not with six students, it sounds wrong.

b.   I-p⟨in⟩asa   ng guro ang isa–ng   mag–aaral.
           PV–⟨INIT⟩.pass GEN teacher NOM one-LNK   student
           'The teacher passed one student.'
           Consultant response: Fine, it sounds like five of them failed. (Collins, 2019:ex.21)

Here, the infelicity of (7a) in contrast to (7b) suggests that *ang isang* obtains an indefinite reading. The failure of (7a) is due to the fact that *ang mag-aaral* 'the student' cannot be uniquely identified within the context. No such requirement is made of the indefinite phrase *ang isang mag-aaral*. The interpretation of (7b) is relatively uncontroversial. However, Collins provides further examples in an attempt to show that not only are NOM+one expressions indefinite but they may also receive an existential interpretation. One piece of evidence he provides is the minimal triplet in (8).

(8)a.   Ma-i~inis   si   Mary kung mag-pa~pa-tugtog si   John ng   rekord
          STA-IMPF~mad NOM Mary if   AV-IMPF~CAU-play NOM John GEN record
          'Mary will be annoyed if John plays a record (any record).'
   b.   Ma-i~inis   si   Mary kung i-pa~pa-tugtog   ni   John ang rekord
          STA-IMPF~mad NOM Mary if   CV-IMPF~CAU-play GEN   John NOM record
          'Mary will be annoyed if John plays the record.'
   c.   Ma-i~inis   si   Mary kung i-pa~pa-tugtog   ni   John ang isang   rekord
          STA-IMPF~mad NOM Mary if   CV-IMPF~CAU-play GEN   John NOM one:LNK   record
          'Mary will be annoyed if John plays a record.'
          Collins: "Any record in general"

While the interpretation of (8a) and (b) should be uncontroversial, the interpretation reported for (c), I believe, is inaccurate. It is, for instance, a perfectly natural follow-up to such a sentence to ask, *Aling rekord iyan?* 'Which record is that?', where the anaphoric demonstrative refers back to the record but this should not be possible if *ang isang rekord* receives an existential interpretation. This is parallel to the infelicity in the English (9), where the object in A's statement would normally be interpreted as an existential indefinite.

(9)   A: Mary will be annoyed if John plays a record.
          B: #Which record is that?

Collins (2019) further presents an array of "naturalistic" examples culled from the internet which are meant to demonstrate the existential indefinite reading of NOM+one phrases. Only some of these are relevant for the distinction between an existential and specific reading as most of the contexts are equally compatible with either reading while others are more compatible with a specific reading. For instance, (10) is easily compatible with the interpretation 'I bought a particular small book at Biola Bookworm' and (11), with the interpretation involving 'a particular parable'.

(10)   *Binili ko ang isang maliit na aklat sa Biola Bookworm.*
       b⟨in⟩ili-∅=ku   aŋ=isa=ŋ   ma-li?it na aklat sa=*Biola Bookworm*
       ⟨INIT⟩buy-PV=1s.GEN NOM=one=LNK ADJ-small LNK book OBL=Biola Bookworm
       'I bought a small book at Biola Bookworm.'

(11)   *Isinalaysay ni Jesus ang isang talinhaga upang ituro sa kanila na dapat silang laging manalangin...*
       i-s⟨in⟩alaysay ni=Jesus   aŋ=isa=ŋ   talinha:ga? upaŋ i-tu:ru? sa=kanila na
       CV–⟨INIT⟩tell GEN=Jesus NOM=one=LNK parable   to   CV-teach OBL=3p LNK

       da:pat=sila=ŋ   la:gi=ŋ   manalaŋin...
       should=3p=LNK always=LNK AV:pray
       'Jesus recounted a parable in order to teach them that they must always pray...'
       (Lukas 18:1)

We can thus set such examples aside and focus on those cases which are incompatible with a specific reading. The examples in (12) and (13) provide such cases. The contexts of both make it clear that an existential interpretation is intended.

(12)   *Subali't hindi ko nakita ang isang larawan ng aking sarili.*
       suba:lit hindi?=ku   na-ki:ta?   aŋ=isa=ŋ   lara:wan naŋ=a:kiŋ   sari:li
       but   NEG=1s.GEN STA.PRF-see   NOM=one=LNK image   GEN=1s.GEN:LNK self
       'But I didn't see a reflection of myself.'

(13)   *Ano ang dapat kong gawin kung nakaligtaan ko ang isang dosis?*
       ano aŋ=da:pat=ku=ŋ   gaw-in kuŋ na-ka-ligta?-an=ku   aŋ=isa=ŋ   dosis
       what NOM=should=1s.GEN=LNK do-PV if   INIT-STA-omit-LV=1s.GEN NOM one:LNK dose
       'What should I do if I miss a dose?'

We find however that (12) and (13) and, indeed, all the examples meant to show existential indefinite interpretations are infelicitous in their intended meaning and rather strongly favor a specific or definite reading; (12), from a religious tract entitled "Hell is Real: I've been there!", describes the protagonist looking into the mirror and not seeing her reflection. However, rather than 'But I didn't see a reflection of myself', the Tagalog translates more accurately to 'But I

didn't see the other image of myself' or 'I didn't see one of my images', in other words, either a definite or partitive (specific) interpretation.

But whence such "naturally ocurring" data? Tracking down the sources of Collins' "naturalistic" evidence, we find that all of the questionable examples are translations from English, produced either by a human translator or a machine.[6] For instance, an example like (14), from a different religious tract, belonging to the Jehovah's Witnesses, was most likely translated by a native speaker. It is revealing to compare the given translation (likely a back translation made with the help of a native speaker) with the original English source of the Tagalog, given beneath.

(14)　Maingat na　p⟨in⟩i~pili-∅　　　　ng　gagamba ang　isang　　dahon,
　　　careful　LNK ⟨INIT⟩IMPRF~choose-PV GEN spider　　NOM one-LNK leaf,

　　　marahil　mula sa　mga nakalapag sa　lupa.
　　　probably from OBL PL　fallen　　OBL ground

　　　Collins (2019) translation: 'Carefully the (leaf-curling) spider chooses one leaf, probably from ones fallen on the ground.'

　　　Original English: 'The spider carefully selects a leaf, perhaps from among those lying on the ground below.'[7]

The point of contention here is the interpretation of *ang isang dahon* as either an existential indefinite ('any leaf') or as presuppositional ('a particular leaf'). The original English clearly presents the object as a partitive of the larger set, 'those lying on the ground below'. Due to referential properties of the oblique case marker *sa,* the only interpretation available to the Tagalog phrase *sa mga makalapag sa lupa* is also definite despite lacking a demonstrative. The patient in the translation was thus likely expressed as *ang isang dahon* due to the partitive meaning in the source, i.e. 'one of the leaves lying on the ground below'. We can conclude that none of Collins' examples successfully demonstrate the possibility of an existential indefinite reading for NOM+one phrases. They do however appear to freely denote specific indefinites but, as we will see, even this reading is highly restricted and requires syntactic licensing within a particular sociolinguistic register. In the next section, I present further arguments for the referential role of the case markers.

4.0　Case markers as case marking determiners

4.1　The role of syntactic context

One of the most serious problems for the proposal that the case markers have no semantic contribution emerges from the simplest contexts of all. The following fragment exclamations are

---

[6] Cases like (12) are so egregious as a translation of the English that it is hard to imagine it could be the work of a native speaker.

[7] From: https://wol.jw.org/en/wol/d/r1/lp-e/102002528?q=leaf+curling+spider&p=par

unambiguous in their definiteness, as indicated in the translations. An exclamation such as (15a) is only possible when there is no identifiable rat and (15b) is only possible when a particular rat has already been established in the discourse. Note that (15c), the *ang isa* construction, cannot obtain a specific indefinite interpretation here but rather takes on a new meaning where *isa* is interpreted as 'other' and *ang* contributes the usual definite reading.

(15)a. *Daga!*   b.   *Ang daga!*   c.   *%Ang isang daga!*
    daga?   aŋ=daga?   aŋ=isa=ŋ   daga?
    rat   NOM=rat   NOM=one=LNK rat
    'A rat!/*The rat!'   'The rat!/*A rat!'   (OK for 'The other rat!')

When the referent in the fragment is uniquely identifiable by virtue of context or general world knowledge, *ang* is obligatory, as seen in (16), where it is forced by the knowledge that 'president' is a unique position held by an individual. We find, therefore, that in an out of the blue context with no plausible ellipsis, only the case marking determiners could be responsible for the attested differences in interpretation.

(16) a. *Ang pangulo!*   b. *%Pangulo!*
    aŋ=paŋu:lu   paŋu:lu
    NOM=president   president
    'The president!'   (For, 'The president!')

As first noted by Adams and Manaster-Ramer (1988), it is not just *isa* which appears to override the definiteness of *ang*, but a range of weak quantifers, as seen in (17).

(17)   *Dumating ang isang/maraming/ilang dyip.*
    d⟨um⟩atiŋ   aŋ=isa=ŋ/ma-da:mi=ŋ/ilaŋ   dyip
    ⟨AV.INIT⟩arrive   NOM=one:LNK/ADJ-many=LNK/few:LNK   jeep
    'A jeep arrived. / Many jeeps arrived. / A few jeeps arrived.'

However, Adams and Manaster-Ramer (1988) also notes that *isa* is distinct from other cardinals, which do not have the same effect. In (18), the same type of nominative phrase can only receive a definite interpretation, unlike (17).

(18)   *Dumating ang dalawang dyip.*
    d⟨um⟩atiŋ   aŋ=dalawa=ŋ dyip
    ⟨AV.INIT⟩arrive NOM two:LNK   jeep
    'The two jeeps arrived.' (NOT: 'Two jeeps arrived.')

Nonetheless, for all these cases, once the *ang* phrase is removed from its clausal context, the

ability to obtain an indefinite reading is lost, as seen in B's responses in (19).

(19) A: *Sino ang dumating kanina?*
　　 sinu　　　aŋ=d⟨um⟩atiŋ　　　kaniːna
　　 who.NOM NOM=⟨AV.INIT⟩arrive earlier
　　 'Who arrived earlier?'

　B: i.　*Ang isang babae.*　　ii.　*Ang maraming babae.*　　iii.　*Ang ilang babae.*
　　　 aŋ=isaːŋ　　babaːʔi　　aŋ=ma-daːmi=ŋ　　babaːʔi　　aŋ=ilaŋ　　babaːʔi
　　　 NOM=one=LNK woman　　NOM=ADJ-many=LNK woman　　NOM=some:LNK woman
　　　 'The other woman.'　　'The many women.'　　　'Some of the women.'
　　　 (NOT: 'One woman.')　　(NOT: 'Many women.')　　(NOT: 'Some women.')

This argues strongly for the indefinite interpretation of pivots as an artifact of certain syntactic contexts rather than a freely available option for weakly quantified phrases. Conversely, this evidence also argues for a definite determiner function of the nominative case marker.[8]

4.2　Voiceless predicates

A small number of common verbs in Tagalog, including *kailangan* 'need' and *gusto* 'want', do not require voice to form a declarative predication. With bare predicates such as these, the experiencer takes the genitive case and the theme takes either *ang* or *ng* purely based on its definiteness, with no regard to role or grammatical function, as seen in (20) and (21).

(20) a.　*Kailangan ko ng susi.*　　　　b.　*Kailangan ko ang susi.*
　　　 kaʔilaːŋan=ku naŋ=suːsi　　　　　 kaʔilaŋan=ku　aŋ=suːsi
　　　 need=1s.GEN　GEN=key　　　　　　 need=1s.GEN　NOM=key
　　　 'I need a key.'　　　　　　　　　 'I need the key.'

(21) a.　*Gusto ko ng asul.*　　　　　b.　*Gusto ko ang asul.*
　　　 gusto=ku　　naŋ=asul　　　　　 gustu=ku　　aŋ=asul
　　　 like=1s.GEN GEN=blue　　　　　 like=1s.GEN NOM=blue
　　　 'I like blue/a blue one.'　　　　 'I like the blue one.'

Nominal predicates offer another context where voice is removed from consideration and here too we find a pattern as above. In canonical copular clauses such as (22), a bare noun phrase predicate like *problema* can only be interpreted indefinitely.

---

[8] As Tom Payne points out (p.c.), there are Bisayan languages that have two varieties of nominative marking which are said to differ in referentiality. It is unclear whether these languages represent real exceptions to the generalization that Philippine nominative/pivot arguments receive a definite interpretation.

(22)   Problema iyan
        problem   that.NOM
        'That is a problem.'

The same NP, when introduced by *ang* can only be interpreted definitely, as shown in (23a). (In this case, the demonstrative must be in the predicate position for reasons that are not relevant here, but see Kaufman 2018). Crucially, as seen in (23b), the addition of *isa* does not derive an indefinite interpretation, as seen in (23b), where the 'other' interpretation again rears its head.

(23)a. *Iyan ang problema.*   b. *Iyan ang isang problema.*
       iyan   aŋ=problema   iyan   aŋ=isa=ŋ   problema
       that.NOM NOM=problem   that.NOM NOM=one=LNK problem
       'That is the problem.'   'That is the other problem.' (Not: 'That's a problem.')

We can conclude from all the above that the default interpretation of *ang* marked phrases is definite, indicating a referent that can be uniquely identified by the hearer. The indefinite readings of *ang* phrases are licensed in particular contexts but even in these contexts we only find specific rather than existential indefinite interpretations. In other words, speaker-based identifiability is common to all contexts whereas hearer-identifiability can be overridden under certain circumstances. This stands in contradistinction to Collins' analysis, in which it is clausal syntax that licenses the definite reading of *ang*. Here we have seen that it is clausal syntax which licenses the indefinite reading, with the definite reading being expressed by the *ang* marker itself.

4.3   The role of discourse context

While in some of the contexts offered by Collins, a specific indefinite reading is possible, it must be noted that other contexts, as in (28), which are similar on the surface, do not allow an indefinite interpretation of any kind. Here, we again find the 'other' meaning of *isa*.

(28)A: *Anong ginawa mo sa tindahan?*
        ano=ŋ   g⟨in⟩awa-ø=mu   sa=tindahan?
        what=LNK ⟨INIT⟩do–PV=2s.GEN OBL=store
        'What did you do in the store?'

   B: *Binili ko ang isang saging.*
        b⟨in⟩ili–ø=ku   aŋ=isa=ŋ   sa:giŋ
        ⟨BEG⟩buy–PV=1s.GEN NOM=one=LNK banana
        'I bought the other banana.' (NOT: 'I bought a banana')

Collins' example, repeated here in abbreviated form as (29), appears to allow a specific indefinite reading within a similar syntactic structure.

(29)   Context: The teacher is running a seminar in which six students signed up:

   Ipinasa ng guro ang isang mag-aaral.
   i-p⟨in⟩asa   naŋ=guruʔ aŋ=isa=ŋ   magʔaʔaral
   CV-⟨INIT⟩pass GEN=teacher NOM=one=LNK   student
   'The teacher passed one student.'   (Collins 2019:ex.21)

The crucial difference here is that the context in (29) naturally sets up a partitive reading in which the teacher passes one of the six previously introduced students. In (28), a partitive reading is not easily accommodated and the definite interpretation is thus essentially obligatory.

It must be emphasized that examples cited in the literature which could be plausibly interpreted with a specific interpretation have been asserted without sufficient evidence to allow an existential reading. For instance, Paul et al. (2015:372) assert that the subject in (30) can be completely novel to both the hearer and listener, but there is nothing in the context that rules out a reading in which the jeep is uniquely identifiable by the speaker, i.e. 'a certain jeep is arriving'.

(30)   *Dumadating ang isang dyip.*
   d⟨um⟩a:~datiŋ   aŋ=isa=ŋ   dyip
   ⟨AV⟩IMPRF~arrive NOM=one=LNK jeep
   'A jeep is arriving.'

Similarly, the same authors claim that *ang isang isda* in (31) can have an existential indefinite reading, but there is nothing in the context that would require it.

(31)   *Kailangang kainin ni Pedro ang isang isda.*
   kailaŋaŋ   kaʔin-in ni=Pedro   aŋ=isa=ŋ   isdaʔ
   need:LNK eat-PV   GEN=Pedro NOM=one=LNK fish
   'Pedro needs to eat a fish' (a specific fish or any fish at all)

When we refine the context to strongly favor an existential reading, we find that speakers clearly distinguish the interpretation of NOM+one phrases from that of a genitive AV object. In (32), we find a typical existential indefinite reading, in which the agent is seeking a spouse without any specific individual in mind.

(32)   *Naghahanap siya ng asawa.*
   nag-ha:~hanap=siya   naŋ=asa:wa
   AV-IMPRF~search=3s.NOM   GEN=spouse
   'S/he's looking for a spouse.'

In contrast, (33) cannot be interpreted similarly. Rather it can only be interpreted with the agent

seeking the other spouse or one of some set of aforementioned spouses. This shows that once we use contexts where the existential and specific readings have very different implications, the semantic contribution of *ang* becomes clear.

(33)   *Hinahanap niya ang isang asawa.*
        h⟨in⟩aː-haːnap-∅=niya   aŋ=isa=ŋ   asaːwa
        ⟨INIT⟩IMPRF~search-PV=3.GEN NOM=one=LNK spouse
        'S/he is looking for the other spouse/one of the spouses.'
        NOT: 'S/he's looking for a spouse.'

Aspect and mood also play a role here. For instance, in a plain episodic context with perfective aspect as in (34a), it is not immediately obvious whether the pivot may refer to any teacher at all or a particular teacher identifiable to the speaker. However, judgments are sharpened in an irrealis context such as (34b), which cannot obtain the interpretation 'A teacher (any teacher) might complain'. Rather, we find either the familiar definite or specific indefinite interpretations.

(34)a. *Nagreklamo ang isang guro.*
        nag-reklamo   aŋ=isa=ŋ   guruʔ
        INIT.AV-complain NOM=one=LNK teacher
        'A teacher complained.'
    b. *Baka magreklamo ang isang guro.*
        bakaʔ   mag-reklamo   aŋ=isa=ŋ   guruʔ
        perhaps AV.INF-complain NOM=one=LNK teacher
        'The other teacher might complain.'/'A particular teacher might complain.'

The key difference between these two contexts is that in (34a), by virtue of having taken place, a particular teacher did complain, whether anyone in the discourse can identify that teacher or not. The difference between an existential and specific reading of this type of simple sentence is thus inherently subtle. In (34b), on the other hand, where the event is irrealis and there is no particular teacher that has already complained, it is clear that the speaker is not referring to any teacher at all but rather has a specific one in mind.

4.4.   Indefinite pronouns and negative *ni*

Despite all the above, certain cases pose real problems to the idea that *ang* phrases always receive at least a specific indefinite interpretation if not the default definite interpretation. The two clearest cases involve indefinite pronouns and *ang* phrases containing the Spanish loan *ni* 'neither'. We can see examples of the first type in (35) and (36), where the two *ang* phrases receive a low scope indefinite interpretation, i.e. 'anyone' and 'anything', rather than the specific 'a particular person/thing'.

(35)   *Hindi ko sinisi ang sinuman.*
        hindiʔ=ku   s⟨in⟩isi-∅   aŋ=sinu-man.
        NEG=GEN.1S   ⟨INIT⟩blame-PV NOM=NOM.who-even
        'I didn't blame anyone.'

(36)   *?Hindi ko kinain ang anuman.*
        hindiʔ=ku   k⟨in⟩aːʔin-∅   aŋ=anu-man.
        NEG=GEN.1S   ⟨INIT⟩eat-PV   NOM=what-even
        'I didn't eat anything.'

Note that neither (35) nor (36) represent the regular, unmarked way of expressing the given propositions and that (36) appears far more unnatural than (35). The more typical way of expressing indefinites under the scope of negation is shown in (37), with the negative existential *wala* as a matrix predicate.

(37)a. *Wala akong sinisi.*
        walaʔ=aku=ŋ   s⟨in⟩isi-∅
        NEG.EXT=1s.NOM=LNK ⟨INIT⟩blame-PV
        'I didn't blame anyone.'
    b. *Wala akong kinain.*
        walaʔ=aku=ŋ   k⟨in⟩aːʔin-∅
        NEG.EXT 1s.NOM=LNK ⟨INIT⟩eat-PV
        'I didn't eat anything.'

Nonetheless, we find examples of both in formal speech and writing. An example of a low scope indefinite *ang* phrase from one of the more recent Bible translations is shown in (38).

(38)   *Huwag ninyong gawin ang anuman dahil sa pansariling hangarin o dahil sa kayabangan.*
        huwag=ninyo=ŋ   gaw-in aŋ=anu-man   daːhil sa=pan-sarili=ŋ   haŋaːd-in
        NEG.IMPER=2p.GEN=LNK do-PV NOM=what-even because OBL=INSTR-self=LNK aim-PV
        o daːhil   sa=ka-yabaːŋ-an
        or because OBL=STA-conceit-LV
        'Do nothing out of selfish ambition or vain conceit.'
                                                    (Ang Bagong Tipan: Standard Filipino Version, Philippians 2:3)

The second case of unambiguous low scope indefinites mentioned above employs the negative element *ni.* We find a clear difference in interpretation between nominative phrases expressed by *ang isa* and *ang ni isa* under negation, as seen in (39) and (40).

(39) *Hindi ko sinisi ang isang estudyante.*
hindiʔ=ku   s⟨in⟩iːsi-∅   aŋ=isa=ŋ   estudyante
NEG=GEN.1S   ⟨INIT⟩blame-PV   NOM=one=LNK   student
'I didn't blame a particular student.'   (∃ > NEG)
NOT: 'I didn't blame a single student.'   (NEG > ∃)

(40) *Hindi ko sinisi ang ni isang estudyante.*
hindiʔ=ku   s⟨in⟩iːsi-∅   aŋ=ni   isa=ŋ   estudyante
NEG=GEN.1S   ⟨INIT⟩blame-PV   NOM=NEG   one=LNK   student
'I didn't blame a single student.'   (NEG > ∃)
NOT: 'I didn't blame a particular student.'   (∃ > NEG)

In (39), only a specific reading is available while in (40) only the low-scope reading is available, yielding two very different interpretations. In this sense, *ang ni isa* behaves like English 'single' under negation while *ang isa* behaves like 'particular'.

I tentatively offer two distinct explanations for the anomalous behavior of *ang* phrases with indefinite pronouns and *ni* in this regard. Recall that, while both are well attested in writing and more formal language, there appears to be a difference in acceptability between the behavior of *sinuman* 'whoever' and *anuman* 'whatever', with *sinuman* appearing more naturally under the scope of negation. (As mentioned above, there is a tendency to avoid such structures altogether in the daily language in favor of clauses taking the negative existential as the matrix predicate.) I attribute this to an inherent definiteness on the part of *sinuman* due to its personal case marker component *si*. In other words, *sinuman* is semantically indefinite but morphosyntactically definite. As shown in (41), personal name arguments are similar to pronouns in that they cannot serve as genitive objects of actor voice verbs. Rather, when a personal name or pronouns appears in this function, it must take the oblique case instead of the expected genitive case.

(41) *Hindi ako humalik kay/*ni Linda.*
hindiʔ=aku   h⟨um⟩alik   kay/ni   Linda
NEG=1s.NOM   ⟨AV.BEG⟩kiss   OBL/GEN Linda
'I didn't kiss Linda.'

Similarly, the indefinite pronoun cannot take the genitive case as the object of an actor voice predicate, as seen in (42), and instead must take the oblique, just like a definite personal or pronominal argument.

(42) *Hindi ako humalik kaninuman/*ninuman.*
hindiʔ=aku   h⟨um⟩alik   kaniːnu-man/niːnu-man
NEG=1s.NOM   ⟨AV.BEG⟩kiss   OBL:who-even/GEN:who-even
'I didn't kiss anyone.'

Note that this contrasts with *anuman* 'whatever/anything', which remains in the genitive case when funcitoning as an actor voice object, as seen in (43).

(43)   *Hindi ako bumili ng/*sa anuman.*
     hindiʔ=aku   b⟨um⟩ili   naŋ/sa   anu-man
     NEG=1s.NOM ⟨AV.BEG⟩buy   GEN/OBL what-even
     'I didn't buy anything.'

It is this morphosyntactic definiteness which plausibly licenses constructions such as (35), with *sinuman*, as opposed to (36), with *anuman*. It is also possible that indefinite expressions such as *sinuman* and *anuman* distinguish presuppositional and non-presuppositional interpretations. In this case, the difference between *anuman* as an AV object and *anuman* as a pivot may be parallel to the English free translations in (44), although exploring this in more depth must be left to further work.

(44)a.   *Kumuha ka ng anumang gusto mo.*
     k⟨um⟩u:ha=ka   naŋ=anu-maŋ   gustu=mu
     ⟨AV⟩take=2s.NOM   GEN=what-even:LNK want=2s.GEN
     'Take whatever you want.' (non-presuppositional)
   b.   *Kunin mo ang anumang gusto mo.*
     kun-in=mu   aŋ=anu-maŋ   gustu=mu
     take-PV=2s.GEN   NOM=what-even:LNK want=2s.GEN
     'Take whatever it is you want.' (presuppositional)

     We can return now to the other exceptional low scope *ang* phrase, occurring with the negator *ni*, as seen earlier in (40). The negator *ni* appears to override the usual scope and definiteness properties associated with the *ang* phrase. As seen below, we obtain much the same reading in both the patient voice, as in (45a), and the actor voice, as in (45b), when an undergoer is marked with *ni*.

(45)a.   *Hindi ko ininom ang ni isang patak ng alak.*
     hindiʔ=ku   ⟨in⟩inom-∅   aŋ=ni   isa=ŋ   patak naŋ=a:lak
     NEG=1s.GEN ⟨BEG⟩drink–PV   NOM=NEG one=LNK   drop   GEN alcohol
     'I didn't drink a drop of alcohol.'
   b.   *Hindi ako uminom ng ni isang patak ng alak.*
     hindiʔ=aku   ⟨um⟩inom   naŋ=ni   isa=ŋ   patak naŋ=a:lak
     NEG=1s.NOM ⟨AV.BEG⟩drink GEN=NEG one=LNK   drop   GEN alcohol
     'I didn't drink a drop of alcohol.'

While the low scope reading of the undergoer is expected in the actor voice, it is unexpected in the patient voice, where the indefinite argument is expressed by an *ang* phrase. Key to understanding the exceptional behavior of *ni* is that it is a Spanish borrowing, one which has made its way into many indigenous languages of the Philippines as well as those of Latin America. In Spanish, *ni* generally co-occurs with a higher negator, such as *no*, as seen in (46), in a pattern that resembles negative concord.

(46)   *No queda   ni   una gota de vino.*
       NEG remain.3s.PRES   NEG one drop of wine
       'There isn't a drop of wine left.'

Many Philippine languages appear to have calqued this structure, hand in hand with borrowing the negator *ni*, as exemplified by (47).

(47)   *Walang natira, ni isang patak ng alak.*
       wala=ŋ   na-tira   ni isa=ŋ   patak naŋ=aːlak
       NEG.EXT STA.BEG-remain NEG one=LNK drop   GEN=alcohol
       'There isn't a drop of alcohol left.'

That the syntax of *ni* was borrowed in tandem with the lexeme can be seen in the ungrammatical (48), which shows that, just as with Spanish postverbal arguments, Tagalog *ni* needs to fall within the scope of a higher negator to be licensed.

(48)   **Ininom ko ang ni isang patak ng alak.*
       ⟨in⟩inom–ø=ku   aŋ=ni   isa=ŋ   patak naŋ=aːlak
       ⟨BEG⟩drink–PV=1s.GEN   NOM=NEG   one=LNK   drop   GEN alcohol

Remarkably, *ni* marked arguments can even be topicalized, as in (49), while maintaining their low scope interpretation.

(49)   *Ni isang salita, wala akong narinig mula sa iyo.*
       ni isa=ŋ   salitaʔ walaʔ=aku=ŋ   na-dinig   mulaʔ sa=iyu
       NEG one=LNK word   NEG.EXT=1s.NOM=LNK STA.INIT-hear from   OBL=2s
       'I didn't hear a single word from you.'

This, too, resembles Spanish, as seen in (50) (although note that in Spanish the fronting of a *ni* marked object obviates the need for a higher negator).

(50)   *Ni una palabra me   han   dicho.*
       NEG one word   1s.ACC PERF.3p say.PRTC
       'They haven't said a single word to me.'

Having established that *ni* was borrowed along with certain syntactic properties, it becomes easier to understand why it can appear as the *ang* phrase of one of the undergoer voices in Tagalog, as this could represent a calque from a Spanish transitive clause, as in (51).

(51)   Bueno, imagino   que   alcohol no   probarás   ni   una gota.
        well   imagine.1s.PRES COMP alcohol NEG try.2s.FUT NEG one drop
        'Well, I imagine that alcohol, you wouldn't try a single drop.'

                                                          (http://elvira.lllf.uam.es/docs_es/corpus/corpus.html)

We can refer to the principle whereby an undergoer clause in a Philippine language is influenced by its transitive counterpart in Spanish and English as "translational congruence", which may play a key role in the distribution of non-canonical low scope indefinites in Tagalog, both with and without *ni*. Recall from (38) that a clause like 'Do nothing' was translated as *Huwag ninyong gawin ang anuman* in one standard translation of the New Testament despite the rarity of this usage in the spoken language. Recall also that the majority of the examples in Collins (2019) containing low scope *ang isa* were found to be translated from English sources. I posit that translational congruence is at play in all these examples.

 Although this provides a vista for explaining much of the above data, we must note that native functional items may behave similarly to borrowed ones in the modern language. For instance, the native word *kahit* 'even' patterns much like *ni*, allowing for a low scope reading inside of an *ang* phrase, as seen in (52a), and requiring a higher negator, as seen in (52b).

(52)a. *Hindi ko ininom ang kahit isang patak ng alak.*
        hindiʔ=ku   ⟨in⟩inom-∅   aŋ=ka:hit   isa=ŋ   patak naŋ=a:lak
        NEG=1s.GEN   ⟨BEG⟩drink-PV   NOM=even   one=LNK   drop GEN alcohol
        'I didn't drink even one drop of alcohol.'
   b. **Ininom ko ang kahit isang patak ng alak.*
        ⟨in⟩inom-∅=ku   aŋ=kahit   isa=ŋ   patak naŋ=a:lak
        ⟨BEG⟩drink-PV=1s.GEN   NOM=even   one=LNK   drop GEN alcohol

However, some Tagalog speakers appear to generally judge examples with *ni* or *kahit* under the plain negator *hindi* as in (45) or (52) as awkward, preferring constructions with the negative existential in their place, as in (47) and (49). I do not know how widespread this pattern is but it may represent a more conservative grammar that does not allow existential readings for pivots even in these exceptional cases.

 Returning to the concept of "translationese" and translational congruence, we find that some of the examples cited in favor of existential pivots display multiple hallmarks of this register. For instance, in (53), cited by Collins (2019) to demonstrate a low scope indefinite *ang* phrase, we find several characteristics that are rare in speech but commonly employed in newspaper headlines: (i) the indefinite argument *unggoy* is atypically topicalized in the first

clause; (ii) the topicalized *unggoy* lacks case marking despite being a primary argument of the clause; (iii) the topic atypically corresponds to a null pronoun in the second clause (cf. Himmelmann, 1999; Nagaya 2006).

(53) *Unggoy nakawala, kinagat ang isang bata.*
      uŋguy   naka-walaʔ   k⟨in⟩agat-∅   aŋ=isa=ŋ   baːtaʔ
      monkey AV.POT–escape ⟨INIT⟩bite–PV NOM=one=LNK child
      'Monkey escapes, bites a child.'

These features mimic non-Philippine syntax in different respects, as can be seen in the structural similarity (53) to the English translation. We leave it to future work to see if all exceptional *ang* indefinites could have plausibly spread from Spanish and English models.

5.0   The textual evidence

5.1   A glimpse at the historical picture

If I am correct in attributing the use and interpretation of indefinite *ang* phrases in part to language contact, we should be able to see its emergence in the historical record, as several Philippine languages are richly attested over a 300-400 year period. I cannot attempt to carry out the complex philological work required here, but only provide a first baby step in that direction, which does appear to provide some encouraging evidence with regard to *ang isa*.

The first work published in Tagalog, the 1593 *Doctrina Cristiana*, contains no instances at all of *ang isa*. Fast forwarding three hundred years to Francisco 'Balagtas' Baltazar's (1838) most famous work, *Florante at Laura*, we find only four instances of *ang isa*, shown below (untranslated):

(54) *At saka madalas ilalâ ng tapang,*
      ay ang guniguning takot ng kalaban,
      ang isang guerrerong *palaring magdiwang*
      mababalita na at pangingilagan.

(55) *Di nag iláng buwan ang sa Reynong tuwà*
      at pasasalamat sa pagka-timawa,
      *dumating **ang isang hukbong** maninira*
      ng taga-Turkyang masakim na lubha.

(56) *Nang gabing malungkot na kinabukasan*
      wakas na tadhanang ako'i pupugutan,
      *sa carcel ay nasok **ang isang general***
      dala ang patawad na laong pamatáy.

(57)   *Nang paghanaping ko'y ikáw ang nataós*
         pinipilit niyaóng táong balakyót,
         hindi ko nabata't bininit sa búsog
         ang isang palasóng *sa lilo'y tumapos…*

None of these examples require a low scope existential interpretation as all are compatible with a specific (i.e. 'a particular') reading. Fast-forwarding once again, we can observe Patricio Mariano's 1911 translation of Jose Rizal's (1891) *El Filibusterismo* from the original Spanish, where we find 220 instances of *ang isa*, including certain cases which can only be interpreted as low scope existentials. In (58), we see Rizal's original sentence, in which an object, 'a single handspan of his land', receives a low scope indefinite interpretation.

(58)   *Y   se   negó   resueltamente á pagar   ni   á ceder*
         CONJ REFL refuse.3s.PRET resolutely   to pay:INF nor to cede:INF

         un palmo siquiera de sus   tierras,
         a   palm   even   of 3s.GEN land:PL
         'And he refused to pay or cede not a single handspan of his land.'

In (59), we find Mariano's Tagalog translation, in which this argument is translated with the phrase *ang isang dangkal*, as shown.

(59)   *At nagmatigás na sa hindî pagbabayad ni ibigáy ang isáng dangkal man lamang ng*
         kaniyáng lupà.
         at   nag-ma-tigas   na   sa=hindiʔ pag-ba~bayad ni   i-bigay
         CONJ   AV.INIT-ADJ-hard LNK OBL=NEG GER-TR~pay   NEG CV-give

         aŋ=isa=ŋ   daŋkal=man=la:maŋ naŋ=kanya=ŋ   lu:paʔ
         NOM one=LNK handspan=even=only GEN=3.OBL=LNK land
         'And he refused to pay or cede not a single handspan of his land.'

Again, in (60), we see Rizal's original Spanish with a low scope indefinite object, *á una madre*, and in (61) we find Mariano's translation employing the phrase *ang isang ina*.

(60)   *¿qué cosa hay más sagrada que   alimentar á una madre?*
         what thing EXT more sacred   COMP feed:INF to a   mother
         'What is more sacred than feeding a mother?'

(61)   *¿mayroon pa bang kabanal-banalang bagay na gaya nğ pakanin ang isáng iná?*
         mayruʔun=pa=ba=ŋ   ka-banal~bana:l-aŋ   ba:gay na   ga:ya
         EXT=still=QM=LNK holy:LNK STA-INTNS~sacred-LV:LNK   thing LNK like

         naŋ=pa-ka:n-in   aŋ=isa=ŋ   ina
         GEN=CAU-eat-PV   NOM=one=LNK mother
         'What is more sacred than feeding a mother?'

Why are clauses such as these rendered in an undergoer voice, forcing a low scope argument into an *ang* phrase, rather than in the actor voice (i.e. *magbigay ng isang dangkal man lamang*, *magpakain ng isang ina*), where they could naturally avail of the low scope reading common to actor voice objects? If the principle of translational congruence equates transitive clauses in English and Spanish with undergoer clauses in Philippine languages, as is plausible, then we may be able to explain such examples and why they are so much more common in writing than in natural speech.

5.2   A glimpse at the cross-linguistic picture

If translational congruence is a major factor behind indefinite pivots in Philippine languages we could furthermore expect such pivots to be even more marginal in languages that have been subject to less Spanish and English influence (either directly or through translation). As one of many possible examples, we can examine Maguindanawn, a Danaw language of western Mindanao with a different history of language contact in which English and Spanish played a lesser role.[9] In Maguindanawn, a novel indefinite subject of an intransitive clause would generally be introduced via the existential *aden* as a matrix predicate, as seen in (62a). However, the NOM+one construction, shown in (62b), is also possible.

(62)a. Aden   babay   a   naka-uma.
      EXT   woman LNK   AV.NVOL-arrive
      'A woman arrived.' (Lit: 'There is a woman who arrived.')

    b. Naka-uma   i   sakataw a   babay.
       AV.NVOL-arrive NOM   one   LNK   woman
       'A woman arrived.'

While Tagalog speakers may tend judge the existential construction and the NOM+one construction as equivalent in simple intransitive clauses such as the above, the Maguindanawn existential in (62a) was judged to be more natural while that in (62b) was seen to be more formal, as might be expected if the NOM+one construction originated or spread through colonial languages.

As noted earlier, the scenario described in (62) above is compatible with both an existential reading (i.e. a woman unknown to both hearer and speaker arrived) as well as a specific reading (i.e. a woman familiar to the speaker but novel to the hearer arrived). When we use a less ambiguous example, we find that Maguindanawn rejects an indefinite interpretion of NOM+one expressions altogether.

First, we can see how Maguindanawn displays the same default indefinite interpretation of AV objects and definite interpretation of pivots that we find in Tagalog and most Philippine

---

[9] Many thanks to Shandra Gonsang for all the Maguindanawn examples discussed here and her detailed judgments on their interpretations. Maguindanawn is also written as Maguindanaoan

languages, as shown in the difference between (63) and (64).[10]

(63) Pang-ilay   sekanin sa   kaluma.
      PROG.AV-seek 3.NOM   OBL   spouse
      'S/he is looking for a spouse.'

(64) Pang-ilay   nin   i   kaluma nin.
      PROG.PV-seek 3.GEN NOM   spouse 3s.GEN
      'S/he is looking for his/her spouse.'

Just as in Tagalog, when we employ a context that clearly favors an existential reading we find that the indefinite reading fails completely, yielding the 'other' interpretation, as seen in (65).

(65) Pang-ilay   nin   i   sakataw a   kaluma.
      PROG.PV-seek 3.GEN NOM   one   LNK   spouse
      'He is looking for his other wife.'

Ibaloy, a Southern Cordilleran language which, like Maguindanawn, has been subject to less contact with Spanish and English in comparison to Tagalog, is also described by Ruffolo (2005) as showing a pattern similar to the above. Nominative/pivot phrases are 'always understood as definite' (Ruffolo, 2005:141) and existential constructions are employed to introduce indefinite arguments when they would normally surface as pivots, as in (66).

(66) *Baray bii ja egmakapiyana on'aseba.*
      wada=j   biʔi   ja   ʔəg=maka-pijan=a   ʔon-ʔasəwa
      EXT=NOM woman LNK NEG=POTPATV/IMPRF-like want=LNK ACTV/IMPRF-marry
      'there is a woman who does not want to get married'   (Ruffolo, 2005:141)

Ruffolo also documents the existence of a NOM+one construction, as in (67), which she crucially notes 'does not often occur in transitive clauses'.

(67) *No metey i sakeya too.*
      no   mə-təj   ʔi   sakəj=a   toʔo
      if/when POTPATV/IMPRF-die NOM   one=LNK person
      'when a person died'   (Ruffolo, 2005:141)

Although we have not focused here on the role of valency in licensing the NOM+one construction it appears very likely that the tendency mentioned by Ruffolo exists in Tagalog and beyond.

---

[10] Note that the AV and PV verbs happen to be identical in the progressive aspect although Maguindanawn otherwise displays the full array of voice/aspect distinctions common to most Philippine languages.

Tagalog speakers are far more likely to accept *Dumating ang isang lalaki* (cf. (5) above) with some type of indefinite interpretation than *Binili ko ang isang saging* (cf. (28) above). The reason for this could be simple. If an actor voice option is available for introducing a novel participant then it is the natural choice. This is the case when the participant is an undergoer and would thus map to the actor voice object position, which is naturally interpreted as indefinite. In other words, the patient voice clause *Binili ko ang isang saging* with a NOM+one pivot is preempted by the actor voice clause *Bumili ako ng saging*. However, when the novel participant is the sole argument of a monovalent predication, no mere manipulation of voice will allow it to escape becoming the pivot as actor voice is the only choice. In this case, a more complex construction is required, employing the existential. In certain registers, the NOM+one stands as an alternative to this more complex option. In other words, *Dumating ang isang lalaki* is not preempted by *May dumating na lalaki*, either because it is more complex or because it is different enough syntactically to not be in the same candidate set. On this view, the difference between intransitive and transitive predication with regard to the NOM+one construction can be reduced to an economy principle whereby voice alternations require less 'effort' than the construction of a biclausal expression.

6.0   Conclusion

I have highlighted some empirical and methodological flaws in the recent literature on Tagalog NP-interpretation and offered the outlines of a more sociolinguistically informed, competition-based analysis that I believe better accounts for Philippine linguistic reality. Two keys to a clearer understanding of the empirical landscape are using diagnostics that clearly disambiguate existential from specific interpretations and clearly distinguishing "translationese" from the grammar of the everyday colloquial language. The resulting picture can be described as a syntax-semantics interface based in *bayanihan* (collective labor/mutual aid); the nominative case marker provides a core referential function while syntactic relations determine the interpretation of genitive marked arguments. The mapping of participants to argument structure via the voice system (and existential predicates) is largely determined by matching the participants' referentiality with the appropriate case markers and syntactic positions.

However, language contact has led some languages to develop a register with a more Indo-European style of NP-interpretation, where most of the work takes place within the NP itself via definite and indefinite determiners rather than through voice selection and existential predicates.[11]

Along the way, I pointed out several areas in need of further investigation, especially

---

[11] On the other hand, there are Austronesian languages like those of the South Sulawesi subgroup (cf. Mamuju in Kaufman, 2017) which maintain a basic opposition between actor voice and patient voice but have lost case marking on NPs altogether. These languages have drifted in the opposite direction from the Indo-European style registers discussed here, as voice and grammatical relations play a dominant role in determining NP interpretation, without the help of case marking determiners commonly found in Philippine languages.

with regard to the diachronic development of the NOM+one construction and its distribution across Philippine-type languages. In addition to these, a comprehensive analysis must also take into consideration the possible "dereferentialization" of *ang* in favor of the case marked demonstrative *yung*, as discussed by Nagaya (2011), and whether this may also account for certain patterns of indefinite *ang*. Finally, there is the intriguing possibility that registers which license indefinite *ang* more freely have inverted the ranking of the parameters determining voice selection, as discussed by Nolasco (2003). For instance, if patient affectedness takes priority over patient definiteness, we may expect a pattern similar to that found in what I have argued to be "translationese". An Optimality Theory style ranking of these parameters could yield a factorial typology that captures voice selection across different registers of Tagalog and different Philippine-type languages. This is clearly an area with great potential for further exploration if we can avoid the methodological pitfalls of the past.

References

Adams, K. L., and Manaster-Ramer, A. (1988). Some questions of topic/focus choice in Tagalog. *Oceanic Linguistics*, 27, 79–101.

Bell, S. J. (1978). Two differences in definiteness in Cebuano and Tagalog. *Oceanic Linguistics*, 17, 1–9.

Bloomfield, L. (1917). *Tagalog Texts with Grammatical Analysis*. University of Illinois Press.

Cena, R. M. (1977). Patient primacy in Tagalog. Paper presented at the LSA Annual Meeting, Chicago.

Collins, J. (2019). Definiteness determined by syntax: A case study in Tagalog. *Natural Language and Linguistic Theory*, 37, 1367–1420.

Foley, W. A. (1998). Symmetrical voice systems and precategoriality in Philippine languages. Paper presented at the 3rd LFG conference, Brisbane.

Himmelmann, N. P. (1991). The Philippine challenge to universal grammar. Institute for Linguistics, University of Köln Working Papers 15.

Himmelmann, N. P. (1997). *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*. Niemeyer.

Himmelmann, N. P. (1999). The lack of zero anaphora and incipient person marking in Tagalog. *Oceanic Linguistics*, 38, 231–269.

Himmelmann, N. P. (2005). Tagalog. In A. Adelaar, and N. Himmelmann (Eds.), *The Austronesian Languages of Asia and Madagascar* (pp. 350–376). Routledge.
Kaufman, D. (2015). Lexical category and alignment in Austronesian. In L. Travis, J. Coon, and D. Massam (Eds.), *Oxford Handbook of Ergativity*, (pp. 589–628). Oxford University Press.

Kaufman, D. (2018). Austronesian predication and the emergence of biclausal clefts in Indonesian languages. In S. Riesberg, A. Shiohara, and A. Utsumi (Eds.), *Perspectives on Information Structure in Austronesian Languages*, (pp. 207–245). Language Science

Press.

Kaufman, D. (2024). The languages of central and southern Philippines. In A. Adelaar and A.
Schapper (Eds.), *Oxford Guide to the Malayo-Polynesian languages of Southeast Asia*,
(pp. 347–373). Oxford University Press.

Kroeger, P. (1993). *Phrase Structure and Grammatical Relations in Tagalog*. Center for the
Study of Language and Information.

McFarland, C. D. (1978). Definite objects and subject selection in Philippine languages. In C.
Edrial-Luzares and A. Hale (Eds.), *Studies in Philippine Linguistics*, *vol. 2*. (pp. 139–
182). Linguistic Society of the Philippines.
Nagaya, N. (2006). Preferred referential expressions in Tagalog. *Tokyo University
Linguistic Papers* (TULIP), 25, 83-106.
Nagaya, N. (2011). Rise and fall of referentiality: articles in Philippine languages. In F.H.
Yap, K. Grunow-Hårsta, and J. Wrona (Eds.), *Nominalization in Asian Languages:
Diachronic and Typological Perspectives*, (pp. 589–626). Amsterdam and Philadelphia:
John Benjamins.

Nolasco, R.M. (2003). Ang pagkaergatibo at pagkatransitibo ng mga wikang pilipino: Isang
pagsusuri sa sistemang bose. Doctoral Dissertation, University of the Philippines,
Diliman.

Paul, I., Cortes, K., and Milambiling, L. (2015). Definiteness without D: The case of ang and ng
in Tagalog. *Canadian Journal of Linguistics*, 60, 361–390.

Pearson, M. (2005). The Malagasy subject/topic as an A′-element. *Natural Language and
Linguistic Theory*, 23, 381–457.

Reid, L. A., and Liao, H. (2004). A brief syntactic typology of Philippine languages. *Language
and Linguistics*, 5, 433.

Reid, L. A. (2002). Determiners, nouns, or what? problems in the analysis of some commonly
occurring forms in Philippine languages. *Oceanic Linguistics*, 41, 295–309.

Richards, N. (2000). Another look at Tagalog subjects. In I. Paul, V. Phillips, and L. Travis
(Eds.), *Formal Issues in Austronesian Linguistics*, (pp. 105–116). Kluwer.

Ruffolo, R. (2005). Topics in the morpho-syntax of Ibaloy, Northern Philippines. PhD
dissertation, Australian National University.

Schachter, P. (1976). The subject in Philippine languages: Topic, actor, actor-topic, or none of
the above. In C. Li (Ed.), *Subject and Topic*, (pp. 491–518). Academic Press.

Schachter, P. (1977). Reference-related and role-related properties of subjects. In P. Cole and J.
Saddock (Eds.), *Grammatical Relations: Syntax and Semantics 8*, (pp. 279–306).
Academic Press.

Schachter, P, and Otanes. F.T. (1982). *Tagalog Reference Grammar*. University of California
Press.

Von Heusinger, K. (2019). Indefiniteness and Specificity. In J. Gundel and B. Abbott (Eds.), *The
Oxford Handbook of Reference*, (pp.146-167). Oxford University Press.