Reasoning about definiteness without articles

James Collins, Stanford University

Introduction: A challenge for a theory of implicature is the question of how pragmatic alternatives are calculated and more specifically, to what extent grammar plays a role in the calculation. This paper refines the notion of implicature calculation with reference to an empirical paradigm from Tagalog. Tagalog indefinites implicate nonuniqueness via competition with definites, much like English indefinites (Heim 1991, 2011 *etc.*). The pragmatic view of the nonuniqueness implication of Tagalog indefinites finds remarkably clear evidence from certain clause types in which the definite form is morphosyntactically blocked. In structures where the definite is impossible, no pragmatic competition arises and an indefinite form may be interpreted as uniquely or non-uniquely instantiated. These data provide evidence that only grammatically well-formed structures are able to enter into pragmatic competition. Though this is an intuitive idea, well-established empirical paradigms demonstrating this notion are uncommon. This paper aims to fill this gap.

The definiteness effect: Like many western Austronesian languages, Tagalog demonstrates a "symmetrical voice" system. Verbs alternate between at least two voice forms, none of which are morphologically unmarked. In focus here are the *patient voice* (PV) forms (as in (1)) and *actor voice* (AV) forms (as in (2)) of transitive verbs, which respectively serve to promote patient and agent NPs to subjecthood. As is typical of symmetrical voice systems crosslinguistically, PV is accompanied by a definite interpretation of the nominative patient NP (1), while AV is accompanied by an indefinite interpretation of the genitive patient (2).

(1)   *t⟨**in**⟩ago*   *ko*   **ang**   *kompyuter*   (2)   **nag***tago*   *ako*   **ng**   *kompyuter*
    ⟨**PV**⟩.hide   GEN.1SG   NOM   computer   **AV**.hide   NOM.1SG   GEN   computer
    'I hid the computer.'   'I hid a computer.'

The use of the indefinite AV variant is infelicitous in contexts where the descriptive content is understood to be uniquely instantiated. In (5), the use of the genitive patient *ng mundo* leads to an unexpected interpretation where there is more than one earth (cf. the English translation with *a*). The PV definite (4) variant is felicitous.

(4)   *pinoprotekta-**han***   *ko*   **ang**   *mundo*   (5)   *?**nag**-poprotekta*   *ako*   **ng**   *mundo*
    protect-**PV**   GEN.1SG   NOM   earth   **AV**-protect   NOM.1SG   GEN   earth
    'I protect the earth.'   '?I protect an earth.'

Following Heim 1991, Percus 2006, *et seq.*, the effect in (5) can be explained by the principle *Maximize Presupposition* (MP): speakers must choose the presuppositionally strongest alternative.

(6)   *Maximize Presupposition:* If S is a **presuppositional alternative** to S', and the context C is such that:

     i. the presuppositions of S and S' are satisfied within C;
     ii. S and S' have the same assertive content relative to C;
     iii. S carries a stronger presupposition than S',
    then S should be preferred to S' (Schlenker 2012: 393)

Non-uniqueness implications as in (2) are cancellable, reinforceable, etc., evidence that they constitute implicatures. The implicature is derivable via pragmatic competition with the corresponding patient voice sentences (1) and (4). Assuming Sp of (5) is obeying MP and that (4) and (5) are presuppositional alternatives, $Addr$ reasons that Sp would have uttered (4) if she believed the

presupposition of (4) ($|\textbf{earth}| = 1$) to be true, and thus must believe it to be false, contradicting common ground knowledge (that there is one earth), and thus leading to infelicity.

Word order: The nonuniqueness implication disappears in clause-types where the definite form is morphosyntactically blocked. In Tagalog, genitive patients only implicate nonuniqueness in verb-initial sentences. In agent-initial sentences, such as relative clauses and *wh*-questions with extracted agents, genitive patients are compatible with either unique or non-unique reference (7).

(7)   *sino*   *ang*   ***nag**-poprotekta* ***ng***   ***mundo***
        NOM.who NOM **AV**-protect   GEN earth
        'Who protects the earth?'

I argue the felicity of (7) is due to a failure of pragmatic enrichment. Agent-initial sentences in Tagalog (as in (7)) *must* have the AV morpheme on the verb (and thus genitive case on the patient) due to the so-called western Austronesian 'Extraction Restriction' (Schachter and Otanes 1972, Georgopoulous 1985, Gerassimova 2005 *a.o.*), which states that extraction of non-nominative NPs is syntactically blocked, thus the corresponding PV version of (7) is ungrammatical.

(8)   **nino*   *ang*   *pinoprotekta-**han** **ang***   ***mundo***
        GEN.who NOM protect-**PV**   NOM earth
        'Who protects the earth?'

I argue that grammatical well-formedness must be a pre-condition for sentence structures to serve as pragmatic alternatives. If (7) and (8) are pragmatic competitors, then the strengthening inference via MP should arise, as (8) is preferable via MP. I argue that the failure to strengthen (7) is a consequence of its pragmatic alternative (8) being ungrammatical by the Extraction Restriction. The hearer does not need to reason about why (7) was chosen over its alternative, as the speaker could not have chosen (8) by the rules of Tagalog grammar. Thus I advocate for grammatical well-formedness being a pre-condition for pragmatic competition, adding the clause (9) to the definition in (6). As the alternatives (7) and (8) fail (9), the strengthening inference due to MP fails to arise.

(9)   iv. *S* and *S′* are syntactically well-formed.

Interpretation of bare NPs: Under MP-based accounts, nonuniqueness implicatures of English indefinites are analyzed via competition between the articles *the* and *a*. However, this analysis does not extend to Tagalog which lacks a definite article. Paul et al (2015) and Collins (2015) have argued that none of the voice markers AV/PV or the nominative/genitive case markers encode the patient's definiteness. All are compatible with indefinite patients (10,11). Thus no morpheme in (1) corresponds to a definite article.

(10) *t⟨**in**⟩ago* *ko*   ***ang*** *isang* *kompyuter*
      ⟨**PV**⟩.hide GEN.1SG NOM one   computer

(11) ***nag**tago* *ako*   ***ng***   *isang* *kompyuter*
      AV.hide NOM.1SG GEN one   computer
      'I hid a computer.'

Collins derives the alternation in (1) and (10) by assuming that PV-verbs like *t⟨**in**⟩ago* combine with *e*-type patient arguments. Bare NP patients like *ang kompyter* in (1) are property denoting and type-shift via *iota* (Partee 1987) inducing a presuppositional definite semantics. In (10), the quantificational determiner lifts the patient to a GQ type, and thus no type-shifting via *iota* is possible, and a definite interpretation is avoided. Thus the presuppositional definite semantics of

(1) is derived via type-shifting and not by a definite article, yet we still observe the same kind of non-uniqueness implicatures derivable by MP.

Without a definite article, how is the notion of "presuppositional alternative" in (6) cashed out? Given a sentence S, interpreters reason about its alternatives $ALT(S)$, calculated with reference to conventionalized sets of lexical items (Horn 1972, Levinson 1983, Hirschberg 1985, *et seq.*). Percus 2006 analyzes MP along these lines, where *a* and *the* are conventionalized alternatives. Given a sentence S, interpreters calculate $ALT(S)$ by swapping occurrences of *a* within S with *the* (and so on for other sets of conventionalized alternatives). I maintain this sort of analysis, though I propose that in Tagalog, the conventionalized alternatives are the voice morphemes, AV and PV.

Relative semantic strength: At what level are the relative semantic strengths of alternative sentences calculated? At the *sentence level*, or only at *the level of the lexical items*? In (6iii), presuppositional strength is measured relative to the sentence level. For Percus, however, presuppositional strength is determined relative to lexical items, meaning that an alternative A, with presuppositionally stronger lexical items, is *always* preferred, regardless of whether A is presuppositionally stronger as a whole. Percus's arguments for this view come from filtering cases like (12).

(12)   a. Everyone with exactly two students assigned the exercise to both of his students.

        b. #Everyone with exactly two students assigned the exercise to all of his students.

(12b) is infelicitous on Percus's account because *all* was used when *both* sufficed. However, by the definition in (6), MP shouldn't create a preference between (12a) and (12b), as neither carries any presupposition at all. Thus (6) fails to provide an explanation for why (12b) is infelicitous.

Percus's account of (12) holds that alternatives with the presuppositionally stronger lexical item *the* are *always* preferable to their counterpart sentences with *a*, regardless of the meaning of the sentence as a whole, so long as the alternative with *the* is felicitous. Singh (2011) argues against this aspect of Percus's account on conceptual grounds, arguing instead that (12) can be explained by a version of (6) which retains the notion that presuppositional strength is checked at the sentence level (6iii), but that MP must be satisfied by all sentence nodes, including embedded sentences (see also Schlenker 2012). Thus, presupposition satisfaction in (6i) is assessed relative to S's local context, rather than the global context/common ground, accounting for the data in (12).

I argue that Tagalog provides new evidence for the view that strength is not assessed at the lexical-item level: we cannot assess strength at the level of AV and PV, because they are not ordered relative to presuppositional strength, neither morpheme inherently encoding presuppositional semantics (10,11). As definiteness arises within the compositional semantics via type-shifting, there is no pair of *lexical items* which can be compared in terms of presuppositional strength. Thus, we must compare presuppositional strength at a level which is syntactically complex enough to allow for type-shifting. Singh's 2011 proposal to check each embedded sentence level (relative to its local context, accounting for data like (12)) is an example of a theory which meets this requirement.

Conclusion: On an utterance of an AV sentence, an interpreter reasons about its alternatives, derived by swapping out the AV morpheme for the PV morpheme (and necessarily making concomitant morphosyntactic alterations such as case assignment). If the PV variant is presuppositionally stronger (as it induces definiteness via *iota* type-shifting) in its local context, the PV variant should be preferred by MP (6), so long as it is grammatically well-formed. I argue the Tagalog data provides evidence that *alternatives must be grammatically well formed to enter into pragmatic competition*, and that *presuppositional strength must be assessed at the sentence level (relative to*

local contexts), and not level of lexical items, in order to incorporate cases where presuppositions are introduced by type-shifting within the compositional semantics, and not by lexical items.

Selected References

Collins, J. N. 2015. Composition without definite articles: A case study in Tagalog. To appear in the *Proceedings of NELS 46*.

Heim, I. 1990. Articles and definiteness. *Semantics: An International Handbook Contemporary Research*, ed. by A. v. Stechow and D. Wunderlich. Berlin: de Gruyter.

Paul, I., K. Cortes, and L. Milambiling. 2015. Definiteness without D: The case of *ang* and *ng* in Tagalog. To appear in the *Canadian Journal of Linguistics*.

Percus, O. 2006. Anti-presuppositions. *Theoretical and empirical studies of reference and anaphora: Toward the establishment of generative grammar as an empirical science*, ed. by A. Ueyama, 52–73. Washington: Japan Society for the Promotion of Science.

Singh, R. 2011. Maximize Presupposition! and local contexts. *Natural Language Semantics* 19: 149–168.

Schachter, P. and F. T. Otanes. 1972. *Tagalog Reference Grammar*. Berkeley: University of California Press.

Schlenker, P. 2012. Maximize Presupposition and Gricean reasoning. *Natural Language Semantics* 20:391–429.