Asking the Right Questions:

Essays in Honor of Sandra Chung

Organized by

Jason Ostrove
Ruth Kramer
Joseph Sabbagh

Kirida Chi'lu-hu Sandy,

Si Yu'us ma'åsi' put i minaolik intension-mu para un istudiåyi yan para un ina i finu' Chamorro. Ginin i istudiu-mu yan i ininå-mu nu i finu' Chamorro, guaha på'gu senma'lak na kandit ni ha i'ina siha i håyi manintirisao tumungu put finu' Chamorro, ispisiåtmenti gi halum i kulehu yan unibetsidåt siha gi halum Amerika, yan gi bula na lugåt siha gi tanu'.

Man sen maguf ham todu nu i manfifinu' Chamorro put i bidå-mu, ya un na'fanbanidosu ham lokkui' ni Mañamorro. Sen bula na ayudu un nå'i ham gi difirentis na manera, ya ti bain fan maleffa lokkui' nu esti.

In kuenta ni manfifinu' Chamorro, hu nånå'i hao ta'lu sen dånkulu na si Yu'us Ma'åsi — in agradesi i bidå-mu, ya in diseseha lokkui' na un gagaigi ha' gi såfu' na manera, gi hinemlu', gi trankilidåt, yan gi minaguf, todu i tiempu.

Kon rispetu yan guinaiya,
I chi'lu-mu as Manny

INTRODUCTION

We are delighted to present this collection of papers to Sandy Chung. We, the organizers and contributors, celebrate and honor her in her many roles: as mentor, as teacher, as researcher, as colleague, and as friend. It is our sincere pleasure to offer this collection of work inspired by her influence on our thinking and writing.

The title of this festschrift, *Asking the Right Questions*, is inspired by Sandy's ability to always know how to pose questions in a stimulating and insightful way. This manifests in her research so that her conclusions are profoundly insightful. We have experienced this in her teaching and advising so that one suddenly sees how the flailing and struggle fall into place. As a colleague, we have all experienced the moment in talks where she asks the linchpin question that, when expanded upon, can make one's thinking seem less secure, while simultaneously providing a path to a deeper understanding. For each of these, we are truly grateful.

Sandy's work is perhaps best known for two things which, at first glance, seem unconnected, but upon further investigation, are interwoven. The first is the clarity of her writing. Many of us are proud converts to the "Chung School of Clear Writing." This focus on clarity of form allows the ideas to blossom. At the same time, the insistence on thorough, careful prose often leads the work to unexpected places that provide connections one would not have seen otherwise.

The second is the content of Sandy's work, centering on the Austronesian family. While she has conducted research on a variety of languages, Sandy is perhaps best known for her in-depth, ongoing investigation of the phonology, morphology, syntax, and semantics of Chamorro. As this description suggests, Sandy also tends to cover a uniquely broad range of topics in her research. This includes her Harvard dissertation on historical and comparative syntax, the semantics of incorporation in her 2003 LI monograph with Bill Ladusaw, the syntax-semantics of sluicing, the existence of lexical categories, the syntax of WH-movement and agreement, the morphophonology of Chamorro clitics, and ongoing groundbreaking psycholinguistics work. Surely there are more that we are forgetting. It is rare to see a scholar move so agilely and insightfully between so many subfields of theoretical linguistics.

Although, perhaps, at the center of this disparate, influential body of work is Sandy's dedication to the Chamorro language and the Chamorro people. As is evident in her 2008 LSA plenary talk, Sandy's commitment to underprivileged languages informs much of how her work has been shaped. This connection goes beyond the theoretical. A major contributor to both the Chamorro dictionary project and the revised orthography, Sandy has managed to do what many academics only dream of: giving back in a tangible way.

On a personal level, Sandy radiates a rare combination of calm, warmth, and openness that allows her students and colleagues to thrive. She can make you feel welcome in virtually any circumstance, whether she is asking you a tough question, or at her home, where she can certainly throw quite the shindig. We are all thankful to have worked with her and to have known her, and we all know that we are much better for it.

Finally, but certainly not least, we would like to thank Maria Zimmer and the Linguistics Research Center for their help publishing, and Scarlett Clothier-Goldschmidt for the cover design.

Jason Ostrove
Ruth Kramer
Joey Sabbagh

February 2017

TABLE OF CONTENTS

ROOTS IN ELLIPSIS AND MULTIDOMINANCE*

Vera Gribanova
Stanford University
gribanov@stanford.edu

Recent research (Veselinova 2006, Siddiqi and Haugen 2103, Harley 2014) supports the conclusion that suppletion of roots is an empirical reality. This finding is consistent with the *Late Insertion* view of roots, wherein their phonological exponence is determined at a post-syntactic stage: roots can compete for lexical insertion (just like functional morphemes), yielding suppletion patterns. Here we compare the suppletive behavior of roots in two syntactic environments: verb stranding ellipsis in Irish and across-the-board movement in Bulgarian nominal phrases. While ellipsis permits suppletion between the root of a stranded verb and an antecedent verb, ATB-movement prohibits suppletive alternations where the ATB-moved element is a suppletive plural but its base positions require a singular form. Existing analyses lead to a paradox: the ellipsis case is consistent with a *Late Insertion* view of roots, while the ATB-movement cases seem to demand that roots undergo *Early Insertion*, their identity already determined in the syntax (Harizanov and Gribanova 2015). I present a re-ananalysis of the relevant ATB-movement cases in multidominance terms, wherein the ban on plural suppletion is a result of the observation that a multidominated root's exponent is (Late) inserted only once, and the inserted form has to be consistent with all of the syntactic environments in which the root appears.

1. Introduction

One of the more productive theoretical developments in recent years has been the idea that roots are morphosyntactically independent units, undergoing composition with other morphological material—categorizing heads, derivational and inflectional morphology—as part of a syntactic derivation (see Arad 2003, Marantz 2001 et seq.). According to one way of understanding this idea, roots have a special status as compared to other functional morphology, in that they are specified with phonological information throughout the derivation (i.e. *Early Insertion*) (Embick 2000, Embick and Halle 2005, Embick and Noyer 2007). An alternative view, which appears to be gaining more prominence in recent work for both theoretical and empirical reasons, takes roots to be equivalent to functional morphemes in having their phonological exponents determined at a point (post-spellout) quite late in the derivation (i.e. *Late Insertion*) (Marantz 1995a,b, Harley 2014, Haugen and Siddiqi 2013).

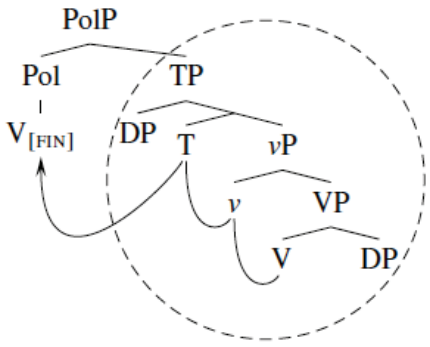
* It is a great honor and pleasure to dedicate this work to Sandy Chung. As with most of my papers, this one has benefitted tremendously from her input and guidance, and was inspired in part by her consideration of related issues in Chung 2012. I have learned a great deal, and continue to be inspired by, the breadth, depth, and unusualness of Sandy's insights. For discussion of the arguments presented here, I thank Pranav Anand, Boris Harizanov, Beth Levin, Jim McCloskey, Jason Ostrove, and the audience at the *Roots IV* workshop. All errors and deficiencies are the author's responsibility alone.

The choice of either position on this question leads to radically different expectations about whether roots should ever undergo suppletion. On the *Early Insertion* view, one expects that roots should not undergo suppletion, as their phonological exponence is determined from the very beginning of the derivation and therefore should not be influenced by other material in their local environment. It follows from the *Early Insertion* view that any putative cases of root suppletion should be analyzable as suppletion of functional morphology; this is the view elaborated in Marantz 1997, who claims that suppletion of *go* vs. *went* in English is suppletion of light verbs (*v*), rather than roots. The *Late Insertion* view, by contrast, takes the phonological exponence of roots to be determined contemporaneously with that of functional morphemes; in Distributed Morphology (DM), this will take place via Vocabulary Insertion. This position predicts that—just like other functional morphemes—roots should be able to undergo suppletion. In confirmation of this view, there is building evidence that root suppletion exists, and that it cannot be boiled down to suppletion of light functional categories. For instance, Veselinova's (2006) crosslinguistic study of suppletion provides extensive examples suppletive verbs with significant lexical content (e.g. *fall-in-water*, *swim*, *bite-off*, etc.). This typological work is supported by recent in-depth case studies of Hiaki suppletive verbs (e.g. *wander*, *enter*, *kill*) (Harley 2014) and Hopi nominal and verbal suppletion (Siddiqi and Haugen 2013), among others.

Taking this evidence at face value, this paper concerns itself with understanding the consequences of *Late Insertion* for the interaction of root suppletion with various morphosyntactic processes that require *matching* or *identity* between roots. A commitment to the *Late Insertion* view will force a particular analysis of at least one of the constructions in question—across-the-board (ATB) movement. The two phenomena under the lens here are ellipsis—in which certain isomorphism requirements are taken to hold between the elided constituent and its antecedent—and ATB phenomena, which are well known to require morphological matching between the element that has extracted ATB and its base positions. In the case of ellipsis, the relevant constructions are those in which a verb moves out of the ellipsis site and is then stranded, as in Irish (1).

(1) Verb-Stranding Ellipsis (VSE):

a. Dúirt siad go dtiocfadh siad, ach ní tháinig ariamh.
say-[PAST] they C come-[COND] they but NEG come-[PAST] ever
'They said that they would come but they never did.' Irish (McCloskey 2012)

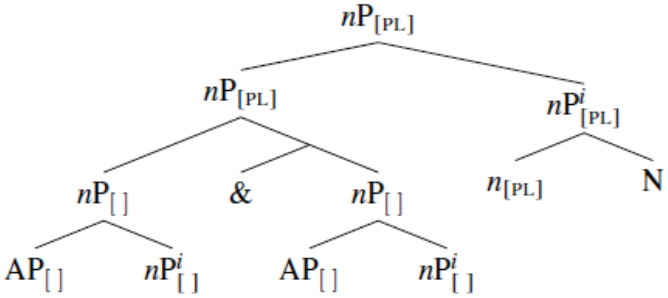
b. 

This operation requires, among other things, a type of isomorphism—in a sense to be defined—between the parts of the verb that originated inside the ellipsis site (V, *v* and T above) and their corresponding antecedent parts. The second phenomenon, illustrated in Bulgarian (2) involves an element high in the syntactic structure corresponding to two base positions; we take the analysis in (2b) (from Harizanov and Gribanova 2015, henceforth H&G) as a starting point for the discussion.

(2) ATB-movement (Bulgarian singular adjectives modifying plural nouns):

a. bălgarsk-**i**-ja i rusk-**i** narod-**i**
bulgarian-SG.M-the and russian-SG.M nation-PL
'the Bulgarian and Russian nations' (= two nations total) (H&G 2015)

b.



Here, too, the ATB operation is known for requiring morphological matching between the ATB-moved element and its base positions (Čitko 2005).

In both case studies, the root may appear in two different syntactic environments, one calling for a suppletive form, and one not. Consistent with previous findings about matching in ellipsis vs. ATB-movement (Barros and Vicente 2011), the two constructions behave differently with respect to whether such a mismatch is permitted. In VSE, a default form and a suppletive form are considered to match as long as there is morphosyntactic (but not necessary phonological) identity between the two roots. By contrast, in ATB-movement of the type in (2), a default singular form in the base positions cannot correspond to a suppletive plural form in the derived position.

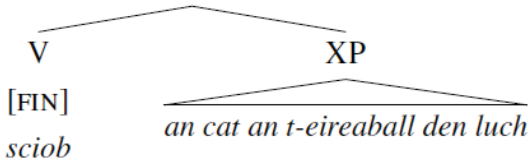
A useful starting point for the discussion is the observation that these two case studies are not equally compatible with a *Late Insertion* view of root exponence. The ellipsis results are well-behaved, in the sense that the isomorphism condition on ellipsis may be taken to apply to semantic representations (Heim 1997, Rooth 1992, Merchant 2001, Takahashi and Fox 2005), where no phonological information would be available (section 2). But the ATB-movement case study presents more of a challenge: H&G conclude that the results require a view of roots in which phonological information is present throughout the syntactic derivation (section 3). The results thus seem, at least initially, to point toward contradictory conclusions with respect to the status of roots and their phonological exponence. I demonstrate here that the ATB-movement case study *is* in fact compatible with the *Late Insertion* view if a multidominance approach to the construction—in which one root is inserted, but associated with several different syntactic environments—is adopted (section 4). The upshot is that the difference in matching requirements stems not from the status of roots—which consistently lack phonological information in the

syntax—but from differences in the nature of the matching requirements involved in these constructions.

2. Isomorphism and suppletion in ellipsis

Evidence for the *Late Insertion* position concerning roots comes from the interaction of ellipsis and root suppletion in languages like Irish, which have the property that the verb moves high enough in the clause to escape a smaller ellipsis domain. The starting point for such a discussion is a syntax for canonical VSO orders in Irish finite clauses that takes the material following the verb to be a constituent to the exclusion of the verb itself (McCloskey 2011).

(3) Sciob an cat an t-eireaball den luch.
cut the cat the tail off-the mouse
'The cat cut the tail off the mouse.' (McCloskey 2012)



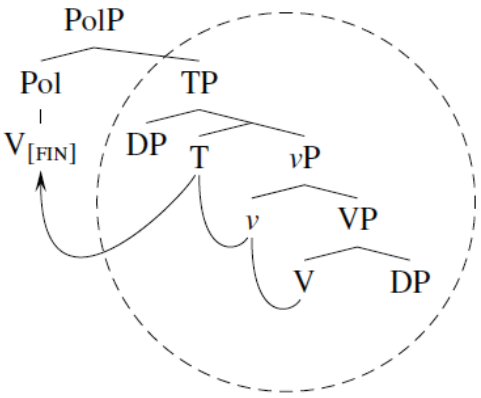
One piece of evidence, among many, for this general schema comes from the availability of RESPONSIVE ELLIPSIS, wherein a verb in isolation may be used to affirm or deny a proposition.

(4) a. Ar chuir tú isteach ar an phost?
INTERR-[PAST] put-[PAST] you in on the job
'Did you apply for the job?'
b. Chuir. / Níor chuir.
put-[PAST] / NEG-[PAST] put-[PAST]
'Yes.' / 'No.' (McCloskey 2012)

(5) Dúirt siad go dtiocfadh siad, ach ní tháinig ariamh.
say-[PAST] they C come-[COND] they but NEG come-[PAST] ever
'They said that they would come but they never did.' (McCloskey 2012)

There is good evidence for the claim that such examples involve ellipsis of the constituent containing minimally the subject and object (i.e., XP in (3)): these constructions support strict and sloppy ambiguities, can take place backwards, and require a linguistic antecedent. Furthermore, Irish lacks object drop; this reduces the analytical possibilities to ellipsis (McCloskey 2012). A more recent instantiation of the schema in (3) comes from McCloskey 2012, where the verb moves to Pol(arity), in conjunction with TP ellipsis. Such configurations may, but need not, involve polarity focus.

(6)



Constituent ellipsis of the type in (4–6) requires identity of the elided material to a linguistic antecedent; the correct characterization of this identity condition has been one of the main focal points of research in this area over the past twenty five (or more) years.[1] Where there is head movement out of the ellipsis site, identity is required between the parts of the verbal complex originating inside the ellipsis site and their corresponding parts in the antecedent (Goldberg 2005b).

(7) Mismatch of terminals originating inside the ellipsis site is impossible (McCloskey 2012):

a.* Níor cheannaigh mé teach ariamh, ach dhíol.
NEG-PAST buy I house ever but sold
'I never bought a house, but I sold one.'

b.* Cé gur mhol an bainisteoir na himreoirí inné cháin inniu.
although C-[PAST] praise the manager the players yesterday criticized today.
'Although the manager praised the players yesterday, he criticized them today.'

c.* Níor éist sí le-n-a cuid daltái ach labhair.
NEG-PAST listen she with-her portion pupils but spoke
'She didn't listen to her pupils but she spoke to them.'

d.* Cháin sé é féin, ach ag an am chéanna chosain.
criticized he him REFL but at the same time defended
'He criticized himself, but at the same time he defended himself.'

Only the parts originating inside the ellipsis site must match: finiteness, mood, and force may be mismatched.[2]

[1] See Rooth 1992, Chung, Ladusaw and McCloskey 1995, Heim 1997, Merchant 2001, Merchant 2013, Fox and Takahashi 2005, Chung 2006, and Chung 2013, among others.

[2] See McCloskey 2012 for a note on why some tense mismatches are permitted even though the diagram in (6) locates T inside the ellipsis domain. The reasoning is that the structure of the left periphery is more fine-grained, with tense features distributed across two projections, only one of which is contained inside the ellipsis domain.

(8) a. Ní theastaíonn sin uaim. Cén fáth a dteastódh?
NEG want that from-me what reason C want[-COND]
'I won't want that. Why would I?'
b. Gabh ar mo dhroim anseo. Chuaigh.
go-[IMPV] on my back here go-[PAST]
'Get up here on my back. He did.'

The putative source for these effects comes from the licensing condition on ellipsis, specifically the ISOMORPHISM clause (adopting here the tradition of Rooth 1992 and Heim 1997):

(9) a. ISOMORPHISM: the elided TP (TP_E) and the antecedent TP (TP_A) must be composed of the same lexical material (variables with distinct indices count as identical) and structurally isomorphic.
b. GIVENNESS: TP_E is contained in some XP_E which is given w.r.t. some constituent XP_A, which contains TP_A. XP_E and XP_A must not overlap. For all assignments *g*, the ordinary semantic value of XP_A w.r.t. *g* is a member of the focus semantic value of XP_E w.r.t. *g*.

The ISOMORPHISM requirement applies in the cases of interest if we understand the relevant parts of the verbal complex (the verb root, V and *v* in Irish) to remain inside the ellipsis site—either because they haven't moved in the narrow syntax or because they have moved but then reconstructed—at the time of ellipsis licensing.

With this much background about Irish VSE in hand, we can now ask what it means to be identical for the purposes for the ellipsis licensing condition. Does the requirement pay attention to morphosyntactic identity or to identity of exponents? What we will see in this section is that the requirement must pay attention to morphosyntax, but not to exponence. This is because we find suppletion of verb roots by complementizers, leading to a situation wherein the root of the stranded verb is morphosyntactically, but not phonologically, identical to its antecedent.

Complementizers in Irish form a prosodic unit with the verb, and syntactic material cannot intervene between the complementizer and the verb. A subset of complementizers triggers suppletive allomorphy in a set of verb roots, which in this case take what is called the *Dependent Form*.

Triggering Complementizers	Non-triggering Complementizers
ní (negation)	a^L (direct relative)
an (interrogative)	*má* (realis conditional)
go (embedded declarative)	
nach (embedded negation)	
a^N (indirect relative)	
dá (irrealis conditional)	

Figure 1: List of suppletion-triggering complementizers in Irish (Ostrove 2015)

Citation Form	**Past Tense Alternations** Independent Form	Dependent Form
bí 'be'	*bhí*	*raibh*
déan 'do'	*rinne*	*dearna*
feic 'see'	*chonaic*	*faca*
téigh 'go'	*chuaigh*	*deachaigh*
Citation Form	**Future Tense Alternations** Independent Form	Dependent Form
faigh 'get'	*gheobhaidh*	*bhfaighidh*
Citation Form	**Present Tense Alternations** Independent Form	Dependent Form
bí 'be'	*tá*	*fuil*
abair 'say'	*deir*	*abair*

Figure 2: List of suppletive verb roots in Irish
(Drawn in part from Ostrove 2015, James McCloskey (p.c.))

As McCloskey (2004) and Ostrove (2015) have demonstrated, it is impossible to relate the Dependent and Independent forms phonologically; consequently, what we see in Irish is a case of genuine root suppletion by complementizers.

The availability of such root suppletion makes it possible to configure VSE environments in which the form of the antecedent root is distinct from the form of the root in the verb stranded outside the ellipsis site. This suppletion would be triggered by distinct complementizers in the two environments, as below.

(10) Antecedent: $C_{non\text{-}triggering}\ V_{independent}\ [_{TP}\ SUBJ\ t_V\ OBJ]$
Ellipsis: $C_{triggering}\ V_{dependent}\ [_{TP}\ \sout{SUBJ\ t_V\ OBJ}]$

The configuration in (10) will force different forms of the root to be realized, in accordance with the different complementizers in the antecedent versus the XP hosting the ellipsis. Such cases are indeed possible and attested:[3]

(11)a. An bhfaca tú Ciarán inné? Chonaic.
Q see.PAST you Ciarán yesterday see.PAST
'Did you see Ciaran yesterday?' 'Yes.'/'I did.'

b. Deireann tú gur inis mé bréag. Ní abraim.
say.PRS you C.PAST told I lie C.NEG say.PRS.1SG
'You say that I told a lie?' 'I do not.'

c. Thug sé dhó é, ach ... dúirt sé nach dtabharfadh
give.PAST he to-him it but ... said he C-NEG give.COND
'He gave it to him but he said that he wouldn't.'

[3] All the examples in (11) were generously shared with me by James McCloskey, who has accumulated a corpus of naturally occurring examples of VSE (among other things) in Irish.

d. Gabh ar mo dhroim anseo. Chuaigh.
go.IMPV on my back here go.PAST
'Get up here on my back. He did.'

Such evidence leads to the conclusion that the licensing condition on ellipsis is concerned only with morphosyntactic, but not phonological, identity between the two verbal roots. This result is completely expected if the ISOMORPHISM clause applies to something like an LF representation, with syntactic information (possibly further manipulated by LF) visible but phonological information inaccessible. Under the *Late Insertion* view of roots, the phonological information associated with roots is not a part of the syntactic derivation, and therefore the ISOMORPHISM condition will have no access to the phonological difference between a dependent and an independent exponent of the same root.[4, 5]

3. Early and late insertion: A paradox

The previous section demonstrated that root suppletion in Irish VSE is grammatical, suggesting that it is the morphosyntactic identity, not the phonological identity, of the root that matters for the ellipsis licensing condition. This follows if the phonological identity of the root is not available at the stage when identity is evaluated for the purpose of ellipsis licensing—consistent with the *Late Insertion* view. Here we discuss a case study that also requires a certain type of matching between roots, but appears initially to require an *Early Insertion* view of roots. It is well known that Across-the-board (ATB) movement also requires matching between the ATB-moved constituent and any requirements that are imposed on the gaps in both conjuncts (Čitko 2005).

(12) a. Kogo$_i$ Jan lubi t_i a Maria podziwia t_i?
who.ACC Jan likes t_{ACC} and Maria admires t_{ACC}
'Who does Jan like and Maria admire?'

[4] A similar argument can be made about English VPE, in which the verb stays inside the ellipsis site and, if it were pronounced, would have a form different than its antecedent counterpart (Warner 1985, Potsdam 1997).

(1) a. Jack **fell** down, and Jill did [~~**fall** down~~] too.

b. If you haven't **told** them yet, you really should [~~**tell** them~~].

The argument is somewhat weaker here, in part because the changes in form may not represent true suppletion.

[5] This empirical result is also compatible with an *Early Insertion* view of roots if, at LF, the ellipsis licensing mechanism ignores the phonological features of the verb root; in such a scenario, ISOMORPHISM is concerned only with matching LF, not matching of other features. There are independent reasons not to adopt this view, however. First, there is the already-mentioned empirical evidence in favor of root suppletion and, consequently, *Late Insertion* of roots. Second, taking on an *Early Insertion* view of roots in this case leads an expectation: although ellipsis identity conditions will have to ignore phonological features of roots on this view, there is at least the weak expectation that some LF processes *will* be able to refer to the phonological identity of roots (but not the phonological identity of any functional morphemes, since there is general agreement that they undergo *Late Insertion*). This expectation is not, to my knowledge, ever borne out.

b.* Kogo$_i$/Komu$_i$ Jan lubi t_i a Maria ufa t_i?
who.ACC/DAT Jan likes t_{ACC} and Maria trusts t_{DAT}
'Who does Jan like and Maria trust?' Polish, Čitko 2005:485

Čitko (2005) showed that in ATB-movement of Polish WH-phrases, ungrammaticality results if different case conditions are imposed on the WH-phrases in their base positions (above, accusative and dative, respectively). What is curious about this effect is that it is ameliorated if there is case syncretism—in this case, between accusative and genitive in animate WH-pronouns.

(13)a.* Czego$_i$/Co$_i$ Jan nienawidzi t_i a Maria lubi t_i?
what.GEN/ACC Jan hates t_{GEN} and Maria likes t_{ACC}
'What does Jan hate and Maria like?'
b. Kogo$_i$ Jan nienawidzi t_i a Maria lubi t_i?
who.ACC/GEN Jan hates t_{GEN} and Maria likes t_{ACC}
'Whom does Jan hate and Maria like?' Polish, Čitko 2005:487

This constitutes evidence that—unlike in ellipsis—the phonological form of the exponent is what matters for the mechanism enforcing matching in ATB-movement. As we will see in the Bulgarian case study in the following sections, certain formulations of ATB-movement make it seem like this matching requirement follows under an *Early Insertion* view of roots; I discuss that kind of system here, and go on to provide an alternative in section 4.

The case study of interest involves Bulgarian coordinated singular adjectives which modify plural nouns. These cases are of particular interest because the suppletion patterns involved led H&G to conclude that an *Early Insertion* view of roots was necessary.

(14)a. bălgarsk-**i**-ja i rusk-**i** narod-**i**
bulgarian-SG.M-the and russian-SG.M nation-PL
'the Bulgarian and Russian nations' (only two nations)
b. bălgarsk-**o**-to i grăck-**o** pravitelstv-**a**
bulgarian-SG.N-the and greek-SG.N government-PL
'the Bulgarian and Greek governments' (only two governments)
c. părv-**a**-ta i posledn-**a** stranic-**i**
first-SG.F-the and last-SG.F page-PL
'the first and last pages' (only 2 pages)

Such examples have an interpretation different from otherwise similar all-plural or all-singular sequences: the meaning in (14) corresponds specifically to two entities, by contrast with all-singular or all-plural sequences (15–16).[6]

[6] See H&G for arguments that such patterns are not derivable by way of ellipsis.

(15) Singular coordinated adjectives modifying a singular noun

a. bălgarsk-**i**-ja i rusk-**i** narod-**ø**
bulgarian-SG.M-the and russian-SG.M nation-SG.M
'the Bulgarian and Russian nation' (one nation)

b. bălgarsk-**o**-to i grăck-**o** pravitelstv-**o**
bulgarian-SG.N-the and greek-SG.N government-SG.N
'the Bulgarian and Greek government' (one government)

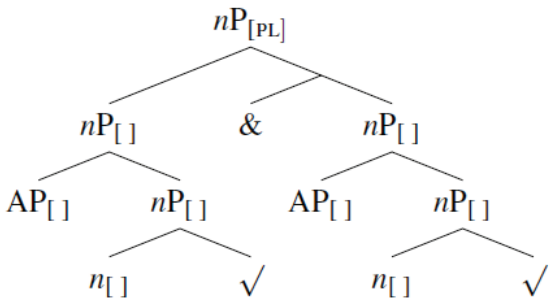
(16) Plural coordinated adjectives modifying a plural noun

a. bălgarsk-**i**-te i rusk-**i** narod-**i**
bulgarian-PL-the and russian-PL nation-PL
'the Bulgarian and Russian nations'
(more than one Bulgarian and more than one Russian nation)

b. bălgarsk-**i**-te i grăck-**i** pravitelstv-**a**
bulgarian-PL-the and greek-PL government-PL
'the Bulgarian and Greek governments'
(more than one Bulgarian and more than one Greek governments)

For the purposes of an initial analysis,[7] H&G assumed, first, that the roots in these examples merge with a categorizing *n* head which itself bears privative number features (absence of features is realized as singular).[8] Consistent with previous analyses of the Bulgarian DP (Dost and Gribanova 2006), APs are taken to be adjuncts (here, to *n*P). In the examples of interest, the coordinate structure involves two coordinated *n*Ps, corresponding to the dual interpretation.

(17)

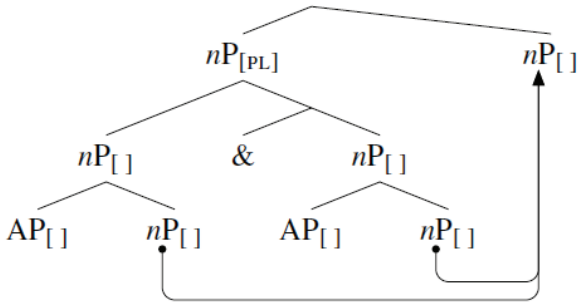


In such a configuration, each conjoined *n*P is non-plural, and APs match these features via the process of nominal concord. H&G's analysis suggests that ATB-movement results in the pronunciation of only one *n*P.

[7] For reasons of time and space, I make no mention here of an alternative analysis of this pattern by Arregi and Nevins (2013); for arguments against this analysis, including the suppletion evidence discussed here, see H&G.

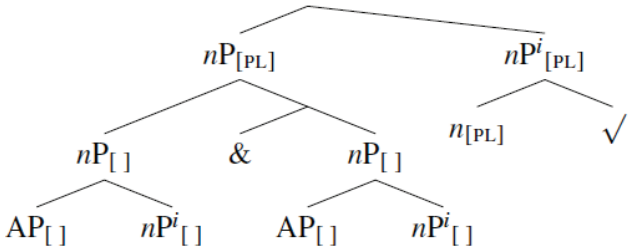
[8] As Harbour (2011) has pointed out, there is good evidence in favor of a binary view of number features, over a privative one. For the purposes of the present analysis, it is convenient to assume a privative view of number features, since ultimately the plural feature of the ATB-moved *n*P will arise as a result of concord with the plural *n*P that it adjoins to. It is more difficult, though probably not impossible, to model this on a view wherein a [+sg] feature would be overwritten by a [-sg] feature in the course of the syntactic derivation.

(18)



The [PL] feature of the *n*P containing the coordinated *n*Ps gets mapped by concord to the head of the ATB-moved *n*P.

(19)



The result is that the adjectives undergo concord with their respective (singular) head *n*Ps, but the moved *n*P undergoes concord with the (plural) node it adjoins to.

As with ellipsis, there are distinct positions in which we find the root in the ATB construction, permitting in principle for a situation in which one environment will trigger suppletion, while the others don't. As H&G pointed out, singular coordinated adjectives cannot modify a plural noun if the form of that noun is suppletive, or involves a stem change.

(20) Suppletive plural:
človek 'person'—*hora* 'people'

a.* nisk-ij-a i visok hora
short-SG.M-the and tall.SG.M people
'the short and tall people'

b.* bulgarsk-ij-a i rusk-i hora
bulgarian-SG.M-the and russian-SG.M people
'the Bulgarian and Russian people'

(21) Plural stem changes:
dete 'child'—*deca* 'children', *oko* 'eye'—*oči* 'eyes', *uho* 'ear'—*uši* 'ears'

a.* naj-nisk-o-to i naj-visok-o deca
most-short-SG.N-the and most-tall-SG.N children
'the shortest and tallest children'

b.* ljav-o-to mi i djasn-o oči
left-SG.N-the my and right-SG.N eyes
'my left and right eyes'

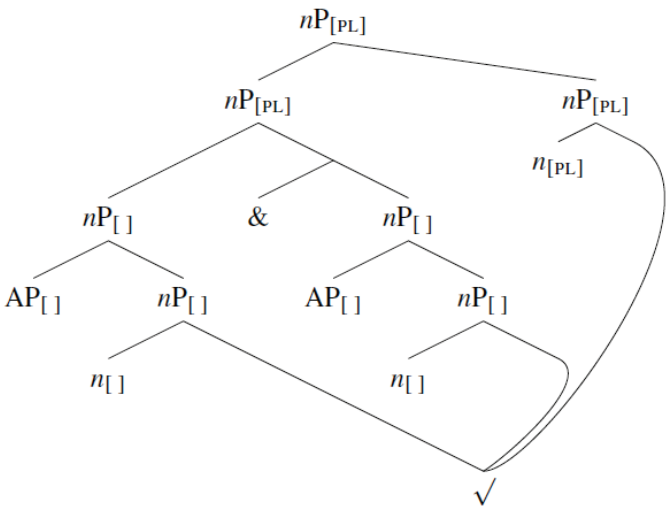
Based on this evidence, H&G concluded that the roots must come with phonological information already realized (e.g. *Early Insertion*), and that this realization can involve no re-writing. This has the consequence that the choice between e.g. *čovek* and *hora* may be made only once, and it must be made when the noun is first merged—that is, when it would be in an environment requiring a singular form. When ATB-movement takes place, the already-inserted singular form cannot further undergo suppletion by plural features, explaining the ungrammaticality of the examples above.

This set of conclusions yields a paradox when compared to the discussion of suppletion in ellipsis in section 2. The identity condition on ellipsis suggests that the exponent of a root plays no role in the syntax, consistent with a *Late Insertion* view of roots. The ISOMORPHISM condition is typically taken to apply to the output of a syntactic derivation, and this condition counts a suppletive and non-suppletive form as identical for its purposes. By contrast, the matching requirement in ATB constructions suggests that the exponent of a root *is* present as part of the syntactic derivation. The phonological exponent of a root must be already a part of the syntax by the time ATB-movement happens, to prevent suppletion of a nominal root by the plural features associated with the landing site of the movement. In what follows, I demonstrate that the conclusion about ATB-movement and the necessity for an *Early Insertion* view of roots is an artifact of H&G's analysis, and provide a re-analysis in multidominance terms, resolving the paradox described here.

4. Exponence in multidominance

There is a well-established line of thought that treats ATB-movement as involving multidominance—a single instance of a syntactic object (daughter) dominated simultaneously by several mothers, with the locus of pronunciation being determined by a postsyntactic linearization algorithm (McCawley 1982, Levine 1985, McCloskey 1986, Moltmann 1992, Muadz 1991, Wilder 1999, Citko 2005, Bachrach and Katzir 2006, and Barros and Vicente 2011). Using this approach not only resolves the *Early* vs. *Late Insertion* paradox, but also provides a better explanation for the ban on suppletive forms in the coordination constructions in section 3.

(22)



In a structure like (22), the multidominated element is the root only (not *n*P, which is the moved constituent in H&G). The plural feature of the coordinated *n*P is still mapped to the adjoined *n*P, as in H&G; but the adjoined *n*P is base-generated high, rather than being ATB-moved. If root exponence is not part of the syntax, the root will be realized phonologically via the usual mechanism of competition (vocabulary insertion). Since there is only one root, there is therefore only one instance of insertion. The different local environments associated with the multidominance structure will produce conflicting requirements on the vocabulary insertion operation:

(23) Conflicting requirements on vocabulary insertion:
$\sqrt{}_{378}$ → [hora] / [PL]
$\sqrt{}_{378}$ → [človek] / [ ]

This rules out the relevant examples—coordinated singular adjectives modifying a suppletive plural noun. It correctly permits, however, cases in which the root does not undergo suppletion by a plural, since in this case the plural will simply be a suffix on a root.[9]

The same result can be derived via H&G's account, if we make certain assumptions about movement: we must adopt the copy theory of movement, in which there can be multiple instances of the same syntactic object; and, crucially, we must assume that lexical insertion happens everywhere there is a copy (despite the fact that only one copy is ultimately pronounced). While there may be a few different ways to implement the idea in (22), the result—given the right set of assumptions—will be that the paradox about the status of roots can be resolved in favor of a *Late Insertion* view, for both case studies in sections 2 and 3.

[9] As H&G show, the construction in question cannot involve mixed number marking on the adjectives (e.g. *A-PL & A-SG N-PL), but nothing about (22) predicts this to be ungrammatical. A possible solution leverages the assumption that number features are privative: if either one of the *n*s in this configuration were to bear a plural feature, that feature would require expression as a prosodically dependent suffix. But given that the root, when it is linearized, will be linearized not adjacent to this plural marker, such a configuration will force a PF crash.

5. Conclusion

We have compared here two different cases in which a) there exists some construction which incorporates a matching requirement and b) the specifics of the construction in question permit—at least in theory—for a situation in which one environment calls for root suppletion while the other does not. The empirical reality is that while elliptical constructions take suppletive and non-suppletive verb roots to be identical to each other for the purposes of the ellipsis licensing condition, ATB-movement constructions care about the phonological exponence of the ATB-moved element and do not permit suppletive plural forms unless all of the relevant environments (the derived position and the base positions) call for a plural form.

The exercise undertaken in this paper takes as its starting point the idea that our evolving understanding of the morphosyntactic status of roots should take into account how roots interact with a diverse range of syntactic environments. In this particular case study, we have seen that keeping our ideas about *Late* vs. *Early Insertion* of roots uniform requires us to take a particular view of what ATB-movement entails. Namely, ATB-movement has to involve one instance of lexical insertion for roots (as in a multidominance approach), and that the resulting exponent must be in harmony with every syntactic environment in which it simultaneously occurs. The observation that ellipsis permits a broader range of mismatches than ATB-movement is not new (Čitko 2005, Vicente and Barros 2011). Here we have tried to pin down just why that might be the case: ellipsis involves comparing two instances of a root to each other; ATB constructions involve insertion of just one element.

References

Arad, Maya. 2003. Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. *Natural Language and Linguistic Theory* 21 (4):737–779.

Arregi, Karlos, and Andrew Nevins. 2013. Contextual neutralization and the elsewhere principle. *Distributed morphology today: Morphemes for Morris Halle*, eds. Ora Matushansky and Alec Marantz, 199–222. Cambridge, Ma: MIT Press.

Bachrach, Asaf, and Roni Katzir. 2006. Right-Node Raising and delayed spellout. *Interphases: Phase-theoretic investigations of linguistic interfaces*, ed. Kleanthes Grohmann, 283–316. Oxford: Oxford University Press.

Barros, Matthew, and Luis Vicente. 2011. Right Node Raising requires both ellipsis and multidomination. *University of Pennsylvania Working Papers in Linguistics* 17, 1–19. Penn Graduate Linguistics Society, Philadelphia.

Chung, Sandra. 2006. Sluicing and the lexicon: The point of no return. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 31, eds. Rebecca T. Cover and Yuni Kim, 73–91. Berkeley Linguistics Society, Berkeley, Ca.

Chung, Sandra. 2012. Are lexical categories universal? The view from Chamorro. *Theoretical Linguistics* 38:1–56.

Chung, Sandra. 2013. Syntactic identity in sluicing: How much and why. *Linguistic Inquiry* 44 (1):1–44.

Chung, Sandra, Bill Ladusaw, and James McCloskey. 1995. Sluicing and logical form. *Natural Language and Linguistic Theory* 3 (3):239–282.

Čitko, Barbara. 2005. On the nature of Merge: External Merge, Internal Merge, and Parallel Merge. *Linguistic Inquiry* 36 (4):475–496.

Dost, Ascander and Vera Gribanova. 2006. Definiteness marking in the Bulgarian. *Proceedings of the 25th West Coast Conference on Formal Linguistics*, eds. Donald Baumer, David Montero, and Michael Scanlon, 132–140. Cascadilla Proceedings Project, Somerville, Ma.

Embick, David. 2000. Features, syntax, and categories in the Latin perfect. *Linguistic Inquiry* 31:185–230.

Embick, David, and Morris Halle. 2005. On the status of stems in morphological theory. *Romance languages and linguistic theory 2003*, eds. Twan Geerts, Ivo von Ginneken, and Haike Jacobs, 37–62. Amsterdam: Benjamins.

Embick, David, and Rolf Noyer. 2007. Distributed Morphology and the syntax/morphology interface. *The Oxford handbook of linguistic interfaces*, eds. Gillian Ramchand and Charles Reiss, 289–324. Oxford: Oxford University Press.

Goldberg, Lotus. 2005a. On the verbal identity requirement in VP ellipsis. Paper presented at the Identity in Ellipsis workshop. Berkeley, Ca.

Goldberg, Lotus. 2005b. *Verb-stranding VP ellipsis: A cross-linguistic study*. Doctoral dissertation, McGill University.

Harbour, Daniel. Valence and atomic number. *Linguistic Inquiry* 42 (4):561–594.

Harizanov, Boris, and Vera Gribanova. 2015. The interaction between across-the-board movement and morphology. *Proceedings of the 49th annual meeting of the Chicago Linguistic Society*. Chicago Linguistics Society, Chicago.

Harley, Heidi. 2014. On the identity of roots. *Theoretical Linguistics* 40 (3/4):225–276.

Haugen, Jason, and Daniel Siddiqi. 2013. Roots and the derivation. *Linguistic Inquiry* 44 (3):493–517.

Heim, Irene. 1997. Predicates or formulas? Evidence from ellipsis. *Proceedings from Semantics and Linguistic Theory VII*, ed. Aaron Lawson, 197–221. Ithaca, NY: CLC Publications.

Levine, Robert D. 1985. Right Node (non-)Raising. *Linguistic Inquiry* 16:492–497.

Marantz, Alec. 1995a. 'Cat' as a phrasal idiom: Consequences of Late Insertion in Distributed Morphology. Ms., MIT, Cambridge, Ma.

Marantz, Alec. 1995b. A late note on Late Insertion. *Explorations in generative grammar*, eds. Young-Sun Kim, Byung-Choon Lee, Kyoung-Jae Lee, Kyun-Kwon Yang, and Jong-Kuri Yoon, 396–413. Hankuk: Seoul.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *Proceedings of the 21st Penn Linguistics Colloquium*, eds. Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark, and Alexander Williams. Philadelphia: University of Pennsylvania Working Papers in Linguistics.

Marantz, Alec. 2001. Words. Paper presented at the West Coast Conference on Formal Linguistics 20. Los Angeles, Ca.

McCawley, James D. 1982. Parentheticals and discontinuous constituent structure. *Linguistic Inquiry* 13 (1):91–106.

McCloskey, James. 1986. Right Node Raising and preposition stranding. *Linguistic Inquiry* 17 (1):183–186.

McCloskey, James. 2004. Mutation and innovation in Northern Irish. Ms., University of California, Santa Cruz, Santa Cruz, Ca. Available at: http://ohlone.ucsc.edu/~jim/PDF/mutation.pdf.

McCloskey, James. 2011. The shape of Irish clauses. *Formal approaches to Celtic linguistics*, ed. Andrew Carnie, 143–178. Cambridge: Cambridge Scholars Publishing.

McCloskey, James. 2012. Polarity, ellipsis and the limits of identity in Irish. Paper presented at the Workshop on Ellipsis, Nanzan University, Nagoya, Japan. Available at: http://ohlone.ucsc.edu/~jim/PDF/nanzan-handout.pdf.

Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands and the theory of ellipsis*. Oxford: Oxford University Press.

Merchant, Jason. 2013. Voice and ellipsis. *Linguistic Inquiry* 44 (1):77–108.

Moltmann, Friederike. 1992. *Coordination and comparatives*. Doctoral dissertation, MIT.

Muadz, Husni. 1991. *Coordinate structures: A planar representation*. Doctoral dissertation, University of Arizona, Tuscon.

Ostrove, Jason. 2015. Allomorphy and the Irish verbal complex. Ms., University of California, Santa Cruz, Santa Cruz, Ca. Available at: http://people.ucsc.edu/%7Ejostrove/QP1.pdf.

Potsdam, Eric. 1997. English verbal morphology and VP ellipsis. *Proceedings of the 27th meeting of the the North East Linguistic Society*, ed. K. Kusumoto, 353–368. GLSA, University of Massachusetts, Amherst.

Rooth, Mats. 1992. Ellipsis redundancy and reduction redundancy. *Proceedings from the Stuttgart Ellipsis Worksho*p, eds. S. Berman and A. Hestvik. Arbeitspapiere des Sonderforschungsbereichs 340, No. 29. Stuttgart.

Takahashi, Shoichi, and Danny Fox. 2005. MaxElide and the re-binding problem. *Proceedings of Semantics and Linguistic Theory XV*, eds. Effi Georgala and Jonathan Howell, 223–240. Ithaca, NY: CLC Publications.

Veselinova, Ljuba. 2006. *Suppletion in verb paradigms: Bits and pieces of the puzzle*. Amsterdam: Benjamins.

Warner, Anthony R. 1985. *The structure of English auxiliaries: A phrase structure grammar*. Bloomington, In: Indiana University Linguistics Club.

Wilder, Chris. 1999. Right Node Raising and the LCA. *Proceedings of the Eighteenth West Coast Conference on Formal Linguistics*, eds. Sonya Bird, Andrew Carnie, Jason D. Haugen, and Peter Norquest, 586–598. Cascadilla Proceedings Project, Somerville, Ma.

SCALAR ALTERNATIVES AND SCALAR INFERENCE INVOLVING ADJECTIVES: A COMMENT ON VAN TIEL, ET AL. 2016[*]

Louise McNally
Universitat Pompeu Fabra
louise.mcnally@upf.edu

Doran, et al. 2009, 2012 and Van Tiel, et al. 2016 have recently presented experimental evidence that gradable adjectives (e.g. *cool*) yield scalar inferences (e.g. to not *cold*) significantly less often than do numerals, quantifiers and modals do and that, moreover, there is also considerable variability *within* the class of adjectives in the frequency with which inferences are likely to be generated. Van Tiel, et al. consider two hypotheses to explain this variation: differences among pairs of adjectives in what they call *scale distinctness* (how easy or difficult it is to differentiate the adjectives), and differences in *scale availability* (how salient the stronger scalar alternative is); they argue that only scale distinctness plays a role, and a rather small one at that; they leave most of the variation unexplained. In this comment, I argue that their measures of scale availability were too crude to detect a role for contextual variation in the scalar alternatives that subjects consider, but that this latter variation is, in fact, a very plausible part of the explanation for the results. I discuss some specific ways adjectives might yield different scalar alternatives in different contexts, under sufficiently sophisticated assumptions about their lexical semantics and about the rhetorical structure of discourse.

1. Introduction

The phenomenon of scalar inference figures prominently in debates within formally-oriented semantics and philosophy of language over where one should draw the line between semantics and pragmatics. The recent history of this debate arguably begins with Grice's (1975) famous proposal that a conversational maxim of Quantity (specifically, "Make your contribution as informative as required (for the purposes of the conversational exchange") induces us, as a rule, to infer that an utterance involving a claim of any given quantitative or scalar strength implicates the denial of claims of greater quantitative or scalar strengths, as exemplified in (1).

(1) a. Sandy bought three pairs of Chi pants. (Sandy did not buy four pairs of Chi pants.)
b. Some of the arguments were dubious (Not all of the arguments were dubious.)
c. You can take High St. to campus. (You are not required to take High St. to campus.)
d. The weather in Santa Cruz is cool. (The weather in Santa Cruz is not cold.)

In the first systematic experimental study of scalar inference patterns across syntactic categories, Doran, et al. 2009, 2012, observed that gradable adjectives (e.g. as in (1d)) yield scalar inferences

[*] Thanks to Bart Geurts for giving me the original impetus to work on this material, to the Festschrift organizers for giving me the impetus to finish it, and to Bob van Tiel and audiences at Radboud University Nijmegen and Utrecht University for comments. Many thanks, Sandy, for your support and interest in my work and career over the years. Congratulations on your 30th anniversary at UCSC!

significantly less often than do numerals, quantifiers and modals. In follow-up experiments using a different task, Van Tiel, et al. (2016; hereafter 'VT+') confirmed this result for adjectives, and moreover found that there is also considerable variability *within* the class of adjectives in the frequency with which inferences are likely to be generated.

VT+ consider various factors in explaining this result, none of which account for all, or indeed even a significant portion, of the data. They tentatively conclude that a combination of idiosyncratic statistical variation between adjectives in rate of scalar inferences, and past experiences with such inferences, also plays a role, but the tone of their conclusions suggests pessimism as to whether anything interesting can be said about these factors. In this comment, I suggest that the high degree of variation they found is exactly what we should expect from scalar inference with adjectives, and that it is a function of entity- and context-dependent assumptions subjects make about the number and nature of the specific scalar alternatives under consideration. There is already evidence for the role of scalar alternatives in influencing inference from studies such as Krifka 2002. Krifka observed that the utterance of a sentence like (2a) is likely to implicate (2b), whereas the utterance of (2c), where the numeral given is likely to be understood on a much coarser level of granularity, is not likely to implicate (2d).

(2) a. The Linguistics Department at UCSC has 15 faculty members.
b. The Linguistics Department at UCSC does not have 16 faculty members.
c. UCSC has 15000 undergraduate students.
d. UCSC does not have 15001 students.

There is no a priori reason to expect that adjectives should be deeply different from numerals in this respect. The principal difference between numerals and adjectives which has masked this similarity is, I will suggest, numerals' lack of polysemy.

In section 2, I provide a summary of VT+'s study and their analysis of the results. In section 3, I discuss the ways in which contextually-determined scalar alternatives could play a role. Finally, in section 4, I conclude by briefly discussing some of the larger implications of their study and the comments presented here.

2. Van Tiel, et al. 2016

VT+ carried out two main experiments. Experiment 1 consisted of questions such as in (3), in which the crucial statement appeared in italics. Subjects were asked whether it was possible to infer from this statement another one that contained a stronger adjective on the same scale:

(3) John says:

She is intelligent.

Would you conclude from this that, according to John, she is not brilliant?

☐ Yes ☐ No

In this experiment, all of the sentences containing adjectives had a simple predicative structure and a pronoun as subject. Experiment 2 had exactly the same design with the same set of adjective pairs (e.g. <*intelligent*, *brilliant*>), except that the italicized test items contained full noun phrases in subject position (e.g., *That professor is intelligent*).[1]

Understanding how VT+ selected their test items is crucial to evaluating their results. The test materials consisted of 32 pairs of adjectives.[2] In order to be sure that these pairs formed true scalar alternatives, they were chosen by first searching the Corpus of Contemporary American English (COCA, Davies 2008), other corpora, and the internet for expressions such as 'X if not Y', 'X or even Y' and 'not just X but Y', where X and Y were adjectives; these expressions make explicit the fact that X and Y stand in a scalar relation, with Y stronger than X. From an initial set of candidate pairs, they made sure to select both pairs in which the weaker adjective on the scale was more frequent than the stronger adjective and vice versa, where frequencies were drawn from word counts in COCA.

VT+ selected the head nouns for the full noun phrases that appeared as the subjects of predication for each pair of adjectives based on a cloze pre-test administered to 10 subjects. These subjects were presented with sentences containing the two adjectives on each scale, as in (4); they had to provide three completions per item.

(4) The _____ is intelligent but she isn't brilliant.

Among the results for each sentence, VT+ selected three nouns. Whenever possible they ensured semantic variety among the nouns (avoiding the choice of e.g. both *singer* and *actress*, which are semantically more similar than e.g. *singer* and *nurse*), and whenever possible they chose two relatively frequent completions among those generated by the pre-test, and one infrequent completion, although the frequencies reported for each test item in their Appendix A suggest that this latter criterion was difficult to satisfy in a very meaningful way: only 20 out of 96 nouns used in Experiment 2 were mentioned more than 3 times in the cloze pre-test. The full list of adjective pairs and noun subjects used in the test items are listed in Table 1,[3] in order of descending frequency with which they generated a scalar inference on Experiment 2 (e.g. an example like *The task is difficult* licensed the conclusion *The task is not impossible* more frequently than e.g. *This child is content* licensed the conclusion *The child is not happy*).

As the results of Experiments 1 and 2, if not identical, were highly correlated, and as Experiment 1 leaves more room for speculation in terms of what subjects might be imagining when completing the task than does Experiment 2, I will focus on the latter. The rate at which scalar inferences were generated in both experiments varied greatly, and in a fairly smooth fashion, from 96% of the time in the case of <*difficult*, *impossible*> in Experiment 2, to 4% of the time in the case of <*content*, *happy*> in the same experiment. The largest drop in rate of inference was only 11%, occurring between <*possible*, *certain*> (inference in 93% of the cases) and <*allowed*, *obligatory*> (82%).

---

[1] The other aspects of their methodology are not relevant to this comment, so I will not discuss them here. See their paper for details.

[2] They also tested other syntactic categories (verbs, determiners, and adverbs), but I will not discuss these here.

[3] All nouns written in lower case in Table 1 appeared with a definite article or demonstrative determiner.

VT+ considered two main factors that could lie behind these results: the availability of the stronger scalar alternative (hereafter, *scale availability*) and the distinctness of the two alternatives (*scale distinctness*). They carried out additional experiments to operationalize these notions so that they could be included in their statistical model, and conclude that only scale distinctness had a significant role to play in the results of Experiment 2. However, the sum total of these two factors

Scalar pair	**Nouns used in Experiment 2**	**Overall rate of inference in Exp. 2 (%)**
<difficult, impossible>	*task/journey/problem*	96
<cheap, free>	*water/electricity/food*	93
<possible, certain>	*Happiness/Failing/Success*	93
<allowed, obligatory>	*Copying/Drinking/Talking*	82
<rare, extinct>	*plant/bird/fish*	79
<low, depleted>	*energy/battery/gas*	79
<hard, unsolvable>	*problem/issue/puzzle*	71
<warm, hot>	*weather/sand/soup*	64
<palatable, delicious>	*food/wine/dessert*	61
<scarce, unavailable>	*recording/resource/mineral*	57
<memorable, unforgettable>	*party/view/movie*	54
<cool, cold>	*air/weather/room*	46
<good, perfect>	*layout/solution/answer*	39
<old, ancient>	*house/mirror/table*	36
<good, excellent>	*food/movie/sandwich*	32
<adequate, good>	*food/salary/solution*	32
<funny, hilarious>	*joke/play/movie*	29
<dark, black>	*fabric/sky/shirt*	29
<hungry, starving>	*boy/dog/elephant*	25
<unsettling, horrific>	*movie/picture/news*	25
<small, tiny>	*room/car/fish*	25
<big, enormous>	*elephant/house/tree*	21
<snug, tight>	*shirt/dress/glove*	21
<attractive, stunning>	*nurse/model/singer*	21
<ugly, hideous>	*wallpaper/sweater/painting*	18
<wary, scared>	*dog/victim/rabbit*	14
<special, unique>	*dress/painting/necklace*	14
<silly, ridiculous>	*song/joke/question*	14
<tired, exhausted>	*quarterback/runner/worker*	14
<pretty, beautiful>	*model/lady/girl*	11
<intelligent, brilliant>	*assistant/professor/student*	7
<content, happy>	*child/homemaker/musician*	4

Table 1. Adjective pairs and test items used in VT+ 2016 (see their Table 3 and Appendix A).

as they operationalized them accounts for only 22% of the variance in their data; by including test item and participant as random variables, they account for only another 30% of the variance. They conclude (2016: 168) that "[i]n the absence of more successful candidates [for explaining the results, LMcN], we are forced to conclude that a major part of the observed variance was unsystematic." VT+ may of course be correct, but here I explore the possibility that the way they chose to operationalize and test for scale availability, in particular, was not optimal.

Scale distinctness was operationalized in two ways. First, they separated the list of pairs in Table 1 into those that associated with bounded scales and whose stronger member represented an endpoint (e.g. *impossible* or *free*), and those for which this was not the case (e.g. *brilliant* or *happy*). Second, they carried out an experiment (their Experiment 4) in which they asked subjects to measure on a 7-point Likert scale the difference in strength (or *semantic distance*) between statements involving the two adjectives in a given pair. Of these two, scale boundedness proved to be far more important, accounting on its own for 10% of the variance in the data, while semantic distance accounted for about 3% of the variance. Note, by way of illustration, that 9 of 11 pairs for which a scalar inference was generated over 50% of the time have a stronger member that marks the endpoint of a scale.

Scale availability was operationalized in four ways, none of which proved to make a significant contribution to accounting for inference patterns.

- **Grammatical category:** "The availability of a lexical scale <α, β> is greater if α and β are from a closed grammatical class." (2016: 157).[4]
- **Word frequency:** "the availability of a lexical scale <α, β> is an increasing function of the frequency of β relative to that of α" (2016: 157), or of the absolute frequency of β; for this measure VT+ used the COCA corpus.
- **Semantic relatedness**, as measured by Latent Semantic Analysis (Landauer and Dumais 1997): "The availability of a lexical scale <α, β> is an increasing function of the semantic relatedness of α and β" (2016: 158). This measure of relatedness consists in comparing the similarity of the distribution of α and β within a corpus.
- **Association strength:** The availability of a lexical scale <α, β> is an increasing function of the strength of association of β with α (2016: 151).

VT+ measured association strength in their Experiment 3 through a cloze test in which they asked subjects to generate the first three alternatives to the underlined word that came to mind when they read a sentence such as *She is <u>angry</u>*. Half of the subjects saw sentences like those in Experiment 1, with pronominal subjects; half saw sentences like those in Experiment 2, with full noun phrase subjects. In addition, half of each group received instructions that potentially facilitated the generation of a stronger scalar alternative, while the other half did not. For example, half were told (for the example in the instructions with *angry*):

> Which words could have occurred instead of the highlighted one? Some of the alternatives that may come to mind are *beautiful*, *happy*, *married*, and so on. We ask you to tell us the first three alternative

[4] VT+ considered grammatical category because their Experiments 1 and 2, as noted above, contained both open and closed class items.

> words that occur to you when you read these sentences. We are interested in your spontaneous responses, so don't think too long about it. (2016: 152-153)

The other half were given *furious* as the first sample alternative, instead of *beautiful*. However, it turned out that subjects provided stronger scalar alternatives only 27% of the time with pronominal subjects, and only 22% of the time with full noun phrase subjects; these figures barely changed when the instructions facilitated the generation of a stronger scalar alternative, though they did raise to 49% and 44%, respectively, under a more permissive coding of the data on which credit was given for the production of a scalar alternative stronger than the test item but weaker than the target scalar alternative (e.g., *probable* for pair <*possible*, *certain*>). In any case, VT+ found no clear correlation between a positive result on the scale association test and scalar inference.

It is not surprising that VT+ did not find the first two of these measures of scale availability useful in accounting for their original experimental results. The degree of variability *within* the class of adjectives – precisely what we are trying to explain – directly rules out grammatical category as a useful measure. Similarly, though relative or absolute word frequency might be a plausible factor in predicting general scale availability, it is not obvious what role it should play in the context of Experiments 1 and 2, which explicitly provided the scalar alternatives and asked subjects to make an offline judgment about them.

More interesting to consider are the measures of semantic relatedness and association strength. Readers unfamiliar with Latent Semantic Analysis should be aware that it is a fairly coarse technique for capturing semantic relatedness, because the distributions constructed for words generally do not take into account disambiguations beyond basic morphosyntactic category. Adjectives are particularly variable in interpretation depending on the nouns with which they combine. To give just one example, though *warm* and *hot* are scalemates for ascribing temperature, *hot*, but not *warm*, is used for popularity (*a hot/??warm product*), temper, and sex appeal (*a hot/??warm body*); while *warm*, but not *hot*, is used for friendliness or empathy (*a warm/??hot personality*). It is therefore to be expected that *warm* and *hot* emerged as only moderately semantically related on this measure.[5] The fact that *warm* yielded the inference *not hot* 64% of the time – more often VT+ would predict, given the semantic relatedness score of the pair – is also no surprise: The context of the task disambiguated the adjectives and rendered irrelevant the other uses that each has and that contributed to lowering the semantic relatedness score.

Finally, although operationalizing association strength via the results of the cloze test might seem reasonable a priori, the procedure VT+ used has two important weaknesses. In the version of the task with pronominal subjects (e.g. finding substitutions for *old* in *It is old*), subjects had virtually no information about the entity that the adjective was being used to describe. This information, as will be explained in the next section, is essential. However, when a full noun phrase subject was provided (e.g. *That house is old*), VT+ note that stronger scalar alternatives were

[5] The specific measure used in Latent Semantic Analysis is cosine similarity. Words are represented as vectors of co-occurrences with other words (perhaps restricted to a given window) in a corpus or "semantic space". The higher the cosine of the angle between the vectors for two words, the more similar the distributions of those two words. VT+ report that the cosine similarity for *warm* and *hot* was 0.51 (with 0 indicating totally orthogonal distributions, and 1 indicating fully overlapping distributions) in the semantic space defined as "General Reading up to 1st year college" available at http://lsa.colorado.edu. See that site for further details.

generated even less often than with pronominal subjects. They suggest that this was due to the fact that the full noun phrase (e.g. *That house*) could facilitate cloze fillers having nothing to do with the house's age (e.g. *beautiful*, *large*, *white*); in the case of the pronoun, the only facilitator of cloze fillers (other than the subjects' imaginations) was the adjective. In other words, the cloze test was not quite able to hone in on the extent to which a given stronger scalar alternative was likely to be considered for a given choice of subject.

3. Scale alternatives and scalar inference

As noted in the introduction, VT+ conclude that much of the variation in the results may be due to idiosyncratic statistical variation between adjectives in rate of scalar inferences, and subjects' past experiences with such inferences. Even if this is so, I am reluctant to further conclude that nothing interesting can be said about this variation. Rather, I take the variation to point to a particular view of lexical meaning and interpretation, one that Geurts 2011 appeals to for selective cases of scalar inference, as we will see momentarily, but which does not seem to be assumed in VT+ (nor is it widely assumed by formal semanticists more generally). This view of lexical meaning renders the variation less surprising.

The view consists of three assumptions. First, lexical entries for open class words are as a rule polysemous and are assigned a specific content only under contextual specification (see Bosch 1995 for a particularly clear presentation of this approach to the lexicon, though the idea has antecedents in his work going back to the early 1970s; see e.g. Hogeweg, to appear, for recent experimental evidence in its favor). Second, property predication can be fruitfully understood as a categorization task; this task can be carried out in some cases according to rule-based criteria, as is standardly assumed in classical entailment-based approaches to predication, while in others, notably with so-called relative gradable adjectives whose standard is context-dependent, it can be modeled as clustering based on similarity (McNally 2011; see Gärdenfors 2000 for a general framework in which this view can be modeled). Finally, the choice of categories under consideration for any predication can vary considerably from one class of entities to another, and, even within a class of entities, from one context to another. For example, in the case of temperature, for some purposes *hot* or *warm* vs. *cold* will be sufficient (e.g., when categorizing *types* of soups, say *caldo de Nadal* vs. *gazpacho*), whereas in others (e.g., when one wants to inform one's guests about temperature the soup they are about to eat), a finer-grained set of distinctions including also *warm* and *cool* might be called for. The set of category alternatives under consideration in any given context will influence how an entity is classified and what inferences are drawn about it.[6]

These assumptions support an analysis of variable scalar inference as follows. Imagine that a subject sees the sentence *The weather is warm*. We of course have no idea of what was in the subject's minds when deciding whether from this to infer that the weather was not hot; however, given that the warm/hot distinction is very often used for weather, it would certainly be unsurprising if the subject made that inference. Notice that this inference depends on the understanding that there is a stronger alternative than *warm* (namely *hot*) that is not being chosen. Inference from *warm* to *not hot* might also be expected in the case of *The sand is warm* because

[6] This last assumption is very much in the spirit of Bidirectional Optimality Theoretic and related probabilistic approaches to semantics. See e.g. Zeevat 2011 and Lassiter and Goodman 2015.

that difference in categorization correlates with other information that is likely to matter to the hearer: hot sand can burn. In contrast, in the case of *The soup is warm*, the subject could have opted equally for an interpretation involving contrast only with cold soups as a type of dish, or for an interpretation focused specifically on (ambient) temperature. Only in the second case would the inference to *not hot* be called for. If only two categories, *cold* and *warm*, are under consideration, there is no stronger alternative than *warm* available and no scalar inference will be possible.

In this scenario, assuming that 100% of speakers behaved uniformly in the first two cases and that speakers divided 50-50% in the third, we would expect a scalar inference 75% of the time. They observed a 64% inference rate, not terribly far from this.

Interestingly, this sort of analysis is very similar to that defended in Geurts 2011 for sentences like (5) (see also Geurts 2009 for earlier discussion of these examples):

(5) If it's warm, we'll lie out in the sun, but if it's VERY warm, we'll go inside and sit in front of the air-conditioner. (Geurts 2011: 140; example originally due to Horn 2006: 27)

Geurts observes that (5) cannot be truthfully uttered if *warm* does not, in the context, entail *not very warm*. He proposes that in this case, the lexical semantics of *warm* is contextually narrowed. This narrowing, though pragmatic in nature, affects what is said in the Gricean sense, that is, the propositional content of the sentence. In other words, if we infer *warm but not very warm* in cases like (5), the scalar inference is due not to a quantity implicature – it is not defeasible – but rather to a logically prior decision about how to carve up the categorization space.

Geurts notes, correctly, that taking polysemy and the context dependence of lexical meaning into account does not undermine a classical pragmatic theory of quantity implicature:[7] it is simply the case that, in order for scalar reasoning to be applied, the set of alternatives under discussion must first be fixed. Thus, returning to VT+'s results, for those pairs Table 1 for which scalar inferences were infrequently generated, we can hypothesize that the categorization space for most subjects and for most test items did not include the stronger alternative in the pair. Of course, VT+'s four tests for scale availability were intended precisely to test to whether or not this was in fact the case, but arguably none of them provided sufficient context sensitivity to do so.

Space precludes an item-by-item discussion of all of the adjective pairs in Table 1. I will instead finish the section by mentioning two additional factors beyond general polysemy that can affect the number or nature of the alternative categories under consideration in any given context, and that could have played a role in reduced scale availability for some of VT+'s pairs, in the hope that future experiments might be designed to control for them.

The first is subjects' assumptions about the (implicit) question(s) under discussion in the test items. VT+ (2016: 166-167) briefly consider the possibility that the question under discussion could have influenced their results. They note that van Kuppevelt 1996 argued that scalar implicatures for numerals arise when the numeral is part of the information focus of the sentence, as in (6), but not when it is part of the background, as in (7):

[7] Among other things, it does not force one to assume that lexical items are *ambiguous*, nor does it require one to adopt a logical-syntactic approach to implicature (see Chierchia, Fox and Spector 2012). See Geurts 2011 for extensive discussion.

(6) A: How many years has Sandy worked at UCSC?
B: Sandy has worked there for [thirty]F years.

(7) A: Who has worked at UCSC for thirty years?
B: [Sandy]F has worked at UCSC for thirty years.

VT+ dismiss this sort of explanation because there is no evidence of any such difference in information structure between their test items. Though this is of course correct insofar as it goes, the assumption that the question under discussion can be determined on the basis of information structure alone is too crude: A richer notion of discourse structure (as modeled in e.g. Mann and Thompson's (1988) Rhetorical Structure Theory or Asher and Lascarides' (2003) Segmented Discourse Representation Theory) must also be taken into account.

To see this, consider the pair <*adequate*, *good*>. *Adequate* is indeed a scalemate of *good*, but adequacy and goodness are generally evaluated with respect to some purpose (what Kagan and Alexeyenko 2010 call, for slightly different sorts of cases, a *functional standard*). For example, food may be adequate or good in quantity or quality, a salary may be judged as adequate to meet expenses, good insofar as it surpasses that of many peers, etc. VT+'s experimental design did not oblige subjects to choose the same standard for the two adjectives in any given pair. For example, a subject reading *The salary is adequate* might imagine this as addressing an implicit question about whether the salary meets someone's needs or not. When then asked whether that means that the salary is not good, the subject might hold the standard constant and infer that it is not. However, she could just as easily infer that no such conclusion is warranted, because there simply is no reason to think that a salary adequate to meet someone's needs could not be considered good with respect to some other standard. In other words, there is no reason the subject could not imagine that the local question under discussion shifts between the initial presentation of the test item and the subsequent question about the second adjective in the pair. Such a shift would be compatible with a rhetorical structure in which the main question under discussion might be something like *What is your salary like?*, with elaborating sub-questions such as *Does it meet your needs?*, *Are you better off than most of your peers?*, etc.[8] But if the local questions under discussion shift as the subjects move from consideration of one adjective to another, it is entirely likely that the alternative categories under consideration will also shift. In other words, what counts as a scalar alternative is potentially highly dynamic. VT+ do not contemplate this possibility.

A second factor that can affect the set of scalar alternatives under consideration is the overall semantic complexity of the adjectives in question. There are pairs, such as <*wary*, *scared*>, which do not differ only, or perhaps even primarily, in strength. Recall that VT+ chose scalar alternatives by looking for occurrences in phrases of the form 'X if not Y', 'X or even Y' and 'not just X but Y.' One might find examples of *wary if not scared* because wariness is often motivated by emotions related to fear. But wariness is a property of behavioral disposition, while fear is an emotional state that merely implies, but does not entail, particular behavioral dispositions. Wariness might be provoked by an emotion that one would not quite describe as fear (such as anxiety), and higher degrees of fear might be correlated with dispositions that are more extreme

[8] See Hunter and Abrusán, to appear, for discussion of how rhetorical structure and the notion of question under discussion can be related.

than wariness (such as a complete inability to act). For this reason, *wary* and *scared* might well serve as alternatives in some contexts. However, wariness can be provoked by feelings distinct from fear, notably mistrust. It is therefore unsurprising that *scared* was not a salient alternative to *wary* in the vast majority of cases. VT+'s method for choosing test items did not control for this factor.

4. General discussion

The discussion in the preceding section has identified three factors not explicitly acknowledged by VT+ that could have played a role in the highly varying degree to which adjectives yielded scalar inference in their experiments: 1) the polysemy of the adjectives (in contrast to e.g. numerals and quantifiers), 2) their semantic complexity, and 3) the nature of the question under discussion. Once these are taken into account, there does not seem to be any reason to think that adjectives, as a category, interact any differently with the general theory of scalar inference than do numerals or other sorts of expressions for which scalar inference has been more widely studied, just as Geurts 2011 concluded. The fact that scalar alternatives have been shown to vary even for numerals, depending on the granularity of the information that is considered pertinent in the context (which is a function, at least in part, of the question under discussion), suggests that the difference in the variability of scalar inference with adjectives vs. numerals and quantifiers is due to a difference in their lexical richness.

This is a welcome conclusion, but it also carries with it some important implications. Preserving a fairly simple and elegant account of scalar inference crucially depends on taking scalar alternatives to be defined in a highly local, context-dependent fashion. However, any hope of developing a theory of when and how these alternatives arise depends, in turn, on incorporating well-developed theories of polysemy and rhetorical structure into the analysis of interpretive phenomena for which they have been, to date, largely ignored. The gain in insight into the special scalar inference behavior associated with adjectives should serve as an argument for changing this situation and extending the interest in these aspects of language among formally-oriented semanticists and pragmaticists.

References

Asher, Nicholas and Alex Lascarides (2003). *Logics of conversation*. Cambridge: Cambridge University Press.

Bosch, Peter. 1995. Meaning and contextual concepts. *Semantic and conceptual knowledge. Papers from a joint workshop of the Max Planck Arbeitsgruppe 'Strukturelle Grammatik' and the IBM Institute for Logic and Linguistics*, eds. M. Bierwisch and P. Bosch, 79-99. Arbeitspapiere des Sonderforschungsbereichs 340, Vol. 71, Universität Tübingen.

Chierchia, Gennaro, Danny Fox, and Benjamin Spector. 2012. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. *Handbook of semantics*, vol 3, eds. C. Maienborn, K. von Heusinger, and P. Portner, 2297-2331. Berlin: Mouton de Gruyter.

Davies, Mark. 2008. The Corpus of Contemporary American English: 450 million words, 1990-present. Available online at http://corpus.byu.edu/coca/.

Doran, Ryan, Rachel E. Baker, Yaron McNabb, Meredith Larson, and Gregory Ward. 2009. On the nonunified nature of scalar implicature: An empirical investigation. *International Review of Pragmatics* 1: 211-248.

Doran, Ryan, Gregory Ward, Meredith Larson, and Yaron McNabb. 2012. A novel experimental paradigm for distinguishing between what is said and what is implicated. *Language* 88: 124-154.

Gärdenfors, Peter. 2000. *Conceptual spaces*. Cambridge, MA: MIT Press.

Geurts, Bart. 2009. Scalar implicature and local pragmatics. *Mind and Language* 24: 51-79.

Geurts, Bart. 2011. *Quantity implicatures*. Cambridge: Cambridge University Press.

Grice, H. Paul. 1975. Logic and conversation. *Syntax and semantics 3: Speech acts*, ed. P. Cole, 41-58. New York: Academic Press.

Hogeweg, Lotte. To appear. Suppression in interpreting adjective noun combinations and the nature of the lexicon. *Journal of Semantics*.

Horn, Laurence R. 2006. The border wars: A neo-Gricean perspective. *Where semantics meets pragmatics*, eds. K. von Heusinger and K. Turner, 21-48. Amsterdam: Elsevier.

Hunter, Julie and Márta Abrusán. To appear. Rhetorical relations and QUDs. *New frontiers in artificial intelligence: JSAI-isAI Workshops LENLS, JURISIN, KCSD, LLLL revised selected papers*. Berlin: Springer.

Kagan, Olga, and Sascha Alexeyenko. 2010. Degree modification in Russian morphology: The case of the suffix *-ovat*. *Proceedings of IATL 26*.

Kennedy, Christopher. 2001. Polar opposition and the ontology of 'degrees'. *Linguistics and Philosophy* 24: 33-70.

Krifka, Manfred. 2002. Be brief and vague! And how bidirectional optimality theory allows for verbosity and precision. *Sounds and systems. Studies in structure and change: A festschrift for Theo Vennemann*, eds. D. Restle and D. Zaefferer, 439-458. Berlin: Mouton de Gruyter.

van Kuppevelt, Jan. 1996. Inferring from topics: Scalar implicatures as topic-dependent inferences. *Linguistics and Philosophy* 19: 393-443.

Landauer, Thomas K., and Susan T. Dumais. 1997. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review* 104: 211-240.

Lassiter, Daniel and Noah Goodman. 2015. Probabilistic semantics and pragmatics: Uncertainty in language and thought. *Handbook of contemporary semantic theory*, second edition, eds. C. Fox and S. Lappin, 143-167. Oxford: Wiley Blackwell.

Mann, William C. and Sandra A. Thompson (1988). Rhetorical Structure Theory: A theory of text organization. *Text* 8: 243-281.

McNally, Louise. 2011. The relative role of property type and scale structure in explaining the behavior of gradable adjectives. In *ViC 2009 (Papers from the ESSLLI 2009 Workshop on Vagueness in Communcation)*, eds. Rick Nouwen, Robert van Rooij, Uli Sauerland, and Hans-Christian Schmitz, 151-168. Berlin: Springer.

van Tiel, Bob, Emiel van Miltenburg, Natalia Zevakhina, and Bart Geurts. 2016. Scalar diversity. *Journal of Semantics* 33: 137-175.

Zeevat, Henk. 2011. Bayesian interpretation and Optimality Theory. *Bidirectional Optimality Theory*, eds. A. Benz and J. Mattausch, 191-220. Amsterdam: John Benjamins.

A PRELIMINARY LOOK AT EXCEPTIVES IN TAHITIAN*

Eric Potsdam
University of Florida
potsdam@ufl.edu

Maria Polinsky
University of Maryland, College Park
polinsky@umd.edu

This paper examines the syntactic expression of exceptive constructions (e.g., *Sandy has worked on every Polynesian language except Tahitian*) in Tahitian (Polynesian). Cross-linguistically, exceptives can be phrasal or clausal. We show that exceptives in Tahitian are clausal, despite the apparently reduced appearance of the relevant construction. Furthermore, exceptives in Tahitian are not expressed by a dedicated construction but rather by a juxtaposition of two clauses in which the first makes a generalization and the second explicitly states an exception to that generalization via a negative clause. We compare this strategy of "conjoined exceptives" to that of conjoined comparatives; a comparison between the two may be helpful for future work on the structure of exceptives.

1. Introduction

This paper is an initial investigation into the syntactic expression of exceptives in the Polynesian language Tahitian. An exceptive is construction used to express exceptions to generalizations. A typical exceptive in English is (1).

(1) Everyone left except/but Bob.

We will use the term EXCEPTIVE CONSTRUCTION to refer to the entire sentence that expresses an exception. The EXCEPTIVE PHRASE is that part of the construction that identifies the sentence as an exceptive; it need not literally be a phrase. *Except Bob* is the exceptive phrase in (1). EXCEPTION XP refers to the exception itself, *Bob* in (1).

Our goals are modest: to present a basic description of the Tahitian patterns and to make some preliminary analytical claims. Our main conclusion is that exceptive phrases in Tahitian are strictly clausal in nature despite their apparently reduced appearance. Section 2 presents the basics of Tahitian word order, its exception constructions, and negation. Section 3 argues that the exceptive phrase in Tahitian is actually a (reduced) clause. Section 4 concludes.

* This work owes a great deal to Sandy Chung, who pioneered the comparative study of the syntax of Polynesian languages, starting with her seminal comparative work, published as Chung 1978. It touches on topics that Sandy has worked on: negation, ellipsis, clause structure, and, of course, Polynesian languages. We are grateful to Sandy's friendship, mentoring, encouragement, and high standards of scholarship over the years.

We would like to thank our Tahitian consultants on Moorea and Raiatea: Mate Mahuta, Noeline Mahuta, Djelma Maono, Tuterai, Maruhi, Odile Meyer, Hinano Murphy, Jean-Luc Tere, and Tevahine Tairua as well as Jacques Vernaudon for engaging discussions. Glossing follows Leipzig Glossing Conventions.

2. Tahitian

Tahitian is a Nuclear Polynesian language spoken by approximately 60,000 people in French Polynesia (Lewis et al. 2016). Its basic word order is VSO, (2), and VOS is ungrammatical. Case marking is nominative-accusative; however, the accusative marker is often dropped.

(2) 'Ua hōhoni te ma'o 'i te tāvana
PFV bite DET shark ACC DET chief
'The shark bit the chief.'

The verb is preceded by tense-aspect-mood (TAM) particles, which vary between matrix and embedded clauses. These are given in Table 1 for perfective and imperfective clauses (see Tryon 1970:32–37, Markey 1976, Académie Tahitienne 1986:201–56, Lazard and Peltzer 2000:124–42).

	PERFECTIVE	IMPERFECTIVE
MATRIX CLAUSE	*'ua*	*tē ... DEIC*
DEPENDENT CLAUSE	*i*	*e*

Table 1. Tahitian aspectual particles

Non-verbal clauses are predicate-initial, subject-final, (3).

(3) a. Mai te fare mai au
PREP DET house DIR 1SG
'I am coming from the house.' (Lazard and Peltzer 2000:42)

b. 'E fa'ehau terā ta'ata
PRED soldier DEM man
'That man is a soldier.' (Lazard and Peltzer 2000:36)

2.1. Exceptives

Exceptives in Tahitian take a number of forms. The three dominant forms of the exceptive phrase offered by our consultants are given in (4a, b, c).[1]

(4) 'Ua tae pauroa mai te mau tamari'i,
PFV come all DIR DET PL child
'All the children came ...'

a. 'o Poe noa 'aita
DET Poe just NEG

b. 'aita rā 'o Poe
NEG but DET Poe

[1] Other constructions were intermittently offered but space considerations prevent us from addressing them here.

c. tē toe ra 'o Poe
IPFV remain DEIC DET Poe
'... except Poe.'

In (4a), the exception XP *Poe* is followed by the particle *noa*, which Walroos 2002 indicates occurs "after nominal sequences, expressing the idea of restriction alone, only, just". This *XP noa* is then followed by *'aita* 'NEG', the marker of sentential negation (see section 2.2). In (4b), the exceptive phrase begins with the sentential negation marker *'aita* and the particle *rā*, which Walroos 2002 translates as 'but, although'. This is followed by the exception XP. In (4c), the exception XP is preceded by the imperfective form of the verb *toe* 'remain, be left over'. We will not consider the form in (4c) here.

We will argue in the following section that the first two constructions are not exceptive specific. They do not parallel the English translation 'All the children came, except for Poe'. Rather, better translations for (4a, b) would be 'All the children came, only Poe didn't' and 'All the children came, but Poe didn't'. That is, there is no genuine exceptive-specific exceptive phrase; the exception is expressed using a juxtaposed negative clause which has the pragmatic force of providing an exception to the general statement made in the first clause. The main clause and the exceptive phrase are conjoined paratactically. Before we present our evidence for this proposal, it is necessary to briefly describe negative clauses in Tahitian.

2.2. Negation

Negation in Tahitian is a predicate (Lemaître 1973:17, Académie Tahitienne 1986:328–34, Lazard and Peltzer 2000:49–59, Peltzer 1996, Tryon 1970:46–48; see also Hohepa 1969 and Chung 1970 on Māori where the situation is comparable). It takes a clausal complement whose subject obligatorily raises to a position immediately following the negative marker. The TAM marker preceding the embedded verb is from the dependent series in Table 1.

(5) a. 'Ua tai'o 'oe 'i terā puta
PFV read 2SG ACC DEM book
'You read that book.'
b. 'Aita 'oe i tai'o 'i terā puta
NEG 2SG PFV.DEP read ACC DEM book
'You didn't read that book.'
c. *'Aita i tai'o 'oe 'i terā puta
NEG PFV.DEP read 2SG ACC DEM book

The sentential negation marker varies in form with tense, aspect, and mood. Three forms are given in Table 2 (to simplify exposition, we have omitted other forms).

PERFECTIVE	'aita
IMPERFECTIVE	'e'ita
PROHIBITIVE	'eiaha

Table 2. Tahitian negative particles

A distinct form, *'ore*, is used for constituent negation and non-finite clauses (Peltzer 1996). The former use is illustrated in (6). Note that *'ore* is not used in exceptives.

(6) a. 'E rave tāua 'i te 'ohipa mai te fa'aea 'ore
FUT do 1DU.INCL ACC DET work with DET rest NEG
'We will do the work without stopping.' (Tryon 1970:47)

b. pinepine 'ore
often NEG
'rarely'

c. nehenehe 'ore
possible NEG
'impossible' (Tryon 1970:48)

3. The Clausal Nature of Tahitian Exceptives

This section presents data supporting the position that the exceptive phrase in Tahitian is a (reduced) negative clause. The clausal status of these exceptive phrases is signaled by *'aita* in examples like (4), which we claim is the ordinary sentential negation marker and not a particle/preposition equivalent to English *except*. The phrases in (4) are negative clauses that have been reduced in some way, leaving an ordinary instance of sentential negation behind. More accurate translations for (4a, b) under our proposal are thus as in (7) and (8), respectively.

(7) 'Ua tae pauroa mai te mau tamari'i, 'o Poe noa 'aita
PFV come all DIR DET PL child DET Poe just NEG
'All the children came, just/only Poe didn't.'

(8) 'Ua tae pauroa mai te mau tamari'i, 'aita rā 'o Poe
PFV come all DIR DET PL child NEG but DET Poe
'All the children came, but Poe didn't.'

The following subsections offer evidence from various domains in support of our proposal.

3.1. Unreduced Exceptives

A straightforward piece of evidence for the clausal status of exceptive phrases is that they can be expressed in their unreduced form. Deletion of the embedded material seems to always be optional. (9) corresponds to (7) and (10) corresponds to (8).

(9) 'Ua tae pauroa mai te mau tamari'i,
PFV come all DIR DET PL child
'o Poe noa 'aita (i tae mai)
DET Poe just NEG PFV.DEP come DIR
'All the children came, just Poe didn't (come).'

(10) 'Ua tae pauroa mai te mau tamari'i,
PFV come all DIR DET PL child
'aita rā 'o Poe (i tae mai)
NEG but DET Poe PFV.DEP come DIR
'All the children came, but Poe didn't (come).'

3.2. Cross-Polynesian Comparison

A similar observation in support of our proposal comes from related Polynesian languages. In at least some of these languages, only the fully expressed versions of the negative clause, as in section 3.1, are possible. The reduction operations available in Tahitian seem to be unavailable. The examples in (11) are from Niuean,[2] and the example in (12) is from Tongan.

(11)a. Kai oti e Mele e tau ika *Niuean*
eat all ERG Mary ABS PL fish
kae nākai/ai kai (e ia) e lahakula
but not/not eat ERG 3SG ABS tuna
lit. "Mary eats all fish but (she) does not eat tuna."
'Mary eats all fish except tuna.'

b. Ne hau (a) au he tau aho oti kae nākai/ai (a au) he tau ahu tapu
PRS work ABS 1SG OBL PL day all but not/not ABS 1SG OBL PL Sunday
lit. "I work all days but I do not work on Sundays."
'I work every day except Sunday.'

(12) 'Oku 'iloa 'e ia 'a e kakai katoa ka *Tongan*
PRS see ERG 3SG ABS DET people all but
'ikai ke 'iloa ('e ia) 'a Mele
NEG SBJ see ERG 3SG ABS Mary
'He saw everybody except Mary.'

This micro-variation will no doubt ultimately provide an important window on the ellipsis operations taking place in Tahitian, as they seem to be unavailable in some related languages.

3.3. Form of the Negative Marker

If the negative marker in exceptives is the clausal negator found in sentential negation contexts, we expect that it should show the morphosyntactic variation in Table 2, which indicates that negation varies in form with different TAM. This is indeed the case. (13) illustrates an exceptive using the imperfective negative particle *'e'ita* in the context of a future event and (14) shows the prohibitive marker *'eiaha* in a directive speech act.

---

[2] We would like to thank our Niuean consultants for help with these and other data: Grace Latoa, Mele Nemaia, Pat and Granby Siakimotu, and Kara Tukuitoga.

(13) 'E tauturu ra te tāvana 'i te ta'ato'ara'a 'e'ita rā 'oe
FUT help DEIC DET chief ACC DET everyone NEG.IPFV but 2SG
'The chief will help everyone but you.'

(14) 'E tāma vau te fare i te mau mahana ato'a
FUT clean 1SG DET house PREP DET PL day all
'eiaha rā te tapati
NEG.PROH but DET Sunday
'You must clean the house every day, except Sunday.'

3.4. Lack of Connected Exceptives

The literature on exceptives makes a distinction between connected exceptives and free exceptives (e.g. Hoeksema 1987, Pérez-Jiménez and Moreno-Quibén 2012). Connected exceptives are DP modifiers in which the exceptive phrase applies to the domain of quantification of the DP. Free exceptives are CP modifiers that are exceptions to propositions expressing a generalization. We tentatively assume that connected exceptives are phrasal modifiers of DPs and free exceptives are clauses (Pérez-Jiménez and Moreno-Quibén 2012). If the Tahitian exceptive phrase is a clause, it will not show behavior unique to connected exceptives, since the latter are phrasal. One characteristic of connected exceptives is that they can appear adjacent to the quantificational nominal that they modify. This is not possible in Tahitian, in contrast to the English translations:

(15)a. 'Ua haere pauroa te ta'ata (*'aita rā 'o Marama)
PFV go all DET people NEG but DET Marama
i te tāmā'ara'a ('aita rā 'o Marama)
PREP DET festival NEG but DET Marama

b. 'Ua haere pauroa te ta'ata (*'o Marama noa 'aita)
PFV go all DET people DET Marama just NEG
i te tāmā'ara'a ('o Marama noa 'aita)
PREP DET feast DET Marama just NEG
'Everyone (except Marama) went to the feast (except Marama).'

3.5. Stripping

Our proposal is that exceptive phrases are simply (reduced) negative clauses. They do not semantically encode an exceptive meaning as in the English *except*. As a result, we do not expect such clauses to be restricted to exceptives. We expect to see such (reduced) negative clauses used elsewhere for a different purpose. This is what we find. These reduced negative clauses are also used in Stripping, (16b) shows that the same element is used to express stripped negative phrases.

(16)a. 'Ua haere ātu vau i te 'oro'a,
PFV go DIR 1SG PREP DET festival
'aita rā 'o Marie 'i haere mai
NEG but DET Mary PFV.DEP go DIR
'I went to the festival but Mary didn't go.'

b. 'Ua haere ātu vau i te 'oro'a, 'aita rā 'o Marie
PFV go DIR 1SG PREP DET festival NEG but DET Mary
'I went to the festival but not Mary.'

Whatever the processes are that create Tahitian exceptive phrases, they are used to generate negative stripping as well.

3.6. Lack of Cross-Clausal Binding

Finally, claiming that Tahitian exceptive phrases are independent clauses means that there will be no c-command relations between the generalization in the first clause and the exceptive phrase because the two are only paratactically related—i.e. they are coordinated. For example, a quantifier in the first clause cannot bind a pronominal variable in the exceptive. In (17), the pronoun *tōna* 'his' in the second clause cannot be coindexed with the quantified DP 'each child' in the first clause. Such a relationship might be possible if the exceptive were a VP modifier attached low in the first clause.

(17) 'Ua rave maita'i te mau tamari'i tāta'itahi i te 'ohipa,
PFV do well DET PL child each ACC DET work

a. 'i tōna mahana fānaura'a noa 'aita
PREP his day birth only NEG

b. 'aita rā 'i tōna mahana fānaura'a
NEG but PREP his day birth
'Each child$_i$ worked hard, except on his$_{*i}$ birthday.'

3.7. By Way of Summary

We conclude that the exceptive in Tahitian is not a dedicated construction but a juxtaposition of two clauses in which the first makes a generalization and the second explicitly states an exception (or exceptions) to that generalization via a negative clause.

If this conclusion is on the right track, Tahitian instantiates "conjoined exceptives". This strategy can be compared to that of conjoined comparatives (Stassen 1985, Beck et al. 2010), where a gradable property is predicated of the subject of one conjunct, and asserted not to hold of the subject of the other conjunct. This is illustrated by (18) from Itelmen (Bobalijk 2012:19), where comparison is expressed indirectly via the juxtaposition of two contrasting clauses. As in the exceptive situation, languages that use conjoined comparatives do not have a dedicated comparative construction but, rather, express the intended meaning by employing clause juxtaposition, often with a negative clause in the second conjunct.

(18) Tinuʔn ʎeŋu-ʔn-č č'eβuz-laχ-aʔn, a xaŋnaʔn qaʔm *Itelmen*
these berries-PL-DIM sweet-ADJ-PL but those not
lit. "These berries are sweet but those are not."
'These berries are sweeter than those.'

We do not have data on exceptives in Itelmen. In the related language, Chukchi, comparatives are formed in a typical way by marking the standard of comparison with locative case, (19). Meanwhile, Chukchi exceptives are formed as in Tahitian, using a negative clause as the exceptive phrase, (20).[3]

(19) Iŋəl̩ çaça-ŋ wa-l̩ʔə-n l̩əɣuunʔ-ək *Chukchi*
bluberry.ABS tasty-ADV AUX-NMLZ-3SG lingonberry-LOC
'Blueberries taste better than crowberries.'

(20) Əməl̩ʔo jet-ɣʔe-t, Ivan l̩uŋ-jetə-l̩ʔə-n *Chukchi*
all.ABS come-PFV-3PL Ivan.ABS NEG-come-NMLZ-3SG
lit. "Everybody came, Ivan did not come."
'Everybody came except Ivan.'

Systematic data on a possible correlation between conjoined exceptives and conjoined comparatives is lacking but there does not appear to be one. Preliminarily we observe languages that have conjoined comparatives but not conjoined exceptives such as Fijian (Pearson 2010) and possibly Itelmen, and languages which have conjoined exceptives but not conjoined comparatives, such as Tahitian and Chukchi. English, Russian, and Malagasy are examples of languages that use neither. We are not yet aware of any languages that use conjoined structures for both comparatives and exceptives, and we do not think that a correlation between two structures will be found.

4. Conclusion and Future Work

This paper has provided an initial investigation into the analysis of exceptive constructions in Tahitian, arguing that they are expressed using reduced negative clauses. The obvious next question is how a full negative clause is reduced to yield the exceptive phrases we see in the examples repeated below.

(21) 'Ua tae pauroa mai te mau tamari'i, 'o Poe noa 'aita
PFV come all DIR DET PL child DET Poe just NEG
'All the children came, just/only Poe didn't.'

(22) 'Ua tae pauroa mai te mau tamari'i, 'aita rā 'o Poe
PFV come all DIR DET PL child NEG but DET Poe
'All the children came, but Poe didn't.'

[3] We are grateful to Marusya Pupynina for her help with the data.

We leave a full answer to this question for future work; any such work should probably be preceded by a more solid empirical exploration of different possibilities across languages.

References

Académie Tahitienne. 1986. *Grammaire de la langue tahitienne*. Papeete: Académie Tahitienne.

Beck, Sigrid, Sveta Krasikova, Daniel Fleischer, Remus Gergel, Stefan Hofstetter, Christiane Savelsberg, John Vanderelst, and Elisabeth Villalta. 2010. Crosslinguistic variation in comparison constructions. *Linguistic Variation Yearbook* 9:1–66.

Bobaljik, Jonathan. 2012. *Universals in comparative morphology*. Cambridge, Ma.: MIT Press.

Chung, Sandra. 1970. Negative verbs in Polynesian. MA thesis, Harvard University.

Chung, Sandra. 1978. *Case marking and grammatical relations in Polynesian*. Austin: University of Texas Press.

Hoeksema, Jacob. 1987. The logic of exception. P*roceedings of the Fourth Eastern States Conference on Linguistics*, eds. Amanda Miller and J. Powers, 100–113. Columbus, Ohio: The Ohio State University.

Hohepa, Patrick W. 1969. *Not* in English and *kore* and *eehara* in Māori. *Te Reo* 12:1–34.

Lazard, Gilbert, and Louise Peltzer. 2000. *Structure de la langue tahitienne*. Paris: Peeters.

Lemaître, Yves. 1973. *Lexique du tahitien contemporain*. Paris: Editions de l'IRD.

Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). 2016. *Ethnologue: Languages of the world, nineteenth edition*. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Markey, Patricia P. 1976. Tense and aspect in Tahitian. *Austronesian Studies: Papers from the Second Eastern Conference on Austronesian Languages*, ed. Buenaventura Naylor-Paz, 51–61. Ann Arbor: Center for South & Southeast Asian Studies, University of Michigan.

Pearson, Hazel. 2010. How to do comparison in a language without degrees: A semantics for the comparative in Fijian. *Proceedings of Sinn und Bedeutung* 14, eds. Martin Prinzhorn, Viola Schmitt, and Sarah Zobel, 356–372. University of Vienna.

Peltzer, Louise. 1996. La négation en tahitien. *Cahiers de l'institut de linguistique* 22-23:375–380.

Pérez-Jiménez, and Moreno-Quibén. 2012. On the syntax of exceptions. Evidence from Spanish. *Lingua* 122:582–607.

Soltan, Usama. 2016. On the syntax of exceptive constructions in Egyptian Arabic. *Perspectives on Arabic linguistics XXVII*, eds. Stuart Davis and Usama Soltan, 35–57. Amsterdam: John Benjamins.

Stassen, Leon. 1985. *Comparison and Universal Grammar*. Oxford: Basil Blackwell.

Tryon, D. T. 1970. *Conversational Tahitian*. Canberra: Australian National University Press.

Walroos, Sven. 2002. *English-Tahitian, Tahitian-English dictionary*. Honolulu: The Mā'ohi Heritage Press.

GENERAL NUMBER NOUNS IN AMHARIC LACK NUMP*

Ruth Kramer
Georgetown University
rtk8@georgetown.edu

In Amharic, a bare noun can be interpreted as referring to one or more entities. This phenomenon is often referred to as general number and the analysis of general number nouns across languages is controversial. Some have argued that general number nouns are incorporated nominals, whereas others have suggested they lack Num(ber)P. In this paper, I investigate the morphosyntax of Amharic general number nouns, arguing that they are not incorporated. Instead, I develop an analysis of these nouns as lacking NumP, and demonstrate how the analysis makes correct predictions about agreement (and how an alternative analysis that relies on a null plural marker does not). Overall, the paper provides new evidence for no-NumP approaches to general number and for general number being a heterogeneous phenomenon across languages.

1. Introduction

In Amharic (Ethiosemitic; Ethiopia), a bare noun can be interpreted as referring to one or more entities (Kapeliuk 1994, Baye 1996). For example, in (1), the noun *ɨidʒ* 'child' can be interpreted as singular ('one child came') or plural ('children came'), resulting in a translation of 'one or more children.'

(1) **ɨidʒ** mät't'-a
child come.PF-3MS
'One or more children came.'[1]

Following Corbett 2000, I call this phenomenon "general number" henceforth; similar nouns are also referred to as "number neutral" in some sources.

The analysis of general number nouns across languages is controversial. General number has been seen as a hallmark of nominal incorporation into a verb (see e.g., Massam 2009), but it has alternatively been analyzed as signaling a lack of Num(ber)P – without any incorporation (see e.g., Déprez 2005 et seq., Wiltschko 2008). In this paper, I investigate the morphosyntax of general number nouns in Amharic, arguing that they are not incorporated and that they lack NumP. I lay

* Many thanks to Girma Demeke, Bezza Ayalew, and Meriem Tikue for judgments and consultation on the Amharic data. For helpful questions and feedback, thanks are due to the students in the NYU S16 seminar on the syntax of number as well as audiences at the 8th WOCAL, the 44th NACAL, and the Berkeley Fieldwork Forum. Finally, all of my research projects show the enduring influence of Sandy Chung. I am deeply grateful to Sandy for suggesting that I work on Amharic in the first place and for being an inspirational researcher, teacher, and leader in the field.

[1] Gloss abbreviations: 3 –third person, ACC – accusative, ADJ – adjectival suffix, AGR – subject/verb agreement (Chamorro), AUX – auxiliary, C – complementizer, COLL – collective, DEF – definite marker, COM – comitative, F – feminine, IMPF – imperfective, IND – indicative, L – linker, M – masculine, .O – object marker, PF – perfect, PL – plural, PROG – progressive, S – singular, UNM – unmarked case, WH[OBJ] – wh-object agreement.

out the basic properties of Amharic general number nouns in Section 2 and develop the argument that they are not incorporated in Section 3. In Section 4, I develop a no-NumP analysis of these nouns and show how this analysis makes correct predictions. Section 5 concludes.

The analysis here is preliminary; the goal is to lay some groundwork because there is almost no previous theoretical linguistic research on Amharic general number nouns.[2] I also hope to stoke some interest in investigating general number in African languages and Semitic languages, which have not been a part of the general number literature (to the best of my knowledge, the only exceptions are Ajíbóyè 2010 and Jenks 2017). Finally, the paper provides some support for a heterogeneous approach to general number nouns (following Paul 2012) where general number must be carefully investigated in order to determine what it indicates about nominal syntax.

2. Amharic General Number Nouns: The Basics

In order to be interpreted with general number in Amharic, a noun must lack any overt marker of (in)definiteness, specificity or number (Kapeliuk 1994:10). For example, in (2), the noun *mäs'haf* 'book' is bare in this way and it can be interpreted with general number.

(2) lɨdʒ-u **mäs'haf** wässäd-ä
child-DEF **book** take.PF-3MS
'The child took one or more books.'

However, in (3)-(5), *mäs'haf* has a definite marker, a plural marker and an indefinite article.

(3) lɨdʒ-u **mäs'haf-u-n** wässäd-ä **Definite Marker**
child-DEF **book-DEF-ACC** take.PF-3MS
'The child took the book.'

(4) lɨdʒ-u **mäs'haf-otʃtʃ** wässäd-ä **Plural Marker**
child-DEF **book-PL** take.PF-3MS
'The child took some books.'

(5) lɨdʒ-u **and mäs'haf** wässäd-ä **Indefinite Article**
child-DEF **a book** take.PF-3MS
'The child took a book.'

Accordingly, a general number interpretation is not available for 'book' in (3)-(5).

For any language with general number nouns, it is important to determine whether the nouns are lexically ambiguous between singular and plural interpretations or whether they are truly unspecified for number. In a lexically ambiguous analysis, there would be two lexical entries for each noun with general number, one for the singular interpretation and one for the plural interpretation. In an unspecified-for-number analysis, the noun would have one lexical entry that is unspecified for number and thus can be used for any number.

[2] Baker 2014 is the only exception, and the discussion of Amharic is very brief. See Section 5 and fn. 15.

There is a helpful diagnostic to distinguish these two analyses: the interpretation of a lexically ambiguous word must be held constant over ellipsis. For example, in English, *pen* can mean either a writing implement or an enclosure for livestock, but when it is part of an ellipsis site, its meaning is fixed by the meaning of *pen* in the antecedent, as shown in (6).

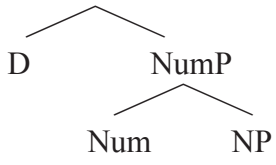
(6) Lee saw a pen, and Sam did, too.
   a. ✓Lee saw a writing implement, and Sam saw a writing implement, too.
   b. ✓Lee saw an animal enclosure, and Sam saw an animal enclosure, too.
   c. ✗ Lee saw a writing implement, and Sam saw an animal enclosure.
   d. ✗ Lee saw an animal enclosure, and Sam saw a writing implement.

Therefore, if a general number noun is lexically ambiguous, then its number interpretation as singular or as plural will be held constant over ellipsis. However, this prediction is not borne out in Amharic: (7) is four-ways ambiguous.

(7) lɨdʒ-u **mäs'haf** wässäd-ä ɨnna Tɨgɨst-ɨmm
   child-DEF **book** take.PF-3MS and Tigist-too
   a. ✓ The child took a book, and Tigist took a book, too.
   b. ✓ The child took books, and Tigist took books, too.
   c. ✓ The child took a book, and Tigist took books, too.
   d. ✓ The child took books, and Tigist took a book too.

Therefore, general number nouns in Amharic are truly unspecified for number, like general number nouns in Turkish (Bliss 2003:39-40), Mandarin (Rullmann and You 2006:177-178), Indonesian (Sato 2009; see also Chung 2000), and Malagasy (Paul 2012:101-102), among other languages.

I close this section with some brief discussion of the approach and assumptions of this paper. This paper will focus almost entirely on the morphosyntactic properties of general number nouns, not on their semantics (however, see Section 5). As for the morphosyntax of nominals, I assume that nominal phrases cross-linguistically are structured as in (8), although they may contain only some of these phrases (see e.g., Alexiadou, Haegeman, and Stavrou 2007):

(8)


= Determiner Phrase
= Number Phrase
= Noun Phrase[3]

Most importantly for present purposes, I assume that NumP is the main syntactic locus of number inflection and number interpretation across languages (Ritter 1991, Carstens 1991, and many others) and more particularly in Amharic (Baye n.d., Kramer 2009, 2016).

[3] I use NP throughout this paper for the purposes of easy comparison with previous work. However, I assume that a noun is decomposed into a category-neutral root and a nominalizing head *n* with the result that the category of a nominal is actually *n*P.

3. General Number Nouns are not Incorporated

Recall that general number nouns have been analyzed in two different ways – either as nominals that have been incorporated into a verb or as DPs that lack NumP. In this section, I argue that general number nouns in Amharic are not morphosyntactically incorporated. I focus on two broad types of incorporation: incorporation of the head N in Section 3.1 (e.g., Baker 1988) and incorporation of NP in Section 3.2 (also called pseudoincorporation; e.g., Massam 2001).

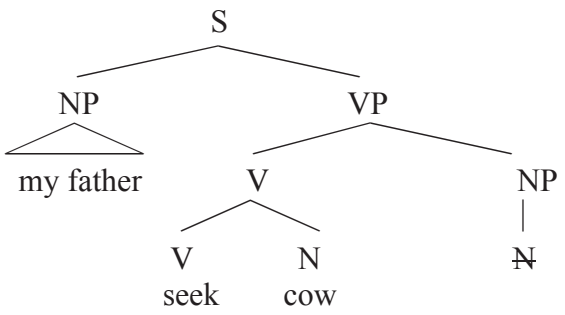
3.1. Incorporation of N

The classic analysis of nominal incorporation is that a noun undergoes head movement and adjoins to (= incorporates into) the verb (see e.g., Baker 1988, 1995, 2009, Baker, Aranovich, and Golluscio 2005; see also the update to this analysis in Baker 2014). I refer to this analysis as N-incorporation henceforth. One language that displays N-incorporation is Mapudungun (Mapudungu; Chile). In the non-incorporated example in (9)a, the object *ta-chi pu waka* 'the cows' has a definite marker and a collective marker; it is also (loosely speaking) in a separate syntactic location from the verb. In contrast, in the incorporated example in (9)b, the object is a bare noun *waka* 'cow' and it is a subcomponent of the verbal stem.

(9) a. Ñi chao kintu-le-y **ta-chi pu waka** — **Non-Incorporated**
my father seek-PROG-IND.3S **the-ADJ COLL cow** — [Mapudungun]
'My father is looking for the cows.' (Baker 2009:149)
b. Ñi chao kintu-**waka**-le-y — **Incorporated**
my father seek-**cow**-PROG-IND.3S — [Mapudungun]
'My father is looking for the cows.' (Baker 2009:149)

The structure of (9)b is roughly as in (10), where the N *waka* 'cow' has moved to adjoin to V.

(10) **N-Incorporation Analysis**
(after Baker 2009:150, (3))



Given this analysis, N-incorporation has several essential properties. First, N and V form a single complex head (possibly with other functional heads to be added, like the progressive and indicative markers in (9)b). Second, N-incorporation is only possible with internal arguments to V due to the Head Movement Constraint. Finally, since the incorporated N is just a head, it cannot be modified by e.g., adjectives.

When these properties are investigated for Amharic general number nouns, it becomes clear that the general number nouns are not incorporated N's. First of all, there is no obvious morphophonological incorporation of N into V. As the data and gloss in (2) suggest, there is a word break after *mäs'haf* in the writing system and (impressionistically) in the prosody. Moreover, general number nominals never intervene between the verb and its inflectional morphology, as they do in Mapudungun. For example, in (11), the general number noun *k'ɨt'äl* 'leaf' does not come between the progressive prefix and the verbal stem.

(11) ɨnnat-e **k'ɨt'äl** ɨyyä-t'ärräg-ätʃtʃ
mother-my **leaf** PROG-sweep.PF-3FS
'My mother, sweeping leaves…' (Kapeliuk 1994:15, orig. translation)

Secondly (and more tellingly), general number nouns are not limited to internal arguments. For example, in (12), the general number noun *lɨdʒ* 'child' is the external argument of a transitive verb.

(12) **lɨdʒ** kek-u-n bäll-a-w
child cake-DEF-ACC eat.PF-3MS-3MS.O
'One or more children ate the cake.'

Third, it is possible to modify a general number noun with an adjective, as in (13).

(13) lɨdʒ-u **k'äyy mäs'haf** wässäd-ä
child-DEF **red book** take.PF-3MS
'The child took one or more red books.' (Cf. Kapeliuk 1994:15)

A nominal phrase containing a general number noun can even be quite syntactically complex, including coordinated nouns and an adjoined PP as in (14).

(14) **ʃukka-na mänkiya kä-bɨrtʃ'ɨk'k'o gar** yɨ-k'atʃ'äl-al
fork-and spoon COM-glass with 3MS-clank.IMPF-AUX.3MS
'Fork[s], spoon[s], glass[es] clank together.' (Kapeliuk 1994:15, orig. translation)

Overall, then, general number nouns in Amharic are not incorporated nouns.[4]

3.2. Pseudoincorporation: Incorporation of NP

Certain languages contain nouns that seem to have some of the properties of N-incorporation (e.g., general number), but to behave like a full phrase otherwise (e.g., adjective modification is licit). This phenomenon is often known as pseudoincorporation (after Massam 2001), and it is most often analyzed as a structurally-reduced nominal phrase (e.g., NP instead of DP) serving as the complement to the verb. At first, this might seem like a more promising approach to Amharic

[4] It is not likely that they are incorporated roots, either (Wiltschko 2008) because, if some Amharic nouns are formed via root and pattern morphology (Leslau 1995, Kramer 2016), the nominal root can be just a string of consonants.

general number nouns, but I show in this section that Amharic general number nouns are not pseudoincorporated.

Psuedoincorporation has perhaps been most thoroughly investigated for Austronesian languages (see e.g., Massam 2001 on Niuean, Chung and Ladusaw 2004 on Chamorro, Ball 2008 on Tongan). However, here I start by comparing Amharic general number nouns to pseudoincorporated nouns in Turkish (Öztürk 2009) because Amharic and Turkish are broadly similar syntactically. (15) contains a non-pseudoincorporated nominal *kitabı* 'the book;' it is case-marked and has singular number.

(15) Ali **kitab-ı** okudu. **Non-pseudoincorporated Nominal**
Ali **book-ACC** read [Turkish]
'Ali read the book.' (Öztürk 2009:335, (1b))

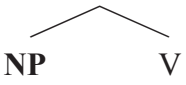
In (16), the pseudoincorporated object *kitap* 'book' does not have case-marking and it has general number.

(16) Ali **kitap** okudu. **Pseudoincorporated Nominal**
Ali **book** read [Turkish]
'Ali did book reading.' (Öztürk 2009:335, (1a))

However, Öztürk (2009) shows that *kitap* 'book' and other nominals like it are not the result of N-incorporation. For example, these nominals can be modified by adjectives.

(17) Ali **ekşi elma** yedi. [Turkish]
Ali **sour apple** ate
'Ali did sour apple eating.' (Öztürk 2009:339, (14a))

Instead, following Massam 2001, Öztürk proposes that the incorporated nominal is merged as the immediate NP sister to V, as shown in (18).

(18) 

(Öztürk 2009:340, (17))

The lack of functional projections in the nominal phrase explains many of the key properties of pseudoincorporated nouns. Since there is no NumP, the noun is unspecified for number and thus has general number. Since there is no KP projection, it does not have case marking. Also, assuming APs are adjoined to NP, APs are licit modifiers of the pseudoincorporated nominal.

Given this analysis of pseudoincorporation, there are a couple of reasons not to treat Amharic general number nouns as pseudoincorporated. Although these nouns (of course) have

general number like Turkish pseudoincorporated nouns, unlike Turkish, they can be part of a DP that is assigned accusative case, as in the possessed general number noun in (19).[5]

(19) lɨdʒ-u **yä-Almaz-ɨn mäs'haf** wässäd-ä
child-DEF **of-Almaz-ACC book** take.PF-3MS
'The child took one or more books of Almaz's.'

It may seem suspicious that the accusative case marker is on the possessor in (19), not the head noun. However, this is the typical placement of the case marker in possessed DP's, as is clear from the non-general number example in (20).

(20) Almaz yä-Gɨrma*(**-n**) wändɨmm ayy-ätʃtʃ-ɨw
Almaz of-Girma*(-ACC) brother see.PF-3FS-3MS.O
'Almaz saw Girma's brother.'

This (somewhat unusual) placement of the accusative case marker follows from the analysis of Amharic case marking developed in some of my previous work in collaboration with Mark Baker (Baker and Kramer 2014). We proposed that all object DPs are assigned accusative case syntactically and this results in a [+ACC] feature being added to each (morphological) word in the DP. However, the [+ACC] feature can only be realized on definite (morphological) words like pronouns, proper names, or definite-marked nouns; it cannot be realized on bare nouns like the head nouns of the DP objects in (19) and (20). This explains why it is necessary to add a possessor to see the case-marking emerge for a general number noun; a bare noun can never be the host for an overt accusative case marker.

The fact that general number nouns in Amharic receive case is important because pseudoincorporated nominals do not generally have case markers and are thus assumed not to receive case (this claim is most clearly articulated in Levin 2015). This is true (as far as I know) for all pseudoincorporation languages that would otherwise overtly case-mark an object including Niuean (Massam 2001), Turkish (Öztürk 2009), Tongan (Ball 2008), and Sakha and Tamil (Baker 2014).[6] So, the fact that Amharic general number nouns can be part of a DP that is assigned accusative case is a strong indication that they are not pseudoincorporated.

Moreover, recent work has argued that, in pseudoincorporation, the pseudoincorporated noun must be immediately adjacent to the verb (Baker 2014, Levin 2015).[7] For example, in Chamorro, nominal modifiers can adjoin on either side of the N, like the relative clause in (21).

---

[5] It is not clear whether the ability to have a possessor in itself indicates that Amharic general number nouns are not pseudoincorporated. While pseudoincorporated nouns cannot have possessors in Niuean (Massam 2001:168) or Chamorro (Chung and Ladusaw 2004:88), there is no data on possessors in Öztürk 2009. It is possible that possessors are within NP in Amharic, in which case they would be expected in a pseudoincorporated nominal.

[6] One exception here is Hungarian (Farkas and de Swart 2003:94) but the relevant nominals are likely not morphosyntactically incorporated (see Baker 2014:20, fn. 12).

[7] There are two apparent exceptions to this generalization. The first is that, in some languages, pseudoincorporated objects can scramble away from the verb (Turkish: Öztürk 2009, Hindi and Amharic: Baker 2014). See Baker 2014 for one approach to this exception. The second is that, in Turkish, the focus particle *bile* 'even' can intervene between the pseudoincorporated object and the verb (Öztürk 2009:346). However, it may be that this focus particle is part of the verbal morphology, and thus the pseudoincorporated object is still adjacent to the verb broadly defined.

(21) a. ädyu i [yä-hu] na lepblu **Noun-Final Order**
that the WH[OBJ].like-AGR L book [Chamorro]
'that book which I like'
b. ädyu i lepblu [ni yä-hu] **Noun-Initial Order**
that the book C WH[OBJ].like-AGR [Chamorro]
'that book which I like' (Chung and Ladusaw 2004:80, (8ab))

However, when a nominal phrase is pseudoincorporated, only the noun-initial order is licit. The idea is that this is because the N is immediately adjacent to the V only in this order (Chung and Ladusaw 2004:143-144).

(22) a. Täi-[amiga ni yä-hu] si Carmen **Noun-Initial Order**
AGR.not.have-friend C WH[OBJ].like-AGR UNM Carmen [Chamorro]
'Carmen has no women friends who I like.'
b. *Gäi-[yä-hu na kareta] si Juan **Noun-Final Order**
AGR.have-WH[OBJ].like-AGR L car UNM Juan [Chamorro]
Intended: 'Juan owns a car that I like' (Chung and Ladusaw 2004:143, (19) and (20))

In Amharic, though, as seen in Section 3.1, a general number noun need not be adjacent to the verb. In fact, it can be in the typical subject position initial in the sentence, separated from the verb by a definite, case-marked DP object, as in (23).[8]

(23) **lidʒ** kek-u-n bäll-a-w
child cake-DEF-ACC eat.PF-3MS-3MS.O
'One or more children ate the cake.'

Thus, in addition to case-marking, this is more evidence that general number nouns in Amharic are not pseudoincorporated. In fact, this particular constellation of facts (presence of case marking, absence of verb adjacency) is exactly what Levin 2015 predicts for a language that lacks pseudoincorporation. Levin proposes that pseudoincorporated nouns must be adjacent to the verb because they are structurally too small to receive Case; thus, they need to adjoin to the verb (and become part of its projection) in order to escape the Case Filter. In Amharic, general number nouns are (apparently) part of a big enough projection to receive case, so they need not be adjacent to the verb.

Overall, I conclude that Amharic general number nouns are neither incorporated N's, nor pseudoincorporated NPs.

4. Analysis: No NumP

[8] While it might also seem remarkable that a transitive subject is pseudoincorporated in this example, this is not the most salient aspect of the data since transitive subjects can also be pseudoincorporated in Turkish (Öztürk 2009). However, importantly, when a transitive subject is pseudoincorporated in Turkish, it **must** be adjacent to the verb (Öztürk 2009:335, (3) and (4)), so it is the lack of verb-adjacency that truly sets apart Amharic general number nouns.

4.1. The Fundamentals of the Analysis

In this section, I develop a preliminary analysis of general number nouns in Amharic as lacking NumP, in support of previous NumP-less approaches to general number nouns (see e.g., Déprez 2005 et seq., Wiltschko 2008, Sato 2009, Loewen 2011). In Amharic, general number nominals behave as if they were contained in typical DP arguments. They have the syntactic distribution of DPs: subject of a transitive verb (23), subject of an intransitive verb (1), direct object (2), object of a postposition (14), and indirect object, as shown in (24).

(24) Almaz kärämela-wotʃtʃ-u-n **lä-lidʒ** sät'-ätʃtʃ
Almaz candy-PL-DEF-ACC to-child give.PF-3FS
'Almaz gave the candy to one or more children.'

General number nouns are also assigned Case ((19)) and they can appear with all kinds of DP-internal phrases including adjectives ((13)), possessors ((19)), and relative clauses, as in (25).

(25) astämari-w **kä-Vermont yä-näbbär-ä-n** **pom** gäzz-a
teacher-DEF **from-VT** **C-be.PF-3MS-ACC** **apple** buy.PF-3MS
'The teacher bought one or more apples which were from Vermont.'

So, in many ways, general number nouns are unremarkable in Amharic; they show the characteristic syntactic behavior of DP's.

However, they are unlike typical DPs in two key respects: they of course have general number and also they are obligatorily indefinite. In (26), the noun *mäs'haf* 'book' cannot be interpreted as definite when it is bare.

(26) mäs'haf
a. 'a book' c. × 'the book'
b. 'books' d. × 'the books'

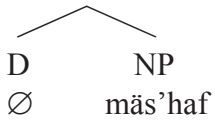
Accordingly, if a noun is definite, then it cannot have general number, as shown in (27).

(27) mäs'haf-u
book-DEF
a. 'the book'
b. × 'the books'

So an analysis of general number nouns will have to explain why they behave like typical DPs syntactically, but have general number and must be indefinite.

To start accounting for these generalizations, I propose that an Amharic general number noun is a DP that lacks NumP. This is shown in (28) for the recurrent example *mäs'haf* 'one or more books.'

(28) **Structure of a General Number Noun**



This predicts that general number nouns will have the syntactic behavior of DPs in Amharic (distribution, case marking, etc.). It also predicts that they will not incorporate because they are structurally too large (DPs, not N's or NP's). The lack of NumP also generates the general number interpretation under the following assumptions (based mostly on Wiltschko 2008:648-650, cf. Link 1983). First, I assume that a bare noun denotes a set comprised of atomic entities and of pluralities. For example, if there were only three books in the world, then the denotation of *mäs'haf* 'book' would be as in (29).[9]

(29) *mäs'haf* 'book' =
{ {b1}, {b2}, {b3}, ← atomic entities
{b1, b2}, {b2, b3}, {b1, b3}, {b1, b2, b3}} ← pluralities

I also assume that a Num head restricts the denotation of a noun – to just atomic entities for Num[-PL] or to just pluralities for Num[+PL]. Since bare nouns in Amharic lack Num, they can be interpreted as 'one or more,' i.e., with general number.

(30) lɨdʒ-u **mäs'haf** wässäd-ä
child-DEF **book** take.PF-3MS
'The child took one or more books.'

(30) is compatible with contexts where 'book' is interpreted as an atomic entity (e.g., {b1}) or as a plurality (e.g., {b1+b2+b3}). Thus, treating general number nouns as DPs that lack NumP explains why they behave like DPs morphosyntactically but they are interpreted as if they lack any specification for number.

However, the distribution of Num must now be carefully constrained such that general number interpretations are only generated in the right contexts. I propose that this is accomplished via syntactic selection by D. The selectional properties of D in Amharic are in (31).

(31) **Selectional Properties of D in Amharic (first pass)**
a. D[+DEF] selects for NumP.
b. D[-DEF] selects for NP.

Definite D selects for NumP whereas indefinite D selects for NP. This derives the key generalizations that (i) general number nouns are always indefinite, and (ii) definite nouns cannot have general number. The NumP in a definite DP forces an interpretation of N as an atomic entity (Num[-PL]) or a plurality (Num[+PL]), but the lack of NumP in indefinite N's allows for an interpretation as either an atomic entity or a plurality.

[9] I represent the denotation of a nominal here as a set, but see Link 1983 for a formalization using join semi-lattices.

However, these selectional properties must be refined because Amharic has indefinite plural-marked nouns:

(32) lɨdʒ-u **mäs'haf-otʃtʃ** wässäd-ä
child-DEF **book-PL** take.PF-3MS
'The child took (some) books.'

Assuming that the plural marker *–otʃtʃ* is a realization of Num (Baye n.d., Kramer 2009, 2016), then an indefinite D must be able to select for a NumP as well. This leads us to the slightly looser restrictions in (33), where an indefinite D selects for a NumP or NP complement.

(33) **Selectional Properties of D in Amharic (final version)**
a. D[+DEF] selects for NumP.
b. D[-DEF] selects for either NumP or NP.

Putting all the claims together, we arrive at the following table of possible DPs in Amharic.

Table 1: Types of DPs in Amharic

***mäs'haf* 'book'**	**Num[-PL]**	**Num[+PL]**	**N**
D[+DEF] *–u*	Definite singular: *mäs'haf-u* 'the book'	Definite plural: *mäs'haf-otʃtʃ-u* 'the books'	*
D[-DEF] -∅	Indefinite singular: *mäs'haf* 'a book'	Indefinite plural: *mäs'haf-otʃtʃ* books'	General number: *mäs'haf* 'one or more books'

In the first row, definite nominals contain either singular or plural Num and are accordingly interpreted as either singular or plural. Indefinite nominals can contain singular Num or plural Num, but they also have the option to lack Num and thus can be interpreted with general number.

Overall, a no-NumP analysis makes the right analytical cuts: general number nouns have the morphosyntax of DPs, the lack of Num leads to a general number interpretation, and Num can only be absent when D is indefinite and can select directly for an NP.[10]

4.2. Ruling out an Alternative

In this section, I briefly develop an alternative analysis to the no-NumP approach from the previous section. I show how the alternative analysis makes a false prediction about agreement, and then demonstrate how the no-NumP analysis makes the right prediction for the agreement data. I take this as further evidence in support a no-NumP approach to general number in Amharic.

[10] The analysis also predicts that general number nouns are bare, since I assume definite markers are realizations of D[+DEF] (Kramer 2010) and plural markers are realizations of Num. However, there is one small sticking point: the indefinite quantifier *and* (see (5)) forces a singular interpretation, but it is indefinite. However, it is homophonous with the numeral *and* 'one' and I suspect it is a type of singular Num head itself.

In the no-NumP analysis, indefinite singular DPs are structurally ambiguous: they contain either a singular NumP or no NumP at all.[11] Therefore, it is not (strictly speaking) required to omit NumP in order to generate a singular interpretation of a bare indefinite noun. To a skeptical reader, this may call into question whether omitting NumP is really ever required at all. Instead, perhaps Amharic also contains a null plural Num, so that a bare noun would either contain a null singular Num or a null plural Num. If so, and if D[-DEF] alone selects for null plural Num, this null-plural approach would generate the basic facts of general number as adequately as the no-NumP analysis.

However, the null-plural approach makes an incorrect prediction about agreement. In Amharic, DPs that contain Num[+PL] trigger plural agreement, regardless of how Num[+PL] is exponed. For example, in (34), the demonstrative must be plural when it is associated with a DP that contains Num[+PL], whether the Num[+PL] is exponed irregularly (via an atypical suffix plus templatic changes) or regularly (via the typical plural suffix *-ot/tʃ*). The same goes for (35) where the copula must have plural agreement when the subject DP contains a Num[+PL].

(34) a. ɨnnäzih mäs'aɦɨft/mäs'haf-otʃtʃ
DEM.PL book.PL/book-PL
'these books'

b. *yɨh mäs'aɦɨft/mäs'haf-otʃtʃ
DEM.S book.PL/book-PL
'*this books'

(35) mäs'aɦɨft-u/mäs'haf-otʃtʃ-u sämayawi natʃtʃäw/*näw.
book.PL-DEF/book-PL-DEF blue be.3PL/*be.3MS
'The books are/*is blue.'

Therefore, if a bare DP contains a null plural Num, this bare DP should trigger plural agreement. However, this prediction is not borne out. Bare nouns that can be interpreted as plural trigger only singular agreement, as shown on the subject agreement on the main verb in (36) and the verb in the relative clause in (37) (Kapeliuk 1994:15, Leslau 1995:180).[12]

(36) **lɨdʒ** mät't'-**a**
child come.PF-3MS
'One or more children came.'

(37) astämari-w kä-Vermont **yä-näbbär-ä-n** **pom** gäzz-a
teacher-DEF from-VT C-be.PF-3MS-ACC apple buy.PF-3MS
'The teacher bought one or more apples which were from Vermont.'

[11] In Table 1, there are also two distinct ways to generate an indefinite plural noun: either as a bare noun (general number) or as a noun with a plural suffix. This may explain why indefinite plural nouns with the plural suffix are reportedly used in particular contexts: for specific indefinites, for humans, for small numbers, etc. (Kapeliuk 1994:25-28, Baye 1996:69-70, Baye n.d.). However, not all speakers share this contrast and it remains to be confirmed what exactly the context is. Cf. Chung 2000 on Indonesian.

[12] Kapeliuk (1994:15) notes two exceptions: adjectives with reduplicated plural forms are licit with general number nouns, and general number nouns that denote humans trigger plural agreement. The latter observation does not hold in all contexts given data like (37), and I have yet to test general number nouns with reduplicated plural adjectives.

In contrast, this agreement pattern is easily predicted in the no-NumP analysis. In this approach, since general number nouns lack NumP, they lack number features. Therefore, when a probe agrees with a general number noun, it will receive no value for number. This is shown schematically in (38) for the verbal functional head that carries subject agreement (which I assume is Asp(ect) since agreement paradigms vary based on aspect; Leslau 1995).

(38)

Asp	…	DP	→	Asp	…	DP
[__ PERS]		[3 PERS]	AGREE	[3 PERS]		[3 PERS]
[__ FEM]		[-FEM]		[-FEM]		[-FEM]
[__ PL]				**[__ PL]**		

Therefore, number agreement with general number nouns will always be the morphological default value for number, i.e., singular, rather than plural (or perhaps, both singular and plural).[13] In sum, then, an alternative analysis based on a null plural Num makes incorrect predictions about agreement, whereas the no-NumP analysis developed in Section 4.1 correctly generates the facts.[14]

5. Conclusion

To recap, Amharic has bare nouns with general number. They must be interpreted as indefinite, but they behave morphosyntactically like DPs (e.g., they are not incorporated). To account for these facts, I proposed that the DP that contains a general number noun lacks NumP and that only indefinite D's select for NP. This analysis makes correct predictions about the agreement patterns of general number nouns, unlike an alternative analysis based on a null plural Num.

In terms of its larger consequences, the analysis provides support for a heterogeneous approach to the analysis of general number nouns in general (as suggested in e.g., Paul 2012). It seems that general number can arise through (or is compatible with) several different syntactic configurations (incorporation, lack of NumP, possibly others). This indicates that general number nouns are not necessarily a natural class. The analysis also provides additional evidence for the existence of DPs that lack NumP across a broad range of language families, as developed in Déprez 2005 et seq. for Haitian Creole, Wiltschko 2008 for Halkomelem (Salish), and Loewen 2011 for Indonesian (see also Chung 2000, Sato 2009).

While Amharic general number nouns do not have the morphosyntax of N- or NP-incorporation, it is a separate question whether or not they have the semantics of incorporation, that is, the semantic properties associated with complex predicate formation (see e.g., Van Geenhoven 1998, Farkas and de Swart 2003, Chung and Ladusaw 2004, Dayal 2011). It may be

[13] As a side benefit, this supports the recent analysis of agreement as an obligatory operation (Preminger 2014), where grammaticality is determined on the basis of whether agreement is attempted, not in whether it is (wholly) successful in the valuation of syntactic features.

[14] There is one additional prediction worth briefly exploring. In my previous work (Kramer 2009, 2016), I argued that Amharic contains a *n*[+PL] which is realized with irregular plural morphology. However, if irregular-plural marked nouns are NPs (=*n*Ps), it is predicted that they could be interpreted with general number (they could be selected for by D[-DEF]). Data is not yet available on whether this is possible. If not, it could be due to blocking: in order to avoid homophony (an irregular plural noun interpreted as either singular or plural), the exponent of a *n*[+PL] could only be inserted in the context of Num[+PL]. This kind of restriction is independently necessary to prevent irregular plural nouns from occurring in singular contexts as well (Kramer 2016).

that certain general number nouns do behave semantically like part of a complex predicate with the verb. For example, Baker (2014) observes that a general number object in Amharic takes low scope with respect to an adverb of quantification, as in (39); this can be an indication of complex predicate formation.[15]

(39) Lämma ɨndägäna ɨndägäna **mäs'haf** gäzz-a.
Lemma again again book buy.PF-3MS
'Lemma repeatedly bought book(s).' (different books, different times) (Baker 2014:39)

Additionally, from a philological perspective, Kapeliuk (1994:12-13) proposes that Amharic has incorporated objects that must be verb-adjacent, and her translations suggest these incorporated objects have general number. However, she considers incorporated objects to be distinct from general number nouns, which she demonstrates can occur in a variety of syntactic positions (Kapeliuk 1994:15-18).

So, there is some reconciliation of disparate facts to be done in future research. Are Amharic general number nouns semantically heterogeneous within the language itself (some form part of a complex predicate whereas others do not)? Or are they semantically homogeneous? If they are homogeneous, they either (i) all form complex predicates (even external argument general number nouns like in (23)), or (ii) none form complex predicates. The answers to these questions might help explain why general number nouns can only be indefinite, which is captured in the present analysis but would be predicted if they were always part of a complex predicate. However these questions pan out, though, the answers will be relevant to determining whether syntactic incorporation is required for complex-predicate interpretation as well as how syntactic structure maps to semantic interpretation generally.

References

Ajíbóyè, Ọládiípọ̀. 2010. Plural strategies in Yoruba. *Studies in African Linguistics* 39:141-181.
Alexiadou, Artemis, Liliane Haegeman, and Melita Stavrou. 2007. *Noun phrase in the generative perspective.* New York: Mouton.

[15] Baker 2014 uses (39) to argue that Amharic has syntactically incorporated general number nouns; I address this approach briefly here. First, as noted above, (39) may indicate complex predicate formation semantically, but not necessarily syntactic incorporation. Second, Baker assumes that it is parametric whether or not incorporated nouns are case-marked. However, except for Hungarian (which probably does not have syntactic incorporation, see fn. 6), there are no attested instances of an incorporated noun receiving case-marking (Levin 2015). Thus, I continue to take the case-marking of Amharic general number nouns to indicate that they are not syntactically incorporated. Finally, Baker provides an example of an Amharic bare noun in subject position that does not have general number and has wide scope, suggesting that general number nouns must be verb-adjacent in Amharic.

(i) wɨʃʃa Almaz-ɨn ɨndägäna ɨndägäna näkkäs-at.
dog Almaz-ACC again again bite.PF.3MS-3FS.O
'A dog bit Almaz repeatedly.' (the same dog in each event) (Baker 2014:39)

However, my consultants found general number nouns in subject position to be grammatical, although none of the sentences contained adverbs of quantification. In this paper, I build on the data that I elicited, but future work will hopefully reconcile the data here with data like (i).

Baker, Mark. 1988. *Incorporation: A theory of grammatical function changing.* Chicago: U. of Chicago Press.
Baker, Mark C. 1995. Lexical and nonlexical noun incorporation. *Lexical knowledge in the organization of language,* eds. U. Egli, P. Pause and C. Schwarze, 3-34. Amsterdam: Benjamins.
Baker, Mark C. 2009. Is head movement still needed for noun incorporation? *Lingua* 119:148-165.
Baker, Mark C. 2014. Pseudo noun incorporation as covert noun incorporation: linearization and cross-linguistic variation. *Language and Linguistics* 15:5-46.
Baker, Mark C., Roberto Aranovich and Lucía A. Golluscio. 2005. Two types of syntactic noun incorporation: noun incorporation in Mapudungun and its typological implications. *Language* 81:138–176.
Baker, Mark and Ruth Kramer. 2014. Rethinking Amharic prepositions as case markers inserted at PF. *Lingua* 145:141-172.
Ball, Douglas. 2008. Clause structure and argument realization in Tongan. Doctoral dissertation, Stanford University.
Baye Yimam. 1996. Definiteness in Amharic discourse. *Journal of African Languages and Linguistics* 17:47-83.
Baye Yimam. Not dated. The Amharic noun phrase.[16] Ms., Addis Abeba University.
Bliss, Heather. 2003. The semantics of the bare noun in Turkish. BA thesis, University of Calgary.
Carstens, Vicki. 1991. The morphology and syntax of determiner phrases in Kiswahili. Doctoral dissertation, UCLA.
Chung, Sandra. 2000. On reference to kinds in Indonesian. *Natural Language Semantics* 8:157-71.
Chung, Sandra, and William A. Ladusaw. 2004. *Restriction and saturation.* Cambridge, MA: MIT Press.
Corbett, Greville. 2000. *Number.* Cambridge: CUP.
Dayal, Veneeta. 2011. Hindi pseudo-incorporation. *Natural Language and Linguistic Theory* 29: 123-167.
Déprez, Viviane. 2005. Morphological number, semantic number and bare nouns. *Lingua* 115: 857-883.
Farkas, Donka and Henriëtte de Swart. 2003. *The semantics of incorporation: From argument structure to discourse transparency.* Stanford: CSLI Publications.
Jenks, Peter. 2017. Numeral classifiers compete with number marking: evidence from Dafing. Paper presented at the Annual Meeting of the Linguistic Society of America, Austin, TX.
Kapeliuk, Olga. 1994. *Syntax of the noun in Amharic.* Wiesbaden: Harrassowitz.
Kramer, Ruth. 2009. Definite markers, phi-features, and agreement: a morphosyntactic investigation of the Amharic DP. Doctoral dissertation, UC Santa Cruz.
Kramer, Ruth. 2010. The Amharic definite marker and the syntax-morphology interface. *Syntax* 13:196-240.
Kramer, Ruth. 2016. A split analysis of plurality: number in Amharic. *Linguistic Inquiry* 47:527-559.
Leslau, Wolf. 1995. *Reference grammar of Amharic.* Wiesbaden: Harrassowitz.

[16] The copy of the manuscript that I have access to is untitled; this is my best guess at a title.

Levin, Theodore. 2015. Licensing without Case. Doctoral dissertation, MIT.
Link, Godehard. 1983. The logical analysis of plural and mass terms: a lattice-theoretical approach. *Meaning, use, and interpretation of language*, eds. R. Bäuerle et al., 302-323. Berlin: Mouton.
Loewen, Gina. 2011. The syntactic structure of noun phrases in Indonesian. Master's thesis, University of Manitoba.
Massam, Diane. 2001. Pseudo noun incorporation in Niuean. *Natural Language and Linguistic Theory* 19:153-197.
Massam, Diane. 2009. Noun incorporation: essentials and extensions. *Language and Linguistics Compass* 3/4:1076-1096.
Öztürk, Balkız. 2009. Incorporating agents. *Lingua* 119:334-358.
Paul, Ileana. 2012. General number and the structure of DP. *Count and mass across languages*, ed. Diane Massam, 99-112. Oxford: OUP.
Preminger, Omer. 2014. *Agreement and its failures*. Cambridge, Mass.: MIT Press.
Ritter, Elizabeth. 1991. Two functional categories in noun phrases: evidence from Modern Hebrew. *Perspectives on phrase structure: heads and licensing,* ed. Susan Rothstein, 37-62. San Diego: Academic Press.
Rullmann, Hotze and Aili You. 2006. General number and the semantics and pragmatics of indefinite bare nouns in Mandarin Chinese. *Where semantics meets pragmatics*, eds. K. von Heusinger and K. Turner, 175-196. Bingley: Emerald.
Sato, Yosuke. 2009. Radical underspecification, general number, and nominal mapping in Indonesian. *Proceedings of AFLA 16*, 197-209. Department of Linguistics: UC Santa Cruz.
Van Geenhoven, Veerle. 1998. *Semantic incorporation and indefinite descriptions*. Stanford: CSLI.
Wiltschko, Martina. 2008. The syntax of non-inflectional plural marking. *Natural Language and Linguistic Theory* 26:639-694.

SYNTACTIC IDENTITY IN SLUICING: SPROUTING IN KASHMIRI CAUSATIVES*

Emily Manetta
University of Vermont
emily.manetta@uvm.edu

Sluicing is understood to be constrained by identity between the elided clause and its antecedent, and a rich research program has centered on an investigation of the precise nature of that identity. This program, led by work including Merchant 2001, Chung 2006, 2013, and Merchant 2013 suggests that in addition to semantic identity, some limited form of syntactic identity must also apply. In this short paper I bring additional evidence from a relatively understudied Indic language, Kashmiri, to support the specific formulation of syntactic identity advanced in Chung 2013. In particular, I argue that sluicing in Kashmiri causative constructions with sprouted nominative causee remnants provides evidence that both the argument structure condition and Case condition are necessary.

1 Introduction

A robust research program has centered around investigating the identity conditions on sluicing and other forms of ellipsis, both syntactic and semantic (Ross 1969; Rooth 1992; Hardt 1993; Chung, Ladusaw, and McCloskey (CLM) 1995, 2011; Merchant 2001; AnderBois 2011; Barros 2014; among many others). Recent work on a diverse range of languages suggests that both semantic and limited syntactic identity must hold between elided material and its antecedent (Chung 2006, 2013; Merchant 2013; Barker 2013). This short paper seeks to provide additional empirical support from a new language family for the specific version of limited syntactic identity advanced in Chung 2013.

Kashmiri is a relatively understudied Indic language, unique within its language family for being verb-second and featuring obligatory wh-fronting. It exhibits an unremarkable version of sluicing: an operation characterized by the ellipsis of a clause-sized constituent (TP), leaving behind a fronted wh-phrase remnant:

(1) RaajI rani kaaNh magar me chu-nI pataa k'a.
Raj cook-FUT something but I AUX-not know what
'Raj will cook something, but I don't know what.'

Chung (2013) argues, based on evidence from English and the Austronesian language Chamorro, that "limited syntactic identity" in sluicing specifically takes the form of two conditions: one on argument structure and one on Case assigning heads.

* I am grateful to Vijay Choudhury and her family for their judgements on the Kashmiri data discussed here. For productive conversations about the causative at various points I thank Omkar Koul, Abir Bazaz, and Chandra Gupta. For productive conversations about sluicing over the last few years I thank Vera Gribanova and Jim McCloskey. Thanks also to the organizers of this volume and especially Ruth Kramer for her helpful comments and questions – though all remaining errors are certainly my own. Of course, none of this work would be possible without the matchless support of Sandy Chung. I am still mindful of, and often find myself returning to, her sharp questions and clear advice. I'm also grateful to Sandy for teaching me how to keep an orchid blooming: fertilizer.

(2) Limited syntactic identity in sluicing
   a. Argument structure condition: If the interrogative phrase is the argument of a predicate in the ellipsis site, that predicate must have an argument structure identical to that of the corresponding predicate in the antecedent clause.
   b. Case condition: If the interrogative phrase is a DP, it must be Case-licensed in the ellipsis site by a head identical to the corresponding head in the antecedent clause. (Chung 2013:30)

In this paper I present new data from sluicing in causative constructions in Kashmiri supporting the formulation in (2), and in particular motivating the Case condition on syntactic identity.

2 Kashmiri causatives

The productive morphological causative in Kashmiri is derived by the addition of a suffix *-aav* or *-Inaav/Iraav* to the inflected verb. It has been described relatively thoroughly (Hook and Koul 1984, 1987; Syeed 1985; Altaha 1992), and elsewhere given a modicum of analytical and comparative attention (Wali 1980, 1981; Bhatt 1999; Bhatt and Embick 2003; Manetta 2014).

(3) bI chus aslam as-Inaav-aan
I.NOM AUX.1SG Aslam.NOM laugh-CAUSE-IMPFV
'I am causing Aslam to laugh.' (Wali and Koul 1998:213)

(4) təm' d'aav-Inoov-us bI mohn-as kitaab
he.ERG gave-CAUS.PST-1SG I.NOM mohan-DAT book.NOM
'He had me give a book to Mohan.' (Wali and Koul 1998:215)

In the interests of space, I will confine our attention to the features of the Kashmiri causative of most interest here: differential case marking found on the causee in causatives of transitive verbs.

Kashmiri is split-ergative and requires ergative marking on transitive subjects in the perfective aspect. Direct objects in these structures consistently appear in the unmarked case, which is variously termed nominative or absolutive in the literature (I will use 'nominative' here, following Anand and Nevins 2006). In Kashmiri, as in Hindi-Urdu, the verb must agree with the most promninent nominative argument; thus in simple perfective transitive clauses, agreement is with the internal argument since the subject is marked ergative. Unsurprisingly, case and agreement possibilities in causative clauses in Kashmiri are more complex. As far as the causee is concerned, two patterns emerge. In Pattern A (in (5)), the causee is nominative. Primary agreement (gender and number) in Pattern A is with the most prominent unmarked argument and thus with the nominative causee, not the more embedded nominative direct object.

(5) PATTERN A: NOMINATIVE CAUSEE
me ran-InaavInoov su oolav
I.ERG cook-CAUSE.PST.MSG he.NOM potatoes.NOMPL
'I had him cook potatoes.' (Wali and Koul 1998:214)

In the second pattern, called Pattern B, the causee is in the dative case, as a complement of the postposition *athI,* sometimes translated as 'by' and sometimes glossed as 'instrumental' (here I

use INSTR).[1] In the perfective aspect, this means that the only nominative (unmarked) argument in the clause is the direct object, which will thus control primary agreement.

(6) PATTERN B: OBLIQUE CAUSEE[2]
me ran-InaavInəəv əmis athi oolav
I.ERG cook-CAUS.PST.MPL he.DAT INSTR potatoes.NOMPL
'I had him cook potatoes.' (Wali and Koul 1998:214)

A third variant is found in the causative of ingesto-reflexive verbs. Pattern A is identical to that for transitive verbs above, but in Pattern B, ingesto-reflexives require the causee to be marked dative only, and to not appear with the post-position *athI*. Primary agreement must still be controlled by the only remaining unmarked argument, the direct object.

(7) INGESTO-REFLEXIVE PATTERN A: NOMINATIVE CAUSEE
asi par-Inəəv kuur hisaab
we.ERG study-CAUSE.PST.FSG girl.F.NOM math.M.NOM
'We taught the girl math'. (Hook and Koul 1984:105)

(8) INGESTO-REFLEXIVE PATTERN B: DATIVE CAUSEE
asi par-Inoov koori hisaab
we.ERG study-CAUSE.PST.MSG girl.F.DAT math.M.NOM
'We taught the girl math'. (Hook and Koul 1984:105)

Important for our purposes is that the differential marking of the causee in Kashmiri seems to produce a meaning difference that is not truth conditional. In Syeed's (1985) detailed study of the Kashmiri causative, he claims that a causee in the nominative case serves to "foreground the causee and background the action" while an oblique or dative causee "signals the reverse" (p. 41). He further suggests that the nominative causee "betrays some degree of direct manipulation…there may be some degree of coercion involved" (p. 43). Butt (2006) asserts that for the ingesto-reflexive class of verbs in the related language Urdu, the meaning difference centers on the degree of affectedness of the causee, comparing it to the dative alternation for English ditransitives (see also Saksena 1980; Alsina and Joshi 1991; and for Italian Folli and Harley 2007).[3] Native speakers confirm that the two versions of the causative with differentially marked causees are synonymous to the degree relevant to the present analysis: a contradiction results when a clause formed with one version of the causative is conjoined with the negation of the other.

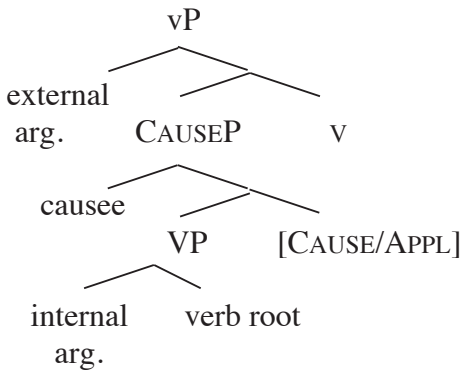
(9) #me ran-InaavInəəv təmis athi oolav magar me
I.ERG cook-CAUS.PST.MPL he.DAT INSTR potatoes.NOM but I.ERG

[1] Complements of postpositions in Kashmiri must typically be marked with either dative or ablative case (determined by the postposition).
[2] Note that this causative structure is not a passive. Kashmiri does have a passive construction in which the ablative form of the infinitive verb is used in combination with the auxiliary verb *yun* 'come'. In the passive the demoted agent is marked with dative case and then followed by the postposition *zəriyi or dəs* 'by' (Wali and Koul 1998:208).
[3] But as Khokhlova (2003:17) claims for modern western Indic languages "non-affectedness of the causee does not automatically imply its instrumental/ablative marking".

ran-InaavInoov-nI su oolav
cook-CAUSE.PST.MSG-NEG he.NOM potatoes.NOM.PL
intended: 'I had him cook potatoes but I did not have him cook potatoes.'

Manetta (2014) proposes an account of Kashmiri causatives following Pylkkänen (2008) in which the causative morpheme is introduced by a functional head distinct from the verbalizing and phase-defining functional head v. A causee may be introduced in the specifier of this head (as schematized in (10)) or adjoined as part of a PP headed by the post-position *athI*.

(10)



For ingesto-reflexive verbs, a version of the CAUSE head is available that assigns inherent dative case to the argument it introduces in its specifier (Folli and Harley 2007; Harley 2008).[4] Following McFadden and Sundaresan (2011), nominative (unmarked) case in Kashmiri is the default case, assigned to any remaining argument that has not received a marked structural or inherent Case from a Case assigning head (e.g. ergative, dative).[5] Thus the causative clauses with nominative and dative causees above will have the same basic architecture, while clauses with the instrumental-marked causee will feature a PP adjunct to CAUSEP (and CAUSEP will lack a specifier). For detailed argumentation supporting this syntax, see Manetta 2014.

3 Sluicing and sprouting in Kashmiri causatives

As CLM (1995, 2011) point out, argument structure mismatches between elided material and its antecedent in ditransitives in English result in ungrammatical sluices:

(11) Ali sent someone an email, but it's unclear who/*to whom.

Chung (2006) and Merchant (2013) further demonstrate that for a range of languages and constructions a predicate in the ellipsis site must have the same argument structure as the corresponding predicate in the antecedent.

[4] As a reviewer points out, it must be selection by the specific predicate that determines which CAUSE head may appear (that which introduces an unmarked or dative specifier or that which introduces no specifier).

[5] The account of causatives developed in Harley 2013, itself building on Pylkkänen 2008 (contra Harley 1995), differs minimally from what is presented here and might also handle the basic facts in Kashmiri. In this view the causative morpheme is introduced in a v head distinct from the Voice head which introduces the external argument. This causative v takes a second VoiceP as a complement, which introduces the causee argument in its specifier (and the causee's case is assigned structurally by the causative v). The argumentation concerning sluicing in the text below stands under either structure.

In addition, Chung 2013 proposes a second condition on syntactic identity involving Case assigning heads. I present here a very brief version of her argumentation centered on data in the Austronesian language Chamorro. I refer the reader to Chung 2013 for a nuanced presentation of this data, acknowledging that many important details will go unmentioned here. The operation termed 'sprouting' in CLM (1995, 2011) describes the process of integration into a sluicing structure of an interrogative DP remnant which does not correspond to an overt DP in the antecedent, as in the English example in (12).

(12) Rohit certainly lied, but we don't know to whom.

Chung (2013) establishes the necessity of both the argument structure and Case conditions on syntactic identity in sluicing using sprouted DPs. Consider the contrast between the grammatical non-elided Chamorro sentence in (13) with the ungrammatical sluice in (14):

(13) Omlat gui', lao ti hu tungu' [håfa ha omlåti]
AGR.fit she but not AGR know what WH.[OBJ].AGR fit.into
'She fits (something) but I don't know what she fits.'
(Chung 2013: (38c))

(14) *Omlat gui', lao ti hu tungu' [håfa _____]
AGR.fit she but not AGR know what
('She fits (something) but I don't know what.')
(Chung 2013: (39c))

In (13), the first instance of the predicate *omlat* 'fit' is intransitive, realizing the thing fit into as the object of a PP headed by a null preposition (see Chung 2013 and references therein for argumentation). The second clause of (13) features the derived transitive form *omlåti* 'fit into', which realizes the thing fit into as a DP internal argument. Crucially, Chung shows that the two predicates are synonymous, so if semantic identity were the only condition on sluicing, the ellipsis in (14) should be licit. The ungrammaticality of (14) must thus be attributed to syntactic conditions on sluicing.

Chung reveals that the clause that has gone missing in (14) can be analyzed in two ways. In the first, schematized in (15a) below, the sprouted DP is the internal argument of the transitive version of the predicate. In the second, in (15b), the sprouted DP is the DP object of the PP that accompanies the intransitive version of the predicate.

(15) a. *[håfa ~~ha omlåti~~]. TRANSITIVE PREDICATE
what WH.[OBJ].AGR fit.into
b. *[håfa ~~omlåt ña [$_{P}$]~~] INTRANSITIVE PREDICATE
what WH.[OBL].fit-AGR

If the elided material includes a transitive predicate as in (15a), then (14) is ruled out due to argument structure mismatch: the antecedent contains the intransitive version of the predicate, and the two predicates have distinct argument structures (even if the PP complement of the intransitive version in the antecedent is implicit and thus unrealized). If the elided material in (14) includes the intransitive version of the predicate as in (15b), there is no problem with argument structure mismatch, but there is no Case assigner in the antecedent clause that

corresponds to the (silent) P head that serves as the Case assigner for the sprouted interrogative DP remnant. This is because the internal argument in the antecedent is implicit and there is thus no PP at all in the antecedent. Under this parse, (14) is ungrammatical due to a violation of the Case condition as the DP remnant is not case licensed by a head identical to a corresponding head in the antecedent clause. Thus the ungrammatical sluice in (14) requires both of the syntactic conditions proposed in (2) to be successfully ruled out.

I now turn to show that sprouting in Kashmiri causative constructions can provide similar support for distinct case and argument structure identity conditions on sluicing. Kashmiri causatives permit both types of non-nominative causees (dative type and instrumental-PP type) to go unexpressed, though nominative causees (detectable by the agreement on the verb) must be overt.

(16) Aslam-an leekh-Inəəv ciTh' INSTRUMENTAL CAUSEE MISSING
Aslam-ERG write-CAUS.FSG letter.F
'Aslam had (someone) write a letter.' (Wali and Koul: 218)

(17) me hech-Inəəv dangal DATIVE CAUSEE MISSING
we.ERG learn-CAUSE.FSG wrestling.F.NOM
'I taught (someone) wrestling.' (Wali and Koul: 218)

(18) *Aslam-an leekh-Inoovus ciTh' NOMINATIVE CAUSEE MISSING
Aslam-ERG write-CAUS.MSG letter.F
Intended: 'Aslam had (someone) write a letter.' (Wali and Koul: 218)

(19) * me hech-Inoov dangal NOMINATIVE CAUSEE MISSING
we.ERG learn-CAUSE.MSG wrestling.F.NOM
'I taught (someone) wrestling.' (Wali and Koul: 216)

Unsurprisingly, non-nominative causees may be sprouted in Kashmiri sluicing:

(20) asi ran-InaavInəəv oolav magar me chu-nI pataa kəmis athi
we.ERG cook-CAUS.PST.MPL potatoes.NOM but I.ERG AUX-NEG know who.DAT INSTR
'We had (someone) cook potatoes but I don't know who.'

(21) asi parInoov hisaab magar me chu-nI pataa kəmis
we.ERG learn-CAUSE.MSG math but I.ERG AUX-NEG know who.DAT
'We taught (someone) math, but I don't know who.'

But compare (20)-(21) to (22)-(23), in which the sprouted interrogative remnant is unmarked (nominative):

(22) *asi ran-InaavInəəv oolav magar me chu-nI pataa kus
we.ERG cook-CAUS.PST.MPL potatoes.NOM but I.ERG AUX-NEG know who.NOM
'We had (someone) cook potatoes but I don't know who.' [6]

[6] Crucially, the ungrammaticality of (22)-(23) cannot be explained away by the constraint observed in Ross 1969 -- that the morphological case of the remnant interrogative DP in a sluiced structure must match the case it would be assigned in the non-elliptical structure -- since a full clause here with fronted nominative causee is grammatical.

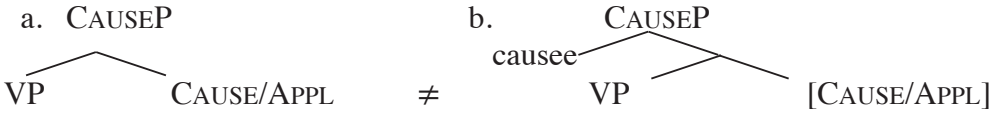
(i) asi ran-InaavInəəv kəNsi athi oolav magar me chu-nI
we.ERG cook-CAUS.PST.MPL somone.DAT INSTR potatoes.NOM but I.ERG AUX-NEG
pataa kus ran-InaavInoov asi oolav

(23) *asi parInoov hisaab magar me chu-nI pataa kus
we.ERG learn-CAUSE.MSG math.M.NOM but I.ERG AUX-NEG know who.NOM
'We taught (someone) math, but I don't know who.'

As discussed above in section 2, clauses with nominative causees and clauses with case-marked causees are synonymous to the required degree. Therefore the sluices in (22) and (23) meet the semantic identity condition: the focus-closures of the elided clause and its antecedent entail each other (Merchant 2001). We must then determine whether the ungrammaticality of these sluices can be attributed to a violation of syntactic identity conditions.

Let us first consider (22), in which the instrumental-type inherent causee in the antecedent would have been introduced in a PP adjunct to a CAUSEP which projects no specifier (following the analysis for this structure in Manetta (2014) reviewed above). The PP adjunct is not present in the antecedent, so we have CAUSEP simply containing the CAUSE head and its complement. By contrast, the sprouted nominative causee in the sluiced clause in (22) would need to have been introduced in the specifier of CAUSEP. The sprouted structure in (22) thus runs afoul of the identity condition on argument structure.

(24)



Now let us turn to (23), in which the causee in the antecedent would have been introduced in the specifier of a CAUSE head capable of assigning inherent dative case to the argument introduced in its specifier. The head has dative case available to assign, even though there is no overt argument to assign it to (in this claim I follow Chung (2013: 40), who makes a similar proposal concerning Chamorro v heads, building on Rothstein (1992)). The sprouted interrogative nominative causee remnant in the sluiced clause would need to be introduced in the specifier of CAUSEP that does NOT assign dative case. Instead, according to the analysis in Manetta (2014), the causee ends up with default nominative case as it receives no structural or inherent case from a head in the derivation (McFadden and Sundaresan 2011). Crucially, (23) is not an argument structure condition violation, since both causees would be introduced into specifiers of CAUSEP and the argument structure is identical. We must instead attribute the ungrammaticality of (23) to Case mismatch. That is, the DP remnant in the elided clause is not case licensed by a head identical to a corresponding head in the antecedent clause. Indeed, the DP remnant is not directly case licensed by a head at all, since it receives no dative case from CAUSE and thus is assigned nominative by default.[7]

(25) a. CAUSE ≠ b. CAUSE
[DAT] [ ]

know who.NOM cook-CAUS.PST.MSG we.ERG potatoes.NOM
'We had someone cook potatoes but I don't know who we had cook potatoes.'

[7] Thanks to Ruth Kramer for her observations on this point.

Identity of Case assigning heads should only be relevant for DP sluicing remnants, and argument structure identity should only apply to arguments. Thus sprouted PP adjunct remnants should be exempt from these conditions, as in English (26).

(26) Mehir met her, but he doesn't remember when.

In Kashmiri causatives, this means that if the predicate permits causees of the instrumental PP type, these should always be felicitous sluicing remnants independent of the case of the causee in the antecedent, since they are PP adjuncts. This prediction is indeed borne out.

(27) Aslam-an leekh-Inoov kaaNh ciTh' magar me chu-nI pataa kəmis athI
Aslam-ERG write-CAUS.MSG someone letter.F but I.ERG AUX-NEG know who.DAT INST
'Aslam had someone write a letter, but I don't know who (by whom).'

(28) Təm' d'aav-Inəəv mohn-as kaaNh kitaab magar me chu-nI pataa
he-ERG give-CAUS.MSG Mohan-DAT someone book.F but I.ERG AUX-NEG know
kəmis athI
who.DAT INST
'He had someone give Mohan a book, but I don't know who (by whom).'

4 Conclusion

In conclusion, this brief exploration of data from Kashmiri causative constructions supports the specific formulation of "limited syntactic identity" in sluicing developed in Chung 2013, repeated here:

(29) Limited syntactic identity in sluicing
a. Argument structure condition: If the interrogative phrase is the argument of a predicate in the ellipsis site, that predicate must have an argument structure identical to that of the corresponding predicate in the antecedent clause.
b. Case condition: If the interrogative phrase is a DP, it must be Case-licensed in the ellipsis site by a head identical to the corresponding head in the antecedent clause. (Chung 2013:30)

To be precise, there is evidence in Kashmiri that both the argument structure condition and the Case condition are needed to rule out ungrammatical sluicing with interrogative remnants that are sprouted nominative causees. The broader argument for these two conditions is made all the stronger by the fact that Kashmiri and Chamorro are from unrelated language families, and that the underlying structures under investigation are of distinct types. Further research should reveal whether the version of limited syntactic identity fleshed out in (29), alongside the condition on semantic identity, is indeed sufficient. As Kashmiri is both morphologically rich and features a number of processes by which argument structure can be manipulated, it remains a fertile testing ground for theories of identity in ellipsis more generally.

References:

Altaha, F. 1992. Kashmiri Causative Constructions and the Antipassive Analysis. Indian Linguistics 55: 1-22.

Alsina, Alex and Smita Joshi. 1991. Parameters in causative constructions. In Lise M. Dobrin, Lynn Nichols, and Rosa M. Rodriguez, editors, *Proceedings of CLS*, volume 27, pages 1-15.
Anand, P. and A. Nevins. 2006. The Locus of Ergative Case Assignment: Evidence from Scope. In Ergativity: Emerging Issues. ed. A. Johns, D. Massam and J. Ndayiragije, Kluwer Academic Publishers. pp. 3—27.
AnderBois, Scott. 2011a. Issues and alternatives. Doctoral dissertation, University of California, Santa Cruz.
Barker, Chris. 2013. Scopability and sluicing. *Linguistics and Philosophy* 36:187–223.
Barros, Matthew. 2014. Sluicing and identity in Ellipsis. Doctoral dissertation, Rutgers University, New Jersey.
Bhatt, Rakesh. 1999. Verb Movement and the Syntax of Kashmiri. Dordrecht: Kluwer.
Bhatt, Rajesh, and David Embick. 2003. Causative derivations in Hindi. Unpublished manuscript, University of Texas at Austin and University of Pennsylvania.
Butt, Miriam. 2006. Theories of Case. Cambridge: Cambridge University Press.
Chung, Sandra. 2006. Sluicing and the lexicon: The point of no return. In Proceedings of the annual meeting of the Berkeley Linguistics Society 31, ed. by Rebecca T. Cover and Yuni Kim, 73–91. Berkeley: University of California, Berkeley Linguistics Society.
Chung, Sandra. 2013. Syntactic Identity in Sluicing: How Much and Why? Linguistic Inquiry 44.1:1-44.
Chung, Sandra, William A. Ladusaw, and James McCloskey. 1995. Sluicing and Logical Form. Natural Language Semantics 3:239–282.
Chung, Sandra, William A. Ladusaw, and James McCloskey. 2011. Sluicing(:) between structure and inference. In Representing language: Essays in honor of Judith Aissen, ed. by Rodrigo Gútierrez-Bravo, Line Mikkelsen, and Eric Potsdam, 31–50. California Digital Library eScholarship Repository. Santa Cruz: University of California, Santa Cruz, Linguistic Research Center.
Folli, R. and H. Harley. 2007. Causation, obligation and argument structure: on the nature of little v. Linguistic Inquiry 38.2: 97-238.
Hardt, Daniel. 1993. Verb phrase ellipsis: Form, meaning, and processing. Doctoral dissertation, University of Pennsylvania, Philadelphia.
Harley, H. 2008. On the causative construction. In Miyagawa, S. and S. Mamoru (eds) The Oxford Handbook of Japanese Linguistics. OUP: Oxford.
Harley, Heidi. 1995. Subjects, events and licensing. Doctoral Dissertation. MIT.
Harley, Heidi. 2013. External arguments and the Mirror Principle: On the distinctness of Voice and v. Lingua 125:34–57.
Hook, P. and O.N. Koul. 1984. On the Grammar of Derived Transitives and Causatives in Kashmiri. In Hook, P. and O.N. Koul (eds). Aspects of Kashmiri Linguistics. Bahri: New Delhi.
Khokhlova, L.V. 2003. Infringement of Morphological and Syntactic Operations' Pairing in "Second Causative" Formation. Indian Linguistics 64.1:1-18.
Manetta, Emily. 2014. The Shape of the Causative Verbal Domain: Evidence from Kashmiri. Syntax 17.3: 235–268.
McFadden, Thomas and Sandhya Sundaresan. 2011. Against the dependence of nominative case on finiteness. Jawaharlal Nehru University, New Delhi, India. January, 2011.
Merchant, Jason. 2001. The syntax of silence: Sluicing, islands, and the theory of ellipsis.

Oxford: Oxford University Press.
Merchant, Jason. 2013. Voice and ellipsis. Linguistic Inquiry 44:77–108.
Pylkkänen, Liina. 2008. Introducing Arguments. Cambridge, MA: MIT Press.
Rooth, Mats. 1992. Ellipsis redundancy and reduction redundancy. In Proceedings of the Stuttgart Ellipsis Workshop, ed. by Steve Berman and Arild Hestvik. Stuttgart: Sonderforschungsbereich 340, University of Stuttgart. Available at http://www.ims.uni-stuttgart.de/mats.
Ross, John Robert. 1969. Guess who? In Papers from the Fifth Regional Meeting of the Chicago Linguistic Society, ed. by Robert I. Binnick, Alice Davison, Georgia M. Green, and Jerry L. Morgan, 252–286. Chicago: University of Chicago, Chicago Linguistic Society.
Rothstein, Susan D. 1992. Case and NP licensing. Natural Language and Linguistic Theory 10:119–139.
Saksena, A. 1980. Causative Relations in Hindi. General Linguistics 20.1: 23–38.
Syeed, S. M. 1985. Morphological Causatives and the Problems of the Transformational Approach. Bloomington. Indiana University Linguistic Club.
Wali, K. 1980. Oblique Causee and Passive Explanation. Linguistic Inquiry 11.1: 258-260.
Wali, K. 1981. Cause, Causer, and Causee: A Semantic Perspective. Journal of Linguistics 17: 289-308.
Wali, Kashi. and O.M. Koul. 1997. Kashmiri. New York: Routledge.

REFERENTIALITY AND RESUMPTION IN *WH*-DEPENDENCIES [*]

Grant Goodall
University of California, San Diego
ggoodall@ucsd.edu

In addition to the many well-known restrictions on *wh*-dependencies, such as islands, there are also two well-known ways to circumvent these restrictions, at least partially: making the *wh*-phrase referential (D-linked) and using a resumptive pronoun instead of a gap. The most common analyses of these phenomena treat the referentiality effect as essentially grammatical, in that referentiality allows for a structure in which grammatical constraints are not violated, and the resumption effect as essentially a processing strategy, in that the pronoun allows the dependency to be processed despite the ungrammatical structure. Here I argue that the true situation is exactly the reverse. Referentiality allows for the dependency to be processed more easily but does not affect grammaticality *per se*, while resumptive pronouns are a grammatical option in English on a par with gaps. This perspective on *wh*-dependencies has a number of empirical advantages. For referentiality, it predicts that referential *wh*-phrases will result in amelioration not just in islands, but in non-islands as well, and acceptability experiments in English and Spanish confirm this. For resumption, it predicts that speakers will spontaneously produce resumptive pronouns in both islands and non-islands, even without production pressure, and that across environments, resumptives will be of relatively low acceptability, due to the processing difficulties that they impose. This latter point is confirmed by an acceptability experiment in English. Overall, the analysis here suggests that languages make use of the means available to them to resolve *wh*-dependencies, without additional stipulations for referentiality or resumption.

1. Introduction

One of the most common ways to form *wh*-questions in natural language is by having a *wh*-phrase in a position indicating the scope of the question and a gap in a position indicating the variable, as in the English examples in (1).

(1) a. What do you think that Mary saw __ ?
    b. I need to ask who they talked to __ yesterday.

[*] I thank Bethany Keffala, Alex Stiller, and the entire Experimental Syntax Lab at UC San Diego, as well as audiences at the CUNY Conference on Human Sentence Processing, the Linguistic Evidence - Berlin conference, the LSA Annual Meeting, Johns Hopkins University, the University of Chicago, the University of California, Santa Cruz, and the University of Potsdam for valuable feedback and discussion. Deepest thanks also to Sandy Chung for being such an inspiration, both when I was a student and in the decades since.

It is well known that there are many restrictions on the dependency formed between the *wh*-phrase and the gap. It is possible to have more than one *wh*-phrase in a clause, for instance, but only the syntactically more prominent one may enter into a dependency of the type in (1). This restriction, known as Superiority (Chomsky (1973)), is illustrated in (2), where a direct object *wh*-phrase illicitly enters into such a dependency despite the presence of a subject *wh*-phrase.

(2) *I wonder what who bought __ .

Another example of a restriction on *wh*-dependencies comes from languages that impose morphological requirements on clauses containing these dependencies. In Chamorro, for instance, *wh*-dependencies in questions generally require special morphology on the verb and are not possible otherwise (see Chung (1982), Chung (1994) and Chung (1998)). This is seen in (3) with a short-distance dependency (from Chung (1998)) and in (4) with a long-distance dependency (from Chung 1994)).

(3) a. Ha-fa'gasi si Juan i kareta.
AGR-wash Juan the car
'Juan washed the car.'
b. Hayi uma' gasi i kareta?
WHO WH.wash the car
'Who washed the car?'

(4) a. Humallum si Maria [na ha-panak si Juan i patgun].
AGR-assume Maria COMP AGR-spank Juan the child
'Maria assumes that Juan spanked the child.'
b. Hayi hinalomnia si Maria [pumanak __ i patgun]?
WHO WH.assume Maria WH.spank the child
''Who does Maria assume spanked the child?'

A third well-known restriction on *wh*-dependencies concerns the position of the gap. Certain domains within the clause, known as *islands*, prohibit gaps, as seen in in (5).

(5) a. *What do you believe the claim that Mary saw __ ? [Complex Noun Phrase Constraint]
b. *What do you wonder who saw __ ? [*wh*-island]

There are also two well-known ways to circumvent these restrictions, at least partially. One involves making the *wh*-phrase *referential* (or *lexically restricted* or *D-linked*) and the other involves using a *resumptive pronoun* in place of the gap. These strategies are illustrated in (6) and (7), respectively, for the case of islands.

(6) a. Which movie do you believe the claim that Mary saw __ ? [CNPC]
b. What city do you wonder who saw __ ? [*wh*-island]

(7) a. ?What do you believe the claim that Mary saw it? [CNPC]
b. ?What do you wonder who saw it? [*wh*-island]

I will restrict attention in this paper to the case of islands, although related effects may be seen with the Superiority and agreement restrictions seen above.

For both referentiality and resumption, there is a commonplace view that forms the point of departure for most discussions. For referentiality, this view is that the phenomenon is at heart a grammatical one. Referential *wh*-phrases are always allowed, but they permit a kind of dependency that bypasses the restrictions that would otherwise lead to island effects (e.g., Cinque (1990), Rizzi (2001, 2004), Szabolcsi and Zwarts (1993, 1997)). For resumption, on the other hand, the commonplace view is that this is an extra-grammatical phenomenon. Resumptive pronouns are not normally allowed in English, but they are recruited by the processor in island structures as a way to express *wh*-dependencies that could not be expressed with a gap (e.g., Chao and Sells (1983), Dickey (1996). Polinsky et al. (2013), Sells (1984)).

I will argue here that these commonplace views of referentiality as a grammatical phenomenon and resumption as a processing phenomenon are exactly wrong. I will present evidence that referentiality has the effect that it does because it lessens the strain on the working memory system as the *wh*-dependency is processed, and that the behavior of resumptive pronouns in English is consistent with their being a grammatically licit option in the language. This complete reversal of the traditional views of these phenomena is made possible by a more careful approach to the relation between grammaticality and acceptability and leads us to a view in which referentiality and resumption follow without stipulation from other properties of the language.

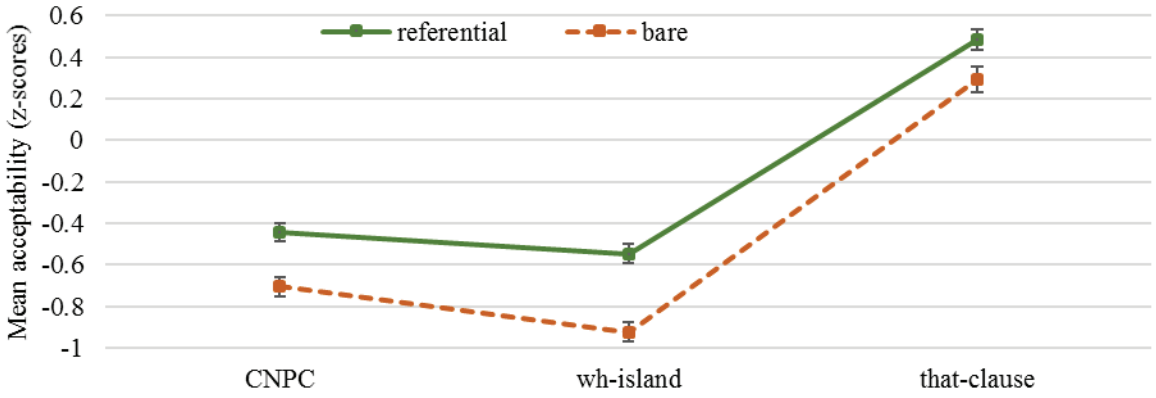
2. Referentiality

In most of the literature on referentiality in *wh*-dependencies, it is assumed that referential *wh*-phrases have some property that allows them to circumvent the grammatical principles leading to the island effect. In Szabolcsi and Zwarts (1997), for instance, referentiality allows the *wh*-phrase to be interpreted as a set of discrete individuals, which in turn allows for a coherent semantic interpretation of the sentence, which would otherwise not be available given the extraction out of the island domain. In Rizzi (2004), the presence of lexical material that is related to previously mentioned elements in the discourse allows for a type of binding that should be immune to many island effects. It is also possible, however, that referentiality is unrelated to the grammar of islands per se and that the amelioration seen reflects instead a lessening of the burden on working memory as the *wh*-dependency is processed. Such a view is plausible, at least, in that a *wh*-phrase that is more referential (and thus more robustly encoded in memory and more distinct from potentially interfering phrases) could reasonably be expected to be more easily retrieved from memory and associated with the gap site, and this view has been argued for extensively (e.g. Kluender and Kutas 1993, Kluender 1998, Hofmeister and Sag 2010, Hofmeister 2011). I what follows, I present and discuss two experimental findings that offer additional evidence for this view.

2.1. English

The working memory account of the effect of referentiality on *wh*-dependencies, unlike the grammatical accounts, clearly predicts that the amelioration will be found generally, not just in

Fig. 1: Bare and referential *wh*-phrases in three contexts



island contexts. Goodall (2015) tests this prediction by measuring the acceptability of *wh*-dependencies where the gap is in an island context (a CNPC or *wh*-island, as in (7) above) or a non-island context (a *that*-clause), with both referential and bare *wh*-phrases. As seen in Fig. 1, the results confirm the prediction of the working memory account: Referential *wh*-phrases result in significant amelioration in both island and non-island contexts. The lack of an interaction between the referentiality of the *wh*-phrase and the structural context suggests that whatever is responsible for the amelioration in the non-island *that*-clause (presumably the working memory effect) is sufficient to account for the amelioration in the two island environments, without additional grammatical mechanisms.

2.2. Spanish

Goodall (2015) offers clear evidence that referentiality leads to amelioration even outside of island environments, but the results should nonetheless be taken with caution. Some other studies have not found this same effect (e.g., Alexopoulou and Keller (2013), Sprouse, Caponigro, Greco and Cecchetto (2016)), but even if it does exist, it could be due to a simple preference for referential *wh*-phrases in general, independently of the dependency between the *wh*-phrase and the gap. Alex Stiller and I have addressed these concerns in an experiment that allows us to replicate the finding from English of a referentiality effect in a non-island environment and to show that the effect is due to the dependency itself, not the simple presence of a referential *wh*-phrase.

2.2.1 Method and Materials

45 participants, all native speakers of Spanish residing in Spanish-speaking countries (excluding the Caribbean region, which is known to differ with regard to the word order factor explored

here) rated sentences on-line on a 7-point scale. Experimental stimuli were all object *wh*-questions, created using a 2 x 2 design, crossing *wh*-phrase (referential vs. bare) and word order (subject – verb (SV) vs. verb – subject (VS)) as in (8).

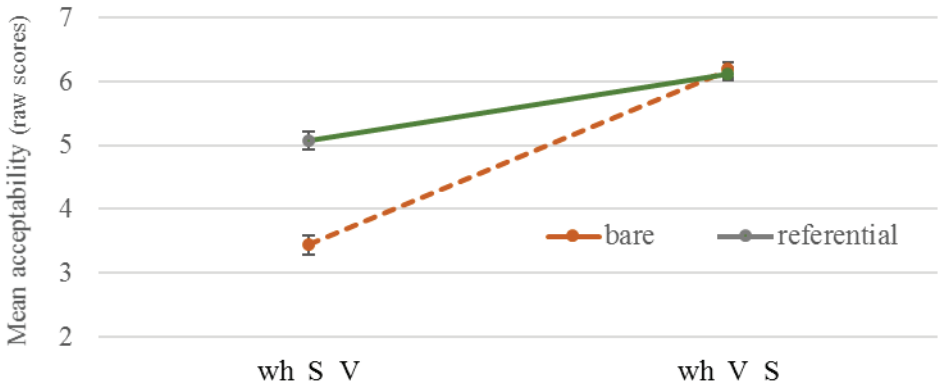
(8) a. ¿Qué la profesora vio en el cine? [bare, SV]
WHAT the teacher saw at the theater
b. ¿Qué vio la profesora en el cine? [bare, VS]
WHAT saw the teacher at the theater
'What did the teacher see at the theater?'
c. ¿Qué película la profesora vio en el cine? [referential, SV]
WHAT movie the teacher saw at the theater
d. ¿Qué película vio la profesora en el cine? [referential, VS]
WHAT movie saw the teacher at the theater
'What movie did the teacher see at the theater?

Participants saw three tokens of each condition, in addition to 36 filler items. Experimental items were counterbalanced using a Latin square and pseudo-randomized.

2.2.2 Results and Discussion

Results are presented in Fig. 2. There are main effects for *wh*-phrase ($p<0.001$) and word order ($p<0.001$), and a significant interaction between the two ($p<0.001$). Crucially, there is a clear effect of referentiality with the SV order ($p<0.001$), but no such effect with the VS order (p-values from an ANOVA analysis presented in Stiller (2014)). Given these results, there are two

Fig. 2: *wh*-phrases in Spanish



conclusions of interest. First, we have another case of amelioration due to referentiality outside of an island environment. That is, object *wh*-questions with SV order are very low in acceptability in Spanish, but they are not a standard island violation. There is no extraction out of an embedded clause here (in fact, the SV order is acceptable in such cases), and unlike in standard islands, adjunct extraction is much better than argument extraction (Goodall (1993, 2010), Torrego (1984)). The referentiality effect that we see here in Spanish thus provides

additional confirmation of the prediction made by the working memory account that amelioration should be found with all *wh*-dependencies, not just those where the gap is in an island.

Second, the amelioration is observed only when there is a non-trivial *wh*-dependency, thus suggesting that the effect is related to the dependency itself, not to the simple presence of a referential *wh*-phrase. That is, amelioration occurs in the SV order, as in (8a) and (8c), but not in the VS order, as in (8b) and (8d). In the SV order, material intervenes between the *wh*-phrase and the gap (posited when the verb is encountered, under standard assumptions), so recovering the *wh*-phrase at the gap site poses a challenge to working memory. In the VS order, on the other hand, the dependency is resolved essentially immediately, since the verb (and thus the gap) occurs right after the *wh*-phrase, so there is little strain on working memory. Crucially, referentiality results in amelioration only in the case where there is a significant burden on working memory.

2.3 Conclusion

We have now seen two cases, one from extraction out of an embedded *that*-clause in English and one from object extraction out of a main clause with SV order in Spanish, where making the extracted *wh*-phrase referential led to a significant increase in acceptability. The first of these cases is clearly not an island and the second is different from standard islands in some crucial respects, so the amelioration seen is unexpected under grammatical accounts of the referentiality effect. This amelioration is exactly what is expected, on the other hand, under an account in which referentiality makes the *wh*-phrase easier to retrieve from memory, thus making the *wh*-dependency easier to resolve and leading to higher acceptability. In addition, the experimental results from Spanish suggested that the amelioration only occurs when there is in fact a dependency that can be made easier. The simple presence of a referential *wh*-phrase is not sufficient.

3. Resumption

We now turn to the other end of the *wh*-dependency, the site of the variable. In all of the examples considered so far, this position has been a gap, but it has long been noted that in certain circumstances and in certain languages, this position may be filled with a pronoun, known as a *resumptive pronoun*. For languages like English, the standard assumption has been that this option is not allowed by the grammar *per se*, but that it is nonetheless chosen by the processor as a last resort when a gap is not possible (see McCloskey (2006) and Asudeh (2011) for overviews of the literature). This view is supported by the fact that resumptive pronouns seem ill-formed in cases where gaps are clearly possible, as in (9), but appear to improve in island environments, where gaps are not possible, as in (10).

(9) a. *Who did you see him/her/them?
    b. *What do you think that Mary saw it?

(10)a. Which girl do you believe the claim that Mary saw her? [CNPC]
    b. What city do you wonder who saw it? [*wh*-island]

Despite this initial plausibility, the view that resumptive pronouns are ungrammatical but recruited by the processor when needed runs into a number of difficulties. First, resumptives appear to be reasonably common in speech, even in cases where a gap would be possible (see Cann et al. (2004), Prince (1990, 1997)), as in (11).

(11) …those little potato things that you put 'em in the oven… (Prince, 1990)

Ungrammatical sentences are typically not frequent in speech, so resumptives like this are hard to explain if they are not allowed by the grammar. Second, and in a similar vein, Ferreira and Swets (2005) found that speakers readily produce resumptives in experimental settings, even when there are no production (i.e., time) pressures. In fact, speakers produce resumptives less often when they have less time to plan their utterances. Again, these results are unexpected if the grammar prohibits resumptives and speakers occasionally produce them only as a last resort. Third, this view of resumptives as driven by processing predicts that in formal acceptability experiments, they should become more acceptable in those cases where they are forced (i.e., where a gap is not possible). This is not what has been found, however. In general, resumptive pronouns seem to be no more acceptable than gaps, whether in island or non-island contexts (Alexopoulou & Keller 2007, Heestand et al. 2011).

3.1. Resumption as a grammatical option in English

Given these difficulties, let us entertain the opposite view, that resumptives are in fact grammatical in English. Given what we know about the behavior of pronouns generally in the language, this view is not implausible. Pronouns may act as bound variables in relation to quantifiers (*wh-* and otherwise), as in (12), or of course they may have referential DPs as antecedents.

(12) Who$_i$ thinks that they$_i$ are intelligent?

With resumptive pronouns in relative clauses, as in (13), both mechanisms are available, which could perhaps explain why resumptives seem particularly natural in this environment (Alexopoulou & Keller 2007, Han et al. 2012).

(13) A man who the girl believes that Mary saw him has just arrived.

Even if resumptive pronouns are allowed by the grammar of English, though, they are still likely to trigger considerable processing costs, despite the traditional idea that they exist in order to facilitate processing. Unlike gaps, pronouns do not unambiguously resolve the *wh*-dependency, so when the processor encounters resumptive pronouns, one would expect it to proceed normally (i.e., attempt to locate an entity in the recent discourse to which the pronoun could refer), while leaving the *wh*-dependency open. Once it is clear that we are outside the scope of the *wh*-phrase and that a gap will no longer be possible (e.g., when *has* is encountered in (13)), the processor must then reanalyze the pronoun so that it is interpreted as resumptive.

Significant processing costs like these are known to lead to large declines in acceptability (see, e.g., Cowart (1997)), so even if the grammar of English permits resumptive pronouns, we still expect them to be of relatively low acceptability. Importantly, the cost associated with reanalysis of the pronoun once it is taken to be resumptive is essentially invariable regardless of the position of the pronoun. We thus do not expect the acceptability of resumptive pronouns to show very much variation depending on whether they are found in island or non-island environments.

This view of resumptive pronouns as a grammatical option in English allows us to address some of the difficulties that we saw earlier with the more traditional view of resumptives as a purely processing phenomenon. The fact that resumptives are relatively frequent in speech, for instance, is no longer surprising. They impose a significant processing cost, as we have seen, but so do many other structures (e.g., embedding or *wh*-movement) and speakers nonetheless use these structures when needed. In addition, we can now understand why speakers produce resumptives even in the absence of any production pressures, since if they are a grammatical option, no particular pressure on production is needed to motivate their use. There is also no reason to expect resumptives to be consistently more acceptable than gaps within islands. As we have seen, resumptives impose their own significant processing costs, and there is unlikely to be a uniform relation between these and whatever grammatical and/or processing costs are associated with gaps.

This last point leads us to an important distinction between gaps and pronouns. Gaps are well known to vary widely in acceptability depending on their exact position (i.e., whether they are inside an island or not), but there is no reason to expect the same of resumptive pronouns. As noted, resumptives impose significant processing costs, but these costs, as well as acceptability, should be relatively uniform across environments. Bethany Keffala and I tested this prediction in an experiment examining relative clauses containing gaps vs. resumptive pronouns in two non-island environments (simple clause and embedded *that*-clause) and two island environments (*wh*-clause and relative clause) (Keffala 2011, Keffala and Goodall 2011, 2013). Sample stimuli are given in (14) with subject gaps/pronouns and in (15) with object gaps/pronouns.

(14)a. This is the chef that __ / he prepared the potatoes.
b. This is the chef that we realized that __ / he prepared the potatoes.
c. This is the chef that we inquired how __ / he prepared the potatoes.
d. This is the chef that we devoured the potatoes that __ / he prepared.

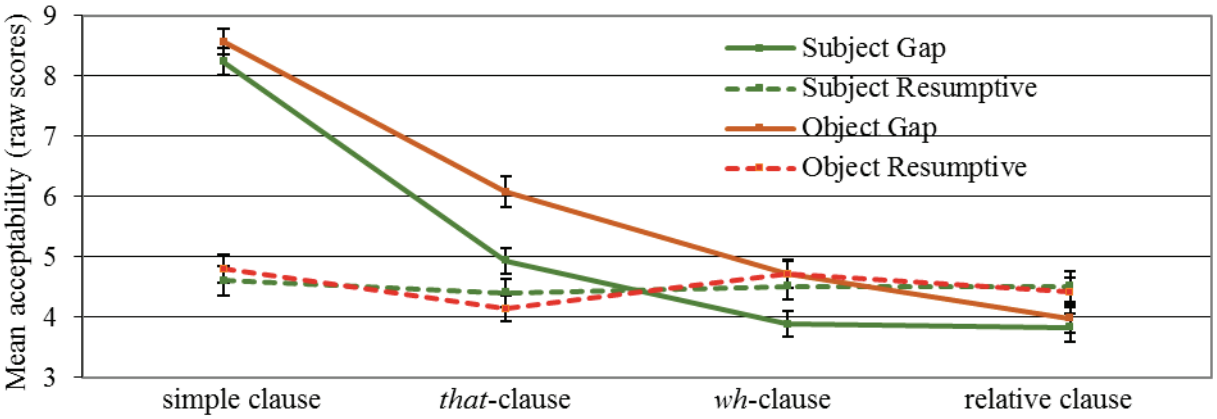
(15)a. These are the potatoes that we prepared __ / them.
b. These are the potatoes that we realized that the chef prepared __ / them.
c. These are the potatoes that we inquired how the chef prepared __ / them.
d. These are the potatoes that we devoured the potatoes that the chef prepared __ / them.

121 participants saw two tokens of each condition, resulting in 64 experimental stimuli (together with 64 filler items), and rated them on an 11-point scale. The results are shown in Figure 3.

As seen in Figure 3, resumptive pronouns, in sharp contrast to gaps, remain at approximately the same level of acceptability regardless of the context (non-island vs. island) or of the specific position (subject vs. object). This is what is predicted by the hypothesis proposed

here, in which resumptive pronouns are allowed by the grammar but incur significant processing costs, but it is not what is predicted by the more common view, in which resumptives are grammatically illicit but sometimes forced for processing reasons. Under such a view, we expect resumptives to show an increase in acceptability when gaps are disallowed, but there is no sign of such an effect here. Note especially that there is no consistent relation between the acceptability of the gaps and that of their resumptive counterparts, in that sometimes the gap is more acceptable than the pronoun (e.g., in main clauses), sometimes it is less (e.g., the subject gap in *wh*-clauses), and sometimes it is about the same (e.g., the object gap in *wh*-clauses) (see also McDaniel and Cowart (1999) and Han et al. (2012)).

Fig. 3: Gaps and resumptive pronouns in relative clauses



3.2. Resumption cross-linguistically

We have now seen that the view of resumptive pronouns as being grammatical in English has a number of advantages. It can explain why speakers produce resumptive pronouns spontaneously in both island and non-island environments, why they do so even in the absence of production pressure, why they nonetheless judge them to be of relatively low acceptability, and why this acceptability bears no consistent relation to the acceptability of gaps in counterpart sentences. One potential problem with this view, however, is that it seems to erase a typological distinction that has been central to analyses of resumptive pronouns for over 30 years. That is, it has been widely assumed since Chao and Sells (1983) and Sells (1984) that in addition to languages like English that have only 'intrusive' resumptive pronouns (termed 'processor resumptives' in Asudeh (2011)), taken to be extra-grammatical and driven by processing pressures, there are also languages like Irish, Hebrew and Lebanese Arabic, in which resumptive pronouns are taken to be a fully integrated part of the grammar. The Irish relative clause in (16), for instance (from McCloskey (2006)), appears to be fully acceptable, unlike English equivalents with resumptives such as (15a).

(16) an ghirseach ar ghoid na síogaí í
the girl C stole the fairies her
'the girl who the fairies stole'

If, as I have proposed here, resumptive pronouns in English are grammatically licit after all, then accounting for the difference between languages like English and languages like Irish would seem to pose a significant obstacle.

A closer look at the analysis proposed here, however, suggests that it may actually be capable of capturing this type of difference in an insightful way. As we have seen, even if resumptive pronouns are grammatical in English, there are still some significant processing factors that discourage their use and lower their acceptability. Gaps are also generally grammatical in English, but as is well known, they also present very significant processing difficulties. In forming *wh*-dependencies in real time, then, speakers must choose between these two options, each of which comes with its own set of problems, but they presumably do the best they can under the circumstances (see Morgan and Wagers (2013), Ackerman et al (2014) and Beltrama and Xiang (2016) for discussion of this choice between resumptives and gaps). As we turn to other languages, it is not difficult to imagine that almost every detail of this decision is going to be different. Languages differ in the nature of their *wh*-phrases, for instance, including the extent to which grammatical features such as number, gender and case are represented. As we saw in the discussion of referentiality, small changes in the *wh*-phrase can lead to significant facilitation in processing the gap, so we would reasonably expect to find differences across languages in the acceptability of gaps and in the choice between gaps and resumptive pronouns. In addition, there is even wider cross-linguistic variation in pronouns, and here too, the amount of grammatical information present on the pronoun and how it relates to the larger pronominal system would reasonably be expected to affect the ease with which antecedents are retrieved from memory and how the choice between gaps and pronouns is made. Finally, the types of intervening structures that will be found between *wh*-phrases and gaps/resumptives varies a lot across languages, and this too could affect the choice between gaps and resumptives, given that gaps are much more sensitive to this structure than pronouns are.

Given all of this, it should not be surprising that a resumptive pronoun might be an optimal choice (and of relatively high acceptability) for a given *wh*-dependency in one language, while being a problematic choice (and of relatively low acceptability) in another language, despite being grammatical in both. In fact, the study of *wh*-dependencies with gaps has already provided examples of how small differences within and across languages can lead to major differences in acceptability. As we have already seen, for instance, making a *wh*-phrase referential can make an island violation seem to disappear, although more rigorous study shows that the island effect is in fact still present (as in Figure 1 above and as discussed in Goodall (2015)). Similarly, certain island violations appear to be absent in some languages, although closer investigation shows that they can still be detected. Namboodiripad and Goodall (2015), for instance, show that although adjunct island violations in Malayalam are of relatively high acceptability and are frequently attested, they nonetheless show a clear island effect when scrutinized experimentally. The status of resumptive pronouns across languages might be analogous to these types of examples with gaps. A whole range of properties might make resumptives more or less acceptable in any given language, even though the fact that they are allowed by the grammar remains constant across languages.

One way to conceive of the situation just described is to imagine that as various properties of the language make resumptives a more or less acceptable choice, the dashed lines in Figure 3 (representing the acceptability of resumptive pronouns) move up or down in parallel. Since the exact acceptability of gaps can also vary across languages, it is difficult to make precise predictions at this abstract level, but we do expect to find languages in which, as in English, resumptives are more acceptable than gaps in some positions, but not in others. This appears to be the case in Swedish and Vata (Engdahl (1985), Koopman (1984), McCloskey (2006)), where resumptives are preferred in embedded subject position (the equivalent of (14b) in English) but are otherwise no better than gaps in standard island environments (such as (15c)-(15d) in English). In other languages, we would expect the acceptability of resumptives to be high enough that they would be essentially interchangeable with gaps in non-island environments, but better than gaps in island environments. This appears to be the case in Irish and a number of other languages (see McCloskey (2006) for discussion).

Whether these differences in the acceptability of resumptives among English, Irish, and other languages can be traced back to the nature of *wh*-phrases, gaps, pronouns and intervening pronouns in these languages, as the analysis here would predict, remains to be seen, of course, and cannot be fully addressed here. It is interesting to note, though, that in at least some cases of resumptive pronouns in Irish, the beginning of the *wh*-dependency gives a cue to the processor to expect a resumptive. In (16), for instance, the complementizer is different from what would be used for a gap (McCloskey (2011)), and it is reasonable to assume that this would remove some of the processing difficulty associated with resumptives that we discussed earlier for languages like English. It is small factors like these that could raise substantially the acceptability of resumptives in particular languages, thus leading to the impression that resumptives differ in their grammatical status across languages. If the analysis proposed here is on the right track, however, this impression is misleading.

4. Conclusion

As we saw at the outset, there are a number of restrictions on *wh*-dependencies, but these restrictions may be at least partially overcome in one of two ways: by adding referential material to the *wh*-phrase or by using a resumptive pronoun in place of the gap. The ameliorating effect of referentiality is standardly taken to be a grammatical phenomenon, while that of resumption is taken to be extra-grammatical. Here we have seen evidence for the opposite view: that referentiality is extra-grammatical (the result of the way that working memory operates) and that resumption is grammatical, even in a language like English.

Beyond these correctives to the way that referentiality and resumption are usually described, the analyses presented here lead us to think about *wh*-dependencies in a new way. If, as claimed here, both gaps and pronouns are grammatical ways to resolve a *wh*-dependency, it is notable that neither is unique to this type of dependency (see McCloskey (2006) for discussion of this point with regard to pronouns). Gaps are presumably part of the general "toolkit" available for A'-movement phenomena and the fact that pronouns are used outside of resumption is of course uncontroversial. The picture that emerges is that *wh*-dependencies present a problem to be solved, in that both ends of the dependency must be represented, and languages use what they have available to solve it. If the *wh*-phrase is in a position indicating the scope of the

dependency, then something else must indicate the position of the variable, and both gaps and pronouns are reasonable ways to do this.

The fact that gaps and pronouns are reasonable solutions, however, does not mean that they are perfect solutions. Gaps have the advantage of being more economical, in that they do not require overt pronunciation, and of indicating the variable position unambiguously, in that they typically cannot be interpreted in any other way. On the other hand, their lack of overt realization can make them hard to detect in real time, a fact that no doubt contributes to the well-known processing burden that gaps impose. Perhaps more importantly, gaps are severely degraded (and thus essentially unavailable) in certain structural environments. Whether this fact is attributable to grammar or to properties of working memory, it means that gaps are not a possible way to resolve *wh*-dependencies in many cases. Some amelioration is possible by using a referential *wh*-phrase, as we have seen, but this is not a very economical solution and in any event, a certain amount of degradation is still present.

Using a pronoun to resolve a *wh*-dependency has the advantage that pronouns are not sensitive to the island effects that restrict the use of gaps. In addition, pronouns have the advantage of being overt and thus easy to detect in real time. The disadvantages of pronouns are considerable, however. They are less economical than gaps, in the obvious sense that pronouns must be pronounced, and they do not show unambiguously that the dependency is being resolved. As we have seen, there is an inherent ambiguity to pronouns in *wh*-dependencies: at the point when they are parsed, they could be taken to resolve the dependency or they could be taken to refer to some other discourse entity. This ambiguity plausibly increases the processing burden and thus decreases acceptability.

Neither of the available ways to resolve a *wh*-dependency is thus cost-free in terms of grammar and/or processing. In fact, some significant cost seems to be an inevitable consequence of a dependency of this type; the only way to avoid it would be to avoid the dependency itself. Many languages do this, adopting instead the *wh*-in-situ strategy, but this induces some significant costs of its own (see Kim and Goodall (2016) for some discussion).

We have seen, then, that *wh*-dependencies present a problem that languages solve, albeit imperfectly, with the mechanisms available to them. In addition to the empirical support that we have seen, this approach has the advantage that nothing needs to be stipulated in the grammar, either about different types of fillers (i.e., referential vs. non-referential) or about resumptive pronouns (including any specification of whether resumptives are intrusive or grammatical), since all resumptives are simply ordinary pronouns under this view. Many problems remain, of course, and others will no doubt emerge, but the general approach appears promising.[1]

This view of *wh*-dependencies has been crucially made possible by the use of formal sentence acceptability experiments. There is an obvious sense in which this is true: the empirical motivation behind the proposals here has come from experimental work. It is also true in a more subtle way, however. As Cowart (1997) and many others have shown, sentence acceptability experiments are significantly influenced both by factors known to be grammatical and others

[1] One interesting challenge remaining is the highest subject restriction on resumptive pronouns that has been reported for some languages (see McCloskey (2006, 2011) for discussion). Another concerns the detailed experimental work that is now emerging on languages in which resumption is traditionally thought to be grammatical (see, e.g., Tucker et al. (2015) and Keshev and Meltzer-Asscher (to appear)).

known to stem from processing (especially working memory). Interpreting an experimental result thus forces us to examine seriously the possible roles of both grammar and processing, and in the cases discussed here, this has led to a reversal of the roles that grammar and processing have typically been thought to play in these phenomena. This in turn allows for a more adequate account of referentiality and resumption, and, it is hoped, a better understanding of *wh*-dependencies more generally.

References

Ackerman, Lauren, Michael Frazier, and Masaya Yoshida. 2014. Resumptive pronouns salvage island violations in forced-choice tasks. Poster presented at the 27th Annual CUNY Conference on Human Sentence Processing, Ohio State University

Alexopoulou, Theodora, and Frank Keller. 2007. Locality, cyclicity, and resumption: At the interface between the grammar and the human sentence processor. *Language*, 110-160.

Alexopoulou, Theodora, and Frank Keller. 2013. What vs. who and which: Kind-denoting fillers and the complexity of whether-islands. *Experimental syntax and island effects*, eds. Jon Sprouse and Norbert Hornstein, 310-340. Cambridge University Press.

Asudeh, Ash. 2011. Towards a unified theory of resumption. *Resumptive pronouns at the interfaces*, ed. Alain Rouveret, 121-187. John Benjamins Publishing.

Beltrama, Andrea, and Ming Xiang. 2016. Unacceptable but comprehensible: the facilitation effect of resumptive pronouns. *Glossa: a journal of general linguistics*, *1*(1): 29.

Cann, Ronnie, Tami Kaplan, and Ruth Kempson. 2004. Data at the grammar-pragmatics interface: The case of resumptive pronouns in English. *Lingua* 115: 1551–1578.

Chao, Wynn and Peter Sells. 1983. On the interpretation of resumptive pronouns. *Proceedings of the Thirteenth Annual Meeting of the North-Eastern Linguistics Society*, 47–61. GLSA, University of Massachusetts, Amherst.

Chomsky, Noam. 1973. Conditions on transformations. *Festschrift for Morris Halle*, eds. Stephen R. Anderson and Paul Kiparsky, 232-286. New York: Holt, Reinhart & Winston.

Chung, Sandra. 1982. Unbounded dependencies in Chamorro grammar. *Linguistic Inquiry* 13: 39-77.

Chung, Sandra. 1994. Wh-agreement and "referentiality" in Chamorro. *Linguistic Inquiry* 25, 1-44.

Chung, Sandra. 1998. *The design of agreement: Evidence from Chamorro*. University of Chicago Press.

Cinque, Guglielmo. 1990. *Types of A'-Dependencies*. Cambridge, MA: MIT Press.

Cowart, Wayne. 1997. *Experimental syntax: Applying objective methods to sentence judgments*. Sage Publications.

Dickey, Michael Walsh. 1996. Constraints on the sentence processor and the distribution of resumptive pronouns. *Linguistics in the laboratory*, eds. M. W. Dickey and S. Tunstall, 157–192. GLSA, University of Massachusetts, Amherst.

Engdahl, Elizabeth. 1985. Parasitic gaps, resumptive pronouns and subject extraction. Linguistics 23: 23–44.

Ferreira, Fernanda, and Benjamin Swets. 2005. The production and comprehension of resumptive pronouns in relative clause 'island' contexts. *Twenty-first century psycholinguistics: Four cornerstones*, ed. Anne Cutler, 263-278. Erlbaum.
Goodall, Grant. 1993. SPEC of IP and SPEC of CP in Spanish *wh*-questions. *Linguistic Perspectives on the Romance Languages*, eds. William J. Ashby, Marianne Mithun, Giorgio Perissinotto, and Eduardo Raposo, 199-209. Amsterdam: John Benjamins Publishing.
Goodall, Grant. 2004. On the syntax and processing of *wh*-questions in Spanish. *Proceedings of the 23rd West Coast Conference on Formal Linguistics*, 101-114. Somerville, MA: Cascadilla Press.
Goodall, Grant. 2010. Experimenting with *wh*-movement in Spanish. *Romance Linguistics 2008: Interactions in Romance*. ed. Karlos Arregi,, 233-248. John Benjamins Publishing.
Goodall, Grant. 2015. The D-linking effect on extraction from islands and non-islands. *Frontiers in psychology* 5.
Han, Chung-hye, Noureddine Elouazizi, Christina Galeano, Emrah Görgulu, Nancy Hedberg, Jennifer Hinnell, Meghan Jeffrey, Kyeong-min Kim, and Susannah Kirby. 2012. Processing strategies and resumptive pronouns in English. *Proceedings of the 30th West Coast Conference on Formal Linguistics*, pp. 153-161. Somerville: Cascadilla Proceedings Project.
Heestand, Dustin, Ming Xiang and Maria Polinsky. 2011. Resumption still does not rescue islands. *Linguistic Inquiry*, 42: 138-152.
Hofmeister, Phillip. 2011. Representational complexity and memory retrieval in language comprehension. *Language and Cognitive Processes*, 26: 376-405.
Hofmeister, Phillp. and Ivan A. Sag. 2010. Cognitive constraints and island effects. *Language*, 86: 366-415.
Keshev, Maayan and Aya Meltzer-Asscher. To appear. Active dependency formation in islands: How grammatical resumption affects sentence processing. *Language*.
Keffala, Bethany. 2011. Resumption and gaps in English relative clauses: Relative acceptability creates an illusion of 'saving'. *Berkeley Linguistics Society* 37, 140-154.
Keffala, Bethany, and Grant Goodall. 2011. Do resumptive pronouns ever rescue illicit gaps in English? Poster presented at the 24th Annual CUNY Conference on Human Sentence Processing, Stanford, California.
Keffala, Bethany, and Grant Goodall. 2013. On processing difficulty and the acceptability of resumptive pronouns. Paper presented at Linguistic Evidence – Berlin Special, Humboldt-Universität, Berlin.
Kim, Boyoung, and Grant Goodall. 2016. Islands and non-islands in native and heritage Korean." *Frontiers in Psychology* 7.
Kluender, Robert. 1998. On the distinction between strong and weak islands: A processing perspective. *Syntax and Semantics* 29: *The Limits of Syntax*, eds. Peter Culicover and Louise McNally, 241-280. Academic Press.
Kluender, Robert., and Marta Kutas 1993. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8, 573- 633.
Koopman, Hilda. 1984. The Syntax of Verbs: From Verb Movement Rules in the Kru Languages to Universal Grammar. Dordrecht: Foris Publications.

McCloskey, James. 2006. Resumption. *The Blackwell Companion to Syntax*, eds. Martin Everaert and Henk C. van Riemsdijk, 94-117. Wiley-Blackwell.
McCloskey, James. 2011. Resumptive pronouns, Ā-binding, and levels of representations in Irish. *Resumptive pronouns at the interfaces*, ed. Alain Rouveret, 65-119. John Benjamins Publishing.
McDaniel, Dana and Wayne Cowart. 1999. Experimental Evidence for a Minimalist Account of English Resumptive Pronouns. *Cognition* 70, B15–B24.
Morgan, Adam Milton and Matthew Wagers 2013. Gap acceptability predicts resumption rate in English. Presented at Architectures and Mechanisms for Language Processing Conference, Marseille, France.
Namboodiripad, Savithry and Grant Goodall. 2015. On the universality of adjunct islands: Evidence from Malayalam. Poster presented at the 28th Annual CUNY Conference on Human Sentence Processing, Los Angeles, California.
Polinsky, Maria, Lauren Eby Clemens, Adam Milton Morgan, Ming Xiang, and Dustin Heestand. 2013. Resumption in English. *Experimental syntax and island effects*, eds. Jon Sprouse and Norbert Hornstein, 341-360. Cambridge University Press.
Prince, Ellen F. 1990. Syntax and discourse: A look at resumptive pronouns. *Berkeley Linguistics Society* 16, 482 – 497.
Prince, Ellen F. 1997. On kind-sentences, resumptive pronouns, and relative clauses. *Towards a social science of language: Papers in honor of William Labov*, eds. Gregory R. Guy, John G. Baugh, Deborah Schiffrin, and Crawford Feagin, 223-235. Cambridge: Cambridge University Press.
Rizzi, Luigi. 2001. Relativized Minimality Effects. *The Handbook of Contemporary Syntactic Theory*, eds. Mark Baltin and Chris Collins, 89-110. Blackwell: Oxford.
Rizzi, Luigi. 2004. Locality and left periphery. *Structures and beyond: The cartography of syntactic structures* 3, ed. Adriana Belletti, 223-251. Oxford University Press.
Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral dissertation, Massachusetts Institute of Technology.
Sells, Peter. 1984. Syntax and semantics of resumptive pronouns. Doctoral dissertation, University of Massachusetts, Amherst.
Sprouse, Jon, Ivano Caponigro, Ciro Greco, and Carlo Cecchetto. 2016. Experimental syntax and the variation of island effects in English and Italian. *Natural Language and Linguistic Theory*.
Stiller, Alex. 2014. The acceptability of inverted and non-inverted *wh*-constructions in Spanish. Comprehensive research paper, Department of Linguistics, University of California San Diego.
Szabolcsi, Anna and Frans Zwarts. 1993. Weak islands and an algebraic semantics of scope taking. *Natural Language Semantics* 1, 235-284.
Szabolcsi, Anna and Frans Zwarts. 1997. Weak islands and an algebraic semantics of scope taking. *Ways of Scope Taking*, edited by A. Szabolcsi, 217-262. Dordrecht: Kluwer.
Torrego, Esther. 1984. On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15, 103-129.
Tucker, Matthew A., Ali Idrissi, Jon Sprouse, and Diogo Almeida. 2015. Grammaticalized Resumption Helps a Little with Islands and D-Linking Helps a Lot: Evidence from

Modern Standard Arabic Acceptability. Poster presented at Architectures and Mechanisms for Language Processing Conference, Malta.

NEW THOUGHTS ON OLD QUESTIONS – RESUMPTION IN IRISH *

Jim McCloskey
University of California, Santa Cruz
mcclosk@ucsc.edu

This paper is concerned with the apparent fact that natural languages build Ā-dependencies either by way of a filler-gap dependency or by way of a resumptive dependency. Its principal empirical goal is to clarify the circumstances under which a choice is made between gaps and resumptive pronouns in Ā-binding constructions in Irish. It is shown that when in competition with gaps pronouns are disfavored to an overwhelming degree and that they are tolerated only in positions where heightened parsing pressures come to bear. The implications of this finding for the theory and typology of resumption are considered. It is argued that, for Irish and English at least, the relevant parameter makes no reference to pronouns but only to properties of the functional head C.

1. A Choice

On the face of things, the syntax of natural language seems to make available two options at least for the construction of binding-relations between a clause-peripheral position α (higher) and a clause-internal position β (lower):

- The creation of a filler-gap dependency between position α and position β, β empty.[1]
- The binding of a pronoun in position β from position α. Pronouns so bound are known as 'resumptive' pronouns.

These options are exemplified in (1), the first illustrating the filler-gap dependency, the second the resumptive dependency – an option deployed fairly frequently, it seems, at least in informal registers (see, for instance, Bennett (2008)).

(1) a. the guy that I was talking to
b. the kind of guy that you never know if he'll be on time or not

Much of the discussion around this pair of options has drawn a distinction between two kinds of language – in one group, the grammar defines both options as well-formed (varieties of Arabic, Hebrew, Irish), while in the other only the filler-gap dependency is well-formed (this class includes English, German, and Greek on most accounts). On this view, (1b) is not a well-formed expression of English but is rather an instance of 'intrusive' resumption (Chao and Sells (1983)). On this

* I am very happy to be able to dedicate this paper to Sandy Chung and to acknowledge the gift of many years of argument, encouragement, scepticism, support, collaboration and intellectual companionship. Discussions over several years with Matt Wagers concerning the theoretical issues dealt with here have been invaluable, as were discussions with Mícheál Hoyne concerning the interpretation of the Irish data. The research reported on here was supported, in part, by NSF Award 1451819 to UCSC (Pranav Anand, PI), Daniel Hardt and James McCloskey, co-PI's).

[1] The syntactic operation which links filler and gap is often taken to be movement (or Internal Merge in recent work); for my purposes here, however, it will mostly not matter how the syntax of that relationship is understood.

view, the use of the resumptive pronoun in (1b) reflects the use of an extragrammatical 'last resort' mechanism under various sorts of performance pressures (for discussion, see Asudeh (2004, 2012), Alexopoulou (2006, 2010), Alexopoulou and Keller (2007), Heestand et al. (2009), Ackerman et al. (2014), Beltrama and Xiang (2016) among many others). For that reason, discussions of resumption often raise difficult and useful questions about the relation between formal grammars and the mechanisms of production.

Irish has been regarded as one of the exemplary members of the 'true resumption' club of languages. And it is indeed very clear why one might conclude that resumption is a grammatically licensed option in that language. Clauses which host filler-gap dependencies are introduced by the 'direct relative' complementizer, while those which host resumption dependencies are introduced by the 'indirect relative' complementizer. The ramifications of THAT choice in turn spread through the morphosyntactic system of the language – determining how verbs are inflected, what verb-stems are used when, what form is taken by certain functional elements (the copula, certain aspectual particles) and so forth (Duffield (1995: chap. 3), McCloskey (2001), Oda (2012), Acquaviva (2014), Ostrove (2015, 2016)). Since the morphological alternations are determined by complementizer choice and since complementizer choice is in turn determined by the choice between a filler-gap dependency and a resumptive dependency, that option in turn, it seems, must be represented in the grammar of the language.[2]

The availability of both options to speakers of Irish is vividly illustrated by the two examples in (2), which were used within minutes of one another by the same radio reporter to describe the same situation (an emergency at sea):

(2) a. an bheirt a bhí siad ag iarraidh a shábháil
the two C-FG be.PAST they PROG try save.NON-FIN
'the two that they were trying to save' RADIO REPORT

b. an bheirt a raibh siad ag iarraidh iad a shábháil
the two C-RP be.PAST they PROG try them save.NON-FIN
'the two that they were trying to save them' RADIO REPORT

(2a) involves a filler-gap dependency; (2b) involves a resumptive dependency. In this case, the difference between the two complementizers is reflected primarily in the different suppletive allomorphs triggered on the finite verb – the 'independent' form *bhí* in (2a) triggered by C-FG, the 'dependent' form *raibh* in (2b) triggered by C-RP (on the mechanisms involved here, see Duffield (1995: chap. 3), McCloskey (1996a, 2001), Oda (2012), Acquaviva (2014) and especially Ostrove (2015, 2016)).

The examples of (2) were spontaneous oral productions; the same pattern of optionality is illustrated in the written medium by the two examples of (3), which were used in the same text by the same author within a page of one another. In this case the syntactic position which hosts either a gap (in (3a)) or a resumptive pronoun (in (3b)) is the subject position of a nonfinite clause, itself a complement to the modal expression *ní mór (do x)* ('must (to x)').

[2] Because the contrast between these two complementizers in Irish will be important for what follows, I will use 'C-FG' to gloss the complementizer which heads clauses which host filler-gap dependencies and 'C-RP' to gloss the complementizer which heads clauses into which a relation of resumptive binding reaches.

(3) a. na tréithe nár mhór a bheith ann
the traits that+must be.NON-FIN in-him
'the traits that it is necessary for him to have' CTP 153

b. na tréithe eile nár mhór don mhúinteoir iad a bheith aige
the traits other that+must to-the teacher them be.NON-FIN at-him
'the other traits that it is necessary for a teacher to have them' CTP 154

The kind of optionality seen in (2) and (3) is available for a fairly broad range of syntactic positions, including at least the following (see McCloskey (1990/2011) for documentation and details):

- direct object position in a matrix clause,
- subject and object positions in complement clauses (finite and nonfinite),
- object of a verb in progressive aspect (see (2) above),
- subject position of finite verbless clauses – so-called 'copula' clauses.

Many of these patterns of optionality (the first and second in particular) will be illustrated in some detail as the discussion proceeds.

Optionality breaks down under two circumstances. There is, in the first place, one position (and only one as far as is currently known) from which resumptive pronouns are excluded – in the highest subject position of a verbal clause (reflecting the so-called Highest Subject Restriction):

(4) *an fear a raibh sé breoite
the man C-RP be.PAST he sick
'the man that (he) was sick'

On the other side of the coin, resumption is the only option for positions out of which movement is impossible: prepositional object position, possessor position, positions within islands, or positions within coordinate structures. See McCloskey (1985, 1990/2011, 2002), Maki and Ó Baoill (2011) for documentation and details; the repertoire of island effects observed is remarkably familiar from theoretical discussion and from discussions of other languages and language-families and the general patterns of obligatoriness, impossibility and optionality for the resumptive are also fairly familiar (see for instance Doron (1982), Borer (1984), Shlonsky (1992), Sichel (2014) on Hebrew).

My focus in this paper will be on this interesting fact – that two distinct mechanisms for establishing Ā-binding relations seem to coexist, in general and within particular languages. My first goal is empirical – to provide a better understanding than has been available to date of the distributional patterns found within one language (Irish), when options like those in (2) and (3) are in play. The second goal will be to use the answers that emerge from that investigation to engage in some more speculative discussion of what linguistic theory should have to say about resumption and the typology of resumption. In pursuing these questions, I will use a data-base of naturally-occurring examples that I have built up over several decades of observation.[3]

[3] Sources for examples are not given here (because of length restrictions). I will gladly supply them on request.

2. A Question

Say one took the view (naive perhaps, but not irrational) that natural language has been so shaped as to allow people to express what they feel they need to express as efficiently as possible – with minimal effort and with a minimum of troublesome ambiguity. Considered in that light, it is very puzzling why the filler-gap option (the movement option) should exist at all as a way of forming Ā-dependencies, given the obvious advantages enjoyed by resumption-based dependencies over filler-gap dependencies.

ONE: EFFABILITY Resumptive-binding, at least of the kind found in Irish, is not island-sensitive. This means that examples like those in (5) are unremarkable and are relatively common.

(5) a. caisleán a mbéifidhe i ndiaidh na scafaill a bhí ag baint an mhaise de
castle C-RP be.COND.IMPERS after the scaffolding C was PROG take the beauty of-it
a chaitheamh anuas
cast.NON-FIN down
'a castle that the scaffolding that was depriving it of its beauty would have recently been taken down ' MRE 253

b. chun an ghoirt úd a mbraithim pé díth sláinte a bhíonn orm ag dul
to the field DEMON C-RP I-feel whatever lack health.GEN C is on-me PROG go
ann dom ag scaradh liom le linn é a fhágaint dom
into-it to-me PROG separate from-me as it leave.NON-FIN to-me
'to that field which I feel whatever ill-health I suffer from as I enter it falling away from me as I leave it' AI 238

c. hata go ligfeadh bligeard sráide fead dá siúlódh bean thairis síos sráid
hat C-RP let.COND blackguard street.GEN whistle if walk.COND woman by-him down street
mhór an Daingin ag caitheamh a leithéid ar a ceann
main the Dingle PROG wear it's like on her head
'a hat that a street-tough would whistle if a woman should walk by him down the main street of Dingle wearing the like of it on her head' PI 54

Attested island-violating structures like those in (5) are often very complex syntactically; in (5b), for example, the resumptive binding relation reaches into an adjunct island which is in turn contained within a relative clause island; in (5c) the binding relation reaches into a possessor position within a nominal, which is in turn within an adjunct island, that island itself then contained within a larger adjunct island (a conditional clause).

Nevertheless, such examples are well-attested in the corpus just described – 165 examples in all. As a point of comparison, there are 439 examples in which an Ā-dependency of one kind or the other reaches into an embedded clause which is not an island – as in the three illustrative examples of (6). The first has a subject gap, the second an object gap, and the third a resumptive pronoun.

(6) a. na fir a d'inis Fionnbhráid damh a tháinig an bealach seo
the men C-FG tell.PAST to-me C-FG come.PAST the way this
'the men that Fionnbhráid told me had come this way' SRNF 51

b. Rúitín a cheap sé a ghortaigh sé.
ankle C-FG think.PAST he C-FG hurt.PAST he
'It was an ankle that he thought he had injured.' RNG 260616

c. an té a gceapann siad go bhfuil airgead aige
the one C-RP think.PRES they C be.PRES money at-him
'the one that they think has money' DGD 216

The relative frequency of such complex structures in our corpus presumably means that speakers find it useful to be able to express the complex properties that they encode. And if relative frequency can stand as a rough proxy for relative usefulness, we can say that the probability that an island example will be pressed into service is 37% of the probability that a biclausal dependency like (6) will be. And it is striking that the island examples are in turn enormously more frequent than examples like (7), which incorporate a dependency which reaches across two clausal boundaries. Of these exactly three examples turn up, by comparison with the 165 island examples:

(7) an rud is dóigh leat ba mhian léi a dhéanfá
the thing C-FG likely with-you C-FG desire with-her C-FG do.COND.S2
'the thing that you think that she would like for you to do' DPB 12

That is, island examples are 55 times more likely to be deployed than are structures like (7).

And indeed there is no reason that I know of to imagine that our cognitive apparatus has any particular difficulty in creating, grasping or manipulating complex properties like those expressed in (5) – the property, say, of being an x such that unspecified people had just removed the scaffolding that was marring the beauty of x (see (5a) above). And such complex properties are easily expressible by way of resumptive binding. They are not so easily expressible using a filler-gap dependency, a fact which becomes immediately clear when one tries to render such examples in grammatical English (as I have many times) without reaching for an intrusive resumptive pronoun. Filler-gap dependencies are hobbled by an array of locality and other kinds of restrictions and constraints which have been one of the major foci of work in theoretical syntax since Ross (1967). Such restrictions considerably reduce the expressive capacity of syntactic systems which rely exclusively on the filler-gap mechanism for negotiating Ā-binding relationships. Resumptive dependencies are not similarly restricted.

TWO: TROUBLESOME AMBIGUITIES: Use of the filler-gap mechanism frequently results (in a VSO language) in ambiguity of a supposedly debilitating kind, one in which it is impossible to tell whether the relativization site is the subject or object of a transitive verb. These ambiguities emerge for Irish because it makes no case distinction between non-pronominal subjects and objects. Given then a relative clause consisting of a transitive verb and a single audible nominal, it is often impossible to tell whether the gap is a subject-gap or an object-gap (see McCloskey (1985), Hoyne (2016)). The examples in (8) illustrate the ambiguity with verbs that select two animate arguments:

(8) a. i ndiaidh bhás an tiománaidhe a mharbh an taoiseach
after death the driver C-FG killed the chieftain
'after the death of the driver who killed the chieftain'

‘after the death of the driver whom the chieftain killed’ IFDT 171

b. na daoine a dhíbir Cromail ó thalamh na hÉireann
the people C-FG expelled Cromwell from land the .GEN Ireland .GEN
‘the people that expelled Cromwell from the land of Ireland’
‘the people that Cromwell expelled from the land of Ireland’ AT 18

This troubling and common-place ambiguity has its source in the awkward fact that, by definition, filler-gap dependencies terminate in phonologically empty positions. The corresponding examples involving resumption are of course unambiguous:

(9) na daoine ar dhíbir Cromail ó thalamh na hÉireann iad
the people C-RP expelled Cromwell from land the .GEN Ireland .GEN them
‘the people whom Cromwell expelled from the land of Ireland’

A language which exclusively used resumptive dependencies in its Ā-binding constructions would not be burdened with potentially troublesome ambiguities like those illustrated in (8). Now of course such subject-object ambiguities are hardly unknown – they are pervasive in the Germanic V2 languages for example (see Kaan (1996), Bader and Meng (1999) for overviews and references). But that literature has revealed widespread garden-path effects in the event that the expected parse (subject precedes object) turns out not to be the parse actually required. There is a strong processing cost entailed by such confounded expectations – one that would not be paid in the present case if all Ā-dependencies were resumptive dependencies and therefore gave rise to no ambiguity. It has in fact often been claimed that object resumption in Irish serves principally to avoid the kind of ambiguities seen in (8), a claim we return to shortly.

THREE: PROCESSING LOAD: There is an old intuition (one which appears in many versions and in many different theoretical frameworks and contexts) that the processing costs associated with resolving resumptive dependencies are less than, or are in some sense preferable to, the processing costs associated with the resolution of filler-gap dependencies (Givón (1975), Keenan and Comrie (1977), Wanner and Maratsos (1978), Maling and Zaenen (1982), Erteschik-Shir (1992), Hawkins (1994, 1999), Ariel (1999), Alexopoulou and Keller (2007)). One of the most important developments in this area in recent years has been that the questions that arise in assessing these ideas have been greatly sharpened in the exciting explosion of work on resumption that the last twenty years or so has seen – in the emerging frameworks of experimental syntax and experimental psycholinguistics in particular.

Most of this work (until very recently at any rate) has been concerned with languages in which resumption is taken to be intrusive in the sense already discussed – not part of the competence grammar *sensu strictu*. A result that has emerged with particular clarity for such languages is that use of a resumptive pronoun does not result in full acceptability for the relevant structures, or even in a measurable increase in acceptability by comparison with identical structures involving a gap (Alexopoulou (2006, 2010), Alexopoulou and Keller (2007), Heestand et al. (2009), Clemens et al. (2012), Han et al. (2012)). This finding has sometimes been presented as being surprising or unexpected, though why this is so is unclear to me. To say that resumption is ‘intrusive’ in a given language is exactly to say that the syntax of that language makes no provision for Ā-binding

relations which terminate in pronouns. That being the case, we should expect that resumptive structures will be judged unacceptable just like any other class of expressions defined as not fully well-formed by the syntactic system of the language. And this is precisely what is observed.

But what is genuinely striking, of course, is that such expressions, though flawed, are produced and used. That they are usable (and used) is evident from both production studies and corpus studies – Prince (1990), Ariel (1999), Ferreira and Swets (2005), Cann et al. (2005), Bennett (2008), Ackerman et al. (2014), Morgan and Wagers (2015). Now it is in no sense paradoxical *per se* that expressions which are ill-formed to some degree should be used by native speakers. Or at least it is not paradoxical if we adopt the kind of framework for investigation urged by Chomsky since at least the middle 1980's – one in which the distinction between E-language (a set of productions) and I-language (an internal symbolic system) has a central place(Chomsky (1986)). Within such a conception, our expectation will be that certain expressions which are defined as fully well-formed will be in practice unusable (maybe they involve many degrees of center-embedding or 35 levels of clausal subordination), and equally that certain expressions defined as not fully well-formed will be interpretable and will, for whatever reason, turn out to be useful and usable. This seems to be exactly the situation that we observe in English or Greek or German with regard to resumptive structures.

But such flawed structures would presumably not be pressed into service if they did not provide some value, despite their ill-formedness, either for those who produce them or for those who must comprehend them. There have been various proposals over the years about what that added value might be in the case of resumptive dependencies. Tony Kroch (1981) suggests that resumption emerges in English in response to poor initial planning on the part of producers; Ash Asudeh suggests (2004; 2012) that they are useful because they express the intended meaning and guarantee local (if not global) well-formedness; Beltrama and Xiang (2016) present evidence that resumptive structures, while they do not improve acceptability, increase comprehensibility by comparison with similar structures containing gaps. And Philip Hofmeister and Elizabeth Norcliffe (2013) argue for a very particular kind of processing advantage linked with resumption. They use the self-paced reading methodology to argue that resumptive pronouns show a measurable processing advantage over gaps – but only in high difficulty contexts. That is, reading times in the region of a resumptive pronoun in English are faster than those measured in the context of a gap – but only if the relevant region is already a region of high difficulty. In this circumstance, resumption is clearly facilitative. Furthermore, resumptive pronouns in such high-difficulty contexts are judged more acceptable than those which occupy less challenging positions. Neither effect, however, is detected in regions where the processing load is low.[4] In such contexts the only measurable consequence of using a resumptive element rather than a gap is lowered acceptability. We will return to this finding in the following section and present some additional evidence in its favor.

For now, though, the general conclusion is that there does seem be evidence of an interesting kind for a processing advantage pertaining to resumptive, as opposed to filler-gap, dependencies. So our naive question remains – why isn't every language a resumption-only language?[5]

[4] For the purposes of their study, what determines heightened difficulty is length of the dependency – cases in which the dependency must cross a complement clause boundary are difficult; monoclausal dependencies are not.

[5] The only case known to me which might qualify as a 'resumption only language' is Palauan, as described by the

There has been a note of embarrassment in this discussion – it could well be that it is fruitless, or worse, given our current level of understanding, to even entertain such naive design questions. But the asking perhaps helps to fend off complacency and may encourage us to be appropriately puzzled by the commonplace (in this case the ubiquity of filler-gap dependencies). And, as it turns out, the concerns and observations that have been to the fore here will help frame the discussion of the more tractable questions that follow – how the choice between the two dependency-types works itself out in the production of Irish sentences.

3. Inside Irish

If it is strange that resumption is not more widespread crosslinguistically as a means of building Ā-dependencies, it is all the more strange that, in a language such as Irish which offers its users a choice in the matter, resumption is massively disfavored in usage, when it is in competition with the gap option. I want to document here the extent to which this is true and to consider what the factors are which push producers of Irish utterances towards one choice or the other. In doing this, I will in part be making good on a trail of promissory notes scattered through my own earlier work on these topics (McCloskey (1985: 64–65), McCloskey (1990/2011: fn. 41, p. 116) for instance). I have claimed that the grammar of Irish makes available a free choice between resumption and filler-gap dependencies and that the choices actually made reflect performance factors. That is probably correct, but I have to confess that I have been shocked to discover, in looking more closely at the facts, how extreme the prejudice against resumption is.

Consider, to begin with, cases like (10), in which a resumptive pronoun appears in the highest object position of a relative clause.

(10) a. jab a-r fhág an oiread sin oibrithe é
job C-RP.PAST leave.PAST so many workers it
'a job that so many workers left (it)' AT 138

b. na tithe seo nár fhág aon duine fós iad
the houses DEMON C-NEG-PAST leave.PAST any person yet them
'those houses that no-one had yet abandoned (them)' LAN 141

66 examples of this type have appeared in my data-base[6] in the course of the three decades or more during which I have been keeping track. This is barely a third of the number of island-violating cases like (5) detected in the same period. I have not kept a count of the corresponding set of examples involving gaps in direct object positions. Given their frequency, recording every such example would have been an enormous task. However it is possible to estimate their frequency – by choosing 100 pages at random from among the texts out of which the data-base was constructed, counting the number of unembedded object gaps found in those pages, and then scaling up to the estimated total page-size of the corpus. That process yields a conservative estimate of around 64,000 examples for the filler-gap strategy in the case of unembedded direct objects. Given the

late Carol Georgopoulos (1991).

[6] Excluding cases in which appearance of the object resumptive is forced by a weak crossover configuration. See below.

uncertainties involved in its calculation, that number is very unlikely to be accurate; but the exercise gives a sense of how enormous the disparity in frequency is between the two patterns.

For more deeply embedded positions, the overall numbers are smaller (the relevant structures being more complex and rarer), but there is also more reason to be confident in the count, since from the start my goal has been to record every example of the various types we'll consider. This effort has yielded 353 examples in which a filler-gap dependency crosses a clause-boundary and terminates in embedded subject or object position (see (6) above). These are positions in which there is again a choice – either a gap or a resumptive pronoun can in principle be used. In fact, there are just 34 examples of the type in (11), in which the pronoun appears:

(11) a. na caiple sin a n-abrann sealgairí go mbíonn siad ag léimnigh agus
the horses DEMON C-RP say.PRES hunters C be.PRES-HABIT they PROG leap and
ag damhsa
PROG dance
'those horses that hunters say leap and dance' SS 238

b. Chonnaic mé iongantais nach gcreidfeadh fear ar bith go bhfeiceadh sé iad.
see.PAST i wonders NEG-C believe.COND man any C see.COND he them
'i have seen wonders that no man would believe that he would see them' UMI 216

That is, gaps are favored over pronouns in this context by a margin of of 91.2% to 8.8%. Why are these patterns as they are? In addressing that question, I will begin with unembedded object resumptives like those of (10) and with a negative conclusion, arguing that a factor which has been claimed to be central is of marginal importance at most. We have seen (at (8) above) that use of the filler-gap syntax can give rise to ambiguity – it is sometimes unknowable from the form of the relative clause itself whether one is encountering a subject gap or an object gap. (12a) is an additional example, one which could have been disambiguated, but was not, by the addition of a single unaccented syllable, an object pronoun.

(12) an t-oifigeach sgannruighthe a tharrtháil mé an oidhche roimhe sin
the officer frightened C-FG save.PAST I/me the night before that
'the frightened officer that I had saved the night before'
'the frightened officer that had saved me the night before' FFF 180

It is claimed in some contemporary pedagogical grammars that the principal function served by the object resumptive pattern is that of avoiding ambiguities like that in (12) (Anonymous (1960: §664, p. 336), Mac Giolla Phádraig (1963: 121), Ó Dónaill (2008: 148–149)). But this seems to be incorrect. There are almost as many attested ambiguous examples of this type (there are 51) as there are cases of object resumption itself. Furthermore, of the 66 attested cases of high object resumption, only 10 would be ambiguous if rendered as a filler-gap dependency. And even if we were to grant that a strategy of ambiguity avoidance is at work in these 10 potentially ambiguous cases, it would remain true that five times as many examples of the same type in the same corpus remain 'un-repaired'. And it would also remain true that the hypothesis of ambiguity-avoidance can account for at most 15% of our observations (10 out of 66 cases). Overall then, ambiguity-

avoidance seems to play at most a very minor role in shaping speaker choices – a conclusion reached independently in a careful recent discussion by Mícheál Hoyne (2016: 65–67).

It turns out moreover that for the ten cases in which ambiguity-avoidance might be thought to play a role, there is another factor which equally well predicts use of the resumptive dependency. Furthermore that factor extends in a natural way to a much larger proportion of the attested examples – 50 out of 66. The factor in question is animacy. Object resumptive pronouns are most frequently attested in the context of an animate head for the relative clause.

The relevant observations are summarized in the table of (13), which provides a breakdown of all examples of unembedded object resumptives attested in the data-base, with respect to three characteristics – animacy of the head, animacy of the object resumptive (these two linked of course), and animacy of the subject of the relative clause. It distinguishes four types of clause by these criteria, all of which are exemplified in the example blocks that follow ((14)–(17)).

(13)

	HEAD	RC-SUBJECT	OBJ-RESUM	
TYPE ONE	-Animate	-Animate	-Animate	7 exs.
TYPE TWO	-Animate	+Animate	-Animate	9 exs.
TYPE THREE	+Animate	+Animate	+Animate	21 exs.
TYPE FOUR	+Animate	-Animate	+Animate	29 exs.

(14) TYPE ONE (7 examples)

Trí rud ná leanfadh aon rath iad
three thing NEG-C follow.COND any good-fortune them
'three things that no good fortune would come of (them)' PF 188

(15) TYPE TWO (9 examples)

bíonn siad ag gearán faoi rud go dtuigeann tú é
be.PRES-HABIT they PROG complain about thing C-RP understand.PRES you it
'they complain about something that you understand (it)' RNG 100914

(16) TYPE THREE (21 examples)

daoine sa cheantar ar mharuigh an t-IRA iad
people in-the district C-RP.PAST kill.PAST the IRA them
'people in the district that the IRA had killed (them)' RNG 120813

(17) TYPE FOUR (29 examples)

an té nach gcorródh gol Phádraig é an oíche sin
the one NEG-C move.COND weeping Patrick him the night DEMON
'the one that Patrick's weeping would not move (him) that night' PI 211

The verb-type which most favors object resumption is the fourth – the class of verbs which take an inanimate external argument and an animate internal argument – a marked alignment pattern. One sub-class of this type is the class of object experiencer verbs (*satisfy*, *shock*, *frighten* and so

on) and these verbs are indeed well-represented in our sample.[7] However other verb-types which exhibit the crucial alignment-pattern (certain causatives for instance) are also well-represented:

(18) a. páiste a ndúiseodh drochbhrionglóid as a shuan é
child C-RP wake.COND bad-dream out-of his sleep him
'a child that a bad dream would waken (him) from his sleep' SG 116

b. an bhean go rúnóch an bhróg a bhí aige í
the woman C-RP fit.COND the shoe C-FG be.PAST at-him her
'the woman that the shoe that he had would fit (her)' SMB 14

But the most striking result here is that 50 out of 66 attested cases of unembedded object resumption – 76% of cases – involve animate heads. This pattern reverses the normal distribution, since in general, in the same corpus, just 30.5% of relative clauses have animate heads. It seems, then, that animacy of the head favors deployment of resumptive pronouns in cases of object relativization. Why should this be so?

Given the results reported in Hofmeister and Norcliffe (2013), our expectation will be that resumptive pronouns will be strongly disfavored in positions which are not loci of high processing difficulty, but will be favored (or less dis-favored) in regions of heightened processing difficulty. The crucial observation now is that there is in fact a large and rich literature in experimental psycholinguistics which shows that the combination of animate subject with object relativization is problematic for processing and, crucially, rare in production. The effect is plain for the English equivalent of our Type Four verbs (inanimate external arguments and animate internal arguments) and is palpable in the discomfort one feels on encountering English examples like those in (19):

(19) a. People that these claims shock should get a life.
b. I haven't met many linguists who this claim surprises.
c. I've worked with many children who this video has disturbed.

The reality of the effect suggested intuitively by (19) has been demonstrated in many studies and by way of an impressively broad range of methodologies – see, among others, Traxler et al. (2002), Mak et al. (2002, 2006), Gennari and MacDonald (2008, 2009), Lowder and Gordon (2014) and Wagers and Pendleton (2015). Gennari and MacDonald (2008, 2009) in particular present the results of two production studies, two corpus studies and two comprehension studies, all of which demonstrate in various ways that examples like (19) are at low probability in terms of production (in the lab and in the wild) and are difficult to comprehend. Roland et al. (2007) and Wagers and Pendleton (2015) provide additional corpus-based evidence establishing similar conclusions. Lowder and Gordon (2014) confirm the core finding with two studies involving eye-tracking while reading and Wagers and Pendleton (2015) add a new kind of evidence. Focusing on English relative clauses, they show, by way of two self-paced reading experiments using the filled-gap paradigm

[7] It is striking in this context that Sichel (2014: 666) reports that object experiencer verbs in Modern Hebrew forbid object gaps and require resumption under object relativization. In Hebrew this is a hard grammatical constraint, it seems, whereas in Irish, as we will see shortly, we are dealing with tendencies and preferences. As she notes, however, (fn. 7, p. 666)) given the proposals of Landau (2009), such apparent direct objects will in fact be objects of a null preposition, in which case the facts more clearly fall into place.

(Crain and Fodor (1985), Stowe (1986), Lee (2004), Wagers and Phillips (2014)), that animate relative clause heads (but not inanimate relative clause heads) lead the comprehender to expect a gap in subject position – an expectation necessarily confounded in the case of object relatives, with the ultimate result of increased reading times at the relative clause subject.

This finding is fairly well understood as far as parsing is concerned. If processing is in general active, probabilistic and predictive (Frazier (1987), Omaki et al. (2015) among many others), then when a comprehender encounters an animate relative clause head an implicit expectation is induced that a subject gap will be encountered and that it will complete the dependency. This is a reasonable expectation, since to a first approximation, subjects will be animate and inanimates will be non-subjects (see for instance Hopper and Thompson (1980), Aissen (1999) among many others). However reasonable such an expectation may be in general, though, it is bound to fail in the cases we care about here – animate heads in the context of object relativization. When the expectation founders and recalibration is called for, the processor struggles. Therefore an object gap in a relative clause headed by an animate nominal is inevitably a position of heightened processing pressure. Given the results of Hofmeister and Norcliffe (2013) then, we should expect that position to be relatively hospitable to resumptive pronouns and the observations summarized in table (13) are understandable.[8]

Or at least we expect them if certain structures which are hard to process are also rarely used. It is hardly obvious why processing and production should be linked in this way, but the fact that they are so linked is well established for the effect we are concerned with (Roland et al. (2007), Gennari and MacDonald (2009), Wagers and Pendleton (2015)). This is why the discussion about animacy and object relativization has been at the heart of recent debates about the mechanisms which connect processing and production (see, for one example, Macdonald (2013) and the various commentaries on that article).

This feels like progress, but the account is not yet complete. Our discussion links patterns in the distribution of resumptives in Irish with a certain class of processing difficulties. As far as I know, however, the same processing issues in English do not result in increased use of resumptive pronouns. The typical response to the difficulties of (19) is that the relative clauses are rendered instead as passives, so that the content expressible as (19) is in fact expressed by (20):

(20) a. People who are shocked by these claims should get a life.
b. I haven't met many linguists who are surprised by this claim.
c. I've worked with many children who have been disturbed by this video.

In the passive structures of (20), animate heads are paired with subject gaps and the problematic pairings are eliminated. It is for this reason that discussions of animacy and object relativization

[8] A question I must sadly leave open is the question of whether or not the animacy of the relative clause subject is a significant factor here. At the level of intuition, it seems to me that the English examples involving inanimate subjects pose more severe difficulties than those with animate subjects; but this is only an intuition. Similarly, in Irish there are more cases involving resumption in the context of an inanimate subject than in the context of an animate subject – this is the difference between TYPE THREE and TYPE FOUR in table (13). However it's not clear that the numerical difference between the two types is significant, especially in the absence of a baseline (on which see below). A similar uncertainty runs through much of the discussion of the English data.

in English have been almost exclusively concerned with issues of voice and argument alignment. And of course it is unsurprising that English speakers would respond as in (20) to the difficulties represented by (19) – in English, resumption is not fully well-formed but passivization is.

But the equivalent of (20) is not available to speakers of Irish, because the language lacks a promotional passive of the English type. The functions served by passive in English are for the most part served in Irish by an impersonal inflection on the finite verb (known as the 'autonomous' form), which licenses a null impersonal subject but triggers no realignment of grammatical relations (Stenson (1989), Nolan (2006), McCloskey (2007, 2010)):[9]

(21) a. Gortaíodh anuraidh é.
hurt.IMPERS.PAST last-year him
'He was hurt last year.'

b. Cuirfear amárach sa reilg áitiúil í.
bury.IMPERS.FUT tomorrow in-the graveyard local her
'She will be buried tomorrow in the local graveyard.'

It follows that the standard English 'repair' is not available to speakers of Irish and in the face of the processing pressure described earlier the prejudice against resumptive pronouns is overcome.

The logic used here extends to other cases of optional resumption. Dependencies terminating in embedded subject and object positions also tolerate resumptive pronouns (see the discussion around (6) above). Such long dependencies are well known to give rise to heightened processing difficulty and so we would expect a degree of tolerance in this context too for the resumptive option. For high objects also, among the 16 (out of 66) attested examples that do not fall under the animacy generalization, five appear either in coordinated relative clauses ((22a), for instance) or in a stacked relative clause ((22b)). In such cases the resumptive pronoun is separated from its binder by a substantial linear distance. But it is well established that increasing the linear distance between filler and gap strains short-term memory resources and leads to processing pressures.

(22) a. scamall dorchadais a d'imigh ar an dtoirt agus ar lean fuarallas é
cloud darkness.GEN C-FG leave.PAST on the moment and C-RP follow.PAST cold-sweat it
'a cloud of darkness that dissipated immediately and which was followed by a cold sweat' CPCC 113

b. ní raibh an t-athrach ba lugha a tháinig ar ghnúis a chomrádaidhe
NEG-PAST be.PAST the change smallest C-FG come.PAST on face his comrade
nach dtug sé fá dear é
NEG-C take.PAST he under-notice it
'there wasn't the smallest change that came over the face of his comrade that he didn't notice' AM 63

If all of this can be maintained, we are left with a residue of 11 (out of 66) examples in which

[9] There are two constructions which show the formal, but not the functional, properties of passive structures and which do involve promotion of the direct object to subject. However these constructions involve the expression of aspect – perfective or progressive – and are in no way semantically equivalent to their non-passive counterparts. See McCloskey (1996a) for discussion of the Perfective Passive and Nolan (2006) for more general discussion.

tolerance of the pronoun cannot be attributed to any processing pressure so far identified.[10]

The patterns observed so far can be summarized in the following terms: when the grammar of Irish seems to offer a choice between using a gap and using a resumptive pronoun, speakers decline the option of using a pronoun by overwhelmingly large margins. The extreme prejudice against pronouns is, however, overcome under the kinds of conditions identified by Hofmeister and Norcliffe (2013) for resumptive pronouns in English – hardly at all under light processing loads, with greater frequency at points of heightened processing pressure. Even in this circumstance, however, if the embedded subject/object condition is representative, pronouns are used only in about 9% of cases in which they might in principle be used.

At this point the supposed distinction between 'intrusive resumption' languages and 'true resumption' languages begins to look suspect. If the grammar of Irish simply makes available a free choice between the two options (and this is how the distinction is usually characterized), it is very puzzling that there should be such a dramatic disparity in frequencies of use between the two options. And the ideas I have been relying on here to describe the Irish facts are exactly those used by Hofmeister and Norcliffe (2013) to describe 'intrusive' resumption in English. In fact the description of the Irish facts in (22) could well have been (apart from the first bullet point) a description of what it means to be an 'intrusive resumption' language.

But there is undeniably SOME difference with respect to resumption which distinguishes the grammar of Irish from the grammar of English. I return to the question of what that difference might be in the next (and final) section. In anticipation of that discussion, though, there are certain other facts which should be highlighted.

The calculations concerning optionality and relative frequency that we have principally been concerned with here so far are simply irrelevant in certain configurations. The list of such configurations is very unsurprising – in positions from which movement (on standard assumptions) is impossible, gaps never appear. So there are, as already reported, 165 examples in our corpus in which an Ā-dependency reaches into an island. In none of those does a filler-gap dependency cross an island-boundary (see (5) above). The same observations hold for what would in the absence of resumption be violations of the coordinate structure constraint. Here too there are no instances of filler-gap dependencies and resumption is the only option. Our corpus yields 20 examples of this type. Even for an effect as apparently delicate as the weak crossover effect (Wasow (1972)) – of which there are 15 examples in our corpus – all show resumption and none show the filler-gap syntax:

(23) fear … ar fhág a bhean é
man C-RP.PAST leave.PAST his wife him
'a man that his wife left him' TC 164

None of this is very surprising – as long as the relevant constraints (island constraints, the coordinate structure constraint, the weak crossover constraint) are part of the grammar of Irish rather than reflections of parsing pressures. For all such cases, resumption is the only option offered by the grammar and the anti-pronominal prejudice cannot be awakened. It should be kept in mind

[10] Inanimacy (or perhaps non-agentivity) of the relative clause subject seems to play a role in some of these cases, both in Irish and in English. See footnote 8 above.

that these apparently categorical effects contrast starkly with the preferences we have mostly been concerned with in the discussion so far. When the grammar allows resumptive pronouns to be in competition with gaps, the anti-pronominal prejudice ensures that pronouns appear only in about 9% or 10% of cases. Even when animate heads are paired with object relativization, for instance, it is easy to find examples in which the facilitative pronoun is not used, as seen in (24).

(24) a. Fear farraige ná scanródh faic
man sea.GEN NEG-C frighten.COND anything
'a seaman that wouldn't be frightened by anything' GLL 9

b. an sagart a mharaigh na Dúchrónaigh ina dhiaidh sin
the priest C-FG kill.PAST the Black and Tans after that
'the priest that the Black and Tans subsequently killed' M 231

4. The Theoretical and Typological Landscape

We are left with the conclusion, then, that the facts are richer and more subtle than a clear-cut distinction between 'true' resumption and 'intrusive' resumption would lead us to expect them to be. The 'true' resumptive pronouns of Irish are disfavored to the same degree and for the same reasons as the 'intrusive' resumptive pronouns of English, and the factors which counter-weigh the anti-pronominal prejudice seem to be parallel in the two languages.

Similarly complex patterns seem to hold for at least some of the other 'true resumption' languages. Ariel (1999) presents a corpus-based study of direct object resumption in Hebrew whose results seem to be very close to those reported here for Irish (10% resumption, 90% gaps in high direct object position). Farby et al. (2010), furthermore, report the results of an acceptability study in Hebrew which mirror the frequency results reported here for Irish in the sense that, outside islands, resumptive pronouns are judged less acceptable (by a small but measurable and reproducible margin) than gaps in the same position, particularly for unembedded direct objects. Meltzer-Asscher et al. (2015) confirm and refine the finding.[11] More recently Tucker et al. (2016) report a very complex set of facts for Modern Standard Arabic (based on two acceptability studies); among their clear conclusions however (p. 36) is 'that resumption is dis-preferred in certain long-distance dependencies,' and that 'in certain grammatical corners, a grammaticalized resumption language can behave like an intrusive resumption language in penalizing the presence of a pronoun'.

There are of course relevant differences between the grammars of Irish and English. But those differences seem to have nothing to do (directly) with pronouns, but rather to center exclusively on the category C. C-RP forces verbs which appear to its immediate right into the 'dependent' form, while C-FG does not. C-FG (but not C-RP) optionally triggers the appearance of WH-forms of verbs to its immediate right (see the brief review at page 80 above):

(25) an luach a shíleas tú a gheobhfas tú
the price C-FG think.PRES-WH you C-FG get.FUT-WH you
'the price that you think you'll get'

[11] Inside islands, both studies found that resumptive pronouns were judged more acceptable than gaps, consistent with what has been reported for Irish over many years and consistent with our corpus-based findings here.

Such morphological effects must be linked with whatever properties give rise to the syntax of filler-gap dependencies (successive-cyclicity, island-effects, across the board effects), all of which are characteristic of clauses headed by C-FG, but not of clauses headed by C-RP. The needed linkages can be made if we take it that C-FG is defined by its finiteness in combination with whatever features drive the syntax of filler-gap dependencies and that those same features are what determine the morphological effects in (25).

On this view, the connection between choice of C-RP and use of a resumptive pronoun is indirect and in an important sense accidental (see also Duffield (1995)). The complementizer which we have called here C-RP appears characteristically, but not exclusively, in C-positions which are implicated in Ā-binding relations but in which no movement is triggered (for reasons of space, I will not here go in to the mechanisms which ensure this outcome; see McCloskey (2002)). It will therefore characteristically appear in contexts which require a binding relation for semantic well-formedness. The variable needed will often be supplied by a pronoun which happens to be within the clause headed by C-RP, since pronouns make good semantic variables. But there is, on this view, no essential relation between C-RP and the pronoun which supplies the variable.

If that variable can be supplied from some other source, then, we expect to observe clauses headed by C-RP which contain no pronoun. There are in fact, many such cases. Among the possibilities allowed is the one illustrated in the examples of (26), in which the variable which is critical for semantic well-formedness is supplied not by a pronoun but rather by the implicit argument of a relational noun (or a noun which can be coerced relatively easily into a relational interpretation):

(26) a. Chuir sé an cheist uirthi a raibh faitíos air roimh an bhfreagra.
put he the question on-her C-RP was fear on-him before the answer
'He put the question to her that he was afraid of the answer.' NGTTS 32

b. obair sheasta aige anois a raibh sé ag déanamh pá mhaith
work steady at-him now C-RP was he make PROG pay good
'and he now had steady work that he was making good pay' IM 123

c. Seo é an fadcheirnín a-r hiarradh ormsa focla na n-amhrán
this it the LP record C-RP.PAST ask.IMPERS.PAST on-me words the.GEN songs.GEN
a scríobh.
write.NON-FIN
'this is the LP that I was asked to write out the words of the songs' SOH 263

Crucially, in such cases the relation between the relative clause head and the variable it binds is not island-sensitive. In (27), the implicit variable associated with the noun *crann* (tree) and bound by the relative clause head *úll* (apples), is contained within a *Wh*-island (a cleft clause):

(27) faoi anam a raghadh ag priocadh úll nach ina ghairdín féin a
about soul C-FG go.COND PROG pick apples NEG-C-COP-PRES in-his garden self C-FG
d'fhás an crann
grow.PAST the tree
'about a spirit that would go picking apples that it wasn't in his own garden that the tree grew' NBN 162

For our purposes here, the most important consequence that should be highlighted is that we now expect no difference in status between the resumptive pronouns of English, say, and the resumptive pronouns of Irish. Neither language has in its grammar a dedicated resumption paragraph.

Their grammars do, of course, differ: in the inventory of C-elements that they possess. In English, all of the complementizers which appear in Ā-binding configurations are elements which force movement into their specifier positions. It follows that for a relative clause which includes a relativization site within an island (or in any position from which movement is impossible) there are no good outcomes. Derivations can either violate the selectional properties of the crucial complementizer, or they can include derivational steps which violate conditions on movement. All outcomes will be ill-formed to one degree or another, depending on the exact calculus by which degrees of ill-formedness for island violations are determined.

Irish will be different. If in a given structure the complementizer C-RP is deployed, movement will be impossible. But basic principles of semantic composition will require that there be a variable within the clause headed by C-RP. That variable may come in the syntactic guise of a pronoun, or it may be found in the implicit variable associated with relational or quasi-relational nouns, as in (26) and (27). On this view, there is no reason to expect any difference in status between English resumptive pronouns and Irish resumptive pronouns. Both should be subject to the prejudice which discriminates against Ā-bound pronouns and both should feel the small amelioration effects identified by Hofmeister and Norcliffe (2013) when they appear in positions or regions associated with heightened processing difficulty. The structures which contain those pronouns, however, will be crucially different in the two languages. An example of high object resumption in Irish (like (28), for instance, repeated from (10) above):

(28) jab a-r fhág an oiread sin oibrithe é
job C-RP.PAST leave.PAST so many workers it
'a job that so many workers left it' AT 138

will be assessed in the following way: it is syntactically and semantically fully well-formed, but in reacting to it, speakers will be aware that because of the anti-pronominal prejudice it should be infrequent (an awareness that their experience of the language will amply bear out), its rarity in this case un-modulated by the kinds of processing effects discussed earlier. The anti-pronominal prejudice, on this view, is a bias influencing the work done by the system of production, one that is finely tuned by the nuances of linguistic experience and finely sensitive to issues of processing pressure. The corresponding English example (the translation of (28)) will be syntactically ill-formed (because it fails to respect the requirements of the complementizer, which demands a movement) and will in addition be subject to the same anti-pronominal prejudice as the Irish example.

The distinctions drawn here are subtle, but they seem to make reasonable sense of the complex of observations so far accumulated. My own experience with native speaker consultants has been that they will respond to examples like (28) by saying something like: 'Well you COULD say it that way, but you probably wouldn't.' It's not clear to me how that reaction would be transduced into a number on a 7-point acceptability scale, but it may be that the reaction described here is the source of the small but significant effect detected for Hebrew by Meltzer-Asscher et al. (2015) and described in the following terms (p. 71):

> gapped versions of the sentences ... received higher ratings than the RP versions even when items were auditorily presented. Although the difference in ratings was small (∼0.5 points on a 7-point scale) it was consistent and reliable. This suggests that in general gaps and RP's are both acceptable, but that nonetheless the alternation between gaps and RP's is not completely free in Hebrew, as there is a slight preference for gaps.

Morgan and Wagers (2015), by contrast, replicate earlier studies in finding English structures containing resumptive pronouns, as we would now expect, 'highly unacceptable and nearly uniformly so across varying syntactic contexts'.

5. Conclusion

The principal empirical goal of this paper has been to clarify the status of Irish with respect to the supposed contrast between 'intrusive resumption' languages and 'true resumption' languages. The main conclusion that emerges is that, as far as the pronouns themselves are concerned, there is no difference between Irish and English. The relevant differences lie elsewhere, as we would expect – in the C-system. In making sense of the empirical landscape, issues about the interactions among grammaticality, frequency, production, and processing quickly come to the fore – in subtle and useful ways. The deepest mystery in all of this, though, it seems to me, is why there should be an anti-pronominal prejudice and why it should have such force. This is the question of section 2 and it remains un-answered.

References

Ackerman, Lauren, Michael Frazier, and Masaya Yoshida. 2014. Resumptive pronouns salvage island violations in forced choice tasks. Poster presented at the CUNY Sentence Processing Conference. Available for download at http://gradstudents.wcas.northwestern.edu/~lma777/research.html.

Acquaviva, Paolo. 2014. The categories of Modern Irish verbal inflection. *Journal of Linguistics* 50:537–586.

Aissen, Judith. 1999. Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory* 17:673–671.

Alexopoulou, Theodora. 2006. Resumption in relative clauses. *Natural Language and Linguistic Theory* 24:57–111.

Alexopoulou, Theodora. 2010. Truly intrusive: Resumptive pronouns in questions and relative clauses. in *Exploring the left periphery*, eds. Kleanthes Grohman and I. M. Tsimpli, 485–505. Special issue of *Lingua* 120.3.

Alexopoulou, Theodora, and Frank Keller. 2007. Locality, cyclicity and resumption: At the interface between the grammar and the human sentence processor. *Language* 83:110–160.

Anonymous. 1960. *Graiméar Gaeilge na mBráithre Críostaí*. Baile Átha Cliath: Mac an Ghoill Teo.

Ariel, Mira. 1999. Cognitive universals and linguistic conventions: The case of resumptive pronouns. *Studies in Language* 23:217–269.

Asudeh, Ash. 2004. Resumption as resource management. Doctoral dissertation, Stanford University.

Asudeh, Ash. 2012. *The logic of pronominal resumption*. Oxford, United Kingdom: Oxford University Press.

Bader, Markus, and Michael Meng. 1999. Subject-object ambiguities in German embedded clauses: An across the board comparison. *Journal of Psycholinguistic Research* 28:121–143.

Beltrama, Andrea, and Ming Xiang. 2016. Unacceptable but comprehensible: the facilitation effect of resumptive pronouns. *Glossa: A Journal of General Linguistics* 1:1–24.

Bennett, Ryan. 2008. English resumptive pronouns and the highest subject restriction. Manuscript, Yale University. Paper presented at the Trilateral Linguistics Weekend (TREND), UCSC, May 10th 2008.

Borer, Hagit. 1984. Restrictive relatives in Modern Hebrew. *Natural Language and Linguistic Theory* 2:219–260.

Borja, Manuel F., Sandra Chung, and Matthew Wagers. 2015. Constituent-order and parser control processes in Chamorro. in *AFLA 21: Proceedings of the Austronesian Formal Linguistics Association*, eds. Amber Camp, Yuko Otsuka, Claire Stabile, and Nozomi Tanaka, ANU Research Publications, 15–32. The Australian National University, Canberra, Australia: Asia-Pacific Linguistics.

Cann, Ronnie, Tami Kaplan, and Ruth Kempson. 2005. Data at the grammar-pragmatics interface: The case of resumptive pronouns in English. *Lingua* 115:1551–1577.

Chao, Wynn, and Peter Sells. 1983. On the interpretation of resumptive pronouns. in *NELS 13: Proceedings of the 13th Annual Meeting of the North-Eastern Linguistics Society*, eds. Peter Sells and Charles Jones, 47–61. Department of Linguistics, University of Massachusetts, Amherst: GLSA.

Chomsky, Noam. 1986. *Knowledge of language*. New York, New York: Praeger Publishers.

Clemens, Lauren, Gregory Scontras, and Maria Polinsky. 2012. English does not have resumptive pronouns: A cross-sentential account of English resumption. Paper presented at the 87th Annual Meeting of the Linguistic Society of America, Boston Massachusetts, January 3rd–6th 2012.

Crain, Steven, and Janet D. Fodor. 1985. How can grammars help parsers? in *Natural language parsing*, eds. David Dowty, Lauri Karttunen, and Arnold Zwicky, 94–128. Cambridge, United Kingdom: Cambridge University Press.

Doron, Edit. 1982. On the syntax and semantics of resumptive pronouns. in *Texas Linguistics Forum*, volume 19, 1–48. Austin, Texas: Department of Linguistics, University of Texas at Austin.

Duffield, Nigel. 1995. *Particles and projections in Irish syntax*. Boston, Massachusetts: Kluwer.

Erteschik-Shir, Nomi. 1992. Resumptive pronouns in islands. in *Island constraints: Theory, acquisition and processing*, eds. Helen Goodluck and Michael Rochemont, 89–108. Dordrecht: Kluwer Academic Publishers.

Farby, Shira, Gabi Danon, Joel Walters, and Michael Ben-Shachar. 2010. The acceptability of resumptive pronouns in Hebrew. in *IATL 26: Proceedings of the 26th Annual Meeting of the Israeli Assocation for Theoretical Linguistics*, ed. Y. N. Falk. Israeli Association for Theoretical Linguistics. Available at http://linguistics.huji.ac.il/IATL/26/TOC.html.

Ferreira, Fernanda, and Benjamin Swets. 2005. The production and comprehension of resumptive pronouns in relative clause 'island' contexts. in *Twenty-first century psycholinguistics: Four cornerstones*, ed. Anne Cutler, 263–278. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Frazier, Lyn. 1987. Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory* 5:519–610.

Gennari, Silvia, and Maryellen MacDonald. 2008. Semantic indeterminacy in object relative clauses. *Journal of Memory and Language* 58:161–187.

Gennari, Silvia P., and Maryellen C. MacDonald. 2009. Linking production and comprehension processes: The case of relative clauses. *Cognition* 111:161–187.

Georgopoulos, Carol. 1991. *Syntactic variables: Resumptive pronouns and $\bar{A}$-binding in Palauan*. Studies in Natural Language and Linguistic Theory 24. Dordrecht: Kluwer Academic Publishers.

Givón, Talmy. 1975. *On the role of perceptual clues in Hebrew relativization*, volume 2.8 of *Afroasiatic Linguistics*. Los Angeles, California: Undena Publications.

Han, Chung-hye, Noureddine Elouazizi, Christina Galeano, Emrah Gorgulu, Nancy Helberg, Jennifer Hinnell, Meghan Jeffrey, Kyeong min Kim, and Susannah Kirby. 2012. Processing strategies and resumptive pronouns in English. in *WCCFL 30: Proceedings of the 30th West Coast Conference on Formal Linguistics*, eds. Nathan Arnett and Ryan Bennett, 153–161. Cascadilla Proceedings Project, Somerville, Massachusetts: Cascadilla Press.

Hawkins, John A. 1994. *A performance theory of order and constituency*. Cambridge, United Kingdom: Cambridge University Press.

Hawkins, John A. 1999. Processing complexity and filler gap dependencies across grammars. *Language* 75:244–285.

Heestand, Dustin, Ming Xiang, and Maria Polinsky. 2009. Resumption still does not rescue islands. Manuscript, Harvard University and University of Victoria.

Hofmeister, Philip, Laura Staum Casasanto, and Ivan A. Sag. 2012. How do individual cognitive differences relate to acceptability judgments? A reply to Sprouse, Wagers and Phillips. *Language* 88:390–400.

Hofmeister, Philip, and Elizabeth Norcliffe. 2013. Does resumption facilitate sentence comprehension? in *The core and the periphery: Data-driven perspectives on syntax inspired by Ivan A. Sag*, eds. Philip Hofmeister and Elizabeth Norcliffe, 225–246. Stanford, California: CSLI Publications.

Hofmeister, Philip, and Ivan A. Sag. 2010. Cognitive constraints and island effects. *Language* 86:366–415.

Hopper, Paul J., and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56:251–299.

Hoyne, Mícheál. 2016. Structural ambiguity and resumptive pronouns: The pragmatics of the transitive 'direct' and 'indirect' relatives in Modern Irish. *Journal of Celtic Linguistics* 17:31–95.

Kaan, Edith. 1996. *Processing subject-object ambiguities in Dutch*, volume 20 of *Groningen Dissertations in Linguistics*. Groningen, the Netherlands: Rijksuniversiteit Groningen.

Keenan, Edward, and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8:63–99.

Keffala, Bethany. 2013. Resumption and gaps in English relative clauses: Relative acceptability creates an illusion of 'saving'. in BLS *37: Proceedings of the 37th Annual Meeting of the Berkeley Linguistics Society*, eds. Chundra Cathcart, I-Hsuan Chen, Greg Finley, Shinae Kang, Clare S. Sandy, and Elise Stickles, 140–154. Berkeley, California: Berkeley, California.

Kluender, Robert, and Marta Kutas. 1993. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8:573–633.

Kroch, Anthony. 1981. On the role of resumptive pronouns in amnestying island violations. in *CLS 17: Papers from the 17th Regional Meeting of the Chicago Linguistic Society*, 125–135. Chicago, Illinois: Chicago Linguistic Society.

Landau, Idan. 2009. *The locative syntax of experiencers*. Cambridge, Massachusetts: MIT Press.

Lee, Ming-Wei. 2004. Another look at the role of empty categories in sentence processing (and grammar). *Journal of Psycholinguistic Research* 33:51–73.

Levy, Roger, and Edward Gibson. 2013. Surprisal, the PDC, and the primary locus of processing difficulty in relative clauses. *Frontiers in Psychology* 4.

Lowder, Matthew W., and Peter C. Gordon. 2014. Effects of animacy and noun-phrase relatedness on the processing of complex sentences. *Memory and Cognition* 42:794–805.

Mac Giolla Phádraig, Brian. 1963. *Réchúrsa gramadaí*. Baile Átha Cliath: Brún agus Ua Nualláin.

Macdonald, Maryellen. 2013. How language production shapes language form and comprehension. *Fron-*

tiers in Psychology 4:Article 228.

Mak, Willem M., Wietske Vonk, and Herbert Schriefers. 2002. The influence of animacy on relative clause processing. *Journal of Memory and Language* 47:50–68.

Mak, Willem M., Wietske Vonk, and Herbert Schriefers. 2006. Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language* 54:466–490.

Maki, Hideki, and Dónall P. Ó Baoill. 2011. *Essays on Irish syntax*. Bunkyo-ku, Tokyo: Kaitakusha Co. Ltd.

Maling, Joan, and Annie Zaenen. 1982. Resumptive pronouns in Swedish. in *Readings on unbounded dependencies in Scandinavian languages*, eds. Elisabet Engdahl and Eva Ejerhed, 223–230. Stockholm, Sweden: Almqvist and Wiksell.

McCloskey, James. 1985. The Modern Irish double relative and syntactic binding. *Ériu* 36:45–84.

McCloskey, James. 1990/2011. Resumptive pronouns, Ā-binding and levels of representation in Irish. in *Syntax of the modern Celtic languages*, ed. Randall Hendrick, volume 23 of *Syntax and Semantics*, 199–248. New York and San Diego: Academic Press. Republished in Rouveret (2011), pp 65–119.

McCloskey, James. 1996a. On the scope of verb raising in Irish. *Natural Language and Linguistic Theory* 14:47–104.

McCloskey, James. 1996b. Subjects and subject-positions in Irish. in *The syntax of the Celtic languages: A comparative perspective*, eds. Robert Borsley and Ian Roberts, 241–283. Cambridge, United Kingdom: Cambridge University Press.

McCloskey, James. 2001. The morphosyntax of Wh-extraction in Irish. *Journal of Linguistics* 37:67–100.

McCloskey, James. 2002. Resumption, successive cyclicity, and the locality of operations. in *Derivation and explanation*, eds. Samuel Epstein and Daniel Seeley, 184–226. London and New York: Blackwell Publishers.

McCloskey, James. 2007. The grammar of autonomy in Irish. *Natural Language and Linguistic Theory,* 25:825–857.

McCloskey, James. 2010. Impersonals in irish and beyond. in *Hypothesis a/hypothesis b: Linguistic explorations in honor of david m. perlmutter*, eds. Donna Gerdts, John C. Moore, and Maria Polinsky, volume 49 of *Current Studies in Linguistics*, 323–344. Cambridge, Massachusetts: MIT Press.

McCloskey, James. 2015. Resumption. in *The blackwell companion to syntax, revised edition*, eds. Martin Everaert and Henk van Riemsdijk, To appear. London and New Tork: Blackwell Publishers. Available at http://ohlone.ucsc.edu/~bloscaidh.

McCloskey, James, and Peter Sells. 1988. Control and A-chains in Modern Irish. *Natural Language and Linguistic Theory* 6:143–189.

McDaniel, Dana, and Wayne Cowart. 1999. Experimental evidence for a minimalist account of English resumptive pronouns. *Cognition* 70:B15–B24.

Meltzer-Asscher, Aya, Julie Fadlon, Kaya Goldstein, and Ariel Holan. 2015. Direct object resumption in Hebrew: How modality of presentation and relative clause position affect acceptability. *Lingua* 166:65–79.

Morgan, Adam Milton, and Matthew Wagers. 2015. English resumptive pronouns are more common where gaps are less acceptable. Manuscript, University of California San Diego and University of California Santa Cruz.

Nolan, Brian. 2006. The passives of Modern Irish. in *Passivization and typology: Form and function*, eds. Werner Abraham and Larisa Leisiö, 132–164. Amsterdam and Phildelphia: John Benjamins Publishing Company.

Oda, Kenji. 2012. Issues in the left periphery of Modern Irish. Doctoral Dissertation, University of Toronto.

Ó Dónaill, Éamon. 2008. *Gramadach gan stró*. Baile Átha Cliath: Gaelchultúr Teoranta, 3rd edition.

Omaki, Akira, Ellen Lau, Imogen Davidson White, Myles L. Dakan, Aaron Apple, and Colin Ohillips. 2015. Hyper-active gap filling. *Frontiers in Psychology* 6:1–18.

Ostrove, Jason. 2015. Allomorphy and locality in the Irish verbal complex. in *American International Morphology Meeting 3*. University of Massachusetts Amherst, Oct. 2nd-4th 2015.

Ostrove, Jason. 2016. Spreading, spanning and linear order in vocabulary insertion. Submitted; under review.

Prince, Ellen. 1990. Syntax and discourse: A look at resumptive pronouns. in *BLS 16: Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics Society: Parasession on the Legacy of Grice*, eds. Kira Hall, Jean-Pierre Koenig, Michael Meacham, Sondra Reinman, and Laurel A. Sutton, 482–497. Berkeley, California: Berkeley Linguistics Society.

Roland, Douglas, Frederic Dick, and Jeffrey Elman. 2007. Frequency of basic english grammatical structures: A corpus analysis. *Journal of Memory and Language* 57:348–379.

Ross, John R. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT. Published as *Infinite Syntax!* Norwood: New Jersey: Ablex (1986).

Rouveret, Alain, ed. 2011. *Resumptive pronouns at the interfaces*. Amsterdam, the Netherlands: John Benjamins Publishing Company.

Sharvit, Yael. 1999. Resumptive pronouns in relative clauses. *Natural Language and Linguistic Theory* 17:587–612.

Shlonsky, Ur. 1992. Resumptive pronouns as a last resort. *Linguistic Inquiry* 23:443–468.

Sichel, Ivy. 2014. Resumptive pronouns and competition. *Linguistic Inquiry* 45:655–693.

Sprouse, Jon, Matthew Wagers, and Colin Phillips. 2012a. A test of the relation between working memory capacity and syntactic island effects. *Language* 88:82–123.

Sprouse, Jon, Matthew Wagers, and Colin Phillips. 2012b. Working memory capacity and island effects: A reminder of the issues and facts. *Language* 88:401–407.

Stenson, Nancy. 1989. Irish autonomous impersonals. *Natural Language and Linguistic Theory* 7:379–406.

Stowe, Laurie. 1986. Parsing wh-constructions: Evidence for on-line gap location. *Language and Cognitive Processes* 1:227–246.

Traxler, M. J., R.K. Morris, and R. E. Seeley. 2002. Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language* 47:68–90.

Tucker, Matthew, Ali Idrissi, Jon Sprouse, and Diogo Almeida. 2016. Resumption ameliorates different islands differentially. Manuscript, Oakland University, Qatar University, NYU Abu Dhabi, and University of Connecticutt. Available at: http://matthew-tucker.github.io/files/papers/resumption-islansds-arabic-acceptability.pdf.

Wagers, Matthew, and Emily Pendleton. 2015. Structuring expectation: Licensing animacy in relative clause comprehension. in *WCCFL 33: Proceedings of the 33rd West Coast Conference on Formal Linguistics*, eds. Kyeong min Kim, Pocholo Umbal, Trevor Block, Queenie Chan, Tanie Cheng, Kelli Finney, Mara Katz, Sophie Nickel-Thompson, and Lisa Shorten, 29–46. Cascadilla Proceedings Project, Somerville, Massachusetts: Cascadilla Press.

Wagers, Matthew, and Colin Phillips. 2014. Going the distance: Memory and control processes in active dependency construction. *The Quarterly Journal of Experimental Psychology* 67:395–433.

Wanner, Eric, and Michael Maratsos. 1978. An ATN approach to comprehension. in *Linguistic theory and psychological reality*, eds. Joan Bresnan and Morris Halle, 119–161. Cambridge, Massachusetts: MIT Press.

Wasow, Thomas. 1972. Anaphoric relations in English. Doctoral dissertation, MIT.

UNSELECTED OBJECTS AND THE ARGUMENT/ADJUNCT DISTINCTION*

Farrell Ackerman
University of California, San Diego
fackerman@ucsd.edu

John Moore
University of California, San Diego
moorej@ucsd.edu

This paper examines locative elements in Thetagovela Moro, a Kordofanian language from Sudan's Nuba Mountains. We argue for a three-way typology of Moro locatives: selected arguments of basic predicates, locative applicative arguments, and unselected locative objects. In Moro these three classes of locatives exhibit identical syntactic behaviors; we argue that they are all syntactic objects. Hence, Moro instantiates a category of unselected objects. This system challenges the traditional argument/adjunct distinction, but fits nicely into a typology predicted by lexicalist proposals.

Introduction

Linguistic theories, in one manner or another, identify at least three independent statuses for syntactic elements: predicator valence slots, semantic roles, and syntactic roles. This is illustrated in (1):

(1)	Predicate valence	X	Y	Z
	Semantic roles	Sem_1	Sem_2	Sem_3
	Syntactic roles	Syn_1	Syn_2	Syn_3

In addition to the elements in (1), which are keyed to lexically selected elements, there is another distinction that has been both instructive and problematic in syntactic theory: the contrast between ARGUMENTS and ADJUNCTS.

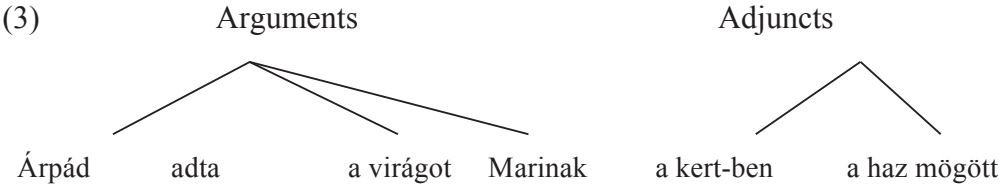
The basic divisions are exemplified for the simple Hungarian sentence in (2), where it is conventionally assumed that *ad* 'give' is a three place predicate with agent, theme and beneficiary semantic roles that are associated with the syntactic roles SUBJ, OBJ, INDIRECT OBJ respectively. These are specified as arguments of the predicate. In contrast, various criteria have been utilized to suggest that the two locative arguments in (2), *the yard* and *the house*, have a different status in clause, specifically, these are adjuncts.

* We thank Elyasir Julima, Ikhlas Elahmer, and Angelo Naser for their native-speaker insights with us. Thanks also to Sharon Rose, Peter Jenks, and other members of the UC San Diego Moro Project for discussions over the years. Finally, this work owes a great debt of gratitude to Sandra Chung, whose career of bringing empirically grounded work to bear on large theoretical questions has been inspirational.

(1) Árpád adta a virágot Marinak a kert-ben a haz mögött
Arpad-nom give-past-3sgOBJ the flower-acc Mary-dat the yard-inessive the house behind
'Arpad gave the flower to Mary in the yard behind the house.'

Thus, one can conventionally partition co-occurring elements into a set of arguments and a set of adjuncts as in (3), where the complements are SELECTED arguments of the predicate and where the adjuncts are UNSELECTED, but simply co-occur.

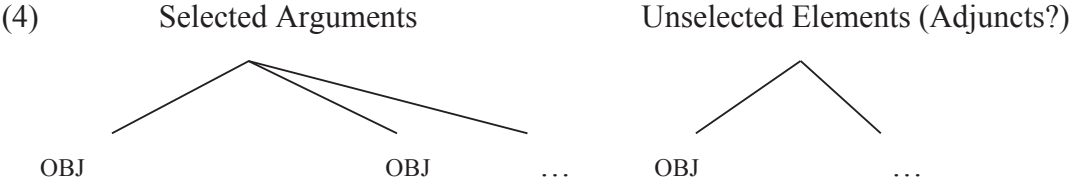
(3)



What is interesting to note is that arguments are assumed to bear core syntactic relations such as SUBJ, OBJ, and OBL(iques),[1] while adjuncts are not. For some languages, such as Hungarian, conventional criteria guiding such partitions often produce straightforward assignments, while the syntactic behaviors in other languages produce problems for such a sharp division.

In the present paper we build on previous work concerning multiple objects in Thetagovela Moro, a Kordofanian language from Sudan's Nuba Mountains. This language allows for several symmetrical selected objects in a single clause. Here we show that unselected elements – that is, constituents one might expect to be adjuncts, also exhibit the relevant object properties; hence, we argue that Moro exhibits UNSELECTED OBJECT phenomena:

(4)



We begin in section 2 by briefly reviewing some of the theoretical issues surrounding the argument/adjunct distinction and presenting a relevant typology of syntactic dependents. In section 3 we introduce relevant basic characteristics of Moro morphosyntax and empirically based arguments for the need to posit unselected objects. Section 5 revisits the typology from section 2 and discusses some of the theoretical implications.

2. The Argument/Adjunct Distinction

Culicover and Jackendoff's (2005: 173) articulation of the notion of ARGUMENT includes the following criteria:

[1] Such relations obtain irrespective of whether they assumed to be grammatical primitives or analogues associated configurational positions.

(5) a. Specified semantic roles, "intrinsically involved in the situation the verb denotes".
b. Expressed in the syntax (either obligatorily or optionally)
c. Aspects of the syntactic category, position, or morphological form potentially stipulated by the verb.

In other words, arguments are closely tied to aspects of a predicate's lexical semantics and may involve idiosyncratic syntactic specification. In contrast, adjuncts, which often denote general time, place, or manner aspects of an event are largely independent of lexically specific details. Another way to characterize this distinction is that arguments are SELECTED, while adjuncts are UNSELECTED.

Nevertheless, when considering the distinction between arguments and adjuncts it is important to keep in mind Dowty's cautionary observations:

> The distinction between complements and adjuncts has a long tradition in grammatical theory, and it is also included in some way or another in most current formal linguistic theories. But it is a highly vexed distinction for several reasons, one of which is that no diagnostic criteria have emerged that will reliably distinguish adjuncts from complements in all cases - too many examples seem to fall into the crack between the two categories, no matter how theorists wrestle with them. (Dowty 2003:34)

The more languages that have been examined over the years, the more evident it has become that a categorical distinction between arguments and adjuncts needs to be theoretically reconceptualized.[2] The essential insight in several lexicalist reconceptualizations is to assimilate argument and adjuncts into an integrative class, i.e., DEPENDENTS, and to permit dependents of both types to appear in the lexical representations of predicates. For example, Bouma, Malouf, and Sag 2001 propose, within an HPSG framework, an ARGUMENT STRUCTURE EXTENSION rule that combines selected elements (ARG-ST) with unselected ones ('adverbials') into a list of DEPENDENTS. While there is a lexical distinction between the two types of dependents, they are undistinguished syntactically. This ARGUMENT STRUCTURE EXTENSION rule is illustrated in (6):

(6) Argument Structure Extension:

$$verb \Rightarrow \begin{bmatrix} \text{ARG} - \text{ST} & \boxed{1} \\ \text{DEPS} & \boxed{1} \oplus list('adverbial') \end{bmatrix}$$

(Bouma, Malouf, and Sag 2001:12)

Somewhat intermediate between selected arguments and unselected adjuncts are applicative arguments. These often involve semantics normally associated with adjuncts, e.g. locatives, instrumentals, and beneficiaries, but behave as (often obligatory) selected arguments. Common

[2] See Przepiórkowski 2016, which provides a detailed summary of the most salient problems identified by other researchers and offers a solution within Lexical Functional Grammar based on the insights of Bouma, Sag, Malouf 2001 and Kathol, Przepiórkowski, and Tseng 2011.

in Bantu and other African languages, they are often encoded as objects. There is a large literature on applicatives in a variety of frameworks: e.g., Relational Grammar (Kimenyi 1980, Dryer 1983), LFG (Bresnan and Moshi 1990, Alsina and Mchombo 1993, Alsina 2001), and Principles and Parameters/Minimalism (Baker 1988, Pylkkänen 2008, McGinnis 2008, Citko 2011). A standard lexicalist approach to applicatives is to posit a valence-increasing MORPHOSEMANTIC rule that adds the applicative element to the predicate's argument structure.[3] Hence, applicative arguments are, in some sense, derived arguments. While these constructions have not been widely discussed in the HPSG literature, Ackerman, Malouf, and Moore 2017 proposes a benefactive applicative lexeme-to-lexeme rule that adds a new argument to the predicate's semantics:

(7) Benefactive Applicative Rule

$$\left[\text{SEM}\ \left[\begin{matrix}\text{INDEX} & e_1\\ \text{RESTR} & \boxed{1}\end{matrix}\right]\right] \Rightarrow \left[\text{SEM}\ \left[\begin{matrix}\text{INDEX} & e_1\\ \text{RESTR} & \left\{\left[\begin{matrix}ben\_rel & \\ \text{INST} & e_1\\ \text{BENEFCT} & b\end{matrix}\right]\right\} \cup\ \boxed{1}\end{matrix}\right]\right]$$

(Ackerman, Malouf, and Moore 2017:39)

Note that this lexeme-to-lexeme rule is distinct from the argument structure extension rule in (6); the rule in (7) has an effect on the semantically selected elements that eventually feed into the ARG-ST, while the rule (6) determines the dependents.

Putting this together, we end up with the following typology of dependents:

(8)a. Selected dependents (arguments)
  i. from the ARG-ST of basic predicates
  ii. added by lexeme-to-lexeme rules
b. Unselected dependents (adjuncts)

While the theory does not make a syntactic distinction between these elements, there may be factors such the grammatical hierarchy and idiosyncratic case that result in different behaviors among dependents (Maling 1993). In the remainder of this paper we present evidence for the typology in (8). In particular, we focus on locatives, as a useful test case; cross-linguistically, we find locatives as arguments of basic predicates, applicatives, and unselected adjuncts. We find locatives in each of these categories exhibiting identical syntactic behaviors in Moro.

[3] Sadler and Spencer 1998 distinguishes between MORPHOSYNTACTIC and MORPHOSEMANTIC lexical rules, where the latter involve a change in semantics, including valence changes. This distinction, under a variety of nomenclatures, has been discussed widely elsewhere, including Ackerman 1992.

3. Moro Morphosyntax and Unselected Objects

The syntax of the Kordofanian language Thetagovela Moro provides empirical challenges to a categorical division between complements and adjuncts. In addition to selected objects, both from basic predicates and applicatives, we argue for a third construct: unselected objects. Unlike traditional adjuncts, these exhibit syntactic characteristics of objects. However, we also will see how these object contrast with objects that are the result of morphosemantic operations. These objects are true objects that fall outside the lexical domain of the predicator – hence, they are unselected objects.

In recent research on Moro morphosyntax Ackerman, Malouf and Moore 2017 demonstrates that this language displays the characteristic properties associated with a multiple object language, i.e., simultaneous object behaviors for multiple co-occurring arguments in a single clause. In addition, as attested in several languages where neither morphological marking nor linear order has designated associated with either syntactic roles or semantic roles, there is pervasive ambiguity of semantic role for all object arguments. The symmetrical behaviors Moro objects poses a challenge for various theoretical proposals that entail asymmetrical syntactic representations among co-arguments (e.g. Stratal Uniqueness, Functional Uniqueness, and UTAH/binary branching; see Ackerman, Malouf, and Moore 2017, section 3 for discussion). The data in the following subsections are taken from Ackerman and Moore 2013 and Ackerman, Malouf, and Moore 2017.

3.1. Moro Objects

We begin by reviewing two basic characteristics of Moro objects: the ability to be referenced by object markers and the ability to passivize. Example (9) illustrates a simple transitive clause; in (10a-b) we see that *ɲogopájá* 'cups' can be referenced by an object marker and can passivize.[4]

(9) kúku g-a-ləvətʃ-ó ɲogopájá
CLg.Kuku CLg.SM-MAIN-hide-PFV CLɲ.cup
'Kuku hid the cups.'

(10) a. kúku g-a-ləvətʃ-ə́-**lo**
CLg.Kuku CLg..SM-MAIN-hide-PFV-**3PL.OM**
'Kuku hid them.'

b. ɲo ɲogopájá ɲ-ʌ-ləvətʃ-**ən**-ú
CLɲ.cup CLɲ.SM-main-hide-**PASS**-PFV
'The cups were hid.

[4] We use the following inter-linear glosses and conventions: SM 'subject marker', CL 'noun class', MAIN 'main clause verb', APL_{BEN} 'beneficiary applicative', APL_{LOC} 'locative applicative', PASS 'passive', PFV 'prefective', IMPFV 'imperfective', ITER 'iterative', LOC 'locative', and OM 'object marker' (1sg, 3sg, 3pl, etc.). High tone is marked with an acute accent; low tone is unmarked.

The example in (11) illustrate a ditransitive predicate. Note that (11) shows two bare noun objects (*óráŋ* 'man' and *ŋerá* 'girl') and the sentence is ambiguous; either object can serve as the theme or the goal.

(11) é-g-a-natʃ-ó óráŋ ŋerá
1SG.SM-CLg -MAIN-give-PFV CLg.man CLŋ.girl
'I gave the girl to the man.' / 'I gave the man to the girl.'

The examples in (12) provide evidence that the two internal arguments of *natʃ* 'give' are, indeed, objects; as indicated by the ambiguity, we see that either the theme or the goal may be referenced by an object marker and either can passivize. In other words, neither object markers nor passivization favor one argument over the other and, therefore, do not serve as a means of disambiguation.

(12) a. é-g-a-natʃ-ə́-**lo** ŋerá
1SG.SM-CLG-MAIN-give-PFV-**3PL.OM** CLŋ.girl
'I gave them to the girl' / 'I gave the girl to them.'
b. óráŋ g-ʌ-nʌtʃ-**ən**-ú ów:á
CLg.man CLg.SM-MAIN-give-**PASS**-PFV CLg.woman
'The man was given a woman.' / 'The man was given to a woman.'

Finally, (13) shows that both internal arguments can show simultaneous object properties; that is, one can passivize while the other is referenced by an object marker.[5] Again, note the ambiguity:

(13) óráŋ g-ʌ-nʌtʃ-**ən**-ə́-**ŋó** *object marking* cum *passivization*
CLg.man CLg.SM-MAIN-give-**PASS**-PFV-**3SG.OM**
'The man was given to her.' / 'She was given to the man.'

3.2. Beneficiary Applicatives

Applicative constructions represent a morphosemantic mechanism by which a predicate may increase its valence by promoting what might otherwise be represented as an adjunct to argument status. Typically, the valence increase is indicated with special verbal morphology. Moro has a beneficiary applicative construction, where a beneficiary argument is encoded as an object, with concurrent verbal morphology. In (14a) we see a beneficiary added to an intransitive predicate, resulting in a transitive construction; in (14b) the addition of a beneficiary to a transitive predicate yields a ditransitive construction (which, again, is ambiguous):

[5] By exhibiting simultaneous object properties as in (11), Moro would be classified as a true SYMMETRIC language, according to the typology discussed in Harford 1991 and Alsina 1996 and 2001. This contrasts with their ALTERNATING ASYMMETRIC languages, where either argument may exhibit object properties, but the two may not do so simultaneously.

(14) a. í-g-ʌləŋ-**ə̱ṯ-**ú ów:á
1SG.SM-CLg-sing-**APL$_{BEN}$**-PFV CLg.woman
'I sang for the woman.'

b. k-ʌ-w:ʌð-i̱ṯ-ú ŋerá um:iə
CLg.SM-MAIN-find-**APL$_{BEN}$**-PFV CLg.girl CLg.boy
'He found the boy for the girl.' / 'He found the girl for the boy.'

(14b) behaves just like other ditransitives – the two internal arguments show all the relevant object properties, including the symmetric simultaneous object properties:

(15) í-g-ʌ-w:ʌð-**itʃ-ən**-ə́-**lo**
1SG.SM-CLg-MAIN-found-**APL$_{BEN}$-PASS**-PFV-**3PL.OM**
'I was found for them.' / 'They were found for me.'

Because beneficiary applicatives result from a valence increasing operation, they do not involve unselected objects. The beneficiary objects are made part of the argument structure by the applicative morphosemantic mechanism. In other words, semantic roles that might otherwise be represented as adjuncts are promoted to argument status.

3.3. Locative Objects

This section will discuss two types of locative objects in Moro and argue that one of these represents an unselected object; that is, an element that is a syntactic object argument, but not part of the predicator's lexical semantics.

We first consider a selected locative argument—one that corresponds to a semantic argument. In (16) we see the three-place verb *ʌ́kk* 'put' that requires both theme and locative arguments, where the latter bears the locative prefix *i-*:

(16) a. k-ʌ-v-ʌ́kk-ʌg-iə eða **í-ðə́dí**
CLg.SM-MAIN-ʔ-ITER-cut-IMPFV CLj.meat **LOC-CLð.hole**
'He is putting the meat in the hole.'

b. *k-ʌ-v-ʌ́kk-ʌg-iə eða
CLg.SM-MAIN-ʔ-ITER-cut-IMPFV CLj.meat
'He is putting the meat.'

Locative arguments behave as objects—they can be referenced by object markers and can passivize:[6]

(17) a. k-ʌ-v-ʌ́kk-ʌg-iə-**Ø-u** eða
CLg.SM-MAIN-ʔ-ITER-cut-IMPFV-**3SG.OM-LOC** CLj.meat
'He is putting the meat in it.'

[6] Note that the third singular object marker is zero.

b. **ðə́diə́** ð-ʌ-v-ʌ́kk-əg-**ən**-iə́-**u** eða
CLð.hole CLð.SM-MAIN-ʔ-ITER-put-**PASS**-IMPFV-**LOC** CLj.meat
'The hole is being put the meat into.'

In both examples in (17), the verb bears the locative *-u* suffix. It is clear that this is not a locative applicative morpheme, as the lexical semantics of the verb is unchanged; the locative argument is already entailed by the predicate. Rather, this seems to be necessary when the syntactic environment—either object markers or passivization—result in the loss of the locative prefix on the locative argument. In other words, this suffix indicates that there is a locative element that is not independently marked as such. Notice also that none of the examples in (16-16) is ambiguous. In (16a) the locative argument is marked with a locative prefix; in (17a-b) the verbal locative suffix indicates that either the object marker (17a) or the passivized subject (17b) is locative.

The example in (18) illustrates an optional locative element that is not related to the predicate's lexical semantics. As it is not part of the predicator's argument structure, it is UNSELECTED. As (18) shows, the locative is optional.

(18) k-a-kə́l-á oʇeá (í-lúgi)
CLg.SM-MAIN-cut-IMPFV CLg.branch LOC-CLl_{PL}.tree
'He is cutting the branches (from the tree).'

The following examples show that the locative exhibits the familiar object properties:

(19) a. k-a-kə́l-á-**l-u** oʇeá
CLg.SM-MAIN-cut-IMPFV-**3PL.OM-LOC** CLg.branch
'He is cutting the branches from them.'
b. **lugi** l-ʌ-kʌ́l-**n**-iə́-**u** oʇeá
CLl_{PL}-tree CLl_{PL}.SM-MAIN-cut-**PASS**-IMPFV-**LOC** branches
'The trees are being cut branches from.'

Thus, there is evidence that these unselected elements are, indeed, objects. Again, we see the *-u* locative suffix registering an otherwise unmarked locative element. The addition of this morphological material might suggest the following applicative analysis:

(20) Locative Applicative Hypothesis
- Locative elements, marked with the *í-* locative prefix are adjuncts.
- In order to show object properties (object markers or passivization), the verb first undergoes a LOCATIVE APPLICATIVE morphosemantic rule that incorporates a locative into the verb's argument structure, and marks the verb with the *–u* suffix.

There are several arguments against this hypothesis. First, we don't find the *-u* suffix unless the locative is an object marker or undergoes passivization. This contrasts with the beneficiary applicative we saw above, where the applicative morphology obligatorily co-occurs with a

beneficiary argument. Secondly, as noted above, the *-u* suffix attaches to *ʌ́kk* 'put' in (15) under exactly the same circumstances. Since this predicate already selects for a locative argument, its argument structure is not augmented by a morphosemantic operation. A third argument comes from the fact that Moro does have a locative applicative construction that crucially contrasts with the unselected object examples discussed here.

3.4. Locative Applicatives

Moro has a locative applicative construction, illustrated in (21), that behaves in a manner that is similar to the beneficiary applicative; the characteristics are summarized in (22).

(21) k-a-kə́l-**ą́ṯ**-a eða **í-k-úgi**
CLg.SM-MAIN-cut-**APL$_{LOC}$**-IMPFV CLj.meat **LOC-CLg.tree**
'He is cutting the meat in the tree.'

(22) Locative Applicatives
a. Add the verbal extension *-ą́ṯ*
b. Increase the valence of the predicate to include a locative argument
c. Have an effect on the aspectual semantics of the predicate (as described below)

These are, of course, characteristics of a morphosemantic rule: additional morphology corresponds to both argument structure augmentation and other semantic modification. Notice that unlike, the *-u* suffix, the locative morphology co-occurs with the locative element. In fact, (23) shows that if there is locative applicative morphology, the locative is obligatory, providing strong evidence for its argument status. Recall that the unselected locative in (18) is optional.

(23) *k-a-kə́l-**ą́ṯ**-a eða
CLg.SM-MAIN-cut-**APL$_{LOC}$**-IMPFV CLj.meat

Locative applicatives behave as objects; e.g. (24) shows they can undergo passivization:

(24) **ugi** k-ʌ-kəl-**ʌtʃ-in**-ú-**u** eða
CLg.tree CLg.SM-MAIN-cut-**APL$_{LOC}$-PASS**-IMPFV-**LOC** CLj.meat
'The tree was cut meat in.'

Notice that as the locative passivizes, and does not bear the locative prefix, the *-u* suffix is added to register an unmarked locative. The fact that this suffix co-occurs with the locative applicative morphology shows that they have distinct functions.

Finally, there are aspectual differences between locative applicatives and unselected locative objects. As we have seen, the addition of an unselected locative does not change the aspectual properties of the predicate. This is also true in (25), where a new locative prefix, *n-* 'on', is used:

(25) k-abə́tw-a n-alé̱ta
CLg.SM-climb-IMPFV LOC_{on}-CLj.wall
'He is climbing on the wall.' (He is simply climbing on the wall)

However, the corresponding example with a locative applicative results in an aspectual change; the activity described in (25) now becomes an achievement in (26):

(26) k-abə́dw-a̱t-a n-alé̱ta
CLg.SM-climb-APL_{LOC}-IMPFV LOC_{on}-CLj.wall
'He is about to climb the wall.'
(e.g., he will clamber over up the wall, e.g. to avoid danger)

Again, the sematic change in (26) is typical of a morphosemantic operation, while the lack of any aspectual change is consistent with an unselected locative in (25).

4. Conclusion

In the previous section we have found three distinct types of locative objects:

(27) a. Locative arguments – selected by the predicate
b. Locative applicative arguments – selected by the predicate as a consequence of an applicative morphosemantic operation
c. Unselected locative objects

Note that none of these fit the usual conception of an adjunct – even in the last case (27c) the locative exhibits the same syntactic behaviors as the locative arguments. The typology in (27) parallels exactly the theoretical paradigm in (8), repeated here:

(28) a. Selected dependents (arguments)
i. from the ARG-ST of basic predicates
ii. added by lexeme-to-lexeme rules
b. Unselected dependents (adjuncts)

The framework that derives (28) involves lexical ARG-ST, ARG-ST modifications that result from morphosemantic lexeme-to-lexeme rules, and, an argument structure extension mechanism that creates a list of syntactically undifferentiated dependents from both the ARG-ST and, at least some, adjuncts. This last mechanism is crucial for accounting for the object behaviors of unselected objects in Moro.
Following previous proposals, Kathol, Przepiórkowski, and Tseng 2011 note that the traditional differences between arguments and adjuncts may be accounted for by a grammatical hierarchy and idiosyncratic marking mechanisms. We clearly need some of this in Moro, as unless passivized or referenced by an object marker, Moro locatives bear a locative prefix, but, again, note this is true of both selected and unselected locatives. Furthermore, again following this

work, it is worth asking whether all unselected, adjunct-like elements become verbal dependents, or whether only some of them do. Moro, in addition to allowing unselected locative objects, allows for unselected instrumental objects (Ackerman and Moore 2013:99-101). Thus, following , Bouma, Malouf, and Sag 2001, we tentatively propose the following rule that adds locatives and instrumentals to a verb's dependents' list:

(29) Moro Argument Structure Extension:

$$verb \Rightarrow \begin{bmatrix} \text{ARG} - \text{ST} & \boxed{1} \\ \text{DEPS} & \boxed{1} \oplus list('locative/instrumental') \end{bmatrix}$$

In this light, we note that applicative constructions generally involve locative, instrumental, or benefactive elements. In (29), we see both unselected locative and instrumental objects. However, we have found no evidence of unselected benefactives, nor does Moro appear to have an instrumental applicative construction. Perhaps unselected objects are limited, cross-linguistically, to semantic elements that can, sometimes, be selected. We leave these and related questions to further inquiry.

References

Ackerman, Farrell. 1992. Complex predicates and morpholexical relatedness: Locative alternation in Hungarian. *Lexical Matters*, eds. Ivan A. Sag and Anna Szabolsci, 55–83. Stanford; CSLI Publications.

Ackerman, Farrell, Robert Malouf, and John Moore. 2017. Symmetrical objects in Moro: Challenges and solutions. *Journal of Lingusitics* 53: 3–50.

Ackerman, Farrell, and John Moore. 2013. Objects in Moro. *Nuba Mountain language studies*, eds. Thilo C. Schadeberg and Roger M. Blench, 83–104. Köln: Rüdiger Köppe Verlag.

Alsina, Alex. 1996. *The role of argument structure in grammar: Evidence from Romance*. Stanford: CSLI Publications.

Alsina, Alex. 2001. On the non-semantic nature of argument structure. *Language Sciences* 23: 355-389.

Alsina, Alex, and Sam A. Mchombo. 1993. Object asymmetries and the Chichewa applicative construction. *Theoretical aspects of Bantu grammar*, ed. Sam A. Mchombo, 17-45. Stanford: CSLI Publications.

Baker, Mark C. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

Bouma, Gosse, Robert Malouf, and Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theor*y 19: 1–65.

Bresnan, Joan, and Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21, 47-85.

Citko, Barbara. 2011. *Merge, move, and labels*. Cambridge: Cambridge University Press.

Culicover, Peter, and Ray Jackendoff. 2005. *Simpler Syntax*. Oxford: Oxford University Press.

Dowty, David. 2003. The dual analysis of adjuncts and complements in Categorial Grammar. *Modifying Adjuncts*, eds. Ewald Lang, Claudia Maienborn, and Catherine Fabricius-Hansen, 33-66. Berlin: Mouton de Gruyter.

Dryer, Matthew S. 1986. Primary objects, secondary objects, and antidative. *Language* 62: 808-45.

Harford, Carolyn. 1991. *Object asymmetries in Kitharaka*. *Berkeley Linguistic Society* 17, *Special session on African language structures,* 98-105. Berkeley Linguistic Society, Berkeley, Ca.

Kathol Andreas, Adam Przepiórkowski, and Jesse Tseng. 2011. Advanced topics in Head-Driven Phrase Structure Grammar. *Non-transformational syntax: Formal and explicit models of grammar*, eds. Robert D. Borsley and Kersti Börjars, 54-111. Oxford: Blackwell.

Kimenyi, Alexandre. 1980. *A relational grammar of Kinyarwanda*. University of California Publications in Linguistics, vol. 91. Berkeley and Los Angeles: University of California Press.

Maling, Joan. 1993. Of nominative and accusative: the hierarchical assignment of grammatical case in Finnish. *Case and Other Functional Categories in Finnish Syntax*, eds. Anders Holmberg and Urpo Nikanne, 51–76. Dordrecht: Mouton.

McGinnis, Martha. 2008. Applicatives. *Language and Linguistics Compass* 2/6: 1225–1245.

Przepiórkowski, Adam. 2016. How not to distinguish arguments from adjuncts in LFG. *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar,* 560-580. CSLI Publications, Stanford.

Pylkkänen, Liina. 2008. *Introducing arguments*. Cambridge Ma: MIT Press, Cambridge.

Sadler, Louisa, and Andrew Spencer. 1998. Morphology and argument structure. *The Handbook of Morphology*, eds. Andrew Spencer, 206-235. Oxford: Blackwell.

PERSONS, PRONOUNS AND PROCESSING ASYMMETRIES[*]

Scarlett Clothier-Goldschmidt & Matt Wagers
University of California, Santa Cruz
scarlettclothier@gmail.com, mwagers@ucsc.edu

Across languages of many types, object relative clauses are characteristically more difficult to understand than subject relative clauses. This S/O asymmetry has been linked to the Accessibility Hierarchy, suggesting a relationship between typological universals or tendencies and online processing complexity in a single language. Here we consider whether the *3>local pronoun constraint, which regulates how semantic arguments are mapped to syntactic positions in many languages, such as Chamorro, can explain the fact that pronouns in subject position often neutralize the SRC/ORC asymmetry.

1. Introduction

1.1. Subject relative clauses: a universal advantage?

The Subject/Object asymmetry in the processing of relative clauses is a classic finding in linguistics. Relative clauses (RCs) with a subject gap, like (1a), are easier to understand than relative clauses with an object gap, like (1b) (Staub, 2010); and subject-gap containing RCs are generally more common (Roland, Dick & Elman, 2007)[1].

(1) a. *Subject relative clause* (SRC)
The healer [who ___ assisted the mayor] received a *bayogu* necklace.

b. *Object relative clause* (ORC)
The healer [who the mayor assisted ___] received a *bayogu* necklace.

The enduring interest in the RC processing asymmetry is grounded, in part, in the fact that it somehow seems linked to a typological universal – namely, the Accessibility Hierarchy (AH) of Keenan & Comrie (1977). This implicational universal, schematized in (2), states that: if a

[*] This paper was adapted from Chapter 2 of SCG's M.A. Thesis (Clothier-Goldschmidt, 2015). The authors thank Adrian Brasoveanu, Pranav Anand, Junko Ito, and the members of the Winter 2015 Linguistics Research Seminar for their feedback throughout the development of this project; the editors of this volume for their feedback; and Chelsea Miller for her technical expertise. And of course we offer our endless gratitude to Sandy Chung for the way in which her teaching and collegiality have inspired, broadened and elevated our own research. *Sen dångkulu na si Yu'us ma'åsi'!*

[1] This claim is true across registers when we consider ordinary transitive RCs with an overt complementizer or relative pronoun. However in two spoken corpora, Switchboard and the British National Corpus, the balance of gap types is more even, with object gaps slightly dominating: this is related to the fact that RCs with null complementizers are more common in speech (Roland, Dick & Elman, 2007) as is the use of local pronominal subjects inside the RC (Fox, 1987).

language can relativize a DP occupying a given position on the hierarchy, then it can relativize a DP in all of the positions which are higher on the hierarchy.

(2) SUBJECT > DIRECT OBJECT > INDIRECT OBJECT > OBLIQUE > GENITIVES > OBLIQUE COMPLEMENTS

Keenan & Comrie proposed that AH "directly reflects the psychological ease of comprehension" (1977:88), and offered some speculation as to why comprehending subject relatives (SRCs) should be easier than comprehending object relatives (ORCs). They proposed that the 'recognition strategy' for relative clauses was essentially grounded in the distribution of how arguments are linked to syntactic positions: lexical predicates almost always require a subject, often require a direct object, sometimes require an indirect object, rarely an oblique, etc., etc. On the basis of this very coarse but universal distribution, comprehenders would do best to assume – by default – that a relativized argument is linked to a RC-internal subject position. In the absence of other evidence, it is the most likely hypothesis to be true.

Another line of thought has linked the relative difficulty of ORCs to working memory. In English, the RC subject intervenes between the head of the RC (or filler) and its gap. If maintaining constituents that are in some sense unintegrated is costly (Wanner & Maratsos, 1978, Gibson, 2000) or, if crossing other constituents promotes interference (Van Dyke & Lewis, 2003), then this would provide a footing for explaining the S/O asymmetry. Correspondingly, Hsiao & Gibson (2003) argued that languages whose word order removes the subject as an intervener should not show the S/O asymmetry – or, if the object intervenes between the filler and a subject gap, the asymmetry should be reversed. Mandarin Chinese, which combines RC-Filler modification order with S-V-O clausal word order, provides such a test case. Hsiao & Gibson's results suggested that ORCs were actually easier in Chinese. A flurry of research since then has problematized this claim, and the balance of evidence now suggests that Chinese too shows an advantage for SRCs (Vasishth et al., 2013, Yun et al., 2015).

The Austronesian language Chamorro has relative clauses that follow the filler, like English, and relative clauses that precede the filler, like Chinese. Compare (3a), a **Filler**-[RC] sentence, with (3b), an [RC]-**Filler** sentence.

(3) *Relative clauses in Chamorro can follow or precede their filler*

a. Impottånti esti na infotmasion put i **hanum** [ni un gigimin __].
important this L information about the water C AGR be.drinking
'This information about the water that you drink is important.'

b. Estagui' i risuttan i CCR … put i [un gigimin __] na **hånum**.
here.is the result.L the CCR about the AGR be.drinking L water
'Here are the results of the CCR...about the water that you drink.'

(Both examples: Commonwealth Utility News, July 2014, p. 8)

Interestingly, ambiguous relative clauses that precede the filler do show a slight preference to be interpreted as ORCs[2]. But although the SRC interpretation is in the minority, it is also the interpretation that arises the fastest, as judged by finger-swipe trajectories in a picture-matching task on tablet computers (Borja, Chung & Wagers, 2015). Thus, even where non-subject relatives may ultimately be easier to process, the 'fingerprints' of a subject advantage can be detected.

The subject advantage in RCs appears to be universal – at least, it appears to reflect a universal pressure that often but not always surfaces as an overall advantage for SRCs (Wagers, Borja, Chung, 2016). One remaining area of controversy surrounds languages with ergative-absolutive alignment, although at least some of these languages give evidence for the primacy of subject gaps (see Clemens, et al., 2015, for discussion; and Fox, 1987, for a related challenge in English).

1.2. Grammatical person and the S/O asymmetry

Fox (1987) analyzed the ORCs contained in a small corpus of relative clauses (n=100) drawn from conversational language and found that they almost all contained pronominal subjects, as in (4).

(4) Have you heard about the orgy [we had __ the other night]? (SN4:5, in Fox, 1987:857)

And relatedly, in language processing studies (Gordon, Hendrick & Johnson, 2001, Warren & Gibson, 2005), it has been found that the usual markers of the S/O asymmetry – increased reading times and decreased comprehension accuracy for ORCs – are neutralized when the subject is a local pronoun. Gordon et al. (2001) compared sentences like (5a), with the local pronoun *you* in object position, to the (5b) with the same pronoun in subject position. In a self-paced reading experiment, they found no reading time differences at critical regions inside the RC *or* the post-RC main verb ('climbed'); there were also no comprehension accuracy differences. This comparison stands in contrast to a sentence pair in which the RC-internal argument is another definite description (like, 'the barber').

(5) a. The banker [that ___ praised *you*] climbed the mountain.
b. The banker [that *you* praised ___] climbed the mountain.

Gordon et al. proposed a memory-based explanation for the SRC/ORC asymmetry when it arises: namely, ORCs are more difficult than SRCs to the extent that they require "intermediate representations to be held in memory and addressed during comprehension" (2001:10). Factors that affect the encoding, storage and retrieval of those intermediate representations thus affect the relative difficulty of the ORC compared to the SRC. In particular, similarity between the two DPs whose syntactic position must be stored leads to greater difficulty or lossiness in encoding and retrieval (Van Dyke & Lewis, 2003). The local pronoun *you* is more distinct from *banker* than is

[2] Chamorro RCs with transitive verbs showing ordinary transitive agreement – not Wh-Agreement (Chung, 1998) – are only ambiguous with respect to the gap location. They are *unambiguously* RCs of some kind. In contrast, irrelevant incremental ambiguities in the processing of some Chinese RCs has muddied the empirical waters there somewhat (see Yun et al., 2015 for a review and more detail)

barber and, as a consequence, ORCs are not appreciably more difficult than SRCs if their subject position if filled with the local pronoun (or, as was demonstrated in other experiments, a name or quantified DP). Bever (1974) had made a related observation about center self-embedding: when the 3 DPs in a center self-embedding sentences are of different "surface lexical type", the sentence is intuitively easier to comprehend. Compare (6a), with three DPs of the same type, versus (6b), with three DPs that are differentiated.

(6) a. The reporter [the politician [the commentator met ___] trusts ___] said the president won't resign.
b. The reporter [everyone [I met ___] trusts ___] said the president won't resign.

Gordon et al. (2001) observed that there are many syntactic, semantic and pragmatic dimensions along which a full DP like *the banker* and the local pronoun *you* differ. In subsequent research (Gordon, Hendrick & Johnson, 2004), other dimensions of dissimilarity were observed that had *no* effect: these include whether or not the two DPs matched in definiteness, genericity, or richness (e.g., a superordinate description *the person* compared to a more specific one like *the barber*). Warren & Gibson (2005) included both first and second person pronouns in their study on processing of object clefts. They found that, consistent with previous observations, object clefts with a pronoun in subject position were processed faster than those with definite DPs in that position.

What remains unclear from the existing literature is whether it is crucial that it be a local person which neutralizes the S/O asymmetry. In Dependency Locality Theory (Gibson, 2000) there is a special status accorded to the indexical pronouns: in some sense, they come "for free" in language processing. In Gordon et al.'s interference based theory, it is plausible (though not necessary) that third person pronouns would cause greater difficulty than first or second person pronouns, because they overlap in their person value with the filler.

A final consideration comes from typology. Many languages effect an alignment between person/animacy and grammatical role (Aissen, 1999). Chamorro is one such language: it requires the subject of a transitive clause to not be outranked by the object on the Person-Animacy Hierarchy (Chung, 1998, 2014):

(7) 2.PERS > 3.PERS ANIMATE *pro* > ANIM > INANIM

For example, a transitive clause cannot have 2.PERS object if its subject is 3.PERS:

(8) *Kao ha kuentusi **hao** åntis di u hånao? *3>2
Q AGR:3.SG speak_to you before AGR:3.SG go
'Did he speak to you before he left?' (Chung, 2014)

(8) needs to be expressed using another form of agreement, like Wh-Agreement, or by passivizing, as in (9).

(9) Kao kuinentusi **hao** åntis di u hånao?

Q PASS.speak_to you before AGR:3.SG go
'Were you spoken to by him before he left?' (Chung, 2014)

As Clothier-Goldschmidt (2015) demonstrates, the sub-constraint *3>2 is never violated in connected Chamorro text. Aissen (1999), following Silverstein (1976), observes that a {1,2}>3 hierarchy is active in many languages and it controls how arguments are mapped to syntactic positions, via case systems, voice systems, etc. Chamorro, though, is somewhat unusual in that first person is not on the Person-Animacy Hierarchy.

Suppose that there is a softened realization of the *3>2 constraint (or *3>local) in English. If local person is in some sense a more optimal subject than third person, then it is possible that the S/O asymmetry is exacerbated when a less optimal subject is present – such as a definite DP. Bresnan et al. (2001) argue that inviolable constraints in some languages may show up as statistical pressures in others. Lummi (Xwlemi'chosen), a Salish language of Washington, exemplifies a "hard" version of the *3>local constraint: in transitive clauses with both a local person and a 3.PERS DP, the local person DP must be in subject position. Like Chamorro (9) v. (8), if a local person is the theme of a transitive verb, the clause must be passivized. Bresnan et al. (2001) argue we can see the same process expressed gradiently in English passivization rates: when there is a 3.PERS agent, there are more passives for local person themes than for 3.PERS themes.

In summary, three different lines of thought predict that the S/O asymmetry should be stronger when the subject is 3.PERS: Dependency Locality Theory (Gibson, 2000), a version of similarity-based interference theory in which grammatical person counts as a dimension of similarity (Gordon et al., 2001, Van Dyke & Lewis, 2003), and a theory grounded in person/animacy hierarchies (Aissen, 1999, Bresnan et al., 2001, Chung, 2014, cf. Kaplan, 2002). We tested this prediction in a self-paced reading experiment contrasting RCs with subject and object gaps, and varying whether or not the RC subject is a 2.PERS pronoun, a 3.PERS pronoun or a definite description.

What we discovered is that both 2.PERS and 3.PERS pronouns patterned alike in the sense that 3.PERS pronouns did not lead to a greater S/O asymmetry than 2.PERS pronouns. In section 2, we report the details of our experiment and, in section 3 we consider the implications of our failure to find a difference between local and non-local pronouns.

2. Experiment

2.1. Materials, Methods and Participants

We created relative-clause containing sentences that systematically varied the GAP SITE (S, O) and RC-internal DP TYPE (*Full DP, 3.PERS, 2.PERS)* in a 2 × 3 design. (10-12) illustrates a sample item set, with the RC-internal DP in bold font. (10-12)(a) show S-gap conditions and (10-12)(b) O-gap.

(10) DP TYPE: *Full DP*
- a. the nurse that ___ welcomed **the mailman** with a smile ran a marathon during the month of July.
- b. the nurse that **the mailman** welcomed ___ with a smile ran a marathon during the month of July.

(11) DP TYPE: *3.PERS*

a. the nurse that ___ welcomed **him** with a smile ran a marathon …
b. the nurse that **he** welcomed ___ with a smile ran a marathon …

(12) DP TYPE: 2.*PERS*

a. the nurse that ___ welcomed **you** with a smile ran a marathon …
b. the nurse that **you** welcomed ___ with a smile ran a marathon …

Relative clauses were always attached to a sentence-initial 2-word subject. Each target sentence was preceded by a screen containing the following transition sentence: "Your friend John/Mary tells you that …" The intent of the transition sentence was both to introduce a potential antecedent for 3.PERS pronouns and to establish a dialogic frame of reference in which 2.PERS was felicitous.

The grammatical gender of the pronoun in 3.PERS conditions was selected to contrast with the stereotypical gender of the RC head, as determined by norms collected in Santa Cruz, CA, for previous research. In doing this, we hoped to minimize any potential interference from the RC head and the pronoun that was not related to Person. When the pronoun was masculine, the transition sentence contained a gender-matching name, 'John', and when it was feminine, 'Mary.' Thus, for (10-12): the stereotypical gender of 'nurse' was feminine, the grammatical gender in 3.PERS conditions was masculine, and the name in the transition sentence was therefore 'John.'

24 item sets were created, combined with 72 fillers and assigned to lists according to a Latin Square. The experiment was administered via the web on Ibex (Drummond, 2016). The transition sentence was presented unmasked on one screen, and participants advanced to the next screen to begin reading the (uncapitalized) target sentence in moving-window, self-paced presentation (Just, Carpenter & Woolley, 1981).

Following each sentence, participants had to answer a comprehension question. Comprehension questions were formulated so that, across the entire experiment, they probed uniformly about the thematic roles associated with the matrix verb, the roles associated with the embedded verb, and any modifiers in the sentences. We thus hoped to limit participants from developing a strategy to engage systematically in shallow processing in some portions of the sentence and not others. The expected correct answer was evenly balanced across trials. For example, the comprehension question for (11) was: "Was it with a smile that the nurse welcomed him?" (expected answer: *yes*). Participants were given feedback on the incorrect trials.

There were 41 participants in the study, aged 18 to 60. Eighteen of these participants were recruited as volunteers through Facebook by the first author; and the rest were given course credit at UCSC for their participation.

2.2. Results

We found that object relative clauses were harder to process than subject relative clauses, but only when the RC subject was a full DP, like 'the mailman'. Crucially, there were no differences between 3.PERS and 2.PERS pronouns.

Reading times across each sentence type are reported in Figure 1. Inside the RC, we compared reading times at the verb across gap types. For full subject RCs, the verb was read approximately 44 ms faster for subject gap sentences compared to object gap sentences (± 20 ms s.e.; $p < .05$). In contrast, for both pronoun conditions, there were no significant differences. For 3.PERS RCs, the verb was actually read numerically slower (-5 ms ± 23 ms) for S-gap sentences compared to O-gap sentences; for 2.PERS RCs, there was a numerical S-gap advantage (15 ms ± 15 ms).

We estimated a linear mixed-effects model, crossing factors of GAP TYPE and DP TYPE, and nesting them within both subjects and items. The RC type factor was Helmert coded: one coefficient, DP TYPE:FORM, contrasted RTs for Full DPs with the average RTs for both pronoun conditions; a second coefficient, DP TYPE:PERSON, contrasted RTs for 2.PERS and 3.PERS pronouns directly. We found no statistically-reliable main effects, but the interaction between GAP TYPE and DP TYPE:FORM was significant ($t = 2.1$, $p < .05$). The RT slow-down for object gaps compared to subject gaps was greater when the RC subject was a Full DP. The DP TYPE:PERSON coefficient was trending marginally toward a crossing interaction ($t = -1.9$, $p < .10$), i.e., the object gap > subject gap asymmetry was reversed for 3.PERS conditions.

When we turned to comprehension question accuracy (Table 1), we found a similar pattern. Comprehension questions were answered less accurately for object gap sentences, but only when there was a full DP subject. A binomial mixed-effects regression, using the same contrast structure as described above, revealed 3 significant effects: firstly, conditions with RC-internal pronouns had greater accuracy than those with a full DP ($z = 3.1$, $p < .005$); secondly, conditions with 2.PERS pronouns had greater accuracy than those with 3.PERS pronouns ($z = 3.0$, $p < .005$); crucially, there was a 2-way interaction between GAP TYPE and DP TYPE:FORM ($z = 2.6$, $p < .001$). This interaction seems to be driven both by the subject advantage for RC-internal full DPs, but also an apparent slight object advantage in 3.PERS conditions – just as we observed in the RT data. However, we inspected 2 pair-wise comparisons on GAP TYPE: for full DP conditions alone, and for 3.PERS conditions alone. Only the subject>object difference in full DP conditions was significant ($z = -2.4$, $p < .05$).

	DP TYPE			
GAP TYPE	DP	2.PERS	3.PERS	
Subject	80	87	75	80
Object	68	88	82	79
	74	87	79	80

Table 1 **Comprehension question accuracy shows a subject gap advantage, but only when the RC subject is a full DPs.** N≅155 per cell.

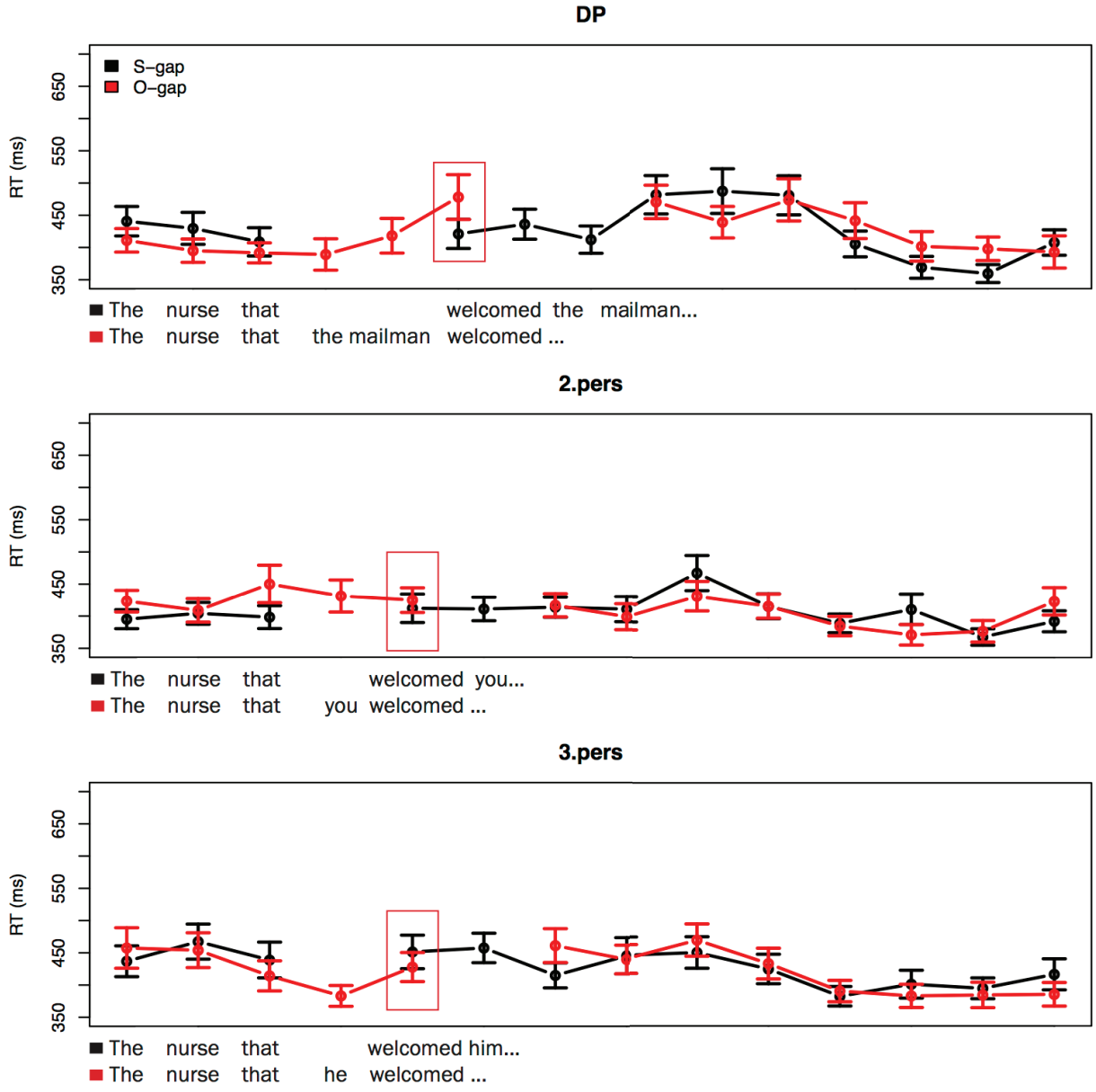


Figure 1 **Object relative clauses are only harder than subject relative clauses when the RC subject is a full DP.** Average reading times, in milliseconds, are given per RC subject type, gap type and sentence region. Error bars indicate standard error of the mean. In each panel, the boxed symbols represent reading times at the critical RC internal verb.

3. Discussion

Our data most straightforwardly support the view that it is the dissimilarity in surface form between the filler and the RC subject that ameliorates the S/O asymmetry. On the assumption that pronouns belong to the same syntactic category – e.g., intransitive determiners – all pronouns should pattern alike. We find no evidence that object relative clauses whose subject is a 3.PERS pronoun are harder to understand than object relative clauses whose subject is a 2.PERS pronouns. If 2.PERS pronouns are more optimal subjects than 3.PERS pronouns, as suggested by a person/animacy hierarchy hypothesis, then this is not reflected in the degree to which they ameliorate the subject>object gap asymmetry.

Because it makes the same predictions, we also fail to support the similarity-based interference hypothesis: 3.PERS pronouns overlap in more features with the RC head than do 2.PERS pronouns. The similarity-based interference hypothesis is powerful enough that it may be appropriately modified to be insensitive to person features, by down-weighting their activation in either an encoding or retrieval structure. This is not *a priori* an especially attractive theoretical move since Person is a grammatically-active feature. Number, for example, is especially prone to cause interference (Wagers, Lau & Phillips, 2009) – so it seems plausible that another agreement-controlling feature should do the same. However, our sentences may not have had the right kind of visible agreement to trigger reference to person features, as all our verbs were in the past tense. Consequently, it may be that Person interference was not induced.

What does any of this say about the suggestion of Bresnan et al. (2001) that inviolable constraints in some languages – like Chamorro's *3>2 constraint – might be seen to have a weak effect in other languages? As with the retreat of the similarity-based interference theory, it may be that the effect is simply too weak to detect. It is worth noting, however, that other experiments on RC processing provide support for the effect of an animate > inanimate hierarchy: inanimate fillers, when paired with animate RC subjects, also neutralize the S/O asymmetry (see Wagers & Pendleton, 2015). Clothier-Goldschmidt (2015) considered whether it were possible to broaden Bresnan et al.'s prediction to other Person-Animacy Hierarchy violations. In the process of creating a Chamorro-English parallel corpus based on the New American Standard bible, she classified the subject and object arguments from 461 transitive clauses according to person, animacy and pronominality. She then analyzed the joint distribution of subject and object features to determine whether the individual subject and object feature sets could be modeled as independent random variables: in a language with a (soft) Person-Animacy Hierarchy, it is predicted this will not be possible because the allowable features on the object depend on the subject. What she found was only a modest departure from independence, and one which ran counter the existence of a soft Person-Animacy Hierarchy in English: there were relatively more Person-Animacy violating (English) clauses than would be predicted by knowing the subject and object distributions alone. In particular, there were *more* 2.PERS objects than there ought to be have been (which might plausibly relate to the epistolary genre of a number of the books in the New Testament).

Finally, it is interesting to speculate whether there is actually an object *advantage* in processing RCs with third person pronominal subjects: we see hints of this both in reading times and in the comprehension data. It would be necessary to establish, firstly, that this were a robust

effect. However, we suspect it may relate to our particular experimental set-up and not to third person in general. Recall that we attempted to simultaneously make a third person and second person pronoun felicitous by introducing the context sentence "Your friend John/Mary tells you that …". One possibility is that pronominal co-reference is more acceptable in the subject position than in the object position, perhaps for reasons of topic continuity – in which case, a countervailing pressure to avoid the third person pronoun in object position could account for detrimental performance in 3.PERS/S-gap conditions.

4. Appendix A: Supplementary materials

A repository of experimental items and aggregated data from the experiment may be downloaded from the second author's web site: http://people.ucsc.edu/mwagers.

References

Aissen, Judith. 1999. Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory,* 17:673-711.

Bever, Thomas G. 1974. The ascent of the specious, or there's a lot we don't know about mirrors. *Explaining linguistic phenomena,* ed. D. Cohen, 173-200. Washington, DC: Hemisphere.

Bresnan, Joan, Shipra Dingare, and Christopher D. Manning. 2001. Soft constraints mirror hard constraints: Voice and person in English and Lummi. *Proceedings of the LFG01 Conference,* ed. Miriam Butt and Tracy Holloway King, 13-32. Stanford: CSLI Publications.

Borja, Manuel F., Sandra Chung, and Matthew Wagers. 2015. Constituent order and parser control processes in Chamorro. *AFLA21: Proceedings of the 21st Meeting of the Austronesian Formal Linguistics Association,* ed. Amber Camp, Yuko Otsuka, Claire Stabile, and Nozomi Tanaka, 15-32. Canberra: Asia-Pacific Linguistics.

Chung, Sandra. 1998. *The design of agreement: evidence from Chamorro*. Chicago: University of Chicago Press.

Chung, Sandra. 2014. On reaching agreement late. *CLS 48, Vol. 1,* ed. Andrea Beltrama, et al, 169-190. Chicago: Chicago Linguistic Society.

Clemens, Lauren E., Jessica Coon, Pedro Mateo Pedro, Adam M. Morgan, Maria Polinsky, Gabrielle Tandet, and Matthew Wagers. 2015. Ergativity and the complexity of extraction: a view from Mayan. *Natural Language and Linguistic Theory,* 33:417-467.

Clothier-Goldschmidt, Scarlett. 2015. The distribution and processing of referential expressions: evidence from English and Chamorro. MA thesis, University of California, Santa Cruz.

Drummond, Alex. 2016. Ibex [Computer software]. Retrieved from https://github.com/addrummond/ibex.

Fox, Barbara A. 1987. The Noun Phrase Accessibility Hierarchy reinterpreted: subject primacy or the Absolutive Hypothesis. *Language*, 63:856-870.

Gibson, Edward. 2000. The dependency locality theory: a distance-based theory of linguistic complexity. *Image, language, brain,* ed. Yasushi Miyashita, Alec Marantz, & Wayne O'Neil, 95-126. Cambridge, MA: MIT Press.

Gordon, Peter C., Randall Hendrick, and Marcus Johnson. 2001. Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 27:1411–1423.

Gordon, Peter, Randall Hendrick and Marcus Johnson. 2004. Effects of noun phrase type on sentence complexity. *Journal of Memory and Language,* 51:97-114.

Just, Marcel A., Patricia A. Carpenter, and Jacqueline D. Woolley. (1982). Paradigms and Processes in Reading Comprehension. *Journal of Experimental Psychology: General*, 111:228-238.

Kaplan, Aaron F. 2002. Patterns of relativization and recent formulations of markedness. Undergraduate thesis, University of North Carolina, Chapel Hill.

Staub, Adrian. 2010. Eye movements and processing difficulty in object relative clauses. *Cognition* 116:71-86.

Roland, Douglas, Frederic Dick, and Jeffrey L. Elman. 2007. Frequency of basic English grammatical structures: a corpus analysis. *Journal of Memory and Language,* 57:348-379.

Van Dyke, Julie, and Richard Lewis. (2003). Distinguishing effects of structure and decay on attachment and repair: a cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language,* 49:285-316.

Wagers, Matthew W., and Emily Pendleton. (2015). Structure expectation: licensing animacy in relative clause comprehension. *Proceedings of the 33rd West Coast Conference on Formal Linguistics*, ed. Kyeong-min Kim et al., 29-46. Somerville, MA: Cascadilla.

Wagers, Matthew W., Ellen F. Lau, and Colin Phillips (2009). Agreement attraction in comprehension: representations and processes. *Journal of Memory and Language,* 61:206-237.

Wanner, Eric, and Michael Maratsos. (1978). An ATN approach to comprehension. *Linguistic Theory and Psychological Reality,* ed. Morris Halle, Joan Bresnan and George A. Miller, 119-161. Cambridge, MA: MIT Press.

Warren, Tess, and Edward Gibson. 2005. Effects of NP type in reading cleft sentences in English. *Language and Cognitive Processes,* 20:751-767.

Yun, Jiwon, Zhong Chen, Tim Hunter, John Whitman, and John Hale. 2015. Uncertainty in processing relative clauses across East Asian languages. *Journal of East Asian Linguistics,* 24:113-148.

ON REACHING AGREEMENT EARLY (AND LATE)*

Maziar Toosarvandani
University of California, Santa Cruz
mtoosarv@ucsc.edu

Many languages constrain how arguments may combine based on their person or animacy. Chung (1998, 2012) argues that such person-animacy effects in Chamorro have a morphological source, arising from language-specific constraints on pronunciation. I explore whether such an account can be extended to certain person-animacy effects in Santiago Laxopa Zapotec, in particular one pattern that parallels the "Ultrastrong" Person Case Constraint (Nevins 2007). While the morphological account that Chung proposes may be appropriate for Chamorro, I argue that a syntactic account is necessary for this person-animacy effect.

1. The grammatical source of person-animacy effects

To what extent, if at all, do morphological patterns reflect deeper syntactic relations? Chung (1998, 2012, 2013) has argued that, at least for verb agreement, they need not. While morphological agreement might parallel the effects of a syntactic operation like Agree, it is, in principle, independent. An important argument comes from certain restrictions on the person or animacy of arguments in Chamorro. A direct object cannot be higher on the hierarchy in (1) than the subject.

(1) *The person-animacy hierarchy in Chamorro*
2 > 3 animate pronoun > 3 animate non-pronoun > 3 inanimate (Chung 1998:34)

Chung argues that this person-animacy effect has a morphological source. Chamorro lacks a pronunciation for verbal agreement that corresponds to the prohibited combinations of arguments.

Many languages exhibit person-animacy effects. In Santiago Laxopa Zapotec (SLZ), it is not possible for the direct object to be first or second person when the subject is third person, if both are pronominal clitics.[1]

(2) a. Ba betw**=a'=ba'**.
already hit.COMP**=1SG=3SG.INF**
'I already hit her/him.' (FA, GZYZ011-s, 19) *1 > 3*

b. *Ba betw**=ba'=a'**.
already hit.COMP**=3SG.INF=1SG**
Intended: 'S/he already hit you.' (RM and FA, GZYZ015, 6:17) *3 > 1*

*I am extremely grateful to Flavio Antonio, Rosa Mendoza, and Fe Silva-Robles for teaching me about their language. I also received helpful comments from Pranav Anand, Steven Foley, Jorge Hankamer, and Nick Kalivoda.

[1]The abbreviations used here are: ANIM = animal, COMP = completive aspect, CONT = continuative aspect, DUB = dubitative, FOR = formal, INAN = inanimate, INF = informal, SG = singular.

(3) a. Ba betw**=o'=ba'**.
already hit.COMP**=2SG=3SG.INF**
'You already hit him.' (FA and RM, GZYZ015, 7:00) *2 > 3*

b. *Ba betw**=ba'=o'**.
already hit.COMP=**3SG.INF=2SG**
Intended: 'S/he already hit you.'(RM and FA, GZYZ015, 7:07) *3 > 2*

This resembles the Person Case Constraint, familiar from many Romance languages, which in its weak version prohibits the direct object from being third person when the indirect object is local (first or second) person (Perlmutter 1968, Bonet 1991).

There are many theories of the Person Case Constraint. Most are syntactic (Anagnostopoulou 2003, Béjar and Rezac 2003, 2009, Adger and Harbour 2007, Nevins 2007, 2011, a.o.), though some are morphological (Bonet 1991, Sturgeon et al. 2012). After providing some background on Santiago Laxopa Zapotec in Section 2, I identify three person-animacy restrictions on pronominal clitic combinations between subjects and direct objects in Section 3. One of these patterns parallels what Nevins (2007) calls the "Ultrastrong" Person Case Constraint.

In Section 4, I explore a morphological account of this pattern, inspired by Chung's (1998, 2012) proposal for person-animacy effects in Chamorro. However, as I show in Section 5, the behavior of pronominal clitics with ditransitive verbs reveals that SLZ cannot simply lack pronunciations for the ill-formed combinations. A syntactic account along the lines advanced in Section 6–7 can derive the person-animacy effects with monotransitives, as well as their absence with ditransitives. As I conclude in Section 8, it may not be so surprising that these effects have a morphological source in Chamorro, but a syntactic one in SLZ.

2. A brief introduction to Zapotec grammar

The Zapotec languages comprise one branch of the Oto-Manguean language family. I present data from Guiloxi and Yalina Zapotec, two mutually intelligible varieties spoken in the southeastern Sierra Norte mountains of Oaxaca. While they differ from one another in some aspects of their phonology and morphology, they exhibit the same person-animacy effects. I will refer to them as a group by a hypernym, Santiago Laxopa Zapotec (SLZ), following the classification of the *Instituto Nacional de Lenguas Indígenas* (http://www.inali.gob.mx/clin-inali/).[2]

Like other Zapotec varieties, SLZ has basic verb-subject-object word order. When arguments follow the verb, they are rigidly ordered.

(4) Dzutw **nu'ule'en bene' xjage'en.**
hit.CONT **woman person man**
'The woman is hitting the man.' (FSR, SLZ57a-s, 1)
Not possible: 'The man is hitting the woman.'

[2]The *Instituto Nacional de Lenguas Indígenas* includes the Guiloxi and Santiago Laxopa varieties in the same dialect area, though the Yalina variety is grouped with *zapoteco serrano del sureste medio*. Impressionistically, all three varieties are close enough to form a single group. Nearly all the data in the paper comes from the Guiloxi and Yalina varieties, and I verified the data in Section 3 with a speaker from Santiago Laxopa proper.

Arguments are cross-referenced on the verb with pronominal clitics, or "syntactically dependent" pronouns in Marlett's (1993) terms.

(5) Bdel**=e'=ba'.**
hug.COMP**=3SG.FOR=3SG.INF**
'S/he (e.g. an elder) hugged her/him (e.g. a non-elder).' (FA, GZYZ012-s, 19)

The full inventory of pronouns in SLZ is provided in Table 1. Alongside the clitics, there are also independent pronouns. Neither set is restricted by grammatical relation.

	INDEPENDENT	CLITICS
1sg.	*neda' (nada')*	*=a'*
2sg.	*lé'*	*=o'*
3sg. formal	*lè'*	*=e'*
3sg. informal	*leba'*	*=ba'*
3sg. animal	*leb*	*=(e)b(a)*
3sg. inanimate	*len*	*=(e)n*

Table 1: Independent and clitic pronouns (singular only) in Santiago Laxopa Zapotec.

At first glance, the clitic pronouns appear to be themselves arguments, since they are in complementary distribution with R-expressions.

(6) Ba bdel**(*=ba')** **bidao' ni** Pedro.
already hug.COMP**=3SG.INF** **child this** Pedro
'This child already hugged Pedro.' (FA and RM, GZYZ014, 27:09)

But, as Kalivoda (2015) observes for Teotitlán del Valle Zapotec, independent pronouns in subject position must be doubled by a clitic. The same is true in SLZ, but with one difference. Only local person independent pronouns must be doubled (7). All third person independent pronouns are, by contrast, in complementary distribution with a clitic (8).

(7) a. Tzxizh***(=a')** **neda'**.
laugh.CONT **=1SG** **1SG**
'I am laughing.' (FA and RM, GZYZ013, 5:27)
b. Tzxizh***(=u')** **lé'**.
laugh.CONT**=2SG** **2SG**
'You are laughing.' (FA and RM, GZYZ013, 6:07)

(8) a. Ba shtas(***=e')** **lè'**.
already sleep.CONT**=3SG.FOR** **3SG.FOR**
'S/he is sleeping.' (FA and RM, GZYZ014, 47:53)
b. Ba bdel**(*=ba')** **leba'** beku'.
already hug.COMP**=3SG.INF** **3SG.INF** dog
'S/he already hugged the dog.' (RM and FM, GZYZ013, 11:35)

c. Shtas**(*=b)** **leb**.
sleep.CONT**=3SG.ANIM** **3SG.ANIM**
'It (an animal) is sleeping.' (RM and FA, GZYZ013, 10:48)

d. Ba bzxup(***=en)** **len**.
already fall.COMP**=3SG.INAN** **3SG.INAN**
'It fell.' (RM and FA, GZYZ014, 49:55)

Kalivoda proposes the pronominal clitics arise through clitic doubling, driven by a probe bearing an unvalued person feature (cf. Béjar and Rezac 2003, Preminger 2014). In his system, R-expressions lack person features entirely, so that they never Agree and hence are never doubled. Independent pronouns, by contrast, possess the relevant features. When the probe Agrees with them, it copies their entire φ-feature bundle, including person (π) and number (#) features.

(9) a. *R-expression*
[π:] . . . DP
[#:**α**]

b. *Independent pronoun*
[π:] . . . DP
$\begin{bmatrix} \pi:\boldsymbol{\beta} \\ \#:\boldsymbol{\alpha} \end{bmatrix}$

This account can be extended to SLZ by treating the third person independent pronouns just like R-expressions, as lacking π-features altogether. However, since third person pronominal clitics do show up when there is no overt argument (8), there would have to be a corresponding null pronoun for each independent third person pronoun that possessed the relevant π-features.[3] While this nonuniformity is unappealing, it is a familiar problem (McCloskey and Hale 1984).

The verb can bear a clitic cross-referencing the direct object as well, as long as a subject clitic is present. A direct object cannot encliticize across an R-expression, across a trace of the subject, or onto the subject itself, as is possible in other Zapotec languages (Marlett 1993).

(10) a. Bdel**=ba'=ba**.
hug**=3SG.INF=3SG.ANIM**
'S/he hugged it.' (RM, GZYZ012-s, 23)

b. *Bdel**=b** **Maria**.
hug**=3SG.ANIM** **Maria**
Intended: 'Maria hugged it.' (FA and RM, GZYZ012, 24:55)

c. ***No$_1$** bet t_1 **=eb**?
who hit **=3SG.ANIM**
Intended: 'Who hit it?' (RM and FA, GZYZ013, 3:34)

[3]The probe cannot simply be endowed with an unvalued [participant] feature, as this would make all clitic doubling with third person pronouns impossible, including null ones. Instead, third person clitics could be actual arguments. Despite their paradigmatic relationship to other clitics, they originate in an argument position and encliticize onto the verb. The person-animacy effects in Section 3 suggest the account in the text might be on the right track.

d. *Bdel **Maria=b**.
hug Maria**=3SG.ANIM**
Intended: 'Maria hugged it.' (RM and FA, GZYZ013, 4:40)

Unlike subject clitics, direct object clitics are always in complementary distribution with an overt argument, regardless of whether it is an independent pronoun or an R-expression.

3. Person-animacy effects

Subject and direct object clitics exhibit specific cooccurrence restrictions. For instance, as shown in (11), if the subject clitic is local person, the object clitic must be third person (b–e). A first or second person object can only be realized as an independent pronoun (a).

(11) a. Ba betw**=a'** **lé'**.
already hit.COMP**=1SG** **2SG**
'I already hit you.' (RM, GZYZ011-s, 18)
b. Ba betw**=a'=ne'**.
already hit.COMP**=1SG=3SG.FOR**
'I hit her/him (e.g. an elder).' (FA, GZYZ011-s, 16)
c. Ba betw**=a'=ba'**.
already hit.COMP**=1SG=3SG.INF**
'I hit her/him (e.g. a child).' (RM, GZYZ011-s, 19)
d. Betw**=a'=ba**.
hit.COMP**=1SG=3SG.ANIM**
'I hit it (an animal).' (RM, GZYZ011-s, 20)
e. E wak=e' gaw**=a'=n**?
Q happen.DUB=3SG.FOR eat**=1SG=3SG.INAN**
'Can I eat it (a thing)?' (RM and FA, GZYZ011, 1:34:06)

The full paradigm of subject and direct object clitic combinations is shown for singular clitics in Table 1. The colored shading, which highlights three important patterns in the data, will be described shortly. There are also local person plural clitics that are not shown for reasons of space.

		DIRECT OBJECT				
		1sg./2sg.	3sg. formal	3sg. informal	3sg. animal	3sg. inanimate
SUBJECT	1sg.	*	*=a'=ne'*	*=a'=ba'*	*=a'=ba*	*=a'=n*
	2sg.	*	*=o'=ne'*	*=o'=ba'*	*=o'=b*	*=o'=n*
	3sg. formal	*	*/=e'=ne'*[4]	*=e'=ba'*	*=e'=b*	*=e'=n*
	3sg. informal	*	*=ba'=ne'*	*	*=ba'=ba*	*=ba'=n*
	3sg. animal	*	*	*	*	*=b=en*
	3sg. inanimate	*	*	*	*	*

Table 2: Subject and direct object pronominal clitic combinations in Santiago Laxopa Zapotec.

[4] This is, in fact, a point of variation. The speaker from Yalina allows the combination of two formal pronominal clitics, while the speakers from Santiago Laxopa and Guiloxi do not.

There are three striking patterns here. First, as shown in (11) above and by the ungrammaticality of the combinations in blue, the direct object clitic must be third person.

(12) In monotransitive clauses, a pronominal clitic cross-referencing the direct object must be third person.

This mirrors the "Strong" Person Case Constraint, which rules out in some Romance languages a local person direct object clitic in ditransitives (Perlmutter 1968:160, Bonet 1991:182).

Second, as shown by the ungrammaticality of the combinations in green, SLZ prohibits the subject and direct object clitics from having the exact same featural specification, if they are both third person.

(13) In monotransitive clauses, if the subject and direct object pronominal clitics are both third person, they may not have exactly the same featural specification.

This effect resembles the ban in Spanish on any combination of third person indirect and direct object clitics: **lo le*, **lo les,* etc. (Perlmutter 1968:140). In SLZ, the relevant constraint is more specific, as just the combinations identical in animacy and formality are banned.

Finally, there is a *relative* constraint on subject and direct object clitics, which can only be seen when both are third person, due to the constraint in (12). As shown by the ungrammaticality of the combinations in pink, the direct object cannot be more animate than the subject.

(14) In monotransitive clauses, the direct object clitic must not outrank the subject on the person-animacy hierarchy.

This parallels the "Ultrastrong" Person Case Constraint, which in Romanian and Classical Arabic prevents a direct object clitic from outranking an indirect object clitic on a person hierarchy: 1 > 2 > 3 (Nevins 2007:297–298).

The relevant person-animacy hierarchy for SLZ is given in (15). Local persons are included for completeness, though strictly speaking they need not be included, as there are no first or second person direct object clitics.

(15) *The person-animacy hierarchy in Santiago Laxopa Zapotec*
1, 2 > 3 human > 3 animal > 3 inanimate

This hierarchy differs from the one for Chamorro in some crucial ways. It creates a more fine-grained division of third person along a scale of animacy: human vs. animal vs. inanimate. There is also no reference to pronouns, as it solely regulates the combination of pronominal clitics.

4. A Chungian account

Moving forward, I leave the first two patterns above aside to seek an account for just the last pattern. Inspired by Chung's (1998, 2012) work on Chamorro, what would a morphological account of the generalization in (14) look like?

Within one realizational theory of morphology, Distributed Morphology, it is not easy to understand why the combinations of pronominal clitics in (16) and (17) are ungrammatical.

(16) a. *Bdi'in**=b=ne'**.
bite.COMP**=3SG.ANIM=3SG.FOR**
Intended: 'It bit her/him.' (RM, GZYZ014, 32:38) *3 animal > 3 human*

b. *Bdi'in**=ba=ba'**.
bite.COMP**=3SG.ANIM=3SG.INF**
Intended: 'It bit her/him.' (RM, GZYZ014, 33:30) *3 animal > 3 human*

(17) a. *Betw**=en=ne'**.
hit.COMP**=3SG.INAN=3SG.FOR**
Intended: 'It hit her/him.' (RM and FA, GZYZ014, 42:21) *3 inanimate > 3 human*

b. *Betw**=en=ba'**.
hit.COMP**=3SG.INAN=3SG.INF**
Intended: 'It hit her/him.' (RM and FA, GZYZ014, 42:54) *3 inanimate > 3 human*

c. *Bxizh**=en=eb.**
strike.COMP**=3SG.INAN=3SG.ANIM**
Intended: 'It struck it.' (RM and FA, GZYZ016, 2:15) *3 inanimate > 3 human*

It is possible for a more specific vocabulary item to block the insertion of more general vocabulary items (Halle and Marantz 1993:120). But to derive the ill-formedness of these combinations, it would be necessary for a vocabulary item—really, for any vocabulary item—to fail to insert in contexts more specific than its featural specification require.

To account for the Person Case Constraint in some Romance languages, Bonet (1991:78–128) proposes a rule of impoverishment, which would eliminate the features that offend the constraint in (14). This would predict incorrectly, however, that the ungrammatical combinations could be repaired by simply omitting one of clitics. Rather, the object must be realized as an independent pronoun.

(18) a. Bdi'in**=b lè'**.
bite.COMP**=3SG.ANIM 3SG.FOR**
'It bit her/him.' (RM, GZYZ014, 32:37)

b. Bdi'in**=b leba'**.
bite.COMP**=3SG.ANIM 3SG.INF**
'It bit her/him.' (RM, GZYZ014, 33:32)

For the ungrammaticality of (16)–(17) to arise through competition, pronominal clitics would have to stand in a blocking relationship with independent pronouns (Bonet 1991:201–209, Cardinaletti and Starke 1999). I do not see how this can be if independent pronouns and pronominal clitics pronounce distinct feature bundles (see also Nevins 2011:948 and Rezac 2011:114–133).

Chung (1998:199–205, 2012:186–187) takes a different approach, countenancing realizational rules that can filter syntactic representations through the satisfaction of their featural specification. In other words, she allows for rules whose outputs are not a morphological formative (a), but a diacritic indicating ill-formedness (b–d).

(19) a. [+participant, –author, –plural]$_{subj}$ → *un*
b. [–participant]$_{subj}$,
[+participant, –author]$_{obj/poss}$ → * *3 > 2*
c. [–pronoun]$_{subj}$,
[–author, +animate, +pronoun]$_{obj/poss}$ → * *non-pronoun > 3 animate pronoun*
d. [–animate]$_{subj}$
[–author, +animate]$_{obj/poss}$ → * *inanimate > animate*
(Chung 2012:187)

Once the features of both subject and direct object have been copied onto a functional head—whether through Agree or through a purely morphological operation—these rules fail to produce a well-formed output for that head, deriving the ungrammatical combinations of arguments.

Of course, in SLZ, it is not combinations of arguments that are ruled out, but rather combinations of clitics. In addition, each clitic realizes a distinct feature bundle—that is, they are not portmanteaux—with the same pronunciation (for the most part) regardless of grammatical relation. Nonetheless, it is possible to formulate a set of realizational rules that derives the ungrammatical combinations of clitics in (16)–(17).

(20) a. [–participant, +animate, +human, +formal] → *=(n)e'*
b. [–participant, +animate, +human, –formal] → *=ba'*
c. [–participant, +animate, –human] → * / ___ [–participant, +animate, +human]
d. [–participant, +animate, –human] → *=(e)b(a)*
e. [–participant, –animate] → * / ___ [–participant, +animate]
f. [–participant, –animate] → *=(e)n*

While human pronominal clitics have just one realizational rule (a–b), the other pronominal clitics have an additional rule (c, e) that produces an ill-formed result if they occur before a clitic higher on the person-animacy hierarchy.[5]

5. A syntactic alternative

What would a syntactic account of these facts look like instead? There are numerous theories of the Person Case Constraint that could be extended to SLZ, and I cannot do justice to them all here (see Anagnostopoulou 2003, Béjar and Rezac 2003, 2009, Nevins 2007, 2011, a.o.). Instead, I simply demonstrate that Nevin's account using Multiple Agree is possible. An account based on Cyclic Agree might also be possible (Walkow 2014).

[5]Either these clitics are all strictly adjacent to one another, because they have moved into this position in the syntax, or the contextual restriction of the rules in (20) are relativized to clitics (Toosarvandani 2016).

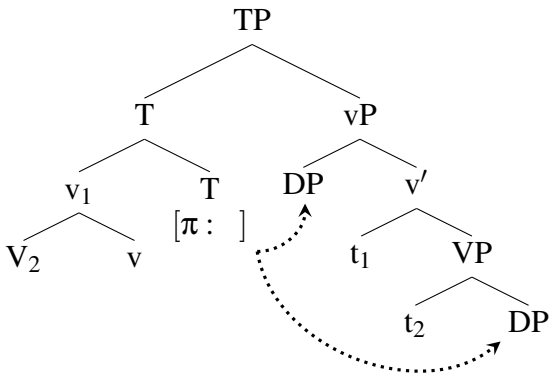
Nevins (2007, 2011) proposes that the various versions of the Person Case Constraint arise as a type of intervention. His account builds on two core assumptions. First, a probe bearing an unvalued feature can Agree with multiple goals bearing a matching feature simultaneously (Hiraiwa 2001). Second, probes can be further relativized, so that they search for a specific value of a feature. Based on this, relativization constraints, such as Contiguous Agree in (21), can be imposed on the Agree relation.

(21) *Contiguous Agree* (Nevins 2007:295)
For a relativization R of a feature F on a probe P, and $x \in Domain(R(F))$, $\neg \exists y$, such that $y > x$ and $p > y$ and $y \notin Domain(R(F))$.

Contiguous Agree requires that no goal intervene between the probe and a goal matching its relativization that does not also match the relativization of the probe.

Following Kalivoda (2015), I assume that the pronominal clitics realize a π-probe on T, which Agrees with every argument in the clause. These Agree relations are established, as Nevins proposes, simultaneously.

(22)



While the probe searches for π-features, it copies the entire φ-feature bundle of any DP it Agrees with, producing clitic doubling (Béjar and Rezac 2003, Preminger 2014). As Kalivoda proposes, some DPs do not have π-features, in which case they do not trigger clitic doubling (or trigger any person-animacy effects). In addition to R-expressions, independent third person pronouns that are phonologically overt would not possess π-features in SLZ.

The forbidden combinations of third person pronominal clitics are ruled out by relativizing T to the marked values of two features: [+animate] and [+human]. To see why, consider the featural specifications for each combination:

(23) a. *3 human > 3 animal*
[–participant, **+animate**, **+human**] > [–participant, **+animate**, –human]
b. *3 animal > 3 inanimate*
[–participant, **+animate**, –human] > [–participant, –animate]
c. **3 animal > 3 human*
[–participant, **+animate**, –human] > [–participant, **+animate**, **+human**]
d. **3 inanimate > 3 human*

[–participant, –animate] > [–participant, **+animate, +human**]

e. **3 inanimate > 3 animal*

[–participant, –animate] > [–participant, **+animate**, –human]

For the well-formed combinations, every goal bearing [+animate] is not c-commanded by another goal bearing [–animate] (a–b), and every goal bearing [+human] is not c-commanded by another goal bearing [–human] (a). By contrast, for the ill-formed combinations, there is such an intervening goal for either the [+animate] feature (d–e) or the [+human] feature (c).

6. Testing a prediction

It is possible, I think, to choose between these morphological and syntactic accounts of person-animacy effects in SLZ. The morphological account predicts that there should be *no syntactic context* where the illicit combinations of pronominal clitics are allowed, as their ungrammaticality is conditioned solely by the featural identity of the following clitic. By contrast, the syntactic account, which refers to the relative hierarchical position of goals, could in principle allow for these combinations in the right syntactic configuration.

The prediction the morphological account makes is not borne out. There is at least one syntactic environment where the ungrammatical combinations of pronominal clitics are attested. With ditransitives, the indirect object can be cross-referenced by a pronominal clitic on the verb. It has the same form as subject and direct object clitics and is located invariantly between them.

(24) a. Ba bia**=a'=ba'=ba**.
already give.COMP**=1SG=3SG.INF=3SG.ANIM**
'S/he already gave it (an animal) to her/him.' (RM and FA, GZYZ014, 1:19:19)

b. *Ba bi**=a'=ba=ba'**.
already give.COMP**=1SG=3SG.ANIM=3SG.INF**
Intended: 'S/he already gave it (an animal) to her/him.' (RM and FA, GZYZ015, 41:24)

While I do not show this here, the Strong Person Case Constraint remains in effect between subject and indirect object clitics, as well as between direct and indirect object clitics.

However, the Ultrastrong Person Case Constraint—which, it might be expected, would prohibit a direct object clitic from outranking the indirect object clitic on the person-animacy hierarchy—is lifted. The grammatical combinations in (26) are the same ones that are ill-formed in (16) above. They form minimal pairs with the combinations in (25).

(25) a. Ba blo'ed=a'**=ne'=b**.
already show.COMP=1SG**=3SG.FOR=3SG.ANIM**
'I already showed it to her/him.' (RM and FA, GZYZ015, 49:45)

b. Ba bi=a'**=ba'=b**.
already give.COMP=1SG**=3SG.INF=3SG.ANIM**
'I already gave it to her/him.' (RM and FA, GZYZ014, 1:19:19) *3 human > 3 animal*

(26) a. Ba blo'ed=a'**=b=ne'**.
already show.COMP=1SG**=3SG.ANIM=3SG.FOR**
'I already showed her/him to it' (RM and FA, GZYZ015, 48:35)

b. E blo'ed=o'**=ba=ba'**?
Q show.COMP=2SG**=3SG.ANIM=3SG.INF**
'Did you show her/him to it?' (RM and FA, GZYZ015, 1:02:00) *3 animal > 3 human*

The complete paradigm of indirect and direct object clitic combinations is given in Table 3.[6]

		DIRECT OBJECT				
		1sg./2sg.	3sg. formal	3sg. informal	3sg. animal	3sg. inanimate
	1sg.	*	*	*	*	*
	2sg.	*	*	*	*	*
INDIRECT	3sg. formal	*	*	*=ne'=ba'*	***=ne'=b***	*=ne'=n*
OBJECT	3sg. informal	*	*=ba'=ne'*	*	***=ba'=b***	*=ba'=n*
	3sg. animal	*	***=b=ne'***	***=ba=ba'***	*	*
	3sg. inanimate	–	–	–	–	–

Table 3: Indirect and direct object pronominal clitic combinations in Santiago Laxopa Zapotec.

It remains to be seen how a syntactic account might deal with these facts. But they clearly suggest a morphological account is not tenable for the Ultrastrong Person Case Constraint in SLZ.

7. Toward a syntactic account of ditransitives

While I cannot offer a complete account here for the behavior of pronominal clitics in ditransitives, I would like to point to some facts suggesting that a syntactic account is on the right track.

The underlying structure of ditransitives in SLZ is not entirely clear. But there is some evidence that the indirect object originates closer to the verb than the direct object. At least one verb, *-e* 'give', exhibits suppletion that is dependent on the person of the goal, as shown in (27): *ben* for local persons (a–b) and *be* for third person (c).

(27) a. Ba **ben**=ba' **nada'** beku'.
already **give.COMP**=3SG.INF **1SG** dog
'Maria already gave the dog to me.' (RM, GZYZ015-s, 7)
b. Ba **ben**=ba' **lé'** beku'.
already **give.COMP**=3SG.INF **2SG** dog
'S/he already gave the dog to you.' (RM, GZYZ015-s, 6)
c. Ba **be**=ba' **leba'** beku'.
already **give.COMP**=3SG.INF **3SG.INF** dog
'S/he already gave the dog to her/him.' (RM, GZYZ015-s, 8)

It is absolutely ungrammatical for a local person indirect object to occur with *be*, or correspondingly a third person indirect object with *ben*.

(28) a. *Ba **be**=ba' **lé'** beku'.

[6]The combinations with an inanimate indirect object are not given, since I was not able to identify a verb for which this was permitted semantically.

already **give.COMP**=3SG.INF **2SG** dog
Intended: 'S/he already gave the dog to you.' (FA, GZYZ015, 31:49)

b. *Ba **ben**=ba' **leba'** beku'.
already **give.COMP**=3SG.INF **3SG.INF** dog
Intended: 'S/he already gave the dog to her/him.' (RM and FA, GZYZ015, 32:15)

This kind of allomorphy must be subject to a locality constraint of some kind, since not just any element can condition suppletion of the verb. It might, for instance, be conditioned strictly locally (Bobaljik and Harley 2013).

(29) *Strict Locality* (Bobaljik and Harley 2013:10)
β may condition the insertion of α in (a), but not (b):
(a) β … $[_{X^0}$ … α
(b) *β … $[_{X^n}$ … α where $n > 0$

Under this view, the indirect object would have to merge as the sister of the verb in order to condition its suppletion.[7] The direct object could then merge as the verb's specifier.

Crucially, the order of clitics, which is fixed, does not reflect this underlying order. This is true across languages in general. Even between closely related languages, there can be variation in the order of clitics, suggesting that this is an idiosyncratic, purely morphological property of languages (Bonet 1995, Miller and Sag 1997). There have been some attempts, though, to derive clitic ordering in individual languages from general grammatical principles (Grimshaw 2001, Sturgeon et al. 2012).

While the position of the direct and indirect objects may not matter for the linear order of pronominal clitics, it does matter, under the syntactic account, for the conditions on Agree. Recall that Contiguous Agree in (21) prohibits the highest goal bearing a feature matching the relativization of the probe from being c-commanded by an intervening goal with a different value. As a consequence, the Ultrastrong Person Case Constraint should only arise in ditransitives if the indirect object *invariantly* asymmetrically c-commands the direct object.

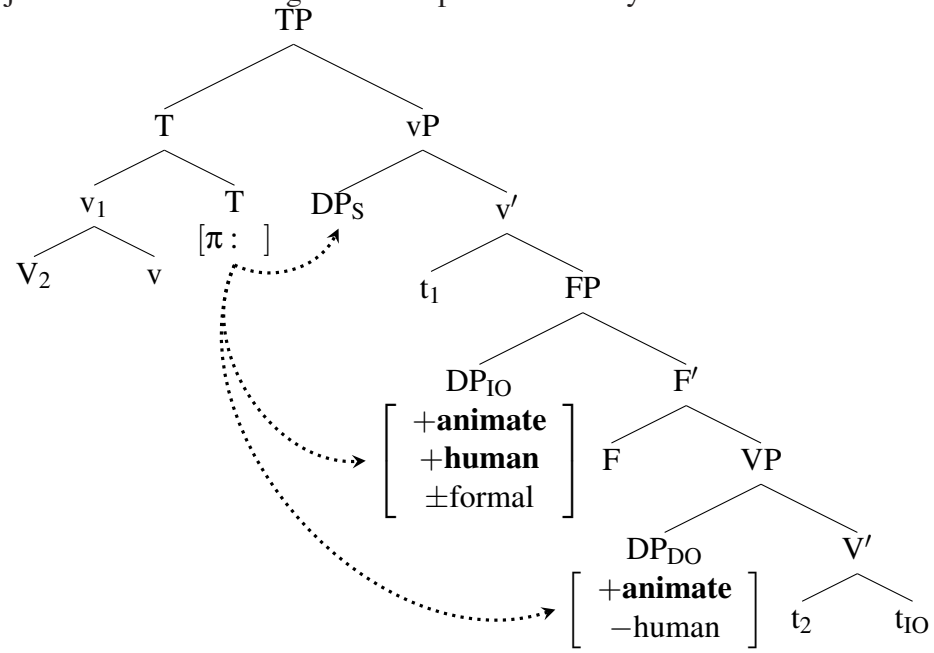
So, where do direct and indirect objects occur in SLZ? In terms of linear order, they are freely ordered (see also Sonnenschein 2004:156–157 on the closely related Zoogocho variety).

(30) a. Ba be Maria **beku' bidau' ni**.
already give.COMP Maria **dog child this**
'Maria gave the dog to this child.' (RM and FA, GZYZ015, 18:13)
b. Ba be Maria **bidau' ni beku'**.
already give.COMP Maria **child this dog**
'Maria gave the dog to this child.' (RM and FA, GZYZ015, 18:46)

[7]Even if the locality condition on verb suppletion were loosened somewhat (Toosarvandani 2016), the indirect object would have to be located closer to the verb than the direct object.

I take this to mean that the indirect object can undergo optional movement to some position above the direct object. This configuration gives rise to the combination in (25), where the indirect object clitic is located higher on the person-animacy scale than the direct object clitic.

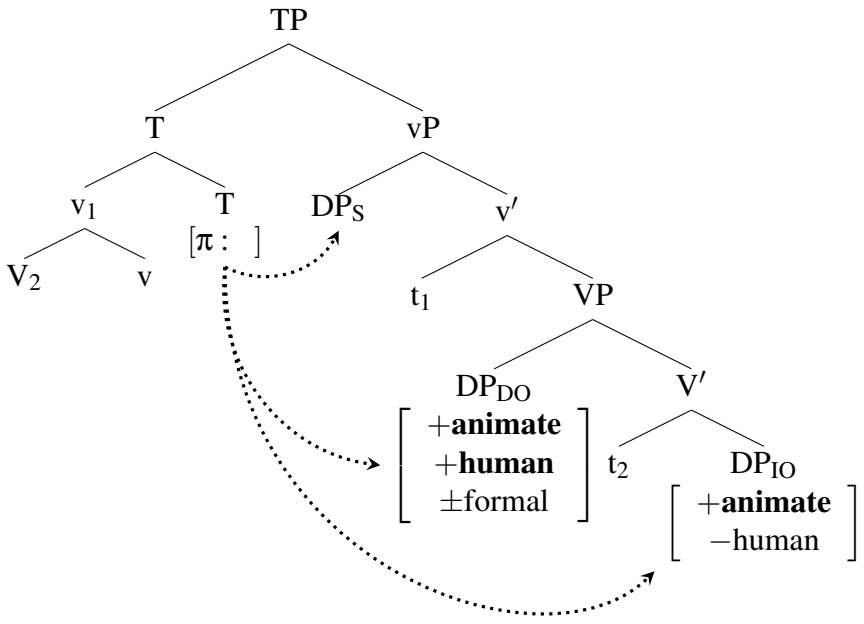
(31) = (25)



Though the indirect object merges below the direct object, it raises, so that the direct object does not intervene between it and the probe. It can thus possess the [+human] feature the probe is relativized to, despite the direct object bearing the [–human] feature.

The opposite order of the same clitics in (26)—the order that appears to violate the Ultrastrong Person Case Constraint—arises when the indirect object does not move.

(32) = (26)



Now the direct object clitic can be located higher on the person-animacy scale than the indirect object clitic without violating Contiguous Agree.

8. Conclusion

While some person-animacy effects can be attributed to morphological constraints, some must, I have argued, arise from syntactic principles. In particular, the Ultrastrong Person Case Constraint in SLZ, which restricts the possible combinations of third person subject and direct object clitics, requires a syntactic account, possibly along the lines that Nevins (2007, 2011) proposes. This is necessary because the constraint is lifted in ditransitives: there are no restrictions between third person indirect and direct objects.

This result might not be particularly surprising if "today's morphology is yesterday's syntax" (Givón 1971:431). We should find similar patterns, which have a morphological source in one language and a syntactic source in another. What is interesting here is not this general finding, but rather that it is possible to discern where person-animacy effects come from in individual languages. While they have a morphological source in Chamorro, as Chung proposes, they have a syntactic source in SLZ.

References

Adger, David and Daniel Harbour. 2007. Syntax and syncretisms of the Person Case Constraint. *Syntax* 10:2–37.

Anagnostopoulou, Elena. 2003. *The syntax of ditransitives: Evidence from clitics*. Berlin: Mouton de Gruyter.

Béjar, Susana and Milan Rezac. 2003. Person licensing and the derivation of PCC effects. *Romance linguistics: Theory and acquisition*, eds. Ana Teresa Pérez-Leroux and Yves Roberge, 49–62. Amsterdam: John Benjamins.

Béjar, Susana and Milan Rezac. 2009. Cyclic Agree. *Linguistic Inquiry* 40: 35–73.

Bobaljik, Jonathan David and Heidi Harley. 2013. Suppletion is local: Evidence from Hiaki. Ms., University of Connecticut and University of Arizona. http://ling.auf.net/lingbuzz/001982.

Bonet, Eulàlia. 1991. Morphology after syntax: Pronominal clitics in Romance. PhD dissertation, Massachusetts Institute of Technology.

Bonet, Eulàlia. 1995. Feature structure of Romance clitics. *Natural Language and Linguistic Theory* 13: 607–647.

Cardinaletti, Anna and Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. *Clitics in the languages of Europe*, ed. H. C. van Riemsdijk, 145–233. Berlin: Mouton de Gruyter.

Chung, Sandra. 1998. *The design of agreement: Evidence from Chamorro*. Chicago: University of Chicago Press.

Chung, Sandra. 2012. On reaching agreement late. *Chicago Linguistic Society (CLS)* 48: 169–190.

Chung, Sandra. 2013. The syntactic relations behind agreement. *Diagnosing syntax*, eds. Lisa Lai-Shen Chen and Norbert Corver, 351–370. Oxford: Oxford University Press.
Givón, Talmy. 1971. Historical syntax and synchronic morphology: An archeologist's fieldtrip. *Chicago Linguistic Society (CLS)* 7: 394–415.
Grimshaw, Jane. 2001. Optimal clitic positions and the lexicon in Romance clitic systems. *Optimality Theoretic Syntax*, eds. Géraldine Legendre, Jane Grimshaw, and Sten Vikner, 205–240. Cambridge, MA: MIT Press.
Halle, Morris and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, eds. Kenneth Hale and Samuel Jay Keyser, 111–176. Cambridge, MA: MIT Press.
Hiraiwa, Ken. 2001. Multiple agree and the defective intervention constraint in Japanese. *MIT Working Papers in Linguistics* 40: 67–80.
Kalivoda, Nick. 2015. Distinguishing doubled clitics and pronominal arguments in Teotitlán del Valle Zapotec. Ms., University of California, Santa Cruz.
Marlett, Stephen A. 1993. Zapotec pronoun classification. *International Journal of American Linguistics* 59: 82–101.
McCloksey, James and Kenneth Hale. 1984. On the syntax of person-number inflection in modern Irish. *Natural Language and Linguistic Theory* 1:487–533.
Miller, Phillip H. and Ivan Sag. 1997. French clitic movement without clitics or movement. *Natural Language and Linguistic Theory* 15: 573–639.
Nevins, Andrew. 2007. The representation of third person and its consequences for persn-case effects. *Natural Language and Linguistic Theory* 25: 273–313.
Nevins, Andrew. 2011. Multiple agree with clitics: Person complementarity vs. omnivorous number. *Natural Language and Linguistic Theory* 29: 939–971.
Perlmutter, David M. 1968. Deep and surface structure constraints in syntax. PhD dissertation, Massachusetts Institute of Technology.
Rezac, Milan. 2011. *Phi-features and the modular architecture of language*. Dordrecht: Springer.
Preminger, Omer. 2014. *Agreement and its failures*. Cambridge, MA: MIT Press.
Sonnenschein, Aaron. 2004. A descriptive grammar of San Bartolomé Zoogocho Zapotec. PhD dissertation, University of California, Los Angeles.
Sturgeon, Anne, Boris Harizanov, Maria Polinsky, Ekaterina Kravtchenko, Carlo Gómez Gallo, Lucie Medová, and Václav Koula. 2012. Revisiting the Person Case Constraint in Czech. *Formal Approaches to Slavic Linguistics (FASL)* 19: 116–130.
Toosarvandani, Maziar. 2016. Vocabulary insertion and locality: Verb suppletion in Northern Paiute. *North East Linguistic Society (NELS)* 46: 247–256.
Walkow, Martin. 2014. Cyclic AGREE derives restrictions on cliticization in classical Arabic. *Perspectives on Arabic Linguistics XXVI*, eds. Reem Khamis-Dakwar and Karen Froud. Amsterdam: John Benjamins.

AGENT FOCUS AND PASSIVE IN TSOTSIL*

Judith Aissen
UC Santa Cruz
aissen@ucsc.edu

Earlier work identified parallels in the function and distribution of AGENT FOCUS and PASSIVE in Tsotsil. This paper argues that AGENT FOCUS in Tsotsil has in fact been reanalyzed as a type of passive, with the option (in some dialects) of WH-AGREEMENT with the agent. Evidence comes in part from its syntactic properties. But most compelling is the syncretism of passive and AF morphology which is a pervasive feature of the language.

1. Introduction

The Mayan languages fall into two classes with respect to A-bar extraction of ergatives (subjects of transitive clauses). One group generally disallows it, while the other permits it. Intimately related to what I will call the ERGATIVE EXTRACTION CONSTRAINT (EEC) is the AGENT FOCUS (AF) construction, whose sole function is to remedy the EEC. That is, it provides the means to express extraction of the agent, when extraction from a canonical transitive clause is precluded by the EEC. Languages which are subject to the EEC have an AF construction while languages which are not lack one.

The EEC itself has been interpreted in several ways. One line of thinking attributes it to syntactic ergativity (Larsen and Norman 1979, Campana 1992, Ordoñez 1995, Coon et al. 2014), seeing languages not subject to the EEC as (just) morphologically ergative. A recent alternative sees it as the consequence of an anti-locality condition on extraction (Erlewine 2016). A third account views it as the consequence of a preference in some languages for the specialized AF construction when the agent is extracted (Stiebels 2006). In Stiebels' analysis, languages not subject to the EEC do not have this preference. See Aissen (to appear) for discussion of some of these approaches.

Tsotsil occupies a sort of intermediate position between the two groups of languages. Ergative extraction is permitted and common, so the language is not subject to any general EEC. However, ergative extraction is precluded in one corner of the grammar and in that corner, we find AF. Thus the complementary relation between ergative extraction and AF is maintained, but the domain in which AF occurs is much smaller than in the other Mayan languages.

In Aissen (1999), I argued that the distribution of AF in Tsotsil was not related to syntactic ergativity, but to the same factors that determine PASSIVE. One of the functions of passive in Tsotsil is to realize clauses in which the external argument (A) is low in topicality and the internal argument (P) is high. When both arguments are 3rd person, active voice is excluded and some

* It is a pleasure to contribute a paper in honor of Sandy Chung, a friend and colleague of many years. The descriptive richness of her work, its depth of analysis, the dissatisfaction with received wisdom, and Sandy's penetrating intelligence are among the qualities that have made her work so influential. The present paper touches on several areas in which Sandy has made seminal contributions, WH-agreement and passive. I would also like to thank Ruth Kramer for her astute comments on an earlier version of this paper.

intransitive paraphrase is required. Passive is not the only remedy, but it is a systematically available one (Aissen 1997). AF has the same function in Tsotsil, but in clauses in which A is extracted. The relation between passive and AF is schematized in Table 1, *x* = low topicality; X = high topicality.

	CLAUSES WITHOUT AGENT EXTRACTION	CLAUSES WITH AGENT EXTRACTION
Ap, ap	transitive	transitive
aP	passive	AF

Table 1: Distribution of voice in Tsotsil

These generalizations and the parallel functions of passive and AF in Tsotsil were established in Aissen (1997, 1999).

This paper considers in more detail the *structure* of AF clauses in Tsotsil. I will suggest that not only do AF clauses have much the same function and distribution as passives, they have been reanalyzed as passive. In some dialects, passive has entirely replaced AF (i.e., AF has been lost). Other dialects have a distinctive AF construction which differs from canonical passive only in two respects: [i] the agent is obligatory (not optional), and [ii] the agent is syntactically licensed by the passive head, permitting it to surface as direct (not oblique). Since this relation involves morphology linked to extraction and Case, I refer to it as *WH-agreement* (Watanabe 1996, Chung 1998). (1) summarizes the proposal.

(1) In Tsotsil, AF is a passive construction endowed with the possibility of WH-agreement with the agent.

The evidence that AF clauses are passive comes in part from their syntactic properties and in part from the relation between passive and AF morphology. In one dialect, historic AF morphology occurs now in what is transparently a passive; in other dialects, passive morphology has replaced historic AF morphology in AF clauses. These developments begin to make sense if AF itself is analyzed as passive. Tsotsil is not unique in showing syncretism of the morphology associated with passive and WH-agreement. It is found also in Chamorro, where the infix *-in-* realizes both passive morphology and WH-agreement with the internal argument. Chung (1998) argues that this is not syntactically significant in Chamorro, that clauses in which *-in* marks WH-agreement are *not* passive. I will suggest that the Tsotsil situation is different, that the appearance of the same morphology in passive and AF clauses reflects the fact that AF clauses *are* passive.

2. Agent Focus in Mayan

There is a large descriptive literature on the morphosyntax of AF constructions in Mayan, going back to the 1970's (see references in Aissen 1999, Stiebels 2006, and Coon 2016). To begin with, consider (2), from Berinstein (1985), which illustrates the situation in Q'eqchi'. (2) is a canonical transitive clause, with two direct 3rd person arguments. It illustrates some of the typological properties common to Mayan: verb-initiality, head-marking, and morphological ergativity. Neither argument is case-marked (the pronominal object is not realized as an independent noun phrase),

but both are indexed on the verb, the subject (A) by an ERG(ATIVE) prefix and the object (P) by an ABS(OLUTIVE) prefix. If A is extracted from such a clause, the expected result is (2). However, (2) is ungrammatical and A–extraction is expressed instead by (2), the AF construction in this language.[1]

(2) a. X-in-x-sak' li wiinq.
Q'EQ REC-ABS1S-ERG3S-hit DET man
'The man hit me.' {Berinstein 1985:162}

b. *Ha' li wiinq k-in-ix-sak'.
FOC DET man PAST-ABS1S-ERG3S-hit
('That's the man who hit me.') {Berinstein 1985:164}

c. Ha' li wiinq ki-sak'-o-k w-e.
FOC DET man PAST-hit-AF-IV GEN1S-OBL
'That's the the man who hit me.' {Berinstein 1985:164}

The details of the AF construction vary across the family, but there are several constant features, visible in (2). One is that the AF verb is detransitivized by an overt suffix (glossed AF).[2] The intransitivity of AF verbs is signalled by the absence of an ERG prefix and often by the presence of an intransitive 'status suffix' (*-k* in (2)). Another is that AF clauses require two syntactically realized arguments. In Q'eqchi' (also Mam) (Berinstein 1985, England 1983), AF clauses have the syntax of canonical antipassives, with demotion of the internal argument. This is also an option in Tz'utujil and for some dialects/speakers of K'ichee' (Dayley 1985, Mondloch 1981).

In another set of languages, both arguments in AF clauses remain 'direct'. This is the situation in all the Q'anjob'alan languages (Craig 1977, Zavala 1992, Pascual 2007), and is an option in some K'ichean languages (Tz'utujil and some dialects of K'ichee') (Dayley 1985, Mondloch 1981). (3a) from Akatek (Q'anjob'alan), is a basic transitive clause, with VSO order. Extraction of A from such a clause should produce (3b). (3b) is not ungrammatical, but it can only be interpreted as involving extraction of P. The AF construction in (3c) expresses the A-extraction reading. (Examples (3a,c,d) are from Zavala (1992:279); (3b) is thanks to Roberto Zavala (p.c.).)

(3) a. [X]-s-ma' ix malin naj xhunik.
AKA ASP-ERG3-hit CLS Maria CLS Juan
'Maria hit Juan.'

b. Maj [x]-s-ma' naj xhunik?
who ASP-ERG3-hit CLS Juan
'Who did Juan hit?' (*not* 'Who hit Juan?')

[1] Abbreviations in glosses: ABS: absolutive, AF: agent focus, APPL: applicative, ASP: aspect, CL: clitic, CLS: classifier, CP: completive, DET: determiner, DIR:directional, ENC: enclitic, ERG: ergative, FOC: focus, GEN: genitive, ICP: incompletive, INCL: inclusive, IV: intransitive status suffix, NEG: negation, OBL: oblique, OCK: Laughlin 1977, PF: perfect, PL: plural, PRO: pronoun, PSV: passive, PT: particle, REC: recent past, RR: reflexive, S: singular, TOP:topic.

[2] The one exception is Yucatec Maya, where the AF verb is not derived via an overt suffix, but through the absence of ergative marking and neutralization of certain TAM distinctions. Further, the general consensus is that AF verbs in Yucatec are transitive (Tonhauser 2003, Bohnemeyer 2009, Gutiérrez Bravo 2015).

c. Maj [x]-ma'-on naj xhunik?
who ASP-hit-AF CLS Juan
'Who hit Juan?'

d. Maj x-in-ma'-on-i?
who ASP-ABS1S-hit-AF-IV
'Who hit me?'

The verb in (3c) is again intransitive (it carries no ERG prefix) but, unlike (2) in Q'eqchi', neither argument is oblique. Although we should consider the possibility that P is a 'covert' oblique, this is not tenable since, if it is 1st or 2nd person, it inflects on the verb, (3d), an observation first made in Craig (1979) for Jakaltek (on Q'anjob'al, see Pascual 2007). I refer to the construction in (3c,d) as *direct* AF. See Stiebels (2006) and Coon et al. (2014) for different views on how direct AF clauses might be analyzed, and Aissen (to appear) for discussion.

Examples (2) and (3c,d) illustrate the two suffixes that are found throughout the family for deriving AF verbs, suffixes that Smith-Stark (1978) reconstructed as *-(V)w* and *-(V)n*. In the K'ichean languages these are allomorphs, with *-(V)w* restricted to root transitive stems (with form CVC) and -(V)*n* to derived transitives. The Q'anjob'alan and Mamean languages use *-(V)n* for all stems.

Given that AF verbs are intransitive, yet require two arguments, it follows that they share core properties with passive. Passive verbs are also derived intransitives throughout Mayan and they are associated with two arguments, though the agent may be syntactically suppressed. It is clear though that in most Mayan languages, AF clauses do not have passive syntax: in languages like Q'eqchi', where AF is a syntactic antipassive, this is obvious. In languages with direct AF, like Q'anjob'al, this is less obvious, but the burden of proof would presumably fall on anyone arguing that they *are* passive.

However, in Tsotsil, the features shared by passive and AF clauses go significantly beyond the two already mentioned. These fall into three classes: [1] as noted above, the distribution of AF is very similar to that of passive, (§3.1); [2] there is evidence that the internal argument (P), not the external argument, is the 'subject' in AF clauses, (§3.2); and [3] the various dialects show syncretism of passive and AF morphology (§3.3).

3. Agent focus in Tsotsil

3.1 The distribution of AF

Zinacantec (Z) Tsotsil has an AF construction which looks very much like the AF construction of Q'anjob'al. The AF verb is formed with the same suffix, *-on* and both arguments (*k'usi* 'what' and *li jbek'ettike* 'our meat') are direct (compare with 3c):[3]

(4) K'usi xu' x-tam-on li jbek'ettik-e?
Z TSO what can ASP-take-AF DET our.meat-ENC
'What could have taken our meat?' {OCK 282}

[3] All Tsotsil examples cited without a source come from my own fieldnotes.

With two non-oblique arguments, AF in Z Tsotsil resembles a transitive clause. However, as in other Mayan languages, AF verbs inflect as intransitives: they carry no ERG marker and in relevant contexts (here, in construction with an auxiliary), they carry the intransitive status suffix *-ik~-uk* (underlined in (5)):

(5) Muk' buch'u x-k'ot ik'-on-uk.
Z TSO NEG who ASP-arrive$_{AUX}$ call-AF-IV
'There was no one to come and call him.' {OCK 42}

Unlike Q'anjob'al and the other Mayan languages, however, extraction of the ergative from a transitive clause is possible in Z Tsotsil and is in fact far more frequent than agent extraction from the AF construction. In contrast to (2) (Q'eqchi'), (6) is grammatical, and in contrast to (3b) (Q'anjob'al), it has an A-extraction reading (as well as the P-extraction reading, i.e., it is ambiguous). (7) is a representative text example; hundreds more could be cited.

(6) Buch'u i-s-kolta ti vinik-e?
Z TSO who CP-ERG3-help DET man-ENC
'Who helped the man?' (*also* 'Who did the man help?')

(7) Pero buch'u s-tam?
Z TSO but who ERG3-take
'But who took it (the ring)?' {OCK 353}

Ergative extraction is not always possible though. As documented in Aissen (1999), there are two contexts in which agent extraction from a transitive clause is blocked in Z Tsotsil and where AF constructions are used instead. Both involve clauses in which P is more salient than A (*aP* settings), where 'salience' involves topicality, either the inherent topicality associated with animacy or the pragmatic topicality associated with the discourse topic. The effect of *animacy* in licensing AF is illustrated by (8). In both cases, P in the relevant clause is human-referring, while A is non-human (compare (7), where A is human and P is inanimate).

(8) a. Mu s-na' [k'usi ti ik'-oj-on-uk ech'el] ti prove tseb-e.
Z TSO NEG ERG3-know what DET carry-PF-AF-IV away DET poor girl-ENC
'The poor girl didn't know what had carried her away.' {OCK 317}

b. k'usi chanul x-ti'-on tajmek ti vo'ne jch'ultottik un-e.
what animal ASP-eat-AF very DET ago our.Lord PT-ENC
'whatever animal would eat our Lord long ago.' {OCK 235}

The effect of *discourse topicality* is illustrated by (9a,b). In both cases, P is the local discourse topic, as suggested by the fact that it is realized by a null pronoun, and A is non-referential.

(9) a. Muk' buch'u x-k'ot ik'-on-uk.
Z TSO NEG who ASP-arrive call-AF-IV
'There was no one to come and call him.' {OCK 42}

b. Pero buch'u x-mil-on?
but who ASP-kill-AF
'But who killed her?' {OCK 230}

In a sense, the use of AF in these contexts can be seen as a device for deflecting what would be the wrong interpretation if A-extraction proceeded from a *transitive* clause. Replacing the AF verbs in these examples with transitive verbs yields only P-extraction readings:

(10) Mu s-na' [k'usi y-ik' ech'el] ti tseb-e.
Z TSO NEG ERG3-know what ERG3-carry away DET girl-ENC
'The poor girl didn't know what s/he had carried away.'

(11) Pero buch'u i-s-mil?
Z TSO but who CP-ERG3-kill
'But who did s/he kill?' (s/he = local topic)

The role of the local context in determining topicality is crucial here. (11) is not ambiguous in its context, but an example like (6) – when presented out of context in an elicitation situation – is. The reason is that (6) has no local topic, and therefore topicality plays no role in its interpretation.

Crucially though, and unlike (2) and (3b), the interpretations of (10) and (11) with P-extraction do not reflect a constraint on ergative extraction. They reflect instead a more general constraint on the interpretation of transitive clauses with two 3rd person arguments (*3–3 clauses*). This constraint forces readings in which the more salient argument is interpreted as the external argument (A) and the less salient one as the internal argument (P) (Aissen 1997).[4]

(12) Obviation Principle: In a transitive 3–3 clause, the more salient argument aligns with A and the less salient argument with P.

The functional motivation for (12) is the fact that when A and P are both 3rd person, they are not formally distinguished in transitive clauses in Tsotsil: they are not case-marked nor can they be distinguished by agreement since they do not differ in grammatical person. Thus the default alignment of topicality with grammatical function serves as a guide to interpretation. To express clauses in which P is more salient than A, Z Tsotsil resorts to various grammatical devices. In the context of agent extraction, that device is the AF construction.

Evidence that the distribution of AF clauses in Z Tsotsil is determined by (12) is the fact that they are possible only when both arguments are 3rd person (Haviland 1981:272; Aissen 1999).

[4] Whether this is a grammatical principle or a heuristic used by speakers to aid in interpretation is not crucial to present concerns. See Zavala (2007), Curiel (2007), Bohnemeyer (2009), and Polian (2013) for discussion of related effects in other Mayan languages.

There simply are no forms with the absolutive morphology that would be required if A (13a) or P (13b) were 1st or 2nd person.

(13) a. *Vo'on l-i-maj-on.
Z TSO PRO.1S CP-ABS1-hit-AF
('It was me that hit him/her/it/them.')

b. *K'usi l-a-tij-on?
what CP-ABS2-awaken-AF
('What woke you up?')

If AF occurs only when the corresponding transitive cannot be interpreted correctly, there is no motivation for it in (13): there is no ambiguity as to grammatical function in transitive clauses where one argument or the other is 1st or 2nd person since these are fully disambiguated by the agreement morphology.

Rather than build these person restrictions into the formal account of AF clauses, I assume that they *emerge* as a consequence of the fact that the relation between AF and TV clauses involves *competition,* a view also espoused in Aissen (2003), Stiebels (2006), and Erlewine (2016). Various constraints can determine the distribution of the two constructions across Mayan, including both pragmatic ones (like the Obviation Principle (12)) and morphological ones (Stiebels 2006, Aissen to appear). The competition between the two in Tsotsil is informally schematized in Tableau 1, OBVIATION refers to (12); *AF is a markedness constraint penalizing the AF construction; TV and AF refer to clauses headed by transitive and AF verbs, respectively:

a=3rd, P=3rd	OBVIATION	*AF
TV	*!	
⇒AF		*

Tableau 1

If the TV candidate is faithful to the input, it violates (12). But if P (or A) is a local person (1st or 2nd person), (12) is irrelevant, since it only references 3–3 clauses. In that case, the AF candidate is excluded by *AF.

a=3rd, P=local	OBVIATION	*AF
⇒ TV		
AF		*!

Tableau 2

In their restriction to *aP* settings, the distribution of AF clauses resembles that of passive, which plays a parallel function in clauses *not* involving agent extraction (see Aissen 1997 for discussion and examples).[5]

[5] Passive has a somewhat wider distribution than does AF (even leaving aside the requirement of A-extraction in AF clauses): it is possible with P's of any person, and the agent may be syntactically suppressed. See §4.3.1 below.

3.2 Object raising

Besides sharing the distribution of passive clauses, AF clauses in Z Tsotsil share some syntactic features with passive. In particular, the internal argument is the 'subject', i.e., it occupies the highest argument position in the clause. Tsotsil does not, to my knowledge, have raising and control constructions that can serve to identify the subject, but agreement morphology provides a useful probe: in AF clauses, the only argument that can be indexed by agreement morphology is P.

Agreement in AF clauses is highly limited since both arguments are 3rd person, and there is no ABSOLUTIVE 3rd person marker. However, the 3rd person plural suffix, *-ik,* does occur in AF clauses. The important observation is that it can index only the internal argument.[6] Consequently, (14a) can only be interpreted with a plural P; the number of A is unspecified. To clarify that A is plural, a plural suffix can be added to the interrogative pronoun, (14b):

(14) a. Buch'u ch-'ik'-on-*ik* ech'el ta poxtael?
Z TSO who ICP-take-AF-PL DIR to be.cured
'Who (sg/pl) is going to take them to the clinic?'

b. Buch'u-tik ch-'ik'-on ech'el ta poxtael?
who-PL ICP-take-AF DIR to be.cured
'Who all will take him/her to the clinic?'

Combining the two plural markers yields (14c) which is interpreted only with both a plural A and a plural P:

c. Buch'u-tik ch-'ik'-on-*ik* ech'el ta poxtael?
who-PL ICP-take-AF-PL DIR to be.cured
'Who all took them to the clinic?'

In passive clauses, plural suffixes also index the internal argument, as do ABS markers (P is not restricted to 3rd person in passive clauses).

(15) a. . . . x-i-mil-e-otik . . .
Z TSO . . . ASP-ABS1-kill-PSV-ABS.1PL.INCL
'we are[n't] killed [for doing nothing at all]' {OCK 230}

b. Ta la x-'ak'-b-at-ik trago.
ICP CL ASP-give-APPL-PSV-PL cane.liquor
'They were given cane liquor.' {OCK 21}

Further, the interpretation of the plural suffix in AF clauses provides evidence that P raises to a position where it c-commands A. This follows from two properties of *-ik*. The first is that it is

[6] This corrects Aissen (1999) where I said that *-ik* could index either A or P in AF clauses, though with a preference for indexing P.

underspecified for CASE and can therefore index ergatives (16a) as well as absolutives, whether P (16b) or S (16c) (the indexed plural is italicized in the translation; no emphasis is implied):

(16) a. I-s-maj-ik (ergative)
Z TSO CP-ERG3-strike-PL
'*They* struck him/her/it/them.'

b. I-j-maj-ik. (absolutive P)
CP-ERG1-strike-PL
'I hit *them*.'

c. I-bat-ik. (absolutive S)
CP-go-PL
'*They* went.'

The second is that while *-ik* is unspecified for Case, it can only index the highest direct 3rd person argument in a clause. *-ik* indexes P in (16b), but this is possible only because the subject is not 3rd person. When both A and P are 3rd person, *-ik* can only index A (Bricker 1977 and confirmed through elicitation). Hence (16a) must have a plural A; the plurality of P is under-determined. I assume then the generalization in (17) for Z Tsotsil:[7]

(17) *-ik* indexes the highest non-oblique 3rd person argument in the clause.

(17) accounts for (16a-c) and is consistent with (15). It also accounts for plural marking in AF clauses like (14), but only if P raises to a position from which it asymmetrically c-commands A. Since it is likely that P in canonical passives likewise raises, this is another property that the two constructions share.

Raising of P in AF clauses is consistent with the fact that A cannot bind P or into P in AF clauses, as A does not c-command P. If A binds P, then extraction of A must proceed from a transitive clause (where A does c-command P), (18a); i.e., there are no reflexive AF clauses, (18b).

(18) a. Oy much'u i-s-jip s-ba ta vo'.
Z TSO ∃ who CP-ERG3-throw GEN3-RR in water
'Someone threw himself into the water.'

b. *Oy much'u i-jip-on s-ba ta vo'.
∃ who CP-throw-AF GEN3-RR in water
('Someone threw himself into the water.') {Aissen1999:474}

Further, if A binds *into* P, then extraction of A must again proceed from a transitive clause, (19a). (19b) is ungrammatical.

[7] Under some morphological conditions, *-ik* can also index a 2nd person plural.

(19) a. Mu'yuk much'u i-s-tam s-tuk'.
Z TSO NEG who CP-ERG3-grab GEN3-rifle
'No one grabbed his rifle.'

b. *Mu'yuk much'u i-tam-on s-tuk'.
NEG who CP-grab-AF GEN3-rifle
('No one grabbed his rifle.') {Aissen 1999:474}

Although raising of the internal argument is shared in Z Tsotsil by AF and passive, that does not entail, of course, that AF clauses *are* passive. There are still salient differences in properties of the agent: in canonical passive clauses, the agent is optional and when it occurs it must be oblique; in AF clauses, the agent is obligatory and is not oblique. Furthermore, in AF clauses, by definition, the agent is extracted, while in canonical passives, it need not be. (20a, b) show some passive clauses in Z Tsotsil. The agent is realized in one of two ways in Z Tsotsil: individuated agents are presented by the relational noun *yu'un*, (20a), and relatively unindividuated ones by the preposition *ta*, (20b).

(20) a. Te la ch-mak-e ta be yu'un li vakax un-e.
Z TSO there CL ICP-stop-PSV on road OBL DET cow PT-ENC
'He was stopped there on the road by a cow.' {OCK 227}

b. . . . ti x-ti'-at ta chon un-e.
. . . that ASP-eat-PSV OBL animal PT-ENC
'[it was known] that he would be eaten by animals' {OCK 81}

What makes a passive analysis of AF clauses compelling in Tsotsil is the fact that AF morphology and passive morphology show significant syncretism: as noted earlier, in one dialect, AF morphology has been extended to passive clauses, and in others, passive morphology has replaced the AF suffix. We document these two developments below, starting with the extension of AF *-on* to passive clauses in Z Tsotsil.

4.3 Voice suffixes

The historic passive and AF suffixes for Tsotsil are listed in Table 2. The passive suffix *-e* is restricted: it attaches only to to monosyllabic (i.e., root) transitive stems and, to my knowledge, it occurs only in Z Tsotsil. The AF suffix, as seen above, is *-on*. This suffix is only found in Z Tsotsil.

PASSIVE	*-at, -e*	*-e* only in Z Tsotsil and only on CVC STEMS
AF	*-on*	*-on* only in Z Tsotsil, not restricted to CVC stems.

Table 2: Historic passive and AF suffixes in Tsotsil

4.3.1 Zinacantec Tsotsil *-on*

The first observation is that while the agent is not oblique in AF clauses like (21a), it is possible to add in the oblique marker (underlined), as in (21b). This is possible for all speakers I have consulted and for some, it is strongly preferred (anticipating the passive analysis for (21b), I gloss *-on* as PSV).

(21) a. A li kremotike$_i$ mu s-na' [much'u$_i$ ta sa'-on-ik].
Z TSO TOP DET boys NEG ERG3-know who ICP seek-AF-PL
'The boys$_i$ don't know who is looking for them$_i$.'

b. A li kremotike$_i$ mu s-na' [much'u$_i$ ta sa'-on-ik <u>yu'un</u> t$_i$].
TOP DET boys NEG ERG3-know who ICP seek-PSV-PL by
'The boys$_i$ don't know who is looking for them$_i$.'
(lit: 'The boys don't know who they are being sought by.')

There are two analytical possibilities for (21b). One is that it is an AF clause and that AF clauses come in two varieties, one in which the agent is direct and one in which it is oblique. In this analysis, *-on* is always associated with A-extraction and there is some flexibility in how the arguments are licensed. The other is that it is a passive clause, i.e., that *-on* has been reanalyzed as a passive suffix, disassociated from A-extraction. In this analysis, (21b) is a passive, with extraction of the oblique agent.

The second analysis appears in fact to be the correct one, as *-on* has been extended to passive functions in Z Tsotsil and is no longer restricted to contexts of A-extraction. It occurs in plain passives:

(22) a. Ta sa'-on *(<u>yu'un</u>) yajnil li Manvel-e.
Z TSO ICP seek-PSV OBL his.wife DET Manuel-ENC
'Manuel$_i$'s wife is looking for him$_i$' (lit: 'Manuel$_i$ is being looked for by his wife.')

b. I-ts'ites-on *(<u>yu'un</u>) Petu' li Maruch-e.
CP-raise-PSV OBL Petrona DET Maria-ENC
'Petrona raised Maria' (lit: 'Maria was raised by Petrona.')

And it occurs in passives involving extraction of a non-agent ((23c) is a text example):

(23) a. Yu'un chopol li pox i-'ak'-b-on *(<u>yu'un</u>) li Pablo-e.
Z TSO because bad DET liquor CP-give-APPL-PSV OBL DET Pablo-ENC
'Because the corn liquor Pedro gave him was bad' = '...that he was given by Pedro...'

b. Li Maruch-e s-jak' k'usi ora i-'il-on *(<u>yu'un</u>) li Petul-e.
DET Maria-ENC ERG3-ask what hour CP-see-PSV OBL DET Petul-ENC
'Maria$_i$ asked when Petul had seen her$_i$' = '... when she had been seen by Pedro.'

c. ...ti much'u vayem ta=x-ta-on yu'un ti j'elek'-e.
...DET who asleep ICP-find-PSV OBL DET thief-ENC
'[You are not like] the one who is found asleep by the thief.'
(But you, brothers, are not in darkness, that this day should overtake you as a thief would.) {1 Thessalonians 5, 4}

In (22a,b) and (23a-c), the agent is not extracted and must be presented as oblique. Given the wide distribution of *-on* as a (pure) passive marker, there is no reason to think then that (21b) involves anything other than extraction of the oblique agent from a passive clause.

In Z Tsotsil then, *-on* occurs in two surface structures which share several core properties: morphosyntactic intransitivity, two arguments, and raising of the internal argument. They are distinct in their realization of the agent: in passives, the agent is realized as an oblique, but in AF it is realized as direct and obligatorily extracted. These features are summarized in Table 3:

	PASSIVE	AF
TRANSITIVITY	intransitive	intransitive
INTERNAL ARG	subject	subject
EXTERNAL ARG	oblique	direct, extracted
SUFFIX	*-on*	*-on*

Table 3: Properties of *-on* clauses in Z Tsotsil

Table 3 suggests the analysis I propose here: that AF clauses are a type of passive,[8] one which has the means to Case-license the external argument, but only if it is extracted. This could be implemented in a feature-checking framework by positing that the head which defines a passive clause optionally carries an unchecked feature associated with extraction. Following Stiebels (2006), whose analysis I build on here, I refer to this feature as [FOC]. This feature is *selectional* in the sense that that it requires that its specifier, the position of the external argument, be filled by a DP with a valued version of the same feature. The head and its specifier enter into an Agree relation which checks the feature on the head and in return Case-licenses the argument. In this respect, the FOC feature functions like a phi-feature. The [FOC] feature on the agent itself either directly or indirectly insures that it will be extracted; being Case-licensed through WH-agreement insures that it will surface as a direct argument. In Z Tsotsil, the head spells out as *-on,* as in (24):

(24) v_{PSV} => *-on*
([*u*FOC])

A plausible historical scenario is that the restricted distribution of AF clauses in Tsotsil triggered their reanalysis as passive. In order to reconcile the surface form of AF clauses with the passive analysis, special provision was required to account for the appearance of the extracted agent as a

[8] An antecedent for this analysis is Laughlin (1975), which identifies *-on* as 'passive (3rd person subject and object)' (p. 26). Laughlin provides no examples or discussion, so it is unclear whether he has in mind AF clauses, transparent passives like (22)-(23), or both.

direct argument, rather than oblique. That provision is the capacity of the [FOC] feature to Case-license the agent.

What gives rise to the transparently passive structures in (21b), (22), and (23) is the fact that the [FOC] feature is only *optionally* associated with *-on* in the contemporary language. Consequently, *-on* is also the spell-out of a passive head without that feature. In that case, the clause is not limited to contexts of agent extraction and, correspondingly, the agent is not Case-licensed by the verbal head. It therefore requires oblique marking.

In identifying AF in Tsotsil with the *-on* passive, I am assuming then that there is a single passive suffix *-on* which occurs in two related structures, one which involves WH-agreement (AF clauses) and one which does not. Evidence for unifying the analysis of the two constructions comes from the fact that they are subject to the same restrictions, restrictions which in fact distinguish them from other passives in the language. In this respect, the situation is crucially different from that of Chamorro, where Chung (1998) shows that despite the use of the same morphology to mark passive and WH-agreement (with the internal argument), the two constructions are subject to different restrictions and are therefore syntactically distinct.

First, in contrast to canonical passives where the agent is an optional adjunct, the agent in *-on* passives is obligatory. Thus, agentless passives with *-e, -at* are standard and unproblematic, but speakers consistently reject or judge as degraded agentless passives in *-on*:[9]

(25) I-maj-{at / -e / *-on} li Manvel-e.
Z TSO

CP-hit-PSV DET Manuel-ENC
'Manuel was hit.'

This property is shared with AF clauses, where the agent must (by definition) be syntactically realized. Second, *-on* passives are restricted to clauses in which both arguments are 3rd person, just as AF clauses are. That the person of the agent is restricted is not surprising since this is true of all passives in the language (Aissen 1997). But the restriction of the internal argument to 3rd person is a feature which distinguishes *-on* passives from passives formed with *-e* and *-at*. As with AF verbs, *-on* passives cannot carry the absolutive markers that would be required if the internal argument were 1st or 2nd person.

(26) L-i-tsak-{at / -e / *-on} ta j'ik'al.
Z TSO

CP-ABS1-grab-PSV OBL spook
'I was grabbed by a spook.'

[9] I have identified a few text examples of agentless *-on* passives (cited in Aissen 1999:481). Such examples suggest there may be speaker variation on this point and that *-on* passives may be evolving to canonical ones.

These two properties then – the obligatory agent and the restriction to 3–3 clauses – support an analysis which unifies the two surface structures in which *-on* appears.

It remains to account for these properties and for the differences between passives formed with *-on* and those formed with the other suffixes. Here I can only sketch a possible approach. Starting with the fact that the agent in *-on* passives is obligatory, I assume, following others, that the external argument in canonical passives (those formed in Tsotsil with *-e, -at*) is existentially bound and that the optional agent phrase is an adjunct which serves to identify that argument. Then the distinction between canonical passives and *-on* passives is one of logical form: in *-on* passives, the external argument is not existentially bound, and therefore must be realized syntactically. This is independently necessary for AF. Hence if AF clauses are a special instance of *-on* passive, per (24), it follows that the agent must be obligatory in *-on* passives as well.

Locating the difference between passives formed with *-on* and those formed with the other suffixes in their logical form provides a way to approach the 3–3 restriction in *-on* passives. I suggested earlier that the person restriction in AF clauses emerges as a consequence of the fact that they compete with active transitives and are optimal only when the Obviation Principle (12) is violated by the corresponding transitive (Tableau 1). If we generalize this approach to all clauses involving *-on,* then *-on* passives too will be limited to clauses in which both arguments are 3rd person. This implies that *-on* clauses of both types exist only to remedy violations of (12), a prediction which needs further exploring, but seems to be on the right track. The fact that other passives are not similarly constrained must then be due to the fact that they do not compete directly with active transitives and would simply be close paraphrases of the corresponding actives. This would follow in a principled way from the assumption that only candidates with the same logical form belong to the same candidate set (Grimshaw 1997, Heck et al. 2002). Per the above discussion, *-on* passives (and AF clauses) are like active transitives in that they do not involve existential binding of the external argument, while passives formed with the other suffixes do.

4.3.2 AF in other dialects

In addition to Z Tsotsil, there are several other major dialect areas, including Chamula, Venustiano Carranza, Chenalho', and Huixtán. As noted earlier, *-on* does not appear to exist as an AF suffix (or a passive suffix) outside of Z Tsotsil. All other dialects use the historic passive suffix *-at* in contexts where Z Tsotsil uses the AF suffix *-on*. In some dialects (or for some speakers of some dialects), extraction of the external argument in an *aP* setting simply proceeds from a canonical passive clause, with the agent marked as oblique.

I have interviewed three Chamulan (C) speakers (from several different towns), and none of them uses the historic AF construction with *-on*.[10] Some say they recognize the construction, but they associate it with other dialects or with the neighboring language Tseltal (which is incorrect), others do not recognize it at all. These speakers instead use the passive suffix *-at* in AF contexts and they require the presence of the oblique enclitic *'o in* such cases:[11]

[10] This is consistent with Robertson (1977), which reports that AF is not used in Chamulan Tsotsil.

[11] *'o* is used widely in Tsotsil to mark extracted obliques, generally ones interpreted as instrument, cause, or agent (Haviland 1981:132).

(27) Buch'u i-tij-at 'o a-me'-e?
C TSO who CP-awaken-PSV OBL GEN2.mother-ENC
'Who woke up your mother?' (lit: 'By whom was your mother awakened?')

The enclitic *'o* is obligatory in (27), indicating that the agent extracts as an oblique, not as a direct argument, i.e., (27) is a passive clause. It appears that these speakers have lost both the historic AF suffix *-on* and WH-Agreement, and the agent is realized as oblique.

I have interviewed one speaker from Venustiano Carranza (V). For him, the agent in an *aP* setting extracts as an oblique flagged by the relational noun *yu'un*. Further, this speaker requires pied piping (with inversion) of *yu'un* (Aissen 1996).

(28) Much'u yu'un maj-ot te tseb-e?
V TSO who OBL strike-PSV DET girl-ENC
"Who hit the girl?' (lit: 'By whom was the girl hit?')

However, in several other dialects, the characteristic features of AF persist, though they are found now in clauses headed by the historic passive suffix *-at*, not *-on*. Both Chenalhó (CH) Tsotsil and Huixtec (H) Tsotsil show this pattern.[12] These dialects all have the single passive suffix (*-at*) and when the agent is not extracted, it is realized as an oblique with *yu'un*.

(29) I-mil-at yu'un ti Herodes-e
H TSO CP-kill-PSV OBL DET Herod-ENC
'they were killed by Herod' {Matthew 2,15}

When the agent *is* extracted, there are two patterns. When A is as topical as P (*Ap* and *ap* settings), extraction of A proceeds from a transitive clause. In (30a), A is human and P is inanimate (=*Ap*); in (30b), both A and P are human, but P is not-topical (=*Ap* or *ap*).

(30) a. Mu'yuk xa much'u ts-tij arpa...
H TSO NEG CL who ICP.ERG3-play harp
'There will be none who play the harp...' {Revelation 18,22}

b. ...ti much'u-tik ch-k'uxubin-ik yantik-e.
DET who-PL ICP.ERG3-pity-PL others-ENC
'[blessed are] the merciful (lit: those who pity others).' {Matthew 5,7 }

[12] The Chenalho' data are from an on-line translation of the New Testament (NT). Data from Huixtec Tsotsil comes from work with two speakers and from the on-line translation of the NT. One could ask whether a translation of the NT is a reliable representation of Tsotsil (in any genre). In some respects, it is quite different. E.g., there are more instances of agent extraction in the translation than otherwise occur in Tsotsil (e.g. *he who, they who, those who* are all translated as headless relatives, with extraction). But this does not make the material irrelevant. What is relevant is this: once the agent is extracted, what morphosyntax is associated with that extraction? In fact, the morphosyntax varies between the AF form and the transitive form just as in forms of Tsotsil which are more authentic and this variation is conditioned by the same factors.

When P is more topical than A, i.e., in *aP* settings, extraction proceeds from a clause in which the verb is suffixed with *-at*. In (31a,b), A and P are both human, but P is the local topic (=*aP*).

(31) a. Mu'yuk boch'o ta=x-contrain-at tey.
CH TSO NEG who ICP-join-AF there
'Nobody joined him there.' {John 4, 44}

b. . . . mu'yuk boch'o xu' ch-poj-b-at.
. . . NEG who can ICP-remove-APPL-AF
'[Maria$_i$ has chosen that good part], thus no one shall take it away from her$_i$.'
{Luke 10,42}

At first glance, it looks like Chenalhó and Huixtec Tsotsil have simply lost AF, replacing it with passive, as in Chamula and Venustiano Carranza. It is true that these dialects have lost the historic AF formative *-on*, but they have retained the syntax associated with *-on*, i.e., WH-agreement with the agent, associating it now though with the historic passive suffix, *-at*. Thus, (32) from Z Tsotsil and (33) from Huixtec Tsotsil are equivalents:

(32) K'usi i-tij-on ame'?
Z TSO what CP-awaken-AF your.mother
'What woke your mother up?'

(33) K'usi i-tij-at ame'?
H TSO what CP-awaken-AF your.mother
'What woke your mother up?'

First, the agent in (31a,b) and (33) is extracted as a *direct* argument, not as an oblique (there is no oblique marker, e.g. *yu'un* or *'o* in those examples). This is only possible because the agent is extracted. When it is not extracted, as in (29), the agent must (if expressed) be oblique. Further, under this analysis, (31a,b) and (33) *are* AF clauses. Since AF clauses only arise through competition with transitives when the transitive would violate the Obviation Principle (Tableau 1), this construction should be possible only when both arguments are 3rd person. Again, the person of A is not at issue since passives in Tsotsil only permit 3rd person agents in any case. But since canonical passives occur with P's of any person, it is significant that in the Tsotsil of Chenalhó and Huixtán, realization of A as a direct argument in a passive clause is possible only if it is extracted and only if P is 3rd person. If P is 1st or 2nd person, the agent must be realized as oblique; omitting the oblique marker in (34) results in ungrammaticality.

(34) K'usi n-a-tij-at *(yu'un)?
H TSO what CP-ABS2-awaken-PSV OBL
'What woke you up?' (Lit: 'What were you woken up by?')

In short, in Chenalhó and Huixtec Tsotsil, the historic passive suffix *-at* has replaced *-on* in AF clauses.

(35) v_{PSV} => *-at* (CH, H Tsotsil)
([FOC])

There are two features then which define AF clauses in Tsotsil: they have all the features of passive except that [i] the agent is an obligatory argument and [ii] it enters into WH-agreement with the voice head. The various dialects of Tsotsil differ along two dimensions: [1] whether they have retained WH-agreement in passive clauses; and [2] how the passive head with WH-agreement is pronounced. In Z Tsotsil it is spelled out by *-on,* in the other dialects, by *-at*. The extension of *-on* to passive clauses without agent extraction and the extension of *-at* to AF clauses suggests strongly that speakers identify AF with passive in this language.

4. Conclusion

The loss of AF in some dialects of Tsotsil appears to have resulted from the reanalysis of AF as passive + WH-agreement, followed by loss of the WH-agreement option. The reanalysis of AF as passive in the first place presumably resulted from the fact that AF has close to the same distribution in clauses with agent extraction that passive has in clauses without it. This evolutionary scenario is shown in Table 4.

Stage 1	Restriction of AF to *aP* settings	all dialects of Tsotsil
Stage 2	Reanalysis of AF as passive + WH-agreement	all dialects of Tsotsil
Stage 3	Loss of WH-agreement (and therefore of AF)	Chamula, Venustiano Carranza, some speakers of Z Tsotsil.

Table 4: Evolution of AF in Tsotsil

Having lost AF in most contexts, but not all (Stage 1), Tsotsil is transitional between the Mayan languages that have AF constructions and those which do not. Some dialects retain an AF construction (those still at Stage 2), while others do not (those which have progressed to Stage 3). Given this variation, the language provides a model of how reanalysis of AF as passive can lead to its loss and as such, it might shed light on how AF was lost more widely in Mayan.

References

Aissen, Judith. 1996. Pied piping, abstract agreement, and functional projections in Tsotsil. *Natural Language & Linguistic Theory* 10:375-414.

Aissen, Judith. 1997. On the syntax of obviation. *Language* 73:705-750.

Aissen, Judith. 1999. Agent focus and inverse in Tzotzil. *Language* 75: 451-485.

Aissen, Judith. 2003. Differential coding, partial blocking, and bidirectional OT. *BLS 29*, eds. Pawel Nowak and Corey Yoquelet, 1-16. Berkeley: Berkeley Linguistics Society.

Aissen, Judith. to appear. Correlates of ergativity in Mayan. *OUP handbook of ergativity,* eds. Jessica Coon, Diane Massam, and Lisa Travis. Oxford: Oxford University Press.

Berinstein, Ava. 1985. *Evidence for multiattachment in K'ekchi Mayan*. New York: Garland.

Bohnemeyer, Jürgen. 2009. Linking without grammatical relations in Yucatec: Alignment, extraction, and control. *Issues in functional-typological linguistics and language theory: A Festschrift for Christian Lehmann on the occasion of his 60th birthday*, eds. Y. Nishina, Y. M. Shin, S. Skopeteas, E. Verhoeven, and J. Helmbrecht, 185-214. Berlin: Mouton de Gruyter.

Bricker, Victoria. 1977. *Pronominal inflection in the Mayan languages*. New Orleans: Middle American Research Institute, Tulane University.

Campana, Mark. 1992. A movement theory of ergativity. PhD dissertation. Montreal: McGill University.

Chung, Sandra. 1998. *The design of agreement. Evidence from Chamorro*. Chicago: The University of Chicago Press.

Coon, Jessica. 2016. Mayan morphosyntax. *Language and linguistics compass* 10(10): 515-550.

Coon, Jessica, Pedro Mateo Pedro, and Omer Preminger. 2014. The role of Case in A-Bar extraction asymmetries: Evidence from Mayan. *Linguistic Variation* 14: 179-242.

Craig, Colette. 1977. *The structure of Jacaltec*. Austin: University of Texas Press.

Craig, Colette. 1979. The antipassive and Jacaltec. *Papers in Mayan Linguistics,* ed. L. Martin, 139-165. Colombia, MO: Lucas Bros. Publishers.

Curiel, Alejandro. 2007. Estructura de la información, enclíticos y configuración sintáctica en tojol'ab'al. MA thesis. Mexico, D.F.: CIESAS.

Dayley, Jon. 1985. *Tz'utujil grammar*. Berkeley: University of California Press.

England, Nora. 1983. *A grammar of Mam, a Mayan language*. Austin: University of Texas Press.

Erlewine, Michael Yoshitaka. 2016. Anti-locality and optimality in Kaqchikel Agent Focus. *Natural Language & Linguistic Theory* 34:429-479.

Grimshaw, Jane. 1997. Projections, heads, and optimality. *Linguistic Inquiry* 28:373-422.

Gutiérrez Bravo, Rodrigo. 2015. *Las cláusulas relativas en maya yucateco*. Mexico, D.F.: El Colegio de México.

Haviland, John. 1981. *Sk'op Sotz'leb: el tzotzil de San Lorenzo Zinacantán*. Mexico, D.F.: UNAM.

Heck, Fabian, Gereon Müller, Ralf Vogel, Silke Fischer, Sten Vikner, and Tanja Schmid. 2002. On the nature of the input in optimality theory. *The Linguistic Review* 19:345-376.

Larsen, Thomas and William Norman. 1979. Correlates of ergativity in Mayan grammar. *Ergativity: Towards a theory of grammatical relations,* ed. Frans Plank, 347-370. New York: Academic Press.

Laughlin, Robert. 1975. *The great Tzotzil dictionary of San Lorenzo Zinacantán*. Washington, D.C.: Smithsonian Institution Press.

Laughlin, Robert. 1977. *Of cabbages and kings*. Washington, D.C.: Smithsonian Institution Press.

Mondloch, James. 1981. Voice in Quiché-Maya. PhD dissertation. Albany: SUNY.

Ordóñez, Francisco. 1995. The antipassive in Jacaltec: a last resort strategy. *CUNYForum: Papers in Linguistics* 19:95-108.

Pascual, Adán. 2007. Transitividad y dependencia sintáctica y discursiva en q'anjob'al. MA thesis. Mexico, D.F.: CIESAS.

Polian, Gilles. 2013. *Gramática del tseltal de Oxchuc*. Mexico D.F.: CIESAS.

Robertson, John. 1977. A proposed revision in Mayan subgrouping. *International Journal of American Linguistics* 43: 105-120.

Smith-Stark, Thom. 1978. The Mayan antipassive: Some facts and fictions. *Papers in Mayan*

linguistics 2, ed. Nora C. England, 169-187. Columbia, Missouri: University of Missouri Miscellaneous Publications in Anthropology.

Stiebels, Barbara. 2006. Agent focus in Mayan languages. *Natural Language & Linguistic Theory* 24: 501-570.

Tonhauser, Judith. 2003. Agent focus and voice in Yucatec Maya. *Proceedings from the 39th Annual Meeting of the Chicago Linguistic Society* 1: 540-558. Chicago: Chicago Linguistic Society.

Watanabe, Akira. 1996. *Case absorption and WH-Agreement*. Dordrecht: Kluwer Academic Publishers.

Zavala, Roberto. 1992. *El kanjobal de San Miguel Acatán*. Mexico, D.F.: UNAM.

Zavala, Roberto. 2007, Inversion and obviation in Mesoamerica. *Endangered languages,* eds. Peter Austin and Andrew Simpson, 267-305. *Linguistische Berichte Sonderheft* 14. Hamburg: Helmut Buske Verlag.

THE INTERACTION BETWEEN INFIXATION AND REDUPLICATION IN CHAMORRO*

Boris Harizanov
Stanford University
bharizan@stanford.edu

1. Introduction

On the basis of Chamorro, an Austronesian language spoken in the Mariana Islands, I provide an argument for an old idea in the study of morphophonology: modeling opacity via serially ordered derivations. The evidence comes from the interaction between infixation and reduplication in Chamorro. I demonstrate that this interaction in the language is opaque (in the sense of Kiparsky 1971, 1976, et seq.) and that it can be understood within a derivational/serial framework where the output of reduplication serves as the input to infixation. The analysis is implemented in Distributed Morphology (Halle and Marantz 1993, 1994, et seq.). Two assumptions, commonly held within this framework but expressible in other frameworks as well, prove essential in the description and analysis of infixation and reduplication in Chamorro: that the association of morphosyntactic terminals with phonological material (i.e. Vocabulary Insertion) proceeds root-outwards and that at least some morphosyntactic terminals delimit cycles for the application of phonological rules that they may trigger. These results do not necessitate a rule-based framework rules and are, in fact, compatible with constraint-based frameworks as long as they adopt multiple derivational stages, such as Stratal Optimality Theory (Kiparsky 2000, 2003, Rubach 2000) and Optimal Interleaving (Wolf 2008). Finally, the analysis presented here relies on a particular conception of aspects of Chamorro clause structure and, in particular, the order in which inflectional material is syntactically composed with a verbal root. Thus, to the extent that the analysis is successful, it provides morphophonological support for structural assumptions about the extended verbal projection that happen, for independent reasons, to be difficult to motivate syntactically in Chamorro.

Section 2 of this paper introduces the phenomenon of infixation as instantiated in Chamorro, including its interaction with a particular vowel fronting alternation, while section 3 outlines the relevant properties of reduplication in Chamorro. Section 4 describes how this type of reduplication interacts with infixation and an analysis of the interaction is offered in section 5. The analysis is followed by a brief note on the difficulties that this interaction, in particular, and opaque interactions, in general, present to non-serial frameworks.

* I would like to express my deepest thanks to Sandy for teaching me everything I know about Chamorro (in addition to a lot more over the years) as well as for giving me the opportunity to contribute to the Chamorro dictionary project – one of the most rewarding experiences I had at UC Santa Cruz. For valuable feedback on this work, I would like to thank Arto Anttila, Lev Blumenfeld, Vera Gribanova, Junko Ito, Paul Kiparsky, Armin Mester, Andrew Nevins, Mark Norris, and the Stanford P-Interest group. Ryan Bennett deserves special thanks for carefully reading numerous early drafts and for providing extensive and extremely helpful feedback.

2. Infixation

2.1. The infixes *-um-* and *-in-*

There are two infixes in Chamorro, *-um-* and *-in-*, each of which can be the exponent of a number of distinct morphosyntactic elements (Topping and Dunga 1973:170ff.):[1]

(1) Grammatical functions of *-um-*
- a. singular or dual subject agreement for intransitive predicates in the realis mood (Chung 2014, ch. 2; cf. Topping and Dunga 1973:185,225 on *-um-* as a "verbalizer")
- b. singular or dual subject agreement for intransitive predicates in the infinitive and invariant marking for transitive predicates in the infinitive (Topping and Dunga 1973:185, Chung 1998:64)
- c. subject *wh*-agreement (Cooreman 1987, Chung 1998:201; cf. Topping and Dunga 1973:184 on -um- as an "actor focus infix")

(2) Grammatical functions of *-in-*
- a. passive with transitive predicates (Chung 1998:37)
- b. object *wh*-agreement with transitive predicates in a nominalization of the predicate (Chung 1998:240; cf. Topping and Dunga 1973:187 in *-in-* as a "goal focus infix")
- c. nominalizer (Topping and Dunga 1973:170,187)

These infixes surface just before the nucleus of the left-most syllable of the stem they combine with. This behavior is observed both when the stem begins with an onsetless syllable and when it begins with a consonant or a consonant cluster (Topping and Dunga 1973:170; Halle 2001:160ff.; Klein 2005:973ff.; Chung 2014, ch. 28):[2]

(3) a. Infixation of *-um-*

Predicate		*-um-* infixation
kɒti	'cry'	k**um**ɒti
peska	'fish'	p**um**eska
dɒŋkulu	'big'	d**um**ɒŋkulu
metgut	'strong'	m**um**etgut
liʔiʔ	'see'	l**um**iʔiʔ
hɒtsa	'lift'	h**um**ɒtsa
tristi	'sad'	tr**um**isti
adzao	'borrow'	**um**adzao

[1] Another infix, -Vl- (where "V" stands for "vowel"), which is unproductive and which I am not concerned with, is found in some onomatopoeic words (Chung 2014, ch. 28): e.g., ***pal****aspas* 'splash' (from *paspas* 'splash').

[2] In these examples, the infixes are bolded. Since the low vowels, /a/ and /ɒ/, are distinguished under primary stress but merge otherwise (to /a/), and since affixation may trigger stress shift, there might be differences in the pronunciation of the root in affixed and unaffixed forms of the same word (Topping and Dunga 1973, Chung 2014).

b. Infixation of *-in-*

Predicate		*-in-* infixation
patmɒda	'slap'	p**in**atmɒda
bisita	'visit'	b**in**isita
tattidzi	'follow	t**in**attidzi
liʔiʔ	'see'	l**in**iʔiʔ
hassu	'think'	h**in**assu
istotba	'disturb'	**in**istotba
hɒlla	'pull'	h**in**alla
sɒngan	'say'	s**in**angan

2.2. A previous analysis

Topping and Dunga (1973) analyze *-um-* and *-in-* in Chamorro as underlying prefixes—an analysis that presupposes multiple derivational stages and cyclic rule application. Their empirical argument is based on word forms like *fantsinemmaʔ* 'forbidden things' and it goes as follows. Suppose that this word is formed via the successive merger of the root *tsommaʔ* 'forbid' with the nominalizing infix *-in-*, and of the resulting unit with the plural marker *fan-*.[3] Now, like many Austronesian languages (e.g. see Zuraw 2010), Chamorro has a nasal substitution rule triggered by the prefixes *man-* and *fan-*: when the /n/ of these prefixes is immediately followed by a voiceless consonant, the /n/ assimilates to that consonant in place of articulation, while the consonant itself is deleted (Topping and Dunga 1973:48ff.; Chung 2014, ch. 29)—e.g. *mampɒliʔ* 'priests' from *man-+pɒliʔ* and *maŋeʔlu* 'siblings' from *man-+tseʔlu*. Topping and Dunga (1973) observe that nasal substitution does not apply in *fantsinemmaʔ*; if it did, the plural form of the nominalization would instead be **faŋinemmaʔ*. The suggested explanation is that *-in-* is initially composed with the root as a prefix and that this blocks the application of nasal substitution once *fan-* is added. It is then only after *fan-* has been added to the stem that *-in-* reaches its surface position, as an infix to *tsommaʔ*. On the surface, the environment for nasal substitution is present, but the rule no longer has a chance to apply due to the following rule ordering:[4]

[3] The change of vowel quality from /o/ in the root to /e/ in the derived word is due to the umlaut alternation (see also the two bottommost rows of (3b)), triggered by the nominalizing infix *-in-* and discussed in section 2.3.

[4] Topping and Dunga (1973) actually suggest not only that Chamorro -VC- infixes are underlying prefixes but also that they are CV- prefixes which undergo metathesis. The latter claim, however, is orthogonal to the question of underlying prefixhood. Note, in addition, that Halle (2001) argues for the underlying CV- nature of these infixes by observing that the /n/ of the infix *-in-* (see steps 3 and 4 in the main text) does not undergo nasal substitution (p. 162). Based on this fact, he concludes that the infix must be represented as *-ni-* at this stage of the derivation. However, this conclusion is unwarranted, as nasal substitution in Chamorro is only triggered by *man-* and *fan-* and no other affixes (see Klein 2005 for a similar point, as well as additional arguments against Halle's (2001) approach).

(4) Derivation of *fantsinemma?* (Topping and Dunga 1973)

a. Combine the nominalizer *in-* with the stem *tsomma?*
   intsomma?
b. Apply any phonological rules triggered by *in-* (umlaut)
   intsemma?
c. Combine the plural marker *fan-* with the stem *intsemma?*
   fanintsemma?
d. Apply any phonological rules triggered by *fan-* (nasal substitution)
   fanintsemma?
   (nasal substitution does not apply here because the triggering context is not present)
e. Apply the infixation rule
   fantsinemma?

However, the failure of nasal substitution to apply in the case described here can receive an alternative explanation and, thus, does not, on its own, provide unequivocal evidence for the underlying prefixhood idea and the need for multiple derivational stages and cyclic rule application. Specifically, Topping and Dunga (1973) themselves suggest (p. 173) that nasal substitution in general applies only if man- and fan- are combined with morphologically simplex stems: "The regular morphophonemic changes caused by man- will be observed except where another type of affixation, e.g., reduplication has already taken place, as in the case of the derived nouns" (Topping and Dunga 1973:235). In this connection, Klein (2005) further notes that nasal substitution fails to apply at the prefix-prefix boundary in other Austronesian languages as well (e.g. Indonesian). Given this much, nasal substation is, in fact, not expected in the case described above because fan- is combined with a morphologically complex stem. Yet, the failure of nasal substitution to apply here is consistent with the analysis of infixation offered in the present paper, which is in many ways a direct descendant of Topping and Dunga's (1973)—see footnote 12 in section 5.

2.3. Umlaut

Before discussing the interaction between infixation and reduplication, a special property of the infix *-in-* must be described. The infix *-in-* is one of a small set of morphemes in Chamorro that cause the vowel of the immediately following syllable to be realized as a front vowel, while its height is preserved: in the context of any one of these morphemes, /u/ is realized as /i/, /o/ as /e/, and /ɒ/ as /a/ (Topping and Dunga 1973, Topping 1980, Chung 1983, 2014, Crosswhite 1996, Klein 2000, Kaplan 2011). The effects of this vowel fronting process can be seen in the two bottommost rows of (3b) and in the examples in (5).[5] These changes in vowel quality do not spread

[5] The other triggers of this vowel fronting process are the proclitic definite article *i*, the proclitic oblique case marker *ni*, the proclitic local-case marker *gi*, the proclitic subject agreement forms *in* (1.EXCLUSIVE.DUAL or 1.EXCLUSIVE.PLURAL) and *en* (2.DUAL or 2.PLURAL; phonemically /in/), the stress-attracting prefix *mi-* 'full of', the directional prefix *san-*, and the prefix *fa?-* 'make (into)'. When the trigger for vowel fronting is not the infix *-in-*, the process is sensitive to stress in that it only affects the following vowel if it bears primary stress. However, when vowel

further rightward and do not reapply upon subsequent affixation (see Kaplan 2008 on the (non-)iterative status of this process in Chamorro). This phenomenon has been likened both to the German morphophonemic alternation umlaut and to vowel harmony (Safford 1903:294, Conant 1911, von Preissig 1918:6, Topping and Dunga 1973, Topping 1980, Chung 1983, Klein 2000, Kaplan 2011).[6]

(5) Chamorro umlaut triggered by *-in-*

Base		Trigger+base
’kɒnnuʔ	‘eat’	k**i’n**annuʔ
’poʔlu	‘put’	p**i’n**eʔlu
’tugiʔ	‘write’	t**i’n**igiʔ
’dulalak	‘chase’	d**i’n**ilalak
tu’tuhun	‘begin’	t**in**i’tuhun
tu’lajka	‘change’	t**in**i’lajka
’goddi	‘knot’	g**i’n**eddi
’tsomma’	‘forbid’	ts**i’n**emma’

Chamorro umlaut might not be entirely morphologically conditioned but phonologically conditioned at least to some extent as well (see Klein 2000 for an overview and discussion). In particular, the fact that most triggering morphemes contain a front vowel suggests that the language might have had a more general phonological process that assimilated a vowel to any front vowel in the immediately preceding syllable (Costenoble 1940). Synchronically, however, Chamorro umlaut is not a general or completely regular alternation because, for example, its application can depend on properties of both the triggering morpheme and the affected base (see Chung 2014, ch. 29 for discussion of this point; cf. Zuraw 2010 on lexical exceptionality in the context of nasal substitution in Tagalog).

3. Reduplication

Predicates in Chamorro can be associated with one of two aspects: progressive or neutral. The neutral aspect is not signaled by any overt morphophonology while the progressive aspect is signaled by reduplication (Topping and Dunga 1973:191,259). Specifically, the primarily stressed vowel of the predicate and any immediately preceding consonants are doubled. As the following examples demonstrate, the resulting open (C)V syllable bears the primary stress and immediately precedes the original syllable, which bears secondary stress.[7]

---

fronting is triggered by the infix *-in-*, it affects the vowel of the immediately following syllable, regardless of whether it is stressed or not. See Chung 1983 and Klein 2000 for details.

[6] A stressed syllable is preceded by ’.

[7] Syllables that bear primary stress are preceded by ’ and the duplicated (C)V sequences are underlined. Since vowel quality in Chamorro depends on syllabification and stress, the roots in related word forms may differ (see for example the reduplicated forms of *kɒti* and *pɒtgun* in (6)). See also footnote 2.

(6) Chamorro reduplication in the progressive aspect

Predicate		Progressive
ʼsoddaʔ	‘find’	ʼsosoddaʔ
ʼkɒti	‘cry’	ʼkɒkati
huʼgɒndu	‘play’	huʼgɒgandu
kimaʼson	‘burn’	kimaʼsoson
maʼpɒtgiʔ	‘pregnant’	maʼpopɒtgiʔ
ʼpɒtgun	‘child’	ʼpɒpatgun
ʼdɒŋkulu	‘big’	ʼdɒdangkulu

Various approaches to the analysis of Chamorro progressive reduplication have been taken.[8] According to what might be called the traditional view (Topping and Dunga 1973, Topping 1980, Broselow and McCarthy 1983), the (C)V reduplicant is inserted just before the original stressed syllable of the base with concomitant shift of primary stress to the reduplicant. In particular, Broselow and McCarthy 1983 (p. 55ff.) analyze Chamorro progressive reduplication as partial reduplication of the stressed syllable, which involves phonological prefixation of the (C)V reduplicant to that syllable. Stress shift to the reduplicant in this case follows from the assumption that the prefixed (C)V reduplicant is a stress-attracting prefix—an otherwise not uncommon occurrence in Chamorro. An alternative, suggested by Yu 2007 (p. 125) and Inkelas 2008 (p. 386), views Chamorro progressive reduplication as an instance of infixation of the (C)V of the stressed syllable after the stressed vowel of the base (see also Clothier-Goldschmidt 2014). In this case, the reduplicant (C)V immediately follows the original stressed syllable, and no stress shift needs to be posited.

4. The interaction between infixation and reduplication

The interaction between infixation and reduplication is such that the generalization about where the infixes *-um-* and *-in-* appear must make reference to the output of reduplication while the generalization about what gets reduplicated must not make reference to the output of infixation. This finding leads to an understanding of word building which involves multiple derivational stages and cyclic rule application whereby reduplication applies at an earlier stage than infixation. This section describes the interaction in detail, while section 5 spells out an analysis couched within a serial/derivational model of word building.

Infixation and reduplication can cooccur within the same word. In this connection, Topping and Dunga 1973:172,191 observe that infixation applies to *sɒga* ‘stay’ to produce ***sum**ɒga*, as in (7). Now, if this is the base to which progressive reduplication applies, one would expect the hypothetical form ⟨*s**um**ɒma**ga*⟩ to be well-formed. However, it is not; the acceptable progressive

[8] I am not concerned here with another type of reduplication found in Chamorro, which is an instance of derivational morphology and which doubles the final CV of the stem, regardless of the position of stress (Topping and Dunga 1973:183, Klein 1997, Inkelas 2008:387).

form is instead ***sumɒsaga***. The attested surface form is predicted if, instead, infixation applies to a representation at which the result of reduplication is already present.

(7) The interaction between *-um-* infixation and reduplication

Predicate		Infix	Reduplication	Infix+reduplication
’sɒga	‘stay’	s**u’m**ɒga	’sɒsaga	s**u’m**ɒsaga *s**um**ɒmaga
’tɒngis	‘cry’	t**u’m**ɒtangis	tɒtangis	t**u’m**ɒtangis *t**um**ɒmangis
’kɒti	‘cry’	k**u’m**ɒti	’kɒkati	k**u’m**ɒkati *k**um**ɒmati
’hɒnao	‘go’	h**u’m**ɒnao	’hɒhanao	h**u’m**ɒhanao *h**um**ɒhanao

Topping and Dunga (1973) discuss *-um-* in this connection but the two infixes, *-um-* and *-in-*, exhibit identical behavior in this respect. For example, infixation of *-in-* into *pulan* ‘watch over’ produces ***pin****ilan*. Here, the affixation of *-in-* triggers umlaut of the vowel that immediately follows *-in-* and, as a result, /u/ is fronted to /i/. If this were the base to which progressive reduplication applies, one would expect the upshot to be the form ⟨***pinin****ilan*⟩ but this form is not well-formed. Instead, the acceptable progressive form is ***pinip****ilan*. The following table contains additional examples of similar nature:

(8) The interaction between *-in-* infixation and reduplication

Predicate		Infix+umlaut	Reduplication	Infix+umlaut+reduplication
’pulan	‘watch over’	p**i’n**ilan	’pupulan	p**i’n**ipilan *p**inin**ilan
’konni?	‘take’	k**i’n**enni?	’kokonni?	k**i’n**ekenni? *k**in**enenni?
’soŋgi	‘set on fire’	s**i’n**eŋgi	’sosoŋgi	s**i’n**eseŋgi *s**in**eneŋgi
’tsuli?	‘bring’	ts**i’n**ili?	’tsutsuli?	ts**i’n**itsili? *ts**in**inili?

As revealed by these examples, when infixation and reduplication cooccur within a single word, the infixes *-um-* and *-in-* still appear before the stressed vowel of the reduplicant but no part of the infix is doubled (Chung 2014, ch. 29). In particular, while reduplication doubles the primarily stressed vowel of the base, it fails to result in the doubling of the preceding consonant (the /m/ of the infix *-um-* or the /n/ of *-in-*). The following section offers an analysis of this interaction within a serial/derivational model.

5. Reduplication precedes infixation

The present analysis of the interaction between reduplication and infixation is couched within Distributed Morphology (Halle and Marantz 1993, 1994, Embick 2010, Bobaljik 2012). The crucial assumption that it relies on is that of *cyclicity*: the association of morphosyntactic terminals with phonological material (i.e. Vocabulary Insertion) proceeds root-outwards (Bobaljik 2000, Adger et al. 2003, Embick 2010), and (at least some) morphosyntactic terminals define cycles that govern the application of (phonological) rules.

Consider first a form like *sumɒsaga* in (7), in which *-um-* is the exponent of singular/dual subject agreement in the realis mood, while reduplication is the expression of progressive aspect. I assume that the former is an Agr head in the syntax while the latter is an Asp head in the syntax, and that they are syntactically combined with the verb *sɒga* as illustrated in (9a).[9]

(9) Derivation of *sumɒsaga*

a. Output of syntax
[Agr [Asp [V]]]

First, the exponent of V (the verbal root) is inserted and parsed into a Prosodic Word, which is subject syllabification and stress assignment (Chung 1983, Kiparsky 1986):

b. Vocabulary Insertion of V
[Agr [Asp ['saga]]]

Then, the progressive Asp head must be spelled out, triggering reduplication. Approaches to reduplication within the present framework of assumptions usually follow the tradition of treating reduplication as a form of affixation (see, e.g. Marantz 1982). Specifically, within Distributed Morphology, at least two implementations of this idea exist (cf. Marantz 1982, McCarthy and Prince 1995, Haugen 2008). One possibility is that, while the reduplicative affix is actually spelled out by a phonologically null Vocabulary Item, it triggers a Readjustment Rule (a type of morphologically triggered phonological rule) that results in reduplication. Another possibility is to spell out the reduplicative affix by the insertion of a Vocabulary Item RED, which acquires its actual phonological content from its context (cf. Marantz's (1982) "skeletal morphemes"). For concreteness, I assume the latter option here: (i) the progressive Asp head is spelled out by RED; (ii) RED copies the vowel and the onset of the primarily stressed syllable of the Prosodic Word that it belongs to; (iii) RED is phonologically affixed to the left of the primarily stressed syllable. The newly formed Prosodic Word is resyllabified and stress is reassigned. At this point, red attracts the primary stress, since it is a stress attracting affix (see section 3).[10]

c. Vocabulary Insertion of Asp
[Agr [RED ['saga]]]

d. Reduplication
[Agr [sa ['saga]]

e. Phonological affixation of RED
[Agr ['sasaga]]

---

[9] As discussed in section 2.1, the infixes *-um-* and *-in-* can each be associated with distinct grammatical functions and, thus, be the exponent of distinct terminals. Thus, the question arises of whether *-um-* and *-in-* are farther away from the root than the progressive Asp head in all their grammatical functions. That this is indeed the case is, in fact, quite plausible for all the grammatical functions of both infixes, with the possible exception of (2c) in which *-in-* functions as a derivational, nominalizing morpheme. For discussion of the predictions that the present analysis potentially makes with respect to the *-in-* nominalizer, see section 7

[10] This is, essentially, an implementation of the approach to Chamorro reduplication found in Topping and Dunga 1973, Topping 1980, Broselow and McCarthy 1983.

Finally, the exponent of Agr, *um*, is inserted. Due to the particular requirements imposed on the distribution of this exponent (either lexically or as the result of general phonological principles active in Chamorro), *um* appears to the immediate right of the initial consonant (cluster) of the stem. This results in the Prosodic Word observed on the surface:

f. Vocabulary Insertion of Agr
[-um ['sasaga]]
g. Phonological affixation of *-um*
[su'masaga]

The derivation of *kinekenni?*, which involves infixation of *-in-* and concomitant umlaut, proceeds in exactly the same way up to the steps responsible for reduplication, shown below. In this form, *-in-* is the exponent of passive voice, while reduplication is again the expression of progressive aspect. I assume that the former is a Voice head in the syntax while the latter is an Asp head in the syntax, and that they are syntactically combined with the verb *konni?* as illustrated in (10a). Once reduplication has applied, the exponent of Voice, *in*, is inserted. Due to the requirements imposed on the placement of this exponent, *in* appears in the familiar stem-medial position. It is this infixation step that applies to the output of reduplication:

(10) Derivation of *kinekenni?*
a. Output of syntax
[Voice [Asp [V]]]
b. Vocabulary Insertion of V
[Voice [Asp ['konni?]]]
c. Vocabulary Insertion of Asp
[Voice [red ['konni?]]]
d. Reduplication
[Voice [ko ['konni?]]]
e. Phonological affixation of RED
[Voice ['kokonni?]]
f. Vocabulary Insertion of Voice
[-in ['kokonni?]]
g. Phonological affixation of *-in*
[ki'nokonni?]

The passive voice morpheme triggers umlaut of the immediately following vowel, which in this case happens to be the nucleus of the reduplicant. As the surface form indicates, however, umlaut affects not just the vowel of the reduplicant but also the corresponding vowel in the base:

h. Umlaut
[ki'nekenni?]

This is then a base-reduplicant "identity effect" whereby a regular phonological alternation applies in an environment where it is not conditioned with the apparent result of maximizing the similarity between the base and the reduplicant—in other words, umlaut "overapplies". In this connection, it should be noted that overapplication of umlaut is only observed in the context of reduplication and does not normally spread to the right (see section 2.3). Such identity effects have traditionally been viewed as a strong argument in favor of Base-Reduplicant Correspondence Theory (McCarthy and Prince 1993, 1995, 1999). Different implementations are possible, however, as long as they ensure base-reduplicant identity (see, in addition, Wilbur 1973, Marantz 1982, Benua 1997, Steriade 1988, Raimy 2000, Frampton 2009, among others).

What is crucial in these derivations is that the steps responsible for infixation (f, g) have access to the output of the steps responsible for reduplication (c, d, e).[11] The Distributed Morphology model adopted here captures this state of affairs by incorporating the assumption that certain phonological operations (supplying red with phonological content, phonological affixation, umlaut) apply in cycles defined by the Vocabulary Insertion of the terminals that triggers them. The order of the cycles is, in turn, given by the order of Vocabulary Insertion, which is itself determined by the relative syntactic positions of the relevant terminals. This model is consistent with the common assumption that only some heads define cycles (Embick 2010) and also with any order (or lack of order) of Vocabulary Insertion within a given cycle (Deal and Wolf 2016).[12]

6. A note on non-serial frameworks

The interactions between infixation and umlaut and between reduplication and umlaut yield surface-true generalizations and do not require reference to intermediate derivational stages. As a result, they can, in principle, be accounted for in a non-derivational/parallel model such as Classical (Parallel) Optimality Theory. The opaque interaction between infixation and reduplication, however, may not be entirely expected under certain kinds of non-derivational/parallel frameworks that eschew (multiple) intermediate derivational stages.

Consider the former two interactions first. The infixation~umlaut interaction (section 2.3) yields the generalization (i) below, which holds of output forms both in the presence and absence of reduplication in (see (5) and (8)). The emergence of this pattern could be treated as the result of the interaction between independent constraints that govern umlaut and the linearization of infixes in Chamorro (e.g. Klein 2000, 2005). Likewise, the umlaut~reduplication interaction (section 4) yields the generalization (ii) below, which also holds of the output forms in (8). As already discussed, this overapplication of umlaut can be seen as the result of pressures towards base-reduplicant identity (McCarthy and Prince 1995; see also Klein 1997).

---

[11] However, the order of (f) and (g) with respect to one another and the order of (c), (d), and (e) with respect to another is not crucial. The steps in each of these two sets can apply in an order different from the one given above or they can apply simultaneously (e.g. it should be possible to obtain the results of the steps within each set in a constrained-based system). Umlaut is, likewise, not crucially ordered with respect to the steps responsible for infixation.

[12] As far as the failure of nasal substitution in forms like *fantsinemma?* (discussed in section 2.2) is concerned, the present analysis must resort to the alternative explanation offered in section 2.2: nasal substitution in the context of the prefixes *man-* and *fan-* applies only if these prefixes are combined with morphologically simplex stems.

(i) The infix *-in-* appears on the surface immediately before the left-most vowel of the stem to which it is attached; the vowel of the immediately following syllable is realized as a front vowel of the same height.

(ii) The vowel in the base and the vowel in the reduplicant have the same quality, so that if one of them is realized as a front vowel due to the presence of an umlaut trigger, the other one is also realized as a front vowel.

In contrast, consider the opaque interaction between infixation and reduplication (section 4). Recall that, in general, the following two generalization hold of reduplication in Chamorro:

(iii) Reduplication doubles the nucleus and the onset of the primarily stressed syllable of the base.

(iv) The reduplicant and the original syllable in the base are each contiguous and adjacent to one another: e.g., *sosodda?* (from *sodda?*) and *hugɒgandu* (from *hugɒndu*)—see section 3 for details and more examples.

When reduplication and infixation cooccur as in (7) and (8), however, these generalizations about the behavior of reduplication are not surface-true: e.g. ***t<u>um</u>ɒtangis*** (from *tɒngis*) and ***s<u>in</u>eseŋgi*** (from *soŋgi*). In this case, a kind of opacity arises, whereby the surface forms are expected to have undergone reduplication according to (iii) and (iv), and appear as ****t<u>um</u>ɒmangis*** and ****s<u>in</u>eneŋgi***. But they do not. Reduplication is expected to double not just the primarily stressed vowel of the base but also the preceding consonant (in this case, the /m/ of the infix *-um-* or the /n/ of the infix *-in-*), contrary to fact.[13]

Within a non-serial framework, it might be tempting to attempt to account for the fact that reduplication in Chamorro does not copy any part of the infixes by excluding affixes in general or infixes in particular from the base of reduplication (e.g. along the lines of Shaw 2005). In general, however, affixal material can readily be copied by reduplication in Chamorro, as examples involving stressed prefixes demonstrate. For instance, the word *a'tsatsamatso?tsu?* 'work together (said of two people)' is composed of the root *matso?tsu?* 'work' and the prefix *a'tsa-*. Progressive reduplication in this case doubles the primarily stressed syllable, which happens to be part of the stress attracting prefix *a'tsa-* (see also Topping et al. 1975, p. xxv).

As McCarthy (2007) points out (p. 108), among many others, cases of opacity which involve non-surface-true generalizations (as opposed to non-surface-apparent generalizations in the sense of McCarthy and Prince 1999) may "supply the best (arguably, the only) evidence for language-particular rule ordering." Such rule ordering can, in turn, be readily implemented in rule-based derivational/serial frameworks. The challenge then is to account for robust instances of opacity, such as the interaction between infixation and reduplication in Chamorro, within frameworks that impose more severe restrictions on the number and kinds of levels of representation. While there may not be consensus about the appropriate approach within such frameworks, there certainly is no shortage of promising ideas: see for example Goldsmith 1993 on

[13] On the phenomenon of phonological opacity, see Kiparsky 1971, 1976 and much subsequent work; for overviews of opacity and relevant related issues (including counterfeeding relations and under-application opacity), see McCarthy 2007 and Baković 2011.

Harmonic Phonology, Lakoff 1993 on Cognitive Phonology, Wilson 2000 on targeted constraints, McCarthy 2002, 2003a on comparative markedness, McCarthy and Prince 1999; McCarthy 2003b on sympathy, Bye 2001 on virtual phonology, among others (see McCarthy 2007 for an overview).

7. Conclusion

The opaque interaction between infixation and reduplication in Chamorro was understood here within Distributed Morphology model which allows the output of reduplication to serve as the input to infixation. The crucial assumptions that the analysis relies on are the following:

(i) the association of morphosyntactic terminals with phonological material (i.e. Vocabulary Insertion) proceeds root-outwards;
(ii) (at least some) morphosyntactic terminals define cycles that govern the application of (phonological) rules;
(iii) the relation between a reduplicant and the base of reduplication is such that it ensures (partial) base-reduplicant identity;
(iv) the progressive Asp head in Chamorro is closer to the root than the heads expressed by the infixes *-um-* and *-in-* (at least in the cases discussed here).

Assumptions (i) and (ii) are two of the central ways in which cyclicity is manifested within Distributed Morphology, but equivalent results can also be achieved in constraint- based frameworks that allow serial derivations (e.g. Kiparsky 2000, Wolf 2008). Assumption (iii) is motivated for Chamorro by the observed overapplication of umlaut in the context of reduplication (Wilbur 1973 and much subsequent work). Finally, independent evidence for the language-specific assumption (iv) and, in particular, for the assumed structures in (9) and (9) is scarce in Chamorro. To the extent that the analysis presented here is successful, the morphophonological behaviors of infixation and reduplication described here can be taken as support for these structures.

In this connection, the analysis makes the prediction that, if, in contrast to (iv), the head expressed by an infix is, in fact, closer to the root than the reduplicative progressive Asp head, reduplication should, in principle, be able to double parts of the infix. That's because in this case the output of infixation will serve as the input to reduplication, due to the relative syntactic prominence of the relevant morphemes. The most promising testing ground for this prediction seems to me to involve the derivational uses of the infix *-in-*, which presumably will be closer to the root than the reduplicative progressive Asp head, which is inflectional.

References

Adger, David, Susana Bejar, and Daniel Harbour. 2003. Directionality of allomorphy: a reply to Carstairs-McCarthy. *Transactions of the Philological Society* 101:109–115.

Baković, Eric. 2011. Opacity and ordering. In *The handbook of phonological theory*, ed. John Goldsmith, Jason Riggle, and Alan Yu, 40–67. Malden, MA: Wiley-Blackwell, 2nd edition.

Bobaljik, Jonathan David. 2000. The ins and outs of contextual allomorphy. In *University of Maryland working papers in linguistics*, ed. Kleanthes K. Grohmann and Caro Struijke, 35–71. Linguistics Department, University of Maryland.

Bobaljik, Jonathan David. 2012. *Universals in comparative morphology: Suppletion, superlatives, and the structure of words*. Current Studies in Linguistics. Cambridge, MA: MIT Press.

Broselow, Ellen, and John McCarthy. 1983. A theory of internal reduplication. *The Linguistic Review* 3:25–88.

Bye, Patrik. 2001. Virtual phonology: Rule sandwiching and multiple opacity in North Saami. Doctoral Dissertation, University of Tromsø.

Chung, Sandra. 1983. Transderivational relationships in Chamorro phonology. *Language* 59:35–66.

Chung, Sandra. 1998. *The design of agreement: Evidence from Chamorro*. Chicago, IL: University of Chicago Press.

Chung, Sandra. 2014. A reference grammar of Chamorro. Ms. University of California, Santa Cruz.

Clothier-Goldschmidt, Scarlett. 2014. Chamorro reduplication in Harmonic Serialism. Ms. University of California, Santa Cruz.

Conant, Carlos Everett. 1911. Consonant changes and vowel harmony in Chamorro. *Anthropos* 6:136–146.

Cooreman, Ann M. 1987. *Transitivity and discourse continuity in Chamorro narratives*. Berlin: Mouton de Gruyter.

Costenoble, Hertha. 1940. *Die Chamoro Sprache*. The Hague: Nijhoff.

Crosswhite, Katherine. 1996. Base-derivative correspondences in Chamorro. In *UCLA working papers in phonology*, ed. C.-S. Hsu, 57–85. UCLA Graduate Linguistics Circle.

Deal, Amy Rose, and Matthew Wolf. 2016. Outwards-sensitive phonologically conditioned allomorphy in Nez Perce. In *The morphosyntax-phonology connection: Locality and directionality at the interface*, ed. Vera Gribanova and Stephanie Shih, 29–60. Oxford University Press.

Embick, David. 2010. *Localism versus globalism in morphology and phonology*. Number 60 in Linguistic Inquiry Monographs. Cambridge, MA: MIT Press.

Frampton, John. 2009. *Distributed reduplication*. Cambridge, MA: MIT Press.

Goldsmith, John. 1993. Harmonic phonology. In *The last phonological rule: Reflections on constraints and derivations*, ed. John Goldsmith, 21–60. Chicago, IL: Chicago University Press.

Halle, Morris. 2001. Infixation versus Onset Metathesis in Tagalog, Chamorro, and Toba Batak. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 153–168. Cambridge, MA: MIT Press.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from Building 20*, ed. Kenneth Hale and Samuel Jay Keyser, 111–176. Cambridge, MA: MIT Press.

Halle, Morris, and Alec Marantz. 1994. Some key features of Distributed Morphology. In *MIT Working Papers in Linguistics: Papers on phonology and morphology*, ed. Andrew Carnie and Heidi Harley, volume 21, 275–288. Cambridge, MA: MITWPL.

Haugen, Jason D. 2008. *Morphology at the interfaces: reduplication and noun incorporation in Uto-Aztecan*. Amsterdam/Philadelphia: John Benjamins.
Inkelas, Sharon. 2008. The dual theory of reduplication. *Linguistics* 46:351–401.
Kaplan, Aaron. 2008. Noniterativity is an emergent property of grammar. Doctoral Dissertation, University of California, Santa Cruz.
Kaplan, Aaron. 2011. Harmonic improvement without Candidate Chains in Chamorro. *Linguistic Inquiry* 42:631–650.
Kiparsky, Paul. 1971. Historical linguistics. In *A survey of linguistic science*, ed. W. O. Dingwall, 579–642. College Park, MD: University of Maryland Linguistics Program.
Kiparsky, Paul. 1976. Abstractness, opacity, and global rules. In *The application and ordering of phonological rules*, ed. Andreas Koutsoudas, 160–184. The Hague: Mouton.
Kiparsky, Paul. 1986. Systematic optionality in the lexical phonology of Chamorro. Ms. Stanford University.
Kiparsky, Paul. 2000. Opacity and cyclicity. *The Linguistic Review* 17:351–366.
Kiparsky, Paul. 2003. Syllables and moras in Arabic. In The syllable in Optimality Theory, ed. Caroline Féry and Ruben van de Vijver, 147–182. Cambridge: Cambridge University Press.
Klein, Thomas B. 1997. Output constraints and prosodic correspondence in Chamorro reduplication. *Linguistic Inquiry* 28:707–715.
Klein, Thomas B. 2000. *"Umlaut" in Optimality Theory: A comparative analysis of German and Chamorro*. Tübingen: Neimeyer.
Klein, Thomas B. 2005. Infixation and segmental constraint effects: UM and IN in Tagalog, Chamorro, and Toba Batak. *Lingua* 115:959–995.
Lakoff, George. 1993. Cognitive phonology. In *The last phonological rule: Reflections on constraints and derivations*, ed. John Goldsmith, 117–145. Chicago, IL: Chicago University Press.
Marantz, Alec. 1982. Re reduplication. *Linguistic Inquiry* 13:435–482.
McCarthy, John, and Alan Prince. 1999. Faithfulness and identity in prosodic morphology. In *The prosody-morphology interface*, ed. René Kager, Harry van der Hulst, and Wim Zonneveld, 218–309. Cambridge, UK: Cambridge University Press.
McCarthy, John. 2002. Comparative markedness [long version]. In *University of Massachusetts occasional papers in linguistics 26: Papers in Optimality Theory II*, ed. Angela Carpenter, Andries Coetzee, and Paul de Lacy, 171–246. Amherst, MA: GLSA.
McCarthy, John. 2003a. Comparative markedness. *Theoretical Linguistics* 29:1–51.
McCarthy, John J. 2003b. OT constraints are categorical. *Phonology* 1:75–138.
McCarthy, John J. 2007. Derivations and levels of representation. In *The Cambridge handbook of phonology*, ed. Paul de Lacy, 99–117. Cambridge: Cambridge University Press.
McCarthy, John J., and Alan Prince. 1993. Prosodic morphology I: Constraint interaction and satisfaction. Technical report, Rutgers University, Center for Cognitive Science. [Also ROA 482].
McCarthy, John J., and Alan Prince. 1995. Faithfulness and reduplicative identity. In *University of Massachusetts occasional papers in linguistics 18: Papers in Optimality Theory*, ed. Jill Beckman, Suzanne Urbanczyk, and Laura Walsh Dickey, 249–384. Amherst, MA: GSLA.

von Preissig, Edward Ritter. 1918. *Dictionary and grammar of the Chamorro language of the island of Guam*. Washington, DC: Government Printing Office.
Raimy, Eric. 2000. *The phonology and morphology of reduplication*. Berlin: Mouton de Gruyter.
Rubach, Jerzy. 2000. Glide and glottal stop insertion in Slavic languages: A DOT analysis. *Linguistic Inquiry* 31:271–317.
Safford, William Edward. 1903. The Chamorro language of Guam. American *Anthropologist* 5:289–311.
Shaw, Patricia A. 2005. Non-adjacency in reduplication. In *Studies on reduplication*, ed. Bernhard Hurch, 161–210. Berlin: Mouton de Gruyter.
Steriade, Donca. 1988. Reduplication and syllable transfer in Sanskrit. *Phonology* 5:73–155.
Topping, Donald, and Bernadita Dunga. 1973. *Chamorro reference grammar*. Honolulu, HI: University of Hawaii Press.
Topping, Donald M. 1980. *Spoken Chamorro*. Honolulu, HI: University of Hawaii Press.
Topping, Donald M., Pedro M. Ogo, and Bernadita C. Dungca. 1975. *Chamorro-English Dictionary*. Honolulu, HI: University of Hawaii Press.
Wilbur, Ronnie B. 1973. The phonology of reduplication. Doctoral Dissertation, University of Illinois.
Wilson, Colin. 2000. Targeted constraints: An approach to contextual neutralization in Optimality Theory. Doctoral Dissertation, Johns Hopkins University.
Wolf, Matthew. 2008. Optimal interleaving: serial phonology-morphology interaction in a constraint-based model. Doctoral Dissertation, University of Massachusetts, Amherst.
Yu, Alan C.-L. 2007. *A natural history of infixation*. Oxford: Oxford University Press.
Zuraw, Kie. 2010. A model of lexical variation and the grammar with application to Tagalog nasal substitution. *Natural Language & Linguistic Theory* 28:417–472.

This collection of papers honors Sandra Chung on the occassion of her 30th anniversary at UC Santa Cruz. We are all delighted to recognize her influence as a thinker, a teacher, a mentor, a colleague, and a friend. The contributions have all been inspired by Sandy's clarity and lucid argumentation and cover a range of topics, including ellipsis, the argument/adjunct distinction, Chamorro morphology, and the syntax of nominals.