James Collins
The Hebrew University of Jerusalem

Mapping meanings to argument structure: The semantics of Samoan case[1]

The Department of Linguistics
University of Toronto
June 9, 2018

1 Introduction

- Some two-argument verbs in Samoan take ergative subjects (henceforth ERG ABS verbs like 1a) but others take dative non-subjects (henceforth ABS DAT verbs like 1b).

(1) a. *sā* ***fuafua*** *e* *le* *mālô* *le* *fausia*
PAST plan ERG the government=ABS the construction
The government planned the construction. ERG ABS verb

b. *sā* ***tago*** *ané* *le* *fōma'i* *'i* *lono* *ulu*
PAST touch DIR=ABS the doctor DAT his forehead
The doctor touched his forehead. ABS DAT verb

  - Ergative and dative marked by prepositional markers *e* and *'i*, respectively.
  - Absolutive marked by a high tone on the preceding mora (Yu 2010, Yu and Stabler 2017).[2]

- Here are some examples from each verb class.

ERG ABS	ABS DAT
fa'a'i'ila 'polish'	*ala* 'be the cause of'
fua 'measure, divide'	*fa'amata'u* 'fear'
ifo 'restrain, control'	*lata* 'be near to'
lau 'read out'	*mulimuli* 'follow'
ta'u 'tell (a story)'	*tepataula'i* 'gaze at'
tausi 'look after'	*telefoni* 'call on the phone'

- Here, I provide an analysis of how verbs map these two classes, based on their meanings, arguing for the following generalization:

  - **ERG ABS verbs** entail initiation by the self-directed action of an individual.
  - **ABS DAT verbs** entail either
    i no initiation by a self-directed action, or
    ii that the dative case-marked argument is a thematic goal/recipient.

[1] With thanks to Lelia Glass, Paul Kiparsky, Beth Levin, Grant Muagututia, Malka Rappaport Hovav, and Todd Snider for helpful discussions

[2] Incorporating this analysis of absolutive, the notation I use throughout is *a* short/no high tone, *ā* long/no high tone, *á* short/high tone, *â* long/high tone. The absolutive high tone is not orthographically represented in Samoan. Where it is included in this handout, it is placed where Yu's theory predicts it should go, though further investigation of the examples in this talk need to be undertaken.

- In spelling out this analysis, this paper provides a general theory of how verbal meanings are mapped to morphosyntactic structure. (2) is an illustration of this question applied to an English verb.[3]

(2) $\lambda x \lambda y.\textbf{eat}(x)(y) \longrightarrow ??? \longrightarrow \begin{bmatrix} \textit{Category}: & \text{V} \\ \textit{Select}: & \langle \text{DP}_{[\text{acc}]}, \text{DP}_{[\text{nom}]} \rangle \end{bmatrix}$

  - How do we know *nominative* and *accusative* are the right cases for arguments of 'eat'?
  - How do we ensure the cases are assigned the right way around?
  - These questions have been a subject of inquiry for decades (e.g., Fillmore 1968, 1970, Dowty 1979, 1991, Foley and Van Valin 1984, Baker 1988, 1997, Bresnan and Kanerva 1989, Harley 1995, Kratzer 1996, Folli and Harley 2005, Levin and Rappaport Hovav 2005, and many more)

- Key questions for this sort of "mapping theory" (see Levin and Rappaport Hovav 2005:§1)
  - What are the components of meaning relevant to argument structure?
  - What is the role of thematic roles and how should they be characterized?
  - How should the mapping be represented formally?

- I propose that this mapping can be understood as constraint satisfaction: pairs of meanings/argument structures like (2) optimally satisfy multiple interacting, violable constraints (spelled out with OT).

- **Roadmap**:
  §2: Previous approaches to the semantics of Samoan case and similar problems
  §3: The generalization pursued here: self-directed action vs. goal/recipient status
  §4: Spelling out the link between lexical semantics and argument structure
  §5: Conclusion

2 Previous approaches to the problem

- Several previous approaches to the distinction between ERG ABS verbs and ABS DAT verbs in Samoan and related languages (Chung 1978, Cook 1993, Blume 1998, Ball 2009, Tollan 2018).

- We also find analyses of related phenomena in unrelated languages, e.g., the accusative/partitive distinction in Finnish (Kiparsky 1998, 2001, 2005, Kratzer 2004, Borer 2005 and others).

- These accounts tend to mark out three semantic factors determining case:
  - Lexical aspect: the temporal boundedness and dynamicity of the event.
  - Affectedness: the (degree of) change undergone by the patientive argument
  - Agentivity: the extent and type of control over the event exhibited by the agentive argument.

[3] Here a particular theory of representing verb meanings and argument structure is assumed, but the question applies no matter what representational choices are made here.

2.1 Does ergative case signal telicity?

- Like Samoan, Finnish demonstrates a case alternation: transitive objects either take accusative or partitive case, depending on various factors including the lexical semantics of the verb.

(3) a. *Etsi-n* *karhu-a* / *kah-ta* *karhu-a*
seek-1SG bear-PART two-PART bear-PART
I'm looking for the/a bear / (the) two bears Kiparsky 2005

b. *Tapo-i-n* *karhu-n* / *kaksi* *karhu-a*
kill-PAST-1SG bear-ACC two-ACC bear-PART
I killed the/a bear / two bears Kiparsky 2005

- Various approaches (e.g., Kratzer 2004, Borer 2005, Poole 2015) assume the relevant factor is telicity:
telic $\leftrightarrow$ *accusative* and *atelic* $\leftrightarrow$ *partitive*.

- Accusative case signals the presence of a [+TELIC] feature on a functional head Asp, which is interpreted as imposing an entailment that the event is culminated, as in (4).

(4) $[+\text{TELIC}] \rightsquigarrow \lambda V.\lambda x.\lambda e.V(x)(e) \wedge \textbf{culminate}(x)(e)$[4]

- Kratzer's analysis additionally assumes that the culmination conditions for verbs are specified by the verb's semantics, as in (5a).

(5) a. $\textit{shoot} \rightsquigarrow \lambda x.\lambda e.\textbf{shoot}(x)(e) \wedge (\textbf{culminate}(x)(e) \leftrightarrow \textbf{hit}(x)(e))$

b. $[+\text{TELIC}](\textit{shoot}) \rightsquigarrow \lambda x.\lambda e.\textbf{shoot}(x)(e) \wedge (\textbf{culminate}(x)(e) \leftrightarrow \textbf{hit}(x)(e)) \wedge \textbf{culminate}(x)(e)$
$= \lambda x.\lambda e.\textbf{shoot}(x)(e) \wedge \textbf{hit}(x)(e)$

- One way of testing Kratzer's notion of telicity for Samoan is with 'before'-phrases.

- Under Kratzer's theory, [+TELIC] verbs encode for a **culminate** property, while [-TELIC] verbs don't.

- Assuming 'before'-phrases temporally orient event descriptions, 'before'-phrases with [+TELIC] verbs should orient the culmination point, but not with [-TELIC] verbs (see Beaver and Condoravdi 2003).

(6) a. $P \textit{ before } Q_{[+\text{telic}]} \rightsquigarrow \textbf{before}(P)(\lambda e.Q(e) \wedge \textbf{culminate}(e))$

b. $P \textit{ before } Q_{[-\text{telic}]} \rightsquigarrow \textbf{before}(P)(Q)$

(7) a. Kim left before [Sandy defeated the final boss]$_{[+telic]}$. (*before the point of culmination*)

b. Kim left before [Sandy wanted croissants]$_{[-telic]}$. (*before the starting point*)

- Using Kratzer's notion of telicity as a guide, we find that telicity is by no means a necessary condition for the ERG ABS case frame in Samoan.

- We find ERG ABS atelic verbs, with no point of culmination, even with definite objects.[5]

[4] Kratzer suggests a generalized, algebraic way of spelling out a semantics for culmination. $[+\text{TELIC}] \rightsquigarrow \lambda V.\lambda x.\lambda e.V(x)(e) \wedge \exists f[\textbf{measure}(f) \wedge \forall x'[x' \sqsubseteq f(x) \rightarrow \exists e'[e' \sqsubseteq e \wedge V(x')(e')]]]$, i.e., a verb V with an object x is [+TELIC] iff V holds of x and of every subpart of x, determined by some (lexically specified) way of measuring out subparts of x.

[5] The ergative suffix *-a/-ina* (ES) is triggered by the negative element *le'i*, see Mosel and Hovdhaugen 1992:§18.9.4.6

(8) a. *...ae le'i* ***tausi****-a* *e* *le* *tamāloá* *le* *paopao*
...before take.care-ES ERG SPEC man=ABS SPEC canoe
...before the man took care of the canoe (⇝ before he started/*finished taking care)

b. *...ae le'i* ***su'e****-a* *e* *le* *teine* *lana* *fagafao*
...before search-ES ERG SPEC girl her pet
...before the girl looked for her pet (⇝ before she started/*finished searching)

c. *...ae le'i* ***tautua****-ina* *e* *le* *tama* *lona* *matai*
...before serve-ES ERG SPEC boy her matai
...before the boy served his matai (⇝ before he started/*finished serving)

- The telicity of the clause is also in part determined by the constitution of the patient (Krifka 1986, 1988, 1992, etc). Telicity alternations triggered by, e.g., number alternations, do not affect the case.

(9) a. *sā* ***lau*** *mái* *lou* *igoa* *e* *le* *fai'āoga*
PAST read DIR=ABS your name ERG SPEC teacher
The teacher read out your name.

b. *sā* ***lau*** *mái* *igoa* *e* *le* *fai'āoga*
PAST read DIR=ABS name ERG SPEC teacher
The teacher read out names.

For this reason, Kratzer's hypothesis that telicity determines case **is unlikely to hold** for Samoan.

2.2 Does ergative case signal affectedness?

- Ball 2009 tackles a similar problem in the closely related Tongan.
- Tongan verbs, analogous to Samoan ABS DAT verbs, are semantically characterized as *participant autonomous entity relations*, following a notion from Blume 1998. These verbs encompass:
  - verbs with a "destination" or "target" (such as *help* or *look at*)
  - verbs with a volitional second argument which helps bring about the event (such as *help*)
  - verbs with a 'point of reference' second argument (such as *look like* and *resemble*).
- As an overarching constraint, these verbs are demonstrate low affectedness of the second argument, "and in some cases, having a first argument which does not cause a change in the situation".
- In order to evaluate the claim that affectedness is a (or the) crucial factor in distinguishing verb classes, we need to pin down a particular notion of affectedness.
- Beavers 2011 spells out a four level hierarchy of affectedness for transitive verbs.
  - Does *V* entail that the patientive argument undergoes a change?
    - If yes, is the degree of change specified?
    - If no, is there a lexically specified potential for change?
- For Beaver, these verb are spelled out such that the categories are ordered by asymmetrical entailment.

(10) **quantized change**
$\models$ **non-quantized change**
$\models$ **specified potential for change**
$\models$ **unspecified potential for change**

- If a language has a grammatical phenomenon which applies to a particular category of affectedness, e.g. **non-quantized change**, then it must also apply to its entailing categories, i.e. **quantized change**.
- Using this theory to spell out Ball's 2009 hypothesis, Tongan (or Samoan) grammar makes some divide along this scale of affectedness entailments, for example, one divide could be:
  - ERG ABS verbs entail **quantized change** or **non-quantized change**
  - ABS DAT verbs entail **specified potential for change** or **unspecified potential for change**
- What would be a counterexample to this theory? If we could find a class of ERG ABS verbs which are *lower on the scale* than a class of ABS DAT verbs.
- Tollan 2018 points out that Samoan verbs of searching are ERG ABS.

(11) a. *ʻoloʻo* ***sāʻili*** *e* *le* *ositāulagá* *sana* *matua*
PROG search ERG SPEC priest=ABS his text
The priest is searching for his text.

b. *sā* ***suʻe*** *e* *le* *tama* *āʻogá* *se* *tali*
PAST look.for ERG SPEC boy school=ABS NSPEC answer
The student looked for an answer.

- This is surprising given an affectedness-based view of case, as verbs of searching do not entail the absolutive argument even exists (see Tollan 2018:§3), thus cannot entail a change of state.
- Likewise they fail Beavers' requirement for **specified potential for change**, i.e., they cannot be "force recipients" (in the sense of Croft 1991) if they do not exist.
- They are thus placed in Beavers' most general category **unspecified potential for change**.[6]
- We find other ERG ABS verbs which do not impose an existential constraint on their second argument, for example, verbs of restraint and preventing.

(12) a. *sā* ***teteʻe*** *e* *puleāʻogá* *le* *puʻeina* *o* *ata*
PAST refrain ERG head.teacher=ABS SPEC take-INA GEN picture
The head teachers refrained from taking pictures. (↛ *there were pictures taken.*)

b. *sā* ***taofi*** *e* *le* *ʻauʻaunagá* *se* *afi*
PAST prevent ERG SPEC service=ABS NSPEC fire
The service prevented a fire. (↛ *there was a fire.*)

- Like verbs of searching, these verbs cannot be understood as imposing an affectedness entailment on their absolutive argument, as the argument need not even exist, let alone be affected.

[6]Beavers' most general category merely entails the patientive individual stands in some relation to an event, i.e., **j** is unspecified for change if $\exists e \exists \theta [\theta(\mathbf{j}, e)]$, i.e., there are hardly any semantic constraints on **j** at all.

- We also find verb which may impose an existential requirement on the absolutive, but nevertheless encode for low affectedness. It cannot be said that these verbs entail that the absolutive is a "force recipient", encoding any kind of **specified potential for change**.

(13) a. *na* ***fa'atali*** *mái* *a'u* *e* *lo'u* *uso* *matua*
PAST wait DIR=ABS me ERG my brother old
My older brother waited for me.

a. *na* ***tatao*** *e* *le* *leoleó* *le* *tagata* *gaoi*
PAST follow ERG SPEC police=ABS SPEC person thief
The police followed in the tracks of the thief.

- In all of the cases above, we can only reasonably classify these verbs as belonging to Beavers' most general category: **unspecified potential for change**.
- Turning to ABS DAT verbs, do we find verbs which should be placed higher in Beavers' hierarchy?
- Non-agentive verbs of soiling/smearing are encoded with ABS DAT. In these cases, both arguments appear to undergo some kind of (non-quantized) change:
  - The absolutive argument becomes covered in the substance denoted by the dative argument.
  - The dative argument is distributed over the location denoted by the absolutive argument.

(14) a. *'Ua* ***panupanú*** *o'u* *lima* *'i* *le* *siamu*
PERF smeared=ABS my.PL hand DAT SPEC jam
My hands are smeared with jam. (Milner 1976:175)

b. *'Ua* ***'ola'olá*** *ipu* *māfolafola* *'i* *le* *ga'o*
PERF soiled=ABS dish flat DAT SPEC fat
The dishes are soiled with grease. (Milner 1976:163)

- We also find ABS DAT verbs encoding events of receiving, in which case the dative should be understood as having **potential for change**. These verbs can either be agentive or non-agentive.
- Verbs like *pisi* 'splash against' fit Beavers canonical type of force recipient verbs, i.e., they are verbs of 'surface contact'.

(15) a. *'ua* ***pisí*** *le* *vai* *'i* *lona* *lima*
PERF splash=ABS SPEC water DAT his hand
Water splashed against his hand.

b. *'ua* *'e* ***talusā*** *'i* *le* *'āiga*
PERF 2SG bring.trouble DAT SPEC family
You have brought trouble to the family.

- Thus we find reasonable cases in which certain ABS DAT verbs should plausibly be ranked higher on Beavers' affectedness hierarchy than certain ERG ABS verbs.
- These cases are systematic counterexamples to the hypothesis that ERG ABS verbs encode for a greater level of affectedness than ABS DAT verbs.

Thus ERG ABS verbs **do not** necessarily encode for more affectedness than ABS DAT verbs.

2.3 Does ergative case signal agentivity?

- Previous work on Samoan makes a link between ergative case and agentivity (e.g., Cook 1988, Duranti and Ochs 1989), other work makes the same association but non-categorically (e.g., Mosel and Hovdhaugen 1992:§18, Tollan 2018).
- In this section, I explore the link between ergativity and agentivity (especially in relation to Tollan's 2018 theory), and propose my own refinement.

2.3.1 Proto high and proto low agents

- Under Tollan's 2018 account, Samoan ABS DAT verbs and ERG ABS verbs are distinguished based on the entailments associated with the subject argument.
- Using a system inspired by Dowty 1991, a verb is classified as ERG ABS if its subject fits better with the prototypical "High Agent", and the verb is ABS DAT if the subject fits better with the "Low Agent".
- A participant's proximity to the prototypical High/Low Agent is determined by how many of the following entailments hold (see Tollan 2018:17).

(16)

Proto High Agent	*Proto Low Agent*
initiator	initiator
experiencer	experiencer
affect on another entity	neither affects an entity nor is affected
brings about a change of state	neither brings about nor undergoes a change of state
effortful	
volitional	
concludes an event	

- Like Ball's analysis, affectedness plays a role in determining case, but the association is not categorical – verbs can fail to entail affectedness but still be encoded as ERG ABS
- For example, using Tollan's proposed lexical entailments, we can derive why the 'searcher'-participant of a verb of searching *su'e* gets ergative case: it is closer to the prototypical "high agent".

(17) the 'searcher'-participant of *su'e* ('search', 'look for'):

Proto High Agent	*Proto Low Agent*
initiator	initiator
effortful	neither affects an entity nor is affected
volitional	neither brings about nor undergoes a change of state
concludes an event	

- Thus under Tollan's system, the notions of *effort* and *volitionality* play a big role, tipping the scales towards the subject receiving ergative case, even without an affectedness entailment.
- However, we also find ABS DAT verbs which encode for effortful/volitional participants.

(18) a. *e* ***'au'auná*** *le* *taule'ale'a* *'i* *le* *matai*
PRES serve=ABS SPEC untitled.man DAT SPEC matai
The untitled man serves the matai.

b. *‘Ia* *‘e* ***fesoasoani*** *‘iate* *ia*
SUBJ 2SG help DAT 3SG
You should help him.

c. *na* ***kisi*** *‘o* *ia* *‘i* *lona* *atali‘i*
PAST kiss ABS 3SG DAT his son
He kissed his son.

d. *Se‘i* *‘e* ***logo*** *atu* *‘i* *le* *ali‘i*
OPT 2SG inform DIR DAT SPEC chief
You should inform the chief.

- When we apply Tollan's system to the predicates above, we find that the "effortful" and "volitional" properties bias the system towards encoding the predicates as ERG ABS, rather than as ABS DAT.
- For example, the following is an assessment of the 'helper' participant of the ABS DAT verb *fesoasoani*. As the verb is ABS DAT, we want the system to encode the 'helper' as a *Proto Low Agent*.
- Although the entailments for each predicate can be tricky to pin down, the "effortful" and "volitional" properties seem to bias the 'helper' argument toward being encoded as a *Proto High Agent*.

(19) the 'helper'-participant of *fesoasoani* ('help'):

Proto High Agent	*Proto Low Agent*
initiator	initiator
affect another entity	
effortful	
volitional	
concludes an event	

Although ERG ABS may be linked to agentivity, we also find agentive subjects of ABS DAT verbs.

2.3.2 Self Directed Initiators (SDIs)

- Leaving aside the question of ABS DAT verbs for now, does the proposed link between ergative and agentivity play out in the data?
- To investigate this, I refer to Cruse's 1973 characterization of four notions of agentivity. Each notion describes an entailment relating to an individual *x*'s participation in an event *e*.

(20) a. **Volitive**: *x*'s participation in *e* is an act of *x*'s will

b. **Effective**: *x*'s participation in *e* is an exertion of force, and *x*'s participation **is not** due to *x*'s internal energy (but by virtue of its position, motion, etc.), e.g., projectiles, instruments, etc.

c. **Agentive**: *x*'s participation in *e* is an exertion of force, and *x*'s participation **is** due to *x*'s internal energy, e.g., animate actors, natural forces, etc.

d. **Initiative**: *x* initiates *e* by virtue of giving a command.

- We can discount **effectivity** (b) and **initiativity** (d) as being necessary conditions for ergative.

- **Against effectivity:** Various ERG ABS predicates require the internal energy of the ergative-argument: *va‘ava‘ai* ‘look after’, *lama* ‘ambush’, *māfaufau* ‘devise (a plan)’, and many others.
- **Against initiativity**: Various ERG ABS predicates don’t require the ergative to have issued a command: *fofoga* ‘sing’, *lau* ‘read out’, *fa‘aita* ‘make angry’, and many others.

- The notion of **volitivity** (a) is more promising.
- However, Mosel and Hovdhaugen 1992 note that natural forces such as weather events, as well as non-human actors like machines and cars are able to be ergative subjects.

(21) a. *na* ***tapuni*** *e* *le* *matagi* *le* *faitoto‘a*
PAST close ERG SPEC wind SPEC door
The wind closed the door. (Mosel and Hovdhaugen 1992:9.68)

b. *e* ***vili*** *e* *le* *masini* *le* *ogalā‘au*
PRES rotate ERG SPEC machine SPEC log
The machine rotates the log. (Mosel and Hovdhaugen 1992:18.275)

- These data suggest volitivity is not a condition for ergative.
- However, we also find that not all speakers accept cases like (21), paraphrasing using locative case.

(22) a. *%‘ua* ***fa‘apa‘u*** *e* *le* *matagí* *le* *lā‘au*
PERF make.fall ERG SPEC wind=ABS SPEC tree
The wind felled the tree.

b. *‘o* *le* *malosi* *o* *le* *matagi* *‘ua* *pa‘u* *ái* *le* *lā‘au*
FOC SPEC strong GEN SPEC wind PERF fall LOC=ABS SPEC tree
The strength of the wind is why the tree fell.

- For this set of speakers, the notion of volitivity might be an additional condition for ergative.
- We similarly find Cruse’s notion of **agentivity** to be relevant. Cruse provides the following minimal pair to isolate the intended semantic notion of “internal energy”.

(23) a. The ball rolled across the floor. (*internal energy unspecified*)
b. The ball rolled itself across the floor. (*internal energy entailed*)

- In (23b), the addition of ‘itself’ provides an entailment that the active participant (‘the ball’) is self-directed: it’s participation in the event is not being propelled by a distinct individual.
- I propose here that this notion of self-directed action is relevant for the encoding of Samoan verbs.

(24) **Self Directed Initiator (SDI) Hypothesis**: Samoan ERG ABS verbs entail that:

a. the ergative-marked argument denotes an initiating participant

b. the ergative-marked argument denotes a self-directed participant

- (24a) entails that the event comes about by the action or mental state of the ergative argument. This condition excludes non-initiator subjects from having ergative case, such as the subjects in (25).

(25) a. *e* ***so'o*** *'uma* *'i* *lo lātou* *tinā*
PRES resemble all DAT SPEC.3PL mother
They all resemble their mother.

b. *'ua* ***tumu*** *le* *ipu* *i* *vai*
PERF full SPEC dish LOC water
The cup is full of water.

- (24b) entails that the participant denoted by the ergative argument operates on its own internal energy. This is general enough to include natural forces and machines.

- However, the definition correctly excludes instrumental and projectile subjects, which are not ergative.

(26) a. **na* ***tatala*** *e* *le* *kî* *le* *faitoto'a*
PAST open ERG SPEC key=ABS SPEC door
The key opened the door

b. **na* ***tipi*** *e* *le* *naifí* *le* *fasipovi*
PAST cut ERG SPEC knife SPEC meat
The knife cut the meat.

(27) a. **'ua* *na* ***tā*** *e* *le* *pulufaná* *le* *tama*
PERF already strike ERG SPEC bullet=ABS SPEC boy
The bullet struck the boy.

b. **'ua* ***nuti*** *e* *le* *papá* *le* *fagu*
PERF smash ERG SPEC rock=ABS SPEC bottle
The rock smashed the bottle.

- The definition also correctly excludes predicates denoting involuntary emotions and states, which are lexicalized with ABS DAT in Samoan.

(28) a. *E* *te* ***alofa*** *'iate* *ia*
2SG PRES love DAT 3SG
You love her/him.

b. *'Ua* *e* ***fa'amoemoe*** *'i* *le* *Atua*
PERF 2SG trust DAT SPEC God
You trust in God.

- We also correctly exclude predicates which are true by virtue of the position or motion of their participants, i.e., not due to their own self-directed initiation.

(29) a. *E* ***latalata*** *le* *fale* *ʻi* *le* *ʻaualа*
PRES near SPEC house DAT SPEC road
The house is near the road. (Milner 1976:98)

b. *ʻua* ***siʻo*** *le* *fanua* *ʻi* *uaea*
PERF surround SPEC land DAT wire
The land is surrounded by wire. (Milner 1976:210)

- The Self-Directed Initator (henceforth SDI) hypothesis shares an insight with Tollan's Dowty-inspired system: agentive notions like effort/volition are determining factors in classifying a verb as ERG ABS.
- Crucially, the hypothesis in (24) takes the SDI role to be only a necessary condition for ergative case.
- It isn't a sufficient condition for ergative: non-ergative subjects (w/ ABS DAT verbs) may be SDIs.

3 Linking thematic roles and case

- The task is to figure out and state precisely the conditions determining which verbs are lexicalized as ERG ABS verbs and which are lexicalized as ABS DAT.
- In doing so, I propose a system of isolating and making generalizations about thematic roles.
- I propose lexical entries like in (30), associating a lexical item with a category (*Cat*), a meaning (*Sem*), and most importantly, morphosyntactic features it assigns to its arguments (*Sel*), including case (K).

(30) a. *siva* 'dance' $\rightsquigarrow \begin{bmatrix} \textit{Cat}: & \text{V} \\ \textit{Sem}: & \lambda x.\lambda e.\textbf{dance}(e)(x) \\ \textit{Sel}: & \text{DP}_x \left[\text{K}: \left\{\text{ERG} \mid \text{DAT} \mid \text{ABS}\right\}\right] \end{bmatrix}$

b. *solo* 'wipe' $\rightsquigarrow \begin{bmatrix} \textit{Cat}: & \text{V} \\ \textit{Sem}: & \lambda x.\lambda y.\lambda e.\textbf{wipe}(e)(x)(y) \\ \textit{Sel}: & \text{DP}_y \left[\text{K}: \left\{\text{ERG} \mid \text{DAT} \mid \text{ABS}\right\}\right] \\ & \text{DP}_x \left[\text{K}: \left\{\text{ERG} \mid \text{DAT} \mid \text{ABS}\right\}\right] \end{bmatrix}$

- This paper's approach is to propose **well-formedness constraints** on lexical entries.
- *A priori*, lexical items may assign any case to any argument, but constraints proposed below will rule out unattested combos, like DAT ERG and so on.
- **A note on the lexicon-syntax interface**
  - The focus of the theory is the well-formedness lexical entries, which can serve as the input to the combinatory syntax. I refrain from making strong assumptions about narrow syntax.
  - The entries in (30) assume (a) the transitive verb selects both its arguments and (b) case features are assigned by the verb. Both of these assumptions are just to keep representations simple.
  - Ideally, the representations could serve as the input for a variety of syntactic frameworks.
  - e.g., see appendix A for a version of the analysis incorporating the argument-selecting head *v* (Kratzer 1996), and incorporates a notion of case assignment via Agree (Chomsky 2001)

3.1 Ranking participants

- The first step in ensuring cases go to the right arguments is to define a notion of thematic roles.
- Like Tollan 2018, I propose to use the framework set out in Dowty 1991.
  - Thematic roles are a set of entailments of a predicate with respect to one of its arguments (e.g., *x* is volitional, *x* undergoes a change of state, ...).
  - Avoids "atomic" notions of thematic roles like Agent, Patient, Instrument, etc, sidestepping the unclear boundaries in defining these terms and their link to syntactic positions.
  - For example, what semantic generalization can we assign to the external argument in the following? What about the internal argument?

(31)
a. I walked a mile.
b. I slept twelve hours.
c. This weighs five pounds.
d. The piano measures 6'5".
e. I paid five dollars
f. The book cost five dollars

- For Dowty, participant roles for individual verbs fit more or less with prototype roles, in particular, Proto Agent and Proto Patient.

(32)

Proto-Agent	*Proto-Patient*
volitional involvement	undergoes change of state
sentience/perception	incremental theme (determines time course of event)
causing the event or change of state	causally affected by another participant
movement	stationary relative to another participant
exists independently	does not exist independently (or at all)

- These entailments provide us with a way to **rank the arguments** of two-argument verbs.
- Participants which are closer to the Proto-Agent in terms of number of entailments satisfied are ranked higher. Conversely, participants closer to the Proto-Patient are ranked lower.[7]

(33)
a. *su'e* $\rightsquigarrow$ searcher $>$ searched-for
b. *alofa* $\rightsquigarrow$ lover $>$ loved-thing
c. *tipi* $\rightsquigarrow$ cutter $>$ cut-thing
d. *mulimuli* $\rightsquigarrow$ follower $>$ followed

- Given the ranking, we can use the features proposed in Kiparsky 1997 (see also Kiparsky 2001, Wunderlich 1997) to isolate lower and higher ranked arguments. Informally:

[7] An open question is whether Dowty-style rankings should have different weightings, e.g., maybe *causing the event or change of state* is a more important factor than any other, suggesting Dowty's system could be reconstrued using weights, or as a set of violable constraints *à la* OT.

(34) a. [+HR]: the highest ranked argument (no argument ranked above me)

b. [+LR]: the lowest ranked argument (no argument ranked below me)

c. [−HR]: the non-highest ranked argument (there is an argument ranked above me)

d. [−LR]: the non-lowest ranked argument (there is an argument ranked below me)

- These features are borne by (i) DP arguments, and (ii) selectional features of verbs.
- We can incorporate these features into our lexical entries like so:
  - The highest ranked argument gets [+HR], everything else gets [−HR].
  - The lowest ranked argument gets [+LR] (even if it already has [+HR], e.g., intransitive subjects), everything else gets [−LR]

(35) a. *siva* 'dance' $\rightsquigarrow \begin{bmatrix} \textit{Cat}: & \text{V} \\ \textit{Sem}: & \lambda x.\lambda e.\textbf{dance}(e) \wedge \textbf{dancer}(e)(x) \\ \textit{Sel}: & \text{DP}_x \begin{bmatrix} \theta: & [+\textbf{HR}], [+\textbf{LR}] \\ \text{K}: & \{\text{ERG} \mid \text{DAT} \mid \text{ABS}\} \end{bmatrix} \end{bmatrix}$

b. *solo* 'wipe' $\rightsquigarrow \begin{bmatrix} \textit{Cat}: & \text{V} \\ \textit{Sem}: & \lambda x.\lambda y.\lambda e.\textbf{wipe}(e) \wedge \textbf{wiper}(e)(y) \wedge \textbf{wipee}(e)(x) \\ \textit{Sel}: & \text{DP}_y \begin{bmatrix} \theta: & [+\textbf{HR}], [-\textbf{LR}] \\ \text{K}: & \{\text{ERG} \mid \text{DAT} \mid \text{ABS}\} \end{bmatrix} \\ & \text{DP}_x \begin{bmatrix} \theta: & [-\textbf{HR}], [+\textbf{LR}] \\ \text{K}: & \{\text{ERG} \mid \text{DAT} \mid \text{ABS}\} \end{bmatrix} \end{bmatrix}$

- We now have a means of associating case and thematic roles.
  - if a DP is marked by ergative, it must bear the [−LR] feature.
  - If a DP is marked by dative, it must bear the [−HR] feature.
  - Absolutives are unspecified for [±HR]/[±LR].

(36) a. $e \rightsquigarrow [-\text{LR}]$ b. *'i* $\rightsquigarrow [-\text{HR}]$[8] c. $^{H} \rightsquigarrow [\,]$

- The system requires that the features of cases must not contradict the selectional features of verbs.
- This is enough to ensure intransitive sole arguments get absolutive case.

(37) *siva* 'dance' $\rightsquigarrow \begin{bmatrix} \textit{Cat}: & \text{V} \\ \textit{Sem}: & \lambda x.\lambda e.\textbf{dance}(e) \wedge \textbf{dancer}(e)(x) \\ \textit{Sel}: & \text{DP}_x \begin{bmatrix} \theta: & [+\text{HR}], [+\text{LR}] \\ \text{K}: & \text{ABS} \end{bmatrix} \end{bmatrix}$

[8]This proposal syncs with Tollan's proposal that Samoan *'i* is actually accusative. [−HR] is the feature associated with accusative in Kiparsky 1997. Similarly, in nominative-accusative aligned Polynesian languages like Maori, accusative case is cognate with Samoan *'i* (see Clark 1976, Chung 1978, Kikusawa 2002, Ball 2007, etc.).

- Both dative case ([−HR]) and ergative case ([−LR]) create contradictions with the intransitive verb's [+HR], [+LR] selectional features, so are ruled out.
- We likewise ensure that for transitives, dative is not assigned to the more agentive argument and ergative is not assigned to the less agentive argument.

(38) *solo* 'wipe' $\rightsquigarrow$
$$\begin{bmatrix} \textit{Cat}: & \text{V} \\ \textit{Sem}: & \lambda x.\lambda y.\lambda e.\textbf{wipe}(e) \wedge \textbf{wiper}(e)(y) \wedge \textbf{wipee}(e)(x) \\ \textit{Sel}: & \text{DP}_y \begin{bmatrix} \theta: & [+\text{HR}], [-\text{LR}] \\ \text{K}: & \{\text{ERG} \mid \cancel{\text{DAT}} \mid \text{ABS}\} \end{bmatrix} \\ & \text{DP}_x \begin{bmatrix} \theta: & [-\text{HR}], [+\text{LR}] \\ \text{K}: & \{\cancel{\text{ERG}} \mid \text{DAT} \mid \text{ABS}\} \end{bmatrix} \end{bmatrix}$$

- However, more needs to be said about what roles associate with ergative and dative.

3.2 Associating the more agentive argument with ergative

- The system doesn't yet (i) ensure ergative case only goes to 'Self Directed Initiators', or (ii) distinguish ERG ABS and ABS DAT verbs.
- The next well-formedness constraint on lexical entries is ERG⇒SDI, defined below.

(39) ERG⇒SDI: If a DP is assigned ergative case, the DP must have the SDI (self-directed initator) role.

- The following lexical entry for *solo* 'wipe' satisfies this constraint. It is specified as combining with an ergative-case DP, and this DP is interpreted as an SDI participant (**wiper** being a subtype of **sdi**).

(40) *solo* 'wipe' $\rightsquigarrow$
$$\begin{bmatrix} \textit{Cat}: & \text{V} \\ \textit{Sem}: & \lambda x.\lambda y.\lambda e.\textbf{wipe}(e) \wedge \textbf{sdi}(e)(y) \wedge \textbf{wipee}(e)(x) \\ \textit{Sel}: & \text{DP}_y \begin{bmatrix} \theta: & [+\text{HR}], [-\text{LR}] \\ \text{K}: & \text{ERG} \end{bmatrix} \\ & \text{DP}_x \begin{bmatrix} \theta: & [-\text{HR}], [+\text{LR}] \\ \text{K}: & \text{ABS} \end{bmatrix} \end{bmatrix}$$

- The constraint also correctly rules out this lexical entry for the ABS DAT predicate *lata* 'be near to'.
- Even though ergative is correctly assigned to a [−LR] argument, it is wrongly assigned to a non-SDI.

(41) *lata* $\not\rightsquigarrow$ *
$$\begin{bmatrix} \textit{Cat}: & \text{V} \\ \textit{Sem}: & \lambda x.\lambda y.\lambda e.\textbf{near}(e) \wedge \textbf{location}(e)(x) \wedge \textbf{theme}(e)(y) \\ \textit{Sel}: & \text{DP}_y \begin{bmatrix} \theta: & [+\text{HR}], [-\text{LR}] \\ \text{K}: & \text{ERG} \end{bmatrix} \\ & \text{DP}_x \begin{bmatrix} \theta: & [-\text{HR}], [+\text{LR}] \\ \text{K}: & \text{ABS} \end{bmatrix} \end{bmatrix}$$

Interim summary

- The [−LR] feature on ergative ensures it is never assigned to intransitive subjects (of, e.g., unergatives).
- The constraint ERG⇒SDI, in (39), additionally rules out ergative case arguments without an SDI role.
- Additionally, the [−LR] feature on ergative may help us understand Tollan's 2018 observation that canonically unergative verbs with cognate objects cannot take ergative.

(42) a. *Sā* ***siva*** *le* *teine* *ʻi* *le* *siva*
PAST dance SPEC girl DAT SPEC dance
The girl danced a dance.

b. *Sā* ***ata*** *le* *teine* *ʻi* *le* *ata*
PAST laugh SPEC girl DAT SPEC laugh
The girl laughed a laugh.

- One potential strategy is to assume that cases like (42) do not involve two participants, so the subject argument takes [+HR][+LR], thus preventing ergative case from emerging.

3.3 Associating the less agentive argument with dative

- The next problem is verbs with SDI subjects which do not take ergative case, such as (43).

(43) a. *e* ***ʻauʻauná*** *le* *taulеʻaleʻa* *ʻi* *le* *matai*
PRES serve=ABS SPEC untitled.man DAT SPEC matai
The untitled man serves the matai.

b. *ʻIa* *ʻe* ***fesoasoani*** *ʻiate* *ia*
SUBJ 2SG help DAT 3SG
You should help him.

- So far, nothing in the theory prevents this lexical entry, which **wrongly assigns ergative** to the subject.

(44) *ʻauʻauna* $\rightsquigarrow$
$$\begin{bmatrix} \textit{Cat}: & \text{V} \\ \textit{Sem}: & \lambda x.\lambda y.\lambda e.\mathbf{serve}(e) \wedge \mathbf{servee}(e)(x) \wedge \mathbf{sdi}(e)(y) \\ \textit{Sel}: & \text{DP}_y \begin{bmatrix} \theta: & [+\text{HR}],[-\text{LR}] \\ \text{K}: & \text{ERG} \end{bmatrix} \\ & \text{DP}_x \begin{bmatrix} \theta: & [-\text{HR}],[+\text{LR}] \\ \text{K}: & \text{ABS} \end{bmatrix} \end{bmatrix}$$

- In these cases, I propose that the non-subject argument **has the properties of a goal** (a category encompassing benefactives, addressees, recipients, etc.).
- To determine whether an argument is a goal, we can use the same kind of strategy as Dowty and define prototypical properties of a Proto-Goal.

Proto-Goal
x is the terminus of the directional path of the event
(45) *x* is the addressee of a speech event
event is for the benefit of *x*
the event causes *x* to have something (abstract or non-abstract)

- Provided we can use properties like (45) to identify goals, we have the tools to give an analysis of agentive verbs which are lexicalized with ABS DAT.
- Next, I propose the following constraint on lexical entries, analogous to ERG⇒SDI.

(46) DAT⇔GOAL:
(i) If a V selects for an dative DP, that DP has a goal role, and
(ii) if a DP has a goal role, it must have dative.

- The constraint DAT⇔GOAL is violated by any lexical entry:
  - which assigns dative to a non-goal
  - or assigns another case to a goal.
- The incorrect lexical entry (44) for *'au'auna* 'serve' is restated below. Now, following the definition of goal above, we take the **servee** role to be a sub-type of **goal**.

(47) *'au'auna* $\not\mapsto *$
$$\begin{bmatrix} \textit{Cat}: & \text{V} \\ \textit{Sem}: & \lambda x.\lambda y.\lambda e.\mathbf{serve}(e) \wedge \mathbf{goal}(e)(x) \wedge \mathbf{sdi}(e)(y) \\ \textit{Sel}: & \text{DP}_y \begin{bmatrix} \theta: & [+\text{HR}],[-\text{LR}] \\ \text{K}: & \text{ERG} \end{bmatrix} \\ & \text{DP}_x \begin{bmatrix} \theta: & [-\text{HR}],[+\text{LR}] \\ \text{K}: & \text{ABS} \end{bmatrix} \end{bmatrix}$$

- This entry is correctly ruled out by DAT⇔GOAL, as dative case is **not assigned to the goal** participant.
- To sum up this section:
  - I take here a constraint-based approach to characterizing lexical entries for verbs.
  - Kiparsky's (1997) system allows us to establish where an argument stands in relation to other arguments with respect to its agentivity.
  - Using this system, we can associate cases with both transitivity and agentivity.
  - More specific linkings of roles and cases are made using more refined constraints like ERG⇒SDI and DAT⇔GOAL, which reference particular ways in which participants interact with events.

4 Lexical Optimality

- In this section, I rule out the remaining illicit case combinations.
- In order to do this, I make precise how exactly well-formedness constraints on lexical entries like ERG⇒SDI and DAT⇔GOAL operate, assuming the principle (48).

(48) **Lexical Optimality**: Lexical entries must optimally satisfy multiple interacting, violable constraints.

- The insight is spelled out using Optimality Theory.

4.1 Ranking thematic constraints

- So far the system ensures ergative only goes to SDIs and dative and goals/recipients are linked.
- But what about dative case-assigning ABS DAT verbs which *do not* have goal participants?
- According to the theory built so far, these are verbs which do not have SDI-subjects, e.g., involuntary actions and states. Here, dative need not be assigned to a goal.

(49) a. *E* ***latalata*** *le* *fale* *‘i* *le* *‘aualа*
PRES near SPEC house DAT SPEC road
The house is near the road. (Milner 1976:98)

b. *‘ua* ***si‘o*** *le* *fanua* *‘i* *uaea*
PERF surround SPEC land DAT wire
The land is surrounded by wire. (Milner 1976:210)

c. *‘Ua* ***aogā*** *le* *vailā‘au* *lenei* *i* *lo‘u* *ma‘i*
PERF benefit=ABS SPEC medicine good DAT my sickness
This medicine has cured my sickness.

- These verbs should violate DAT⇔GOAL as we have dative assigned to a non-goal.
- The constraints which determine meaning-argument structure mappings (ERG⇒SDI and DAT⇔GOAL) are ranked as in (50). Lexical entries optimally satisfy the constraints, given this ranking.

(50) **Role-Case Constraint Ranking:**
ERG⇒SDI ≫ DAT⇔GOAL

- The key insight in (50) is that DAT⇔GOAL is a less important constraint, so dative on a non-goal is tolerated only if the subject is not a self-directed initator.

4.2 Markedness constraints on case frames

- The thematic constraints ERG⇒SDI and DAT⇔GOAL will link dative/ergative with the right thematic roles. But we additionally want to rule out other illicit case frames.
- So far the system says nothing about the illicit frames ERG DAT and ABS ABS, as in (51).

(51) a. **e* ***‘au‘auna*** *e* *le* *taule‘ale‘a* *‘i* *le* *matai*
PRES serve ERG SPEC untitled.man DAT SPEC matai
The untitled man serves the matai. (ERG DAT)

b. **e* ***vilí*** *le* *masiní* *le* *ogalā‘au*
PRES rotate=ABS SPEC machine=ABS SPEC log
The machine rotates the log. (ABS ABS)

- What is the status of “ABS ABS”?:
  - The system in this paper intends that (51b), with a double high tone, should be ruled out.

– However, cases of "ergative drop" (observed in Ochs 1982, see Collins 2014 for an analysis), similar to (51b), but without a high tone on the subject[9], lack an explanation in this system.

(52) *Ia* ***ʻai*** *loa* *Koʻokoʻo* *falaoa*
SUBJ eat now Koʻokoʻo bread
Ok, Koʻokoʻo is going to eat the bread now. (Ochs 1982:653)

– More work is required to determine how to characterize these kinds of examples morphosyntactically, especially with respect to the presence/absence of the high tone.
– For now, we assume that examples like (52) are not examples of ABS ABS.

- In order to rules out ERG DAT and ABS ABS, we need two additional constraints.
- The first penalizes the "marked" cases, which are instantiated with a prepositional marker. See Legendre et al. 1993, Aissen 1999, 2003, Anttila and Kim 2017.

(53) *MC: assign a violation to "prepositional" cases (ERG and DAT).

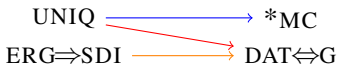
- The next constraint penalizes case frames which assign the same case to the subject and non-subject. See Toivainen 1993, Anttila and Fong 2000, and Collins 2010 for a precursor analysis for Samoan.

(54) UNIQ: assign a violation to case frames with the same case on both arguments.

- Using these two constraints, we can rule out ABS ABS and ERG DAT.

4.3 Lexical optimality in action

- This section demonstrates how the four proposed constraints (ERG⇒SDI, DAT⇔GOAL, *MC, UNIQ) interact to generate the right results.
- The inputs are verb meanings. These meanings vary on whether the higher argument is ±SDI, and whether the lower argument is ±GOAL. Outputs are case assignment patterns.
- The four proposed constraints allow for 24 (4!) possible rankings.
- I propose the following constraints on orderings, which narrow the 24 orders down to just 5.

(55)


- Any of the following rankings gets the right results.

(56)
1. UNIQ ≫ *MC ≫ ERG⇒SDI ≫ DAT⇔G
2. UNIQ ≫ ERG⇒SDI ≫ *MC ≫ DAT⇔G
3. UNIQ ≫ ERG⇒SDI ≫ DAT⇔G ≫ *MC
4. ERG⇒SDI ≫ UNIQ ≫ *MC ≫ DAT⇔G
5. ERG⇒SDI ≫ UNIQ ≫ DAT⇔G ≫ *MC

[9]Thanks to Kristine Yu for this observation.

- Below is a demonstration, arbitrarily using ranking number 1.

  Looking at the first two columns only

  - UNIQ/*MC only look at the outputs (case frames), so the violations are the same for any input.
  - Even though ABS ABS does best on *MC, it is ruled out by the higher ranked UNIQ.
  - *MC rules out ERG DAT due to the double marked cases.

- The crucial action, deciding between ERG ABS and ABS DAT takes place in the grayed-out cells.

(57) a. Verbs with SDI subjects and non-goal non-subjects (e.g., *fa'aleaga* 'destroy', *ifo* 'restrain')

⟨+SDI, -G⟩	UNIQ	*MC	ERG⇒SDI	DAT⇔G
⇒ERG ABS		*		
ABS DAT		*		!*
ABS ABS	!*			
ERG DAT		!**		*

  - ABS DAT satisfies ERG⇒SDI vacuously. It's eliminated by DAT⇔G as it assigns dative to a non-goal.

  b. Verbs with non-SDI subjects and non-goal non-subjects (e.g., *lata* 'be near', *pāgamālie* 'suit, fit')

⟨-SDI, -G⟩	UNIQ	*MC	ERG⇒SDI	DAT⇔G
ERG ABS		*	!*	
⇒ABS DAT		*		*
ABS ABS	!*			
ERG DAT		!**	*	*

  - ERG ABS is ruled out as ergative is assigned to a non-SDI.
  - As DAT⇔G is ranked lowest, the assignment of dative to a non-goal is tolerated.

  c. Verbs with SDI subjects and goal non-subjects (e.g., *fesoasoani* 'help', *'au'auna* 'serve')

⟨+SDI, +G⟩	UNIQ	*MC	ERG⇒SDI	DAT⇔G
ERG ABS		*		!*
⇒ABS DAT		*		
ABS ABS	!*			*
ERG DAT		!**		

  - ERG ABS is ruled out as dative is not assigned to a goal.

  d. Verbs with non-SDI subjects and goal non-subjects (e.g., *taotua* 'come after', *pisi* 'splash against')

⟨-SDI, +G⟩	UNIQ	*MC	ERG⇒SDI	DAT⇔G
ERG ABS		*	!*	
⇒ABS DAT		*		
ABS ABS	!*			*
ERG DAT		!**	*	

  - ERG ABS is ruled out as ergative is assigned to a non-SDI.

Summing up

- The generalization is that the ergative-absolutive case frame is associated with verbs whose:
  i. more agentive arguments are self-directed initiators, and
  ii. less agentive arguments are not goals
- The absolutive-dative case frame on the other hand is associated with verbs whose:
  i. less agentive arguments are goals, or
  ii. more agentive arguments are non-SDIs

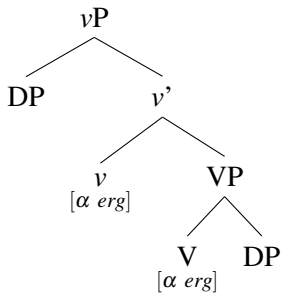
5 Conclusion

- This study provides us with many questions still to explore:
  - How do we account for Ochs' cases of "ergative drop"?
  - What's the right characterization of *fa'a-* and *-Cia*? (see Appendix B for ideas)
  - What about Mosel and Hovdhaugen's (1992) class of "labile verbs", which alternate between ABS DAT/ERG ABS? Are they independent lexical entries or a productive alternation or both?
- Taking stock, this paper proposes a way in which verbal meanings are paired with argument structures, providing an analysis of the link between morphological case assignment and verb semantics.
- The mapping defines a notion of possible argument structures given a verb's meaning, based on optimal satisfaction of violable constraints.
- The well-formedness constraints on lexical entries are determined "pre-syntactically", or independently of the narrow syntax, simply defining what it means to be a possible argument structure.
- For this reason, the proposal can be integrated with a variety of syntactic frameworks. Determining how the lexical entries proposed here interact with the narrow syntax promises to be a productive future research avenue.

Appendix A: Incorporating *v*P and Agree

- The paper proposes well-formedness constraints on lexical entries. However, it refrains from making too many assumptions about how these lexical items interface with the narrow syntax.
- Below is a demonstration of how the insights of the paper could be integrated with a specific set of assumptions: the *v*P hypothesis (Kratzer 1996) and Case assignment via Agree (Chomsky 2001).
- This is intended as a demonstration of the flexibility of the general approach: analogous proposals could be made incorporating differing syntactic assumptions, e.g. dependent case.
- Assume (58) is the general structure of transitive predicates. We take *v* and V to bare a categorial feature [±erg]. The values on the two heads have to match (via selection).

(58)



- The two *v* heads ([+erg] and [-erg]) are specified below, each selects for a matching V in its complement, and assigns ERG or no case to its specifier. DAT cannot be assigned due to its [−HR] feature.

(59) a. $v_{\text{+erg}} \rightsquigarrow \begin{bmatrix} \textit{Cat}: & v \\ \textit{Sem}: & \emptyset \\ \textit{Sel}: & \text{V}_{\text{+erg}} \\ & \text{DP}\begin{bmatrix} \theta: & [+\text{HR}],[-\text{LR}] \\ \text{K}: & \text{ERG} \end{bmatrix} \end{bmatrix}$

b. $v_{\text{-erg}} \rightsquigarrow \begin{bmatrix} \textit{Cat}: & v \\ \textit{Sem}: & \emptyset \\ \textit{Sel}: & \text{V}_{\text{-erg}} \\ & \text{DP}\begin{bmatrix} \theta: & [+\text{HR}],[-\text{LR}] \\ \text{K}: & \emptyset \end{bmatrix} \end{bmatrix}$

- Verbs only select for one DP, which denotes the less agentive of their two arguments.
- Vs can assign either nothing or DAT to this argument. ERG is ruled out because it is [−LR].

(60) a. *solo* 'wipe' $\rightsquigarrow \begin{bmatrix} \textit{Cat}: & \text{V}_{\text{+erg}} \\ \textit{Sem}: & \lambda x.\lambda y.\lambda e.\mathbf{wipe}(e) \wedge \mathbf{sdi}(e)(y) \wedge \mathbf{wipee}(e)(x) \\ \textit{Sel}: & \text{DP}_x\begin{bmatrix} \theta: & [-\text{HR}],[+\text{LR}] \\ \text{K}: & \emptyset \end{bmatrix} \end{bmatrix}$

b. *fesoasoani* 'help' $\rightsquigarrow$ $\begin{bmatrix} \textit{Cat}: & \text{V}_{\text{-erg}} \\ \textit{Sem}: & \lambda x.\lambda y.\lambda e.\textbf{help}(e) \wedge \textbf{sdi}(e)(y) \wedge \textbf{rcpt}(e)(x) \\ \textit{Sel}: & \text{DP}_x \begin{bmatrix} \theta: & [-\text{HR}], [+\text{LR}] \\ \text{K}: & \text{DAT} \end{bmatrix} \end{bmatrix}$

- The two entries in (60) are determined by the same four constraints, which are rephrased in order to fit the syntactic assumptions.
- Basically, constraints referencing case on the more agentive argument now reference whether or not V is [+erg] or [-erg].

(61) a. ERG⇒SDI: If a V is [+erg], then its external argument is an SDI

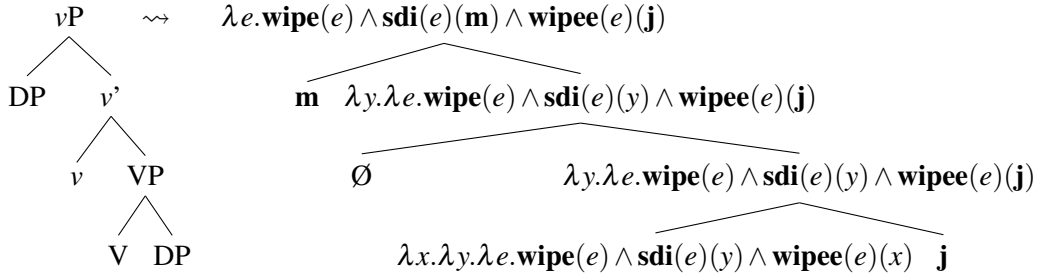
b. DAT⇔G: Assign violations to (i) goals with no dative case and (ii) dative case on non-goals

c. UNIQ: Assign violations to ABS ABS, i.e., Ø-case on DP_x selected by $\text{V}_{\text{-erg}}$ (which also assigns Ø-case to DP_y)

d. ∗MC: Assign violations to marked cases: DAT, and $\text{V}_{\text{+erg}}$ (which assigns ERG).

- Rankings remain the same, permitting lexical entries of the form in (60).
- Semantic composition proceeds as in (62).

(62)



- Where V or *v* assign no case to a DP, this means the DP is unvalued for case, bearing a $[u\text{K}]$ feature.
- Following the general approach of Chomsky 2001, the unvalued case feature can be valued by the closest c-commanding functional head *v* or T, in which case the $[u\text{K}]$ feature is evaluated as ABS.

Appendix B: Relation changing affixes

- Koopman 2012, contra Tollan, denies the link between agentivity and ERG, based on these data.

(63) a. *‘Ua ‘āmata ona* ***iloa*** *mea e le pepe*
PERF start COMP perceive.CIA thing ERG SPEC baby
The baby begins to notice things.

b. *ae* ***si‘omia*** *‘uma Apolima e le papa maualuga*
but surround.CIA all Apolima ERG SPEC rock high
but all of Apolima is surrounded by high rocks. (Mosel and Hovdhaugen 1992:4.233)

c. *Na* ***ufitia*** *atoa le motu e le pogisa*
PAST cover.CIA complete SPEC island ERG SPEC darkness
Darkness covered the whole island. (Mosel and Hovdhaugen 1992:4.235)

- But Koopman’s counterexamples to the generalization that ergatives are agentive have the *-Cia* suffix.
- Mosel and Hovdhaugen 1991 observe that the *-Cia* suffix[10], has an ‘ergativizing’ function, yielding verbs with an ergative-absolutive case frame, regardless of the case assignment of the root.

(64) a. *e* ***alofá*** *le fafine ‘i le pepe*
PRES love=ABS SPEC woman DAT SPEC baby
The woman loves the baby.

b. *e* ***alofagia*** *e le fafiné le pepe*
PRES love.CIA ERG SPEC woman=ABS SPEC baby
The woman treats the baby well.

- Looking at additional examples, we find that *-Cia* affixed verbs do not observe the SDI-generalization cited above: the ergative argument of a *-Cia* affixed verb may be a non-SDI (see the cases in (63)).
- Another systematic class of exceptions are clause-embedding verbs.
- Some clause-embedding verbs, e.g., *lagona* ‘feel, hear’ and *mafai* ‘can, able’, take ergative subjects, despite these subjects not fitting into the semantic category of SDIs.

(65) a. *Na* ***lagona*** *e le tinā ‘olo‘o tagi le pepe*
PAST feel ERG SPEC mother PROG cry SPEC baby
The mother heard that the baby was crying.

b. *‘Ua* ***mafai*** *nei e Ruta ona iloa mea*
PERF can now ERG Ruta COMP know thing
Ruta can now recognize things.

- Both *-Cia*-verbs and clause-embedding verbs which take ergative subjects are not well understood phenomena. I introduce them here to point out that they form systematic exceptions to the SDI-generalization, revised below.

[10]The form of the *-Cia* suffix is lexically specified, usually consisting of a thematic consonant (specific to the root) plus *-ia*.

(66) **Self Directed Initiator Hypothesis**: Samoan ERG ABS verbs which (i) are not affixed with *-Cia*, and (ii) select for a nominal object, entail that:

a. the ergative-marked argument denotes an initiating participant

b. the ergative-marked argument denotes a self-directed participant

- Excluding both *-Cia*-affixed verbs and clause-embedding verbs from the dataset, we do observe a link between ergative case and agentivity, contra Koopman 2012.
- How then should we incorporate relation changing affixes like *-Cia* into the general approach.
- The system pursued here gives us a way of distinguishing relation changing affixes.
  - Class 1: Affixed forms observe lexical optimality with respect to the proposed constraints.
  - Class 2: Affixed forms override lexical optimality with respect to the proposed constraints.
- We can distinguish two Samoan affixes. *fa‘a-* is Class 1, while *-Cia* is Class 2.
- *fa‘a-* is often glossed as a causative. But *fa‘a*-affixed verbs don't always have causative meanings.
- Furthermore, verbs with *fa‘a-* don't always have an ERG ABS case frame. Many take a ABS DAT case frame, depending on the semantics of the affixed form. The root is always intransitive or ABS DAT.

(67) ABS DAT *verbs with* fa‘a-

leo	sound	*fa‘aleo*	hold in high esteem
logo	perceive	*fa‘alogo*	hear
tusa	equal to	*fa‘atusa*	be like
manatu	think	*fa‘amanatu*	remind
lata	be near to	*fa‘alata*	draw near to
alofa	love	*fa‘aalofa*	regret

(Assign [ABS DAT])

(68) ERG ABS *verbs with* fa‘a-

ita	be angry	*fa‘aita*	make angry
mamafa	be important	*fa‘amamafa*	stress, emphasize
o‘o	arrive, happen	*fa‘ao‘o*	convey
afe	enter	*fa‘aafe*	invite
talitonu	annoy	*fa‘atalitonu*	convince
mātamata	look at	*fa‘amātamata*	make s.o. see

(Assign [ERG ABS])

- Although a more systematic study needs to be undertaken, initial sets of examples reveal that *fa‘a-* verbs fit into the same semantic generalization as we've observed for unaffixed forms above.
- Roughly, the *fa‘a*-affixed ABS DAT verbs above fit the general pattern of ABS DAT verbs: their subjects are non-SDIs or they have goal non-subjects.
- On the other hand, two-argument verbs affixed with *-Cia* are always ERG ABS. The root can be either ERG ABS or ABS DAT.

(69) *Verbs with* -Cia

nofo	stay	*nōfoia*	occupy
alofa	love	*alofagia*	treat well
ita	angry at	*itagia*	hate
mana'o	want	*mana'omia*	want, need
o'o	arrive, happen	*o'otia*	move, touch, strike
talisapai	approve	*talisapaia*	trust, have confidence in
ifo	restrain, control	*ifogia*	worship

- As stated earlier, two-argument *-Cia*-affixed verbs always have an ERG ABS frame, even if the ERG case ends up being assigned to a non-SDI. (70) repeats earlier examples.

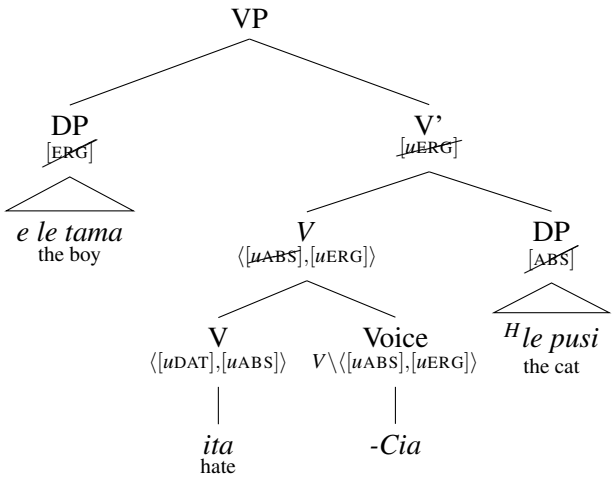
(70) a. *'Ua 'āmata ona* ***iloa*** *mea e le pepe*
PERF start COMP perceive.CIA thing ERG SPEC baby
The baby begins to notice things.

b. *ae* ***si'omia*** *'uma Apolima e le papa maualuga*
but surround.CIA all Apolima ERG SPEC rock high
but all of Apolima is surrounded by high rocks. (Mosel and Hovdhaugen 1992:4.233)

c. *Na* ***ufitia*** *atoa le motu e le pogisa*
PAST cover.CIA complete SPEC island ERG SPEC darkness
Darkness covered the whole island. (Mosel and Hovdhaugen 1992:4.235)

- While I stop short of providing an analysis of the syntactic-semantic contributions of *fa'a-* and *-Cia*, we can observe how they interact with case frames.
- We can understand forms affixed with *fa'a-* as being subject to the same constraints as unaffixed roots.
- Conversely, we can understand *-Cia* as applying to the output of the mapping process.
  - Via the constraints proposed above, the root associates with the case frame ERG ABS/ABS DAT.
  - On combining with *-Cia*, the case frame is overridden with ERG ABS.
- Below is a sketch of one way this could be implemented.
- *Cia* is a functional head which combines with the verb (e.g., pre-syntactically, or via head movement), with the effect of overriding its case frame. The complex affixed form must assign the ERG ABS frame.

(71)



References

Aissen, J. 1999. Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory* 17: 673–711.

Aissen, J. 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21: 435–483.

Anttila, A. and V. Fong. 2000. The partitive constraint in optimality theory. *Journal of Semantics* 17: 281–314.

Antilla, A. and J.-B. Kim. 2017. Locality and variation in Finnish structural case. *Natural Language and Linguistic Theory* 35: 581–634.

Baker, M. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: The University of Chicago Press.

Baker, M. 1997. Thematic roles and syntactic structure. *Elements of Grammar*, ed. by L. Haegeman, 73–137. Dordrecht: Kluwer.

Ball, D. 2007. On ergativity and accusativity in Proto-Polynesian and Proto-Central Pacific. *Oceanic Linguistics* 46: 128–153.

Ball, D. 2009. Clause structure and argument realization in Tongan. Ph.D. thesis, Stanford University.

Beaver, D. and C. Condoravdi. 2003. A uniform analysis of *before* and *after*. *The Proceedings of SALT* 13, ed. by R. Young and Y. Zhou, 37–54. Ithaca, NY: Cornell University.

Beavers, J. 2011. On affectedness. *Natural Language and Linguistic Theory* 29: 335–370.

Blume, K. 1998. A contrastive analysis of interaction verbs with dative complements. *Linguistics* 36: 253–280.

Borer, H. 2005. *Structuring Sense*. Oxford: Oxford University Press.

Bresnan, J. and J. Kanerva. 1989. Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic Inquiry* 20: 1–50.

Chomsky, N. 2001. Derivation by phase. *Ken Hale: A Life in Language*, ed. by M. Kenstowicz, 1–52. Cambridge, MA: MIT Press.

Chung, S. 1978. *Case Marking and Grammatical Relations in Polynesian*. Austin, TX: University of Texas Press.

Clark, R. 1976. *Aspects of Proto-Polynesian Syntax*. Auckland: Linguistic Society of New Zealand.

Collins, J. N. 2010. Syntactic derivations of Samoan predicates. B.A. thesis, University of Sydney.

Collins, J. N. 2014. The distribution of unmarked cases in Samoan. *Argument Realisations and Related Constructions in Austronesian Languages: Papers from the 12th International Conference on Austronesian Linguistics*, Vol. 2, ed. by I. W. Arka and N. L. K. Mas Indrawati, 93–110. Canberra: Asia-Pacific Linguistics.

Collins, J. N. 2016. Samoan predicate initial word order and object positions. *Natural Language and Linguistic Theory* 35: 1–59.

Cook, K. W. 1988. A cognitive analysis of grammatical relation, case, and transitivity in Samoan. Ph.D. thesis, UC San Diego.

Cook, K. W. 1993. A cognitive account of Samoan case marking and cliticization. *Studi Italiani di Linguistica Teorica e Applicata* 22: 509–530.

Croft, W. A. 1991. *Syntactic Categories and Grammatical Relations*. Chicago: University of Chicago Press.

Cruse, D. A. 1973. Some thoughts on agentivity. *Journal of Linguistics* 9: 11–23.

Dowty, D. 1979. *Word Meaning and Montague Grammar*. Dordrecht: Kluwer.

Dowty, D. 1991. Thematic proto-roles and argument selection. *Language* 67: 547–619.

Duranti, A. and E. Ochs. 1990. Genitive constructions and agency in Samoan discourse. *Studies in Language* 14: 1–23.

Fillmore, C. J. 1968. The case for case. *Universals of Linguistic Theory*, ed. by E. Bach and R. T. Harms, 1–88. New York: Holt, Rinehart and Winston.

Foley, W. A. and R. D. Van Valin. 1984. Functional syntax and universal grammar. Cambridge: Cambridge University Press. 339–354.

Folli, R. and H. Harley. 2005. Flavors of v. *Aspectual Inquiries*, ed. by P. Kempchinsky and R. Slabakova, 95–120. Springer.

Harley, H. 1995. Subjects, events, and licensing. Ph.D. thesis, MIT.

Hopper, P. and S. A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56: 251–299

Jackendoff, R. S. 1983. *Semantics and Cognition*. Cambridge: MIT Press.

Jackendoff, R. S. 1987. The status of thematic relations in linguistic theory. *Linguistic Inquiry* 18: 369–411.

Kikusawa, R. 2002. *Proto Central Pacific Ergativity: Its Reconstruction and Development in the Fijian, Rotuman, and Polynesian Languages*. Canberra: Pacific Linguistics.

Kiparsky, P. 1997. The rise of positional licensing. *Parameters of Morphosyntactic Change*, ed. by A. van Kemenade and N. Vincent, 460–494. Cambridge: Cambridge University Press.

Kiparsky, P. 2001. Structural case in Finnish. *Lingua* 111: 315–376.

Kiparsky, P. 2005. Absolutely a matter of degree: The semantics of structural case in Finnish. Paper presented at *the Chicago Linguistics Society*.

Koopman, H. 2012. Samoan ergatives as double passives. *Functional Heads: The Cartography of Syntactic Structures*, ed. by L. Brugé et al., 168–180. New York, NY: Oxford University Press.

Kratzer, A. 1996. Severing the external argument from its verb. *Phrase Structure and the Lexicon*, ed. by J. Rooryck and L. Zaring, 109–137. Dordrecht: Kluwer.

Krifka, M. 1989. Nominal reference, temporal constitution and quantification in event semantics. *Semantics and Contextual Expressions*, ed. by R. Bartsch, 75–115. Dordrecht: Foris Publications.

Krifka, M. 1992. Thematic relations as links between nominal reference and temporal constitution. *Lexical Matters*, ed. by I. A. Sag and A. Szabolcsi, 29–53. Stanford, CA: CSLI Publications.

Krifka, M. 1998. The origins of telicity. *Events and Grammar*, ed. by S. Rothstein, 197–235. Dordrecht: Kluwer Academic Publishers.

Kratzer, A. 2004. Telicity and the meaning of objective case. *The Syntax of Time*, ed. by J. Guéron and J. Lecarme, 389–424. Cambridge, MA: MIT Press.

Legendre, G., W. Raymond, and P. Smolensky. 1993. An optimality-theoretic typology of case and grammatical voice systems. *The Proceedings of BLS 19*. Berkeley, CA.

Levin, B. and M. Rappaport Hovav. 2005. *Argument Realization*. Cambridge: Cambridge University Press.

Milner, G. B. 1976. *Samoan Dictionary*, first published 1966. London: University Press.

Mosel, U. and E. Hovdhaugen. 1992. *Samoan Reference Grammar*. Oslo: Scandinavian University Press.

Ochs, E. 1982. Ergativity and word order in Samoan child language. *Language* 58: 646–671.

Poole, E. 2015. A configurational account of Finnish case. *The Proceedings of PLC* 38, ed. by D. Wilson.

Rappaport Hovav, M. and B. Levin. 2008. The English dative alternation: The case for verb sensitivity. *Journal of Linguistics* 44: 129–167.

Seiter, W. 1978. On the syntactic character of middle objects in Polynesian. *The Proceedings of ICAL* 2, ed. by S. A. Wurm and L. Carrington, 1289–1306. Canberra: Pacific Linguistics Series C.

Toivainen, J. 1993. The nature of the accusative in Finnish. *Case and Other Functional Categories in Finnish Syntax*, ed. by A. Holmberg and U. Nikanne, 111–128. Berlin: de Gruyter.

Tollan, R. 2018. Unergatives are different: Two types of transitivity in Samoan. *Glossa* 3(35): 1–41.

Wunderlich, D. 1997. Cause and the structure of verbs. *Linguistic inquiry* 28: 27–68.

Yu, K. 2011. The sound of ergativity: Morphosyntax-prosody mapping in Samoan. *The Proceedings of NELS 39*, ed. by S. Lima et al., 825–838. Amherst, MA: GSLA, University of Massachusetts.

Yu, K. M. and E. Stabler. 2017. (In)variability in Samoan prosody and consequences for syntactic parsing. To appear in *Laboratory Phonology*.