

“Verbal case” in Ashti Dargwa

Oleg Belyaev

Lomonosov Moscow State University/Institute of Linguistics RAS/Pushkin State Russian Language Institute

Proceedings of the LFG’20 Conference

On-Line

Miriam Butt, Ida Toivonen (Editors)

2020

CSLI Publications

pages 26–46

<http://csli-publications.stanford.edu/LFG/2020>

Keywords: argument structure, agreement, person hierarchy, East Caucasian languages, Dargwa, Glue Semantics

Belyaev, Oleg. 2020. “Verbal case” in Ashti Dargwa. In Butt, Miriam, & Toivonen, Ida (Eds.), *Proceedings of the LFG’20 Conference, On-Line*, 26–46. Stanford, CA: CSLI Publications.



Abstract

A few peripheral verbal forms in Ashti Dargwa (East Caucasian), which has hierarchical person agreement, use the markers *-i-* or *-u-* before the person agreement suffix. At first glance, these markers seem to indicate the grammatical function of the controller of person agreement: *-i-* is used when it is a transitive subject (ergative, *A*), while *-u-* is used when it is a transitive object or an intransitive subject (absolutive, *S/P*). The actual distribution, however, is more complex and cannot be easily described by a single rule: *-i-* is also used with subjects of unergative intransitives, reflexives, and, most puzzlingly, absolutive arguments of verbs with dative experiencer subjects. I show that this distribution cannot be described in terms of morphosyntactic features or GF configurations, and argue that, while an analysis in terms of argument structure is possible, a semantic analysis that connects the use of *-i-* and *-u-* with semantic role specifications of the arguments captures the data in the most straightforward manner. The analysis is formalized in Glue Semantics.

1 Introduction

Ashti is a variety of Dargwa,¹ a branch of East Caucasian notable, among other things, for its rich verbal morphology and the coexistence of gender and person agreement. While gender agreement is virtually always with the absolutive argument of the clause (*S/P*), person agreement follows a hierarchical pattern: in transitive clauses, agreement is with either *A* or *P* depending on which argument ranks higher on the person hierarchy 1, 2 > 3. That is, if one of the core arguments is a speech act participant (SAP) while the other is 3rd person, the verb agrees with the SAP argument. If both arguments are speech act participants, the verbs agrees in person with the absolutive. 3rd person agreement is only possible when both arguments are SAPs.

A few verb TAM paradigms (Generic Present and modal forms morphologically derived from it) include synthetic person markers which are preceded by one of the vowels *-i-* and *-u-*. The functions of these vowels are mysterious; while they are traditionally described as marking transitivity (cf. e.g. Magometov 1963 for Kubachi), Sumbatova and Mutalov (2003) show that in Itsari Dargwa, their distribution is closer to a kind of direct-inverse morphology, but the exact functions vary widely across different Dargwa languages. In this paper, I will show that the functions of *-i-* and *-u-* defy simple characterization in terms of syntax or argument structure, and propose a semantic account that captures their distribution.

¹This research has been supported by the Russian Science Foundation, project no. 18-18-00462. I am grateful to three anonymous reviewers of the LFG conference, two anonymous reviewers of the LFG Proceedings, and to the audience of LFG2020, especially Miriam Butt, for healthy criticism, insightful comments, and helpful advice. Special thanks are due to Ida Toivonen and Miriam Butt as editors of the Proceedings. Most of the data in this paper, but not the analysis, were earlier published in Belyaev (2016). All errors are mine.

Person agreement rules for various Dargwa varieties have been extensively described in the works of Nina Sumbatova, in particular Sumbatova (2011a) and Sumbatova (2011b). An OT-LFG analysis of Dargwa person agreement has been proposed in Belyaev (2013); in this paper, I will use a later version of this analysis in Belyaev (2017), which abandons the idea of m-structure features in favour of an agreement sharing approach. For the purposes of this paper, I will assume an accusative GF mapping for Dargwa ($\mathcal{A} \rightarrow \text{SUBJ}$, $\mathcal{P} \rightarrow \text{OBJ}$); if a version of Falk’s (2006) approach is adopted, as in Belyaev (2013) and Belyaev (2017), $\widehat{\text{GF}}$ can be used instead of SUBJ , and PIV identified with one of the core arguments; all the key parts of the analysis will be unaffected.

2 The distribution of *-i-* and *-u-*

2.1 Morphology

The synthetic TAM series whose endings are derived from the Generic Present set (Generic Present, Realis and Irrealis Conditional, Prohibitive) include one of the vowels *-i-* or *-u-* before the 1st and 2nd person markers:

	SG	PL
1	<i>-u/i-d</i>	<i>-u/i-d-a</i>
2	<i>-u/i-t</i>	<i>-u/i-t-a</i>
3	tr: <i>-u, -an</i> ; intr: <i>a, -an</i>	

Table 1: Generic Present endings

The main function of these markers seems to be to indicate the controller of person agreement (\mathcal{A} , \mathcal{P}), but upon closer scrutiny, the actual distribution is more complex. In what follows, I will describe the main contexts in which *-i-* and *-u-* tend to be used.

2.2 Core functions

With transitive verbs, *-i-* and *-u-* behave as a kind of “verbal case” affixes: *-i-* is used if the controller is ergative, *-u-*, if it is absolutive:

- (1) a. \mathcal{A} agreement: *-i-*

u-dil *čitu b-u:s-i-t*
 thou-ERG cat N-catch.IPFV-I-HAB.2[SG]
 ‘You (sg.) catch the cat.’

- b. \mathcal{P} agreement: *-u-*

čitu.l-dil *u* *u:s-u-t*
 cat-ERG thou [M]catch.IPFV-U-HAB.2[SG]
 ‘The cat catches you.’

In accordance with the ergative case marking pattern, intransitives generally use *-u-*:

- (2) *nus:a* *d-ax-u-d-a-l* ...
 we 1PL-go.IPFV-U-1-PL-COND
 ‘If we go...’

Thus, the initial generalization is that the two vowels indicate which of the core arguments is the controller of person agreement: *-i-* is used when the controller is the transitive subject (that is, the ergative, or \mathcal{A}), while *-u-* is used with objects and intransitive subjects (\mathcal{P}). This is an unusual morphological feature that does not have a common typological name; functionally, the closest equivalent is case. Another counterpart is the category of direct/inverse, which marks the relative position of arguments on a prominence hierarchy; since in Ashti, the choice of agreement controller is determined by the person hierarchy, in transitive clauses *-i-* may be said to be a direct marking, while *-u-* is the inverse. However, this generalization does not apply to intransitive verbs; this is why Sumbatova and Mutalov (2003), while describing a similar pattern in Itsari Dargwa, apply the “inverse” name to *-u-* with reservations.²

This general pattern, however, is violated in three contexts: with agent-like intransitives, with affective (dative experiencer subject) verbs, and with reflexives.

2.3 Split intransitivity

While the normal form of the “verbal case” affix with intransitive verbs is *-u-*, *-i-* can also be used with most intransitive verbs (and with some of them, preferably so). In this case, the subject is interpreted as somehow being more agent-like, or at least more in control of the situation. Thus, in (3a), with the “intransitive” *-u-*, the imperative is interpreted in a kind of admonitive sense: “be careful lest you fall”. In (3b), with *-i-*, the interpretation is rather that the subject should make an effort, perform specific actions so as not to fall.

- (3) a. *-u-* with intransitives: lack of control
ka-mma-w-i:k-u-t
 DOWN-PROH-M-fall.IPFV-U-2[SG]
 ‘[be careful,] do not fall [by accident]’

²The distribution of *-i-* vs. *-u-* in Itsari is different: *-i-* is used whenever \mathcal{A} is higher than \mathcal{P} on the prominence hierarchy $1,2 > 3$, while *-u-* is used when \mathcal{P} is lower than or equal to \mathcal{A} on the same hierarchy. The hierarchy for person agreement in Itsari, in contrast, is $2 > 1 > 3$. Thus, the pattern of *-i-/-u-* marking in Itsari, unlike in Ashti, is detached from person agreement and cannot be described in the same terms.

- b. *-i-* with intransitives: agentivity

ka-mma-w-i:č-i-t
 DOWN-PROH-fall.IPFV-I-2[SG]
 ‘do not fall [, make an effort]’

It is even possible to use *-i-* with verbs that clearly do not involve any agentivity in the direct sense of the word, such as ‘die’ in (4). In this case a control interpretation is “coerced”, in a way: The agent is interpreted as somehow potentially being responsible for their death via the actions that they are about to perform; e.g., (4) could be uttered when the speaker embarks on a dangerous journey. This example also illustrates the fact that the change in *-i/-u-* marking does not influence case marking on the subject, unlike fluid-S languages (Dixon, 1979) where similar semantic effects correlate with ergative/absolute marking of \mathcal{S} .

- (4) **du** *w-ibč-i-lli*, *qal gal.li-j d-ik:-a*
 I M-die.PFV-A-COND.1SG house son-DAT NPL-give.PFV-IMP.SG
 ‘If I die, give the house to (my) son.’

The distribution of *-i-* and *-u-* with intransitives is also reminiscent of the unergative/unaccusative distinction (Perlmutter, 1978; Hout, 2004), although in Ashti it is less lexically conditioned than in more prototypical cases. Interestingly, the effect of using *-i-* and *-u-* with movement verbs (5) is exactly the same as the effect of the choice between ‘be’ and ‘have’ in Romance, cf. (6) from Italian: the former enforces a telic interpretation (while mainly occurring with unaccusatives), while the latter, atelic (while occurring with unergatives). In general, the use of *-u-* vs. *-i-* with intransitives in Ashti agrees with the Auxiliary Selection Hierarchy proposed in Sorace (2000) based on Romance data.

- (5) a. *-u-* with intransitives: telicity
pat’imat.li-š:u w-ax~max-u-t
 P.-APUD[LAT] M-go.IPFV~PROH-U-2[SG]
 ‘do not go to Patimat’
- b. *-i-* with intransitives: atelicity
w-aš~maš-i-t
 M-go~PROH-I-2[SG]
 ‘do not go [anywhere]’

- (6) Italian

- a. ‘be’ with intransitives: telicity
sono corso a casa
 am run to house
 ‘I have run home.’

- b. ‘have’ with intransitives: atelicity

ho corso ore e ore
 I.have run hours and hours
 ‘I have run for hours and hours.’

2.4 Affective verbs

Ashti, like most other East Caucasian languages, has verbs with dative-marked experiencer subjects and absolutive stimuli; verbs with such valency frames are traditionally called “affective verbs”. The specific range of affective verbs varies across East Caucasian languages and even within the Dargwa branch; in Ashti, it includes perception verbs like ‘see’, ‘hear’; volitional and emotional predicates like ‘want’ or ‘like’ (the latter two are actually the same verb); cognition verbs such as ‘think’, ‘understand’, and ‘know’.

For the purposes of most syntactic phenomena (person agreement, control, binding etc.), these dative-marked subjects behave just as ordinary *As*:

- (7) dam *murad ʔulħ-i^a-d*
 I.DAT M. [M]see.PFV-PRET-1[SG]
 ‘I saw Murad.’

- (8) *dam* u *ʔulħ-i-t:i*
 I.DAT thou [M]see.PFV-PRET-2[SG]
 ‘I saw you.’

But with these verbs, *-i-* has to be used regardless of the grammatical function of the controller:⁵

- (9) a. dam *pat’imat j-ulħ-i-d* / **j-ulħ-u-d*
 I.DAT P. F-see.IPFV-A-1[SG]
 ‘I see Patimat.’
- b. *pat’imat.li-j* du *ʔulħ-i-d* / **ʔulħ-u-d*
 P.-DAT I [M]see.IPFV-A-1[SG]
 ‘Patimat sees me.’

This is unexpected, because, if anything, affective verbs are expected to be closer to intransitives, with the dative phrase lacking some of the subject properties. But the dative subject behaves just like the ergative subject, controlling agreement and

⁴The morpheme *-i-* in this example is one of the markers of the Preterite paradigm (*-a-* in other conjugation classes) which is a different morpheme than the *-i-* used in the habitual and modal paradigms. It does not change depending on the agreement controller.

⁵The verb ‘to see’ belongs to the *-un-* conjugation class, where the invariable marker *-a-* can optionally be used instead of both *-i-* and *-u-*. However, *-i-* is also invariably used with “affective” verbs of other conjugations, which do not have this morphological trait.

selecting *-i-* in the suffix. The absolutive, however, gets *-i-* as well, even though it has nothing agentlike about it – there is no direct parallel to unergatives, where this distribution is expected.

2.5 Reflexives

Another challenge to the \mathcal{A}/\mathcal{P} generalization comes from reflexives. These can only use *-i-*:

- (10) *di-l du w-aq[˘]-aq[˘]-i-lli* ...
 I-ERG I M-wound.PFV-CAUS-A-COND.1[SG]
 ‘If I wound myself..’

The general principles of Ashti agreement suggest that \mathcal{P} “wins” if both have the same rank. Otherwise, reflexives clauses could be expected to be detransitivized in some sense. In both cases, the expected affix choice is *-u-*,⁶ or at least free variation between *-i-* and *-u-*. Yet the verb here behaves as if the agreement controller is always \mathcal{A} (selecting *-i-*).

2.6 Summary

To summarize the data of this section, the distribution of *-i-* and *-u-* with transitive verbs and intransitives is given in Table 2.

\mathcal{A}	\mathcal{P}		
	1	2	3
1	-i-d	-u-t	-i-d
2	-u-d	-i-t	-i-t
3	-u-d	-u-t	(-u/-an)
S_{unacc}	-u-d	-u-t	(-a/-an)
S_{unerg}	-i-d	-i-t	(-a/-an)

Table 2: *-i-* and *-u-* with transitives and intransitives

With affective verbs, the distribution is different in that *-i-* is used throughout, see Table 3.

\mathcal{A}	\mathcal{P}		
	1	2	3
1	-i-d	-i-t	-i-d
2	-i-d	-i-t	-i-t

⁶This is apparently the case in Itsari (Sumbatova and Mutalov, 2003), although the data are only given in a table, with no examples.

3 -i-d -i-t (-u/-an)
 Table 3: *-i-* and *-u-* with “affective” verbs

3 Analysis

The distribution of Ashti *-i-* vs. *-u-* is problematic because it cannot be easily tied to any one specific parameter, and different mismatches work in different directions. The initial data on transitive and intransitive verbs indicate that *-i-* is associated with \mathcal{A} (ergative subjects), while *-u-* is associated with absolutive arguments (direct objects and intransitive subjects). The split in intransitive verbs can be described within the same logic: *-i-* marks “agentive” arguments, while *-u-* marks arguments that are more patient-like. However, the opposite logic seems to work in affective verbs: *-i-* is used for both experiencer and stimulus, although the absolutive argument of affective verbs is not agent-like in any way. Finally, the use of *-i-* in reflexive contexts is difficult to explain in the general logic of ergative vs. absolutive agreement.

In what follows, I will evaluate several possible solutions to this problem in the LFG architecture. I will not provide a definitive conclusion, because the phenomenon requires further study, but the discussion herein can serve as the basis for a more developed approach.

3.1 Case

One option is to follow through with the analogy between *-i-* vs. *-u-* and case, treating the distribution of *-i-* and *-u-* as largely idiosyncratic and not directly connected to any semantic or syntactic features. The feature `CASE` has to be introduced to generalize over the mapping between grammatical functions and nominal “flags” (in the terminology of Haspelmath, 2019); if a “flag” directly reflects grammatical function, the syntactic feature `CASE` is not needed, but this is rarely the case (Spencer and Otoguro, 2005; Spencer, 2009). In a similar fashion, one could introduce a separate feature `VCASE` (for “verbal case”) to indicate whether an NP, when indexed by person agreement, should be marked as *-i-* or *-u-*. As seen from (4), dependent case marking in Ashti is independent from *-i-* vs. *-u-* marking on the verb: the intransitive subject stays absolutive even if verbal marking changes. Thus, `VCASE`, if it is introduced, should be kept distinct from `CASE`.

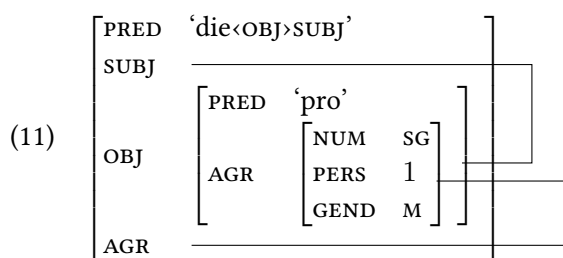
If such a feature is used, *-i-* and *-u-* form a typologically common two-term system (Arkadiev, 2006; Arkadiev, 2009): *-u-* is the direct/nominative “case”, specialized in marking direct objects and non-agentive intransitive subjects, while *-i-* is a typical polyfunctional oblique, used in all other contexts, that do not form a natural class. Such a solution allows us to avoid trying to capture the distribution of *-i-* in a homogeneous way: affective verbs are “double oblique” verbs simply because they are lexically defined as such. Like in many other languages

with reduced case systems, VCASE would then be restricted to 1st and 2nd person pronouns.

Even though such an analysis may well represent the diachronic origin of *-i-* and *-u-* marking (if the person suffixes go back to incorporated pronouns),⁷ it is hardly adequate synchronically, as it stipulates a feature for which all verbal arguments (or at least all verb subcategorization frames) have to be marked, but which only surfaces in very forms in the verbal paradigm.⁸ Another problem with this approach is that it fails to account for the behaviour of reflexives: If “ergative” pronouns are assumed to have the “oblique” feature (*-i-*) and “absolutive” pronouns are assumed to be “direct” (*-u-*), there is no motivation for *-i-* marking in (10), since competition between two speech argument participants normally results in agreement with the absolutive. One would then have to assume that verbs with reflexive objects are also double obliques, or behave in the same way as agentive intransitives; neither assumption has any basis, because reflexives occupy normal argumental positions and there is no detransitivizing morphology on the verb (the 3rd person marker of some paradigms distinguishes transitive and intransitive verbs).

3.2 Syntax

A purely syntactic account would require finding something in common between \mathcal{A} , unergative \mathcal{S} , and both arguments of affective verbs. It is clear that in the standard LFG view of GFs and their distribution, no such common features can be found. If this idea is taken seriously, one would have to assume that unaccusatives, where the verb is marked with *-u-*, are underlyingly transitive; they would then have an OBJ thematic argument and a non-thematic SUBJ argument structure shared with the object, like in (11). This allows us to uniformly describe *-u-* as marking the OBJ controller, and *-i-* as marking the SUBJ controller.



⁷Based on the similarity with auxiliary selection in Romance, one could speculate that verbal endings go back to different auxiliaries instead. Split auxiliary selection is not attested in East Caucasian, but neither is “verbal case” of this kind in general, or hierarchical person agreement outside Dargwa, for that matter. The origin of Dargwa person agreement morphology remains obscure, see Sumbatova (2011a) for an overview.

⁸An anonymous reviewer wonders whether a single stipulation for a subset of grammar could be better than a grammar-wide generalization if it only accounts for a relatively obscure phenomenon. This is true in principle, but in this case, the stipulation would not be minor: it has to cover all clauses.

This solution is clearly less-than-adequate, primarily because there is no independent motivation for such structures; significant changes to the f-structure of numerous verbs have to be introduced purely on the basis of *-i-* vs. *-u-* marking. Capturing the behaviour of affective verbs requires even more complicated solutions: the stimulus would have to be promoted to SUBJ status if it controls verbal person agreement, but not in other contexts. Finally, this analysis does not solve the problem of reflexives, because these would have to be assumed to involve structure sharing between SUBJ and OBJ, which has no syntactic motivation as the reflexive is expressed by a separate NP. To conclude, *-i-* and *-u-* cannot be tied to specific grammatical functions, although it is interesting that an attempt to do this ends up recreating the core assumptions of the Unaccusative Hypothesis.

3.3 Argument structure

3.3.1 Core proposal

Another option is to characterize the distinction between *-i-* and *-u-* in terms of differences in argument structure, i.e. in the module of grammar that determines the mapping between semantic roles and grammatical functions. Lexical Mapping Theory, or LMT (Bresnan and Kanerva, 1989) in LFG typically operates with the decomposition of the core grammatical functions according to two features, [+/-o] and [+/-r], in the way shown in 4.

	-r	+r
-o	SUBJ	OBL _θ
+o	OBJ	OBJ _θ

Table 4: Cross-classification of grammatical functions in LFG

Thematic roles have inherent (under)specifications such as [-o] for agents, [-r] for patients; these are then sequentially mapped to the lowest compatible position on the markedness hierarchy:

- SUBJ \succ OBJ, OBL_θ \succ OBJ_θ

Unaccusativity has been explained in LMT (Bresnan and Zaenen, 1990) in terms of different underspecification patterns for different kind of intransitive subjects: unergative ones are [-o] (akin to transitive subjects) while unaccusative ones are [-r] (akin to direct objects). The standard specifications of transitive verbs and two kind of intransitives are given in (12).

- (12) a. agentive transitive catch \langle *ag* *pt* \rangle
[-o] [-r]
- b. unergative run \langle *ag* \rangle
[-o]

- (16) a. *pat'imat.li-j b-ulh-in*
 P.-DAT N-see.IPFV-PRET.3
 ‘Patimat used to see.’
- b. *#pat'imat j-ulh-in*
 P. F-see.IPFV-PRET.3
 *‘Patimat used to see.’, #‘Someone used to see P.’

Antipassivization in Ashti is basically an instance of Unspecified Object Deletion, as in (17): The [-r] argument is suppressed, and the [-o] becomes the sole argument of an intransitive verb, which, in an ergative language, gets absolutive marking. The “ergative” patient that can be optionally expressed can either be treated as an adjunct or as an oblique.

- (17) eat < *ag pt* >
 [-o] [-r]
 ↓
 ∅

If the stimulus of affective verbs is treated as [+o], it cannot be deleted due to the generalization that only unmarked arguments can be suppressed (Alsina, 1990; Bresnan, Asudeh, et al., 2016, p. 333).

3.3.3 Reflexives

In clauses with reflexive direct objects, both arguments, of course, have the same rank; therefore, according to the general rules, the agreement controller should be the patient, not the agent — thus, *-u-* is expected. But *-i-* is used instead. The only way to solve this contradiction in an argument-structure-based approach is to assume that reflexive clauses are actually intransitive: that is, the reflexive acts as a detransitivizer of sorts. Such an analysis would be very artificial, however, since I am aware of no data that points to reflexive clauses actually being intransitive.

3.3.4 Referring to a-structure features

The final remaining problem is that there is no mechanism in classic LMT for morphosyntactic elements to refer directly to a-structure underspecifications. The analysis of Bresnan and Zaenen (1990) only applies to argument structure derivation (resultatives); it was not meant to capture purely f-structure / syntactic phenomena. Kibort (2014) further develops LMT by dispensing with semantic roles, only leaving positions *arg₁*, *arg₂*, etc., which are associated with feature underspecifications. Findlay (2016) further reduces the role of a-structure, by treating *arg_n* as s-structure features inherently associated with certain feature specifications ([-o], etc.); but these feature specifications themselves are nothing more than disjunctions of GFs that correspond to this specification in classic LMT, e.g. [+o]

is the disjunction $\text{PLUSO} \equiv \{\text{OBJ} \mid \text{OBJ}_0\}$. This leaves us with no way to refer to features like [+o], neither in classic LMT nor in its more recent versions.

A possible way of “saving” such an analysis, at least technically, is to replace feature specifications with ARG_N positions. In these terms, transitive verbs in Ashti operate with arg_1 and arg_2 (just like standard transitive verbs in other languages), while affective verbs have only arg_1 and arg_3 :

- (18) a. agentive transitive
 catch $\langle \text{arg}_1 \text{arg}_2 \rangle$
 [-o] [-r]
- b. “affective”
 see $\langle \text{arg}_1 \text{arg}_3 \rangle$
 [-o] [+o]

In this system, the definition of the 1st person agreement marker with *-u-* can be as in (19). Agreement markers with *-i-* would then be described as the negative version of *-u-*, i.e. describing the negation of the conjunction of the last two equations in (19), or the disjunction of two negative equations. For a more generalizing analysis, these two equations could be put into a template; the definition of *-i-* would then negate this template.

- (19) *-u-d* %GF = $\{(\uparrow \text{SUBJ}) \vee (\uparrow \text{OBJ})\}$
 (%GF AGR PERS) =_c 1
 (%GF AGR NUM) =_c SG
 (\uparrow AGR) = (%GF AGR)
 %GF _{σ} =_c (\uparrow_{σ} ARG₂)

That predicates can be freely associated with different argument slots, and that these may be non-contiguous, is explicitly acknowledged in Findlay (2016): “These argument positions are ordered, and a predicate can select any combination of them – that is, not necessarily a contiguous subsection: a predicate could select an arg_1 and an arg_4 , for example” (p. 301). At least formally, then, an analysis along these lines is possible. However, the deeper question – *why* stimuli get assigned to arg_i is left unresolved. Latest versions of LMT explicitly avoid an association between arg_n positions and semantic roles, so trying to explain it in this way would be at odds with the general logic of the theory. This analysis would find stronger support if several pieces of evidence conspired to motivate it, but at present, the only evidence is *-i-* vs. *-u-* and the impossibility of antipassive mappings.

The issue of reflexives having *-i-* is also still difficult to resolve in terms of LMT, and it is not clear if stipulating that verbs with reflexive direct objects are intransitive would solve more problems than it would introduce.

3.4 Semantics

3.4.1 Core proposal

A final possibility is to analyze the distribution of *-i-* and *-u-* in semantic terms, i.e. in terms of the semantic roles that the corresponding agreement controller maps to. The following generalization could be proposed:

- (20) *-u-*: the agreement controller is the Patient or Theme;
-i-: used elsewhere.

However, it is problematic for morphology to distinguish between Patients / Themes and Stimuli (despite both being mapped to Absolutive OBJ), because stimuli are not usually viewed as a separate theta-role (although they are indeed often discussed when considering more fine-grained issues); of course, semantic roles can be viewed as language-specific, but then this analysis would not be much different from the argument structure analysis, which refers to language-specific, highly idiosyncratic argument structure assignments. A semantic analysis would only have an advantage if it could make use of some intrinsic semantic property of “affective” predicates, but it is clear that traditional conceptions of theta-roles do not involve such subtle distinctions. In particular, stimuli are still Proto-Patients in terms of Dowty (1991), and certainly they are not any more agent-like than patients proper. Furthermore, this analysis requires that all subjects of intransitive verbs that trigger *-i-* are treated as agents. For verbs like ‘die’, this is implausible: with *-i-*, the subject is indeed interpreted as causing the situation in some way, but in the end, it still remains primarily a patient.

Another problem is the way one could refer to semantic roles in the syntax. In some approaches to argument structure, this is simple enough. For example, if the a-structure projection of Butt et al. (1997) is adopted, we can use definitions like the following:

- (21) *-u-d* %GF = $\{(\uparrow \text{SUBJ}) \vee (\uparrow \text{OBJ})\}$
(%GF AGR PERS) =_c 1
(%GF AGR NUM) =_c SG
(\uparrow AGR) = (%GF AGR)
(PATIENT λ^{-1} (%GF))

However, with the advent of Glue Semantics, there is a growing consensus in LFG that there are good reasons to treat semantic roles as belonging to meaning representation and not to a- or s-structure (Asudeh and Giorgolo, 2012). In the system of Asudeh & Giorgolo, further developed in Asudeh, Giorgolo, and Toivonen (2014) (and which Findlay, 2016 builds upon), s-structure only has abstract argument slots, not direct representations of thematic roles; the latter are only represented in the meaning languages using Neo-Davidsonian predicates like *patient(e) = x*.

I believe that both of these problems can be overcome, and a generally coherent semantic analysis of *-i-* and *-u-* can be provided, if one adopts a different

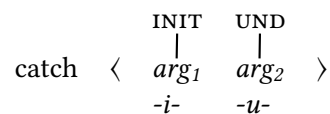
system of semantic roles and introduces constraints on them in the meaning language itself via the appropriate meaning constructors.

3.4.2 Additional semantic distinctions

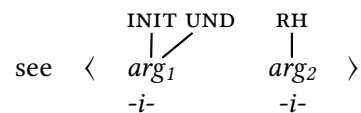
In order to explain the unusual behaviour of “affective” verbs, it is instructive to compare the analysis of a rather similar phenomenon – dative experiencer subjects in Icelandic – in Schätzle (2018). Schätzle provides a compelling account of the syntactic behaviour and case-marking of various types of arguments in Icelandic by combining insights from several different areas of grammar and research traditions. One way in which this work is of significance to the analysis of Ashti is that Schätzle essentially adapts the theory of Ramchand (2008) to LFG. Ramchand’s approach crucially depends on the positions of the arguments in the syntactic structure that she postulates: Initiators, in her system, are the specifiers of *initP* (i.e. “subjects” of the initiating subevent); Undergoers are specifiers of *procP* (“subjects” of the process subevent), Rhemes and Paths are complements of *procP*, and so on. Schätzle instead reinterprets notions like *INITIATOR* and *UNDERGOER* as semantic roles in an approach that combines the main ideas of Butt et al. (1997) with the newer developments in LMT described above: semantic roles map to *arg_n* positions, which, in their turn, are associated with feature specifications that map them to GFs. This leads to a system where semantic roles, arguments structure, and grammatical functions are neatly separated.

What is crucial for this paper is that Schätzle’s approach allows us to incorporate the core insight of Ramchand’s analysis of argument structure. Instead of the traditional – diverse and often confusing – inventory of semantic roles, Ramchand operates with a restricted set of primitive roles (Initiator, Undergoer, Result, Rheme, etc.) that can be combined – i.e. mapped to one argument – in various ways to yield different verb classes. Ramchand’s specifications for each of the verb classes discussed in this paper (in Schätzle’s representation) are as follows:

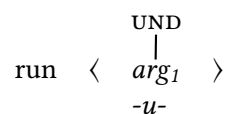
(22) a. agentive transitive



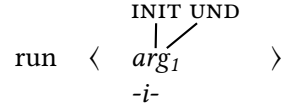
b. “affective”



c. unaccusative



d. unergative



Thus, in agentive transitive verbs, the subject is the Initiator and the object is the Undergoer. In “affective” verbs with experiencer subjects, the subject is simultaneously the Initiator and the Undergoer, while the object is a Rheme;⁹ in unaccusatives, the subject is the Undergoer, while in unergatives, the subject is both the Initiator and the Undergoer, like in “affective” verbs. The latter specification is significant, as it naturally captures the use of *-i-* with verbs like ‘die’: using *-i-* interprets the verb as having an initiating subevent, but keeps the subject an Undergoer of the process subevent.

It can be readily seen from (22) how the distribution of *-i-* and *-u-* can be described: *-u-* marks “exclusive” Undergoers that are not shared with any other semantic role, while *-i-* is the “default” option that marks all other argument types, including Undergoers that simultaneously act as Initiators. Thus, Ramchand’s theory, in the interpretation of Schätzle (2018), allows capturing the relevant generalizations without any additional stipulation, which is a significant advantage over the argument structure approach.

3.4.3 Semantic interpretation

Without going into the details of the system of Asudeh, Giorgolo, and Toivonen (2014), its core ideas can be described as follows. Verbal lexical entries do not directly encode their valency, like in standard Glue analyses, but have a generic meaning constructor like $\lambda e. \text{laugh}(e) : (\uparrow_\sigma \text{EVENT}) \multimap \uparrow_\sigma$. Arguments are then introduced via calling additional templates like AGENT, which do two things: first, define the mapping from GFs to ARG_N positions at s-structure via templates such as ARG1, etc. (following standard LMT principles, discussed in detail in Findlay 2016); second, introduce the specific semantic roles via separate meaning constructors such as $\lambda P \lambda x \lambda e. P(e) \wedge \text{agent}(e) = x : [(\uparrow_\sigma \text{EVENT}) \multimap \uparrow_\sigma] \multimap (\uparrow_\sigma \text{ARG}_1) \multimap (\uparrow_\sigma \text{EVENT}) \multimap \uparrow_\sigma$. Ramchand’s system can be adopted to this approach by replacing predicates like *agent(x)* with predicates corresponding to Ramchand’s roles (traditional theta-roles are effectively redundant in this system); the meaning constructors can be formulated in such a way that one argument carries more than one role. For example, agentive transitive verbs might have meaning constructors as in (23a), while unergatives, where the same argument is the Initiator and the Undergoer, as in (23b).

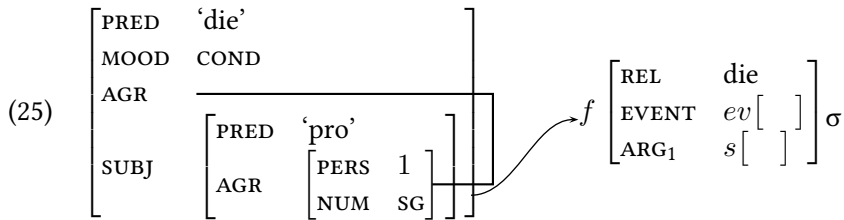
⁹Another option for stative verbs with experiencer subjects is to have the subject as the Initiator and the object as a Rheme; Ramchand analyses ‘fear’ in this way, while ‘see’ is treated akin to verbs with incremental themes. For my purposes, this distinction is not important; crucially, in both classes of verbs stimuli are Rhemes, not Undergoers, which described the Ashti distribution.

- (23) a. $\lambda P \lambda x \lambda e. P(e) \wedge \text{initiator}(e) = x : [(\uparrow_\sigma \text{EVENT}) \multimap \uparrow_\sigma] \multimap (\uparrow_\sigma \text{ARG}_1) \multimap$
 $(\uparrow_\sigma \text{EVENT}) \multimap \uparrow_\sigma$
 $\lambda P \lambda x \lambda e. P(e) \wedge \text{undergoer}(e) = x : [(\uparrow_\sigma \text{EVENT}) \multimap \uparrow_\sigma] \multimap (\uparrow_\sigma \text{ARG}_2) \multimap$
 $(\uparrow_\sigma \text{EVENT}) \multimap \uparrow_\sigma$
- b. $\lambda P \lambda x \lambda e. P(e) \wedge \text{initiator}(e) = x \wedge \text{undergoer}(e) = x : [(\uparrow_\sigma \text{EVENT}) \multimap$
 $\uparrow_\sigma] \multimap (\uparrow_\sigma \text{ARG}_1) \multimap (\uparrow_\sigma \text{EVENT}) \multimap \uparrow_\sigma$

The behaviour of *-i-* and *-u-* can then be described by making person agreement markers “modify” the semantic specification of the agreement controller by adding additional semantic role predicates. Consider the proposed contribution of the 1st person singular marker *-ud* in (24).

- (24) *-u-d* %AGR = {(\uparrow SUBJ)|(\uparrow OBJ)}
 $(\uparrow \text{AGR}) = (\% \text{AGR AGR})$
 $(\% \text{AGR AGR PERS}) = 1$
 $(\% \text{AGR AGR NUM}) = \text{SG}$
 $\lambda P \lambda x \lambda e. P(x)(e) \wedge \text{undergoer}(e) = x \wedge \text{initiator}(e) \neq x :$
 $(\% \text{AGR}_\sigma \multimap (\uparrow_\sigma \text{EVENT}) \multimap \uparrow_\sigma) \multimap (\% \text{AGR}_\sigma \multimap (\uparrow_\sigma \text{EVENT}) \multimap \uparrow_\sigma)$

In this f-description, %AGR is the local name for the agreement controller, which is freely identified with either SUBJ or OBJ, to be later filtered by OT constraints according to the person hierarchy: only higher-ranking controllers, or patients in transitive clauses with two SAPs, are licensed (see Belyaev, 2017 for more detail). The AGR feature of this argument is shared with the AGR feature of the clause. The meaning constructor adds two more statements via conjunction to the entailments introduced by meaning constructors as in (23). Consider the f- and s-structures for (4) in (25): SUBJ maps to ARG1 at s-structure and is associated with the semantic role Undergoer (in the unaccusative interpretation). Instantiating the meaning constructors in (23) and (24) gives us the proof in (26). The resulting meaning constructor must then be combined with a tense or mood operator (in this case, conditional) to yield the resource *f*, but this last step is not important for this paper.



(26)

$$\begin{array}{c}
\lambda P \lambda x \lambda e. P(x)(e) \wedge \text{und}(e) = x \wedge \text{init}(e) \neq x : \\
\frac{(s \multimap ev \multimap f) \multimap (s \multimap ev \multimap f)}{\lambda x \lambda e. \text{die}(e) \wedge \text{und}(e) = x \wedge \text{und}(e) = x \wedge \text{init}(e) \neq x :} \\
\frac{\lambda P \lambda x \lambda e. P(e) \wedge \text{und}(e) = x : \quad \lambda e. \text{die}(e) :}{\frac{(ev \multimap f) \multimap s \multimap ev \multimap f \quad ev \multimap f}{\lambda x \lambda e. \text{die}(e) \wedge \text{und}(e) = x :}} \\
\frac{s \multimap ev \multimap f}{\lambda e. \text{die}(e) \wedge \text{und}(e) = me \wedge \text{und}(e) = me \wedge \text{init}(e) \neq me :} \quad \begin{array}{l} me : \\ s \end{array} \\
ev \multimap f
\end{array}$$

Using *-u-* here makes the sentence grammatical, but it is clearly incompatible with examples where the agreement controller is also the initiator, because these would contain a logical contradiction.

The definition of *-i-d* is semantically essentially the negation of *-u-d* (being logically equivalent to $\neg[\text{undergoer}(e) = x \wedge \text{initiator}(e) \neq x]$):

$$\begin{array}{l}
(27) \quad -i-d \quad \% \text{AGR} = \{(\uparrow \text{SUBJ}) | (\uparrow \text{OBJ})\} \\
\quad \quad \quad (\uparrow \text{AGR}) = (\% \text{AGR AGR}) \\
\quad \quad \quad (\% \text{AGR AGR PERS}) =_c 1 \\
\quad \quad \quad (\% \text{AGR AGR NUM}) =_c \text{SG} \\
\quad \quad \quad \lambda P \lambda x \lambda e. P(x)(e) \wedge [\text{undergoer}(e) \neq x \vee \text{initiator}(e) = x] : \\
\quad \quad \quad (\% \text{AGR}_\sigma \multimap (\uparrow_\sigma \text{EVENT}) \multimap \uparrow_\sigma) \multimap (\% \text{AGR}_\sigma \multimap (\uparrow_\sigma \text{EVENT}) \multimap \uparrow_\sigma)
\end{array}$$

If this is applied to a verb where the first disjunct is always false, such as ‘die’, the second disjunct must necessarily be true. This causes the “coercion” effect which we observed above: Even verbs which are not lexically agentive are interpreted as having some kind of an initiating subevent, precisely because the Initiator semantic role is in fact introduced by the suffix *-i-*.

Another advantage of this approach is that it explains why verbs with reflexive objects use *-i-*. Indeed, reflexives involve bound variable anaphora (Reinhart, 1983), thus the equations $[\text{initiator}(e) = x]$ and $[\text{undergoer}(e) = x]$ will necessarily hold for the same x . This precludes the use of *-u-*, because this requires the agreement controller to be an Undergoer while not being an Initiator. Hence the only option is to use *-i-*, which is consistent with the empirical data.

4 Conclusions

In this paper, I have described an interesting phenomenon in Ashti Dargwa where in certain verbal forms, one of the two suffixes, *-i-* or *-u-*, appears before the person agreement marker. These suffixes, which, at first glance, seem to indicate the grammatical function of the agreement controller (\mathcal{A} vs. \mathcal{P}), actually have a more complex distribution. While *-u* seems to be restricted to patient-like arguments, the distribution of *-i-* is less clear in that it is licensed with a seemingly heterogeneous class of arguments: transitive subject, unergative subjects of intransitive

verbs, reflexives and, most puzzlingly, both experiencers and stimuli of “affective” verbs (transitive with dative experiencer subjects). I sketch four possible analyses of this phenomena: in terms of a special “verbal case” feature, syntax, argument structure, and semantics. The “verbal case” analysis and the syntactic analysis can be immediately rejected, as they introduce too many stipulations that have no independent empirical confirmation. An argument-structure analysis is technically possible, but requires introducing a language-specific mapping of stimuli, which, again, has little empirical motivation. Finally, I propose a semantic analysis which uses Ramchand’s (2008) analysis of argument structure, inspired by the analysis of Icelandic dative subjects in Schätzle (2018). I show that Ramchand’s approach, combined with the theory of valency in Asudeh, Giorgolo, and Toivonen (2014), allows for an elegant and natural analysis of *-i-* and *-u-* in semantic terms: *-u-* is licensed by Undergoers that do not share the semantic role of Initiator, while *-i-* is licensed with all other roles (i.e. in contexts that constitute the negation of the definition of *-u-*). This analysis captures the relevant data and allows for a natural treatment of an “agentivity coercion” effect that arises from using *-i-* with seemingly purely patient-like arguments like the subject of the verb ‘die’. Unlike all other analyses, it also correctly predicts the behaviour of reflexives.

There are still several open questions to be resolved. First, the unavailability of antipassive argument mapping with “affective” verbs is predicted by the argument structure analysis, but is as of yet unexplained in the semantic approach. Second, the use of *-i-* and *-u-* in Ashti should be systematically investigated for all the major classes of predicates identified in Ramchand (2008) and elsewhere. Third, the theory of Ramchand (2008) itself should be more fully adapted for LFG and Glue: I currently use notions like Initiator and Undergoer as primitive labels for semantic roles, whereas Ramchand’s concept of subevents might be more directly incorporated into the semantic component. Finally, the distribution of *-i-* and *-u-* should be compared with corresponding markers in other Dargwa varieties, both for hypothesizing their origin and for achieving a better understanding of their synchronic functions.

References

- Alsina, Alex (1990). “Where’s the mirror principle? Evidence from Chicheŵa.” 13th Generative Linguistics in the Old World Colloquium (St. John’s College, University of Cambridge, Apr. 6–8, 1990).
- Arkadiev, Peter (2009). “Poor (two-term) case systems: Limits of neutralization.” In: *The Oxford handbook of case*. Ed. by Andrej Malchukov and Andrew Spencer. Oxford Handbooks in Linguistics. Oxford: Oxford University Press, pp. 686–699.
- Arkadiev, Peter M. (2006). “Tipologija dvuxpadežnyx sistem.” [Typology of two-term case systems]. PhD thesis. Institute of Slavic Studies RAS.

- Asudeh, Ash and Gianluca Giorgolo (2012). “Flexible composition for optional and derived arguments.” In: *Proceedings of the LFG '12 Conference*. Ed. by Miriam Butt and Tracy Holloway King. Stanford, CA: CSLI Publications, pp. 64–84. URL: <http://users.ox.ac.uk/~cpgl10036/pdf/lfg12asudehgiorgolo.pdf> (visited on 01/23/2014).
- Asudeh, Ash, Gianluca Giorgolo, and Ida Toivonen (2014). “Meaning and valency.” In: *Proceedings of the LFG '14 Conference*. Ed. by Miriam Butt and Tracy Holloway King. Stanford, CA: CSLI Publications, pp. 68–88. URL: <https://web.stanford.edu/group/cslipublications/cslipublications/LFG/19/lfg14.html> (visited on 09/13/2020).
- Belyaev, Oleg (2013). “Optimal agreement at m-structure.” In: *Proceedings of the LFG '13 Conference*. Ed. by Miriam Butt and Tracy Holloway King. Stanford, CA: CSLI Publications, pp. 90–110. URL: <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/18/abstracts/lfg13abs-belyaev.html>.
- Belyaev, Oleg I. (2016). “Funkcija tematičeskix glasnyx v aštynskom: perexodnost', soglasovanie ili kontroliruemost'?” [Thematic vowels in Ashti: Transitivity, agreement or control?] In: *Acta Linguistica Petropolitana* 12.1, pp. 27–39.
- Belyaev, Oleg (2017). “Information structure conditions on the agreement controller in Dargwa.” In: *Proceedings of the LFG'17 Conference*. Ed. by Miriam Butt and Tracy Holloway King. Stanford, CA: CSLI Publications, pp. 66–82. URL: <https://web.stanford.edu/group/cslipublications/cslipublications/LFG/LFG-2017/abstr-belyaev.shtml> (visited on 10/28/2018).
- Bresnan, Joan, Ash Asudeh, Ida Toivonen, and Stephen Wechsler (2016). *Lexical-functional syntax*. Second edition. Blackwell Textbooks in Linguistics 16. Oxford: Wiley Blackwell.
- Bresnan, Joan and Jonni M. Kanerva (1989). “Locative inversion in Chichewa: A case study of factorization in grammar.” In: *Linguistic Inquiry* 20.1, pp. 1–50.
- Bresnan, Joan and Annie Zaenen (1990). “Deep unaccusativity in LFG.” In: *Grammatical relations: A cross-theoretical perspective*. Ed. by Katarzyna Dziwirek, Patrick Farrell, and Errapel Meijas-Bikandi. Stanford, CA: CSLI Publications, pp. 45–57.
- Butt, Miriam, Mary Dalrymple, and Anette Frank (1997). “An architecture for linking theory in LFG.” In: *Proceedings of the LFG '97 Conference*. Ed. by Miriam Butt and Tracy Holloway King. Stanford, CA: CSLI Publications. URL: <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/LFG2-1997/lfg97butt-dalrymple-frank.pdf> (visited on 06/19/2020).
- Dixon, R. M. W. (1979). “Ergativity.” In: *Language* 55, pp. 59–138.
- Dowty, David (1991). “Thematic proto-roles and argument selection.” In: *Language* 67.3, pp. 547–619.
- Falk, Yehuda N. (2006). *Subjects and universal grammar: An explanatory theory*. Cambridge: Cambridge University Press.
- Findlay, Jamie Y. (2016). “Mapping theory without argument structure.” In: *Journal of Language Modelling* 4.2, pp. 292–338.

- Haspelmath, Martin (2019). "Indexing and flagging, and head and dependent marking." In: *Te Reo* 62.1 (Special issue in honour of Frantisek Lichtenberk), pp. 93–115.
- Hout, Angeliek van (2004). "Unaccusativity as telicity checking." In: *The unaccusativity puzzle: Explorations of the syntax–lexicon interface*. Ed. by Artemis Alexiadou, Elena Anagnostopoulou, and Martin Everaert. Oxford Studies in Theoretical Linguistics 5. Oxford: Oxford University Press, pp. 60–83.
- Kibort, Anna (2014). "Mapping out a construction inventory with Lexical Mapping Theory." In: ed. by Miriam Butt and Tracy Holloway King. Stanford, CA: CSLI Publications, pp. 262–282. URL: <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/19/papers/lfg14kibort.pdf> (visited on 06/19/2020).
- Magometov, Aleksandr A. (1963). *Kubačinskij jazyk. (issledovanie i teksty)*. [Kubachi: Grammar and texts]. Tbilisi: Mecniereba.
- Perlmutter, David (1978). "Impersonal passives and the unaccusative hypothesis." In: *Papers from the Annual Meeting of the Berkeley Linguistic Society*. Vol. 4, pp. 157–189.
- Ramchand, Gillian Catriona (2008). *Verb meaning and the lexicon: A first phase syntax*. Cambridge Studies in Linguistics 116. Cambridge University Press.
- Reinhart, Tanya (1983). "Coreference and bound anaphora: A restatement of the anaphora questions." In: *Linguistics and Philosophy* 6.1, pp. 47–88.
- Schätzle, Christin (2018). "Dative subjects: Historical change visualized." PhD thesis. Universität Konstanz.
- Sorace, Antonella (2000). "Gradients in auxiliary selection with intransitive verbs." In: *Language* 76, pp. 59–89.
- Spencer, Andrew (2009). "Case as a morphological phenomenon." In: *The Oxford handbook of case*. Ed. by Andrej Malchukov and Andrew Spencer. Oxford Handbooks in Linguistics. Oxford: Oxford University Press, pp. 185–199.
- Spencer, Andrew and Ryo Otoguro (2005). "Limits to case: A critical survey of the notion." In: *Competition and variation in natural languages: The case for case*. Ed. by Mengistu Amberber and Helen de Hoop. Amsterdam: Elsevier, pp. 119–145.
- Sumbatova, Nina (2011a). "Person hierarchies and the problem of person marker origin in Dargwa: Facts and diachronic problems." In: *Tense, aspect, modality and finiteness in East Caucasian languages*. Ed. by Gilles Authier and Timur Maisak. Diversitas Linguarum 30. Bochum: Universitätsverlag Dr. N. Brockmeyer, pp. 131–160.
- Sumbatova, Nina R. (2011b). "On some grammatical features and the issue of the origin of personal markers in Dargwa." In: *Journal of Language Relationship*, pp. 110–129.
- Sumbatova, Nina R. and Rasul O. Mutalov (2003). *A grammar of Içari Dargwa*. München: LINCOM EUROPA.