



Polyfunctional argument markers in Ket

Implicative structure within the word

Matthew Carter¹ 

Received: 16 February 2023 / Accepted: 24 February 2023 / Published online: 16 March 2023
© The Author(s) 2023

Abstract

Polyfunctionality refers to cases in which the same formal material is systematically reused with different functions. It represents a type of *complexity of exponence*, wherein there is a non-one-to-one mapping between function and form. Much recent work in morphology has emphasized the role of *implicative structure* in resolving the communicative challenges associated with complex form-function mappings. However, previous work has focused almost entirely on complexity of exponence as a challenge for the speaker predicting novel forms (knowing how to encode information), and very little on the challenge for the listener in decoding novel forms. It also has focused almost exclusively on inflectional paradigms, and has not explored how other types of implicative structure might be important in languages with diverse morphological systems. This paper investigates the role of syntagmatic (i.e. word-internal) implicative structure in Ket (Yeniseian, Siberia), a polysynthetic language with numerous polyfunctional markers. It is shown that such markers are organized into networks of implicative relations with one another, wherein less polyfunctional markers along a scale of polyfunctionality disambiguate the function of more polyfunctional markers. This allows uncertainty with regard to their function in any particular instance to remain low. The Ket data make wider typological predictions for the relationship between complexity of exponence and implicative structure.

Keywords Ket · Siberian languages · Polyfunctionality · Implicative structure · Morphological complexity · Syntagmatics · Old polysynthesis

1 Introduction

Ket is the last still spoken of the indigenous Yeniseian languages of Krasnoyarsk Krai in central Siberia (Russia). The Ket verb is often described as polysynthetic (Vajda, 2017b), and exhibits many traits characteristic of what Fortescue (2013) terms

✉ M. Carter
mccarter@ucsd.edu

¹ Department of Linguistics, University of California San Diego, 9500 Gilman Dr., La Jolla, San Diego, CA 92093, USA

an *old polysynthetic language*: the verbal system is the product of what Fortescue calls ‘successive historical layering, with fossilization’, wherein many chronologically distinct waves of new material have been grammaticalized on top of existing material, as older components of the verb have been switched around, fused together, or had their functions reanalyzed (Vajda, 2010, 2013). The end result is a system in which inflectional categories are often expressed discontinuously, redundantly, and with tremendous variability from one verb to another.

One example of this is the way that Ket indexes verbal arguments. In Ket, both subjects and direct objects¹ are indexed on transitive verbs. However, the way that this is done varies significantly from one verb to another. For example, the verbs in (1) and (2) mark their subjects in the same way (*k*), but use different markers, occurring at different positions in the string, for the object (*d* and *ba* respectively).²

Necessary background on Ket verbal morphology, including an explanation of the given segmentations, is provided in §3. For Ket’s usually discontinuous stems, this paper follows one of the practices suggested in the Leipzig Glossing Rules, glossing the first lexical morph in the string with the meaning of the entire lexeme and the remaining lexical morphs with STEM.³

- (1) kusqadda
k-us-q-a-d-da
2SG.SBJ-warm.up-STEM-PRES-1.SG.OBJ-STEM
 ‘you warm me up’ [VZ 133]
- (2) kejbaɣava
k-ej-ba-ɣ-a-v-a
2SG.SBJ-throw-1.SG.OBJ-STEM-N.PST-EM-STEM
 ‘you throw me’ [VZ 139]

This difference between the two verbs holds not only for 1st person singular objects, but for any object, cf. *kusq-aj-da* ‘you warm him up’, *kusq-ij-da* ‘you warm her up’, etc. vs. *kej-a-ɣava* ‘you throw him’, *kej-i-ɣava* ‘you throw her’, and so on.

Now consider another verb, shown in (3). This verb marks its object in the same way as (1) (*d*), but marks its subject twice (*k* and *ku*) redundantly.

¹ Some, but not all, syntactically trivalent verbs index the indirect object instead of the direct object.

² Abbreviations: JOINT = special form of a marker when “jointly indexed” with another marker – when the two mark the same argument; EM = empty morph; KN = form is taken from Kotorova and Nefedov (2015), *A Comprehensive Dictionary of Ket*; PROS=prosecutive case; VZ = form is taken from Vajda and Zinn (2004), *Morfologičeskij slovar’ ketskogo glagola (na osnove južno-ketskogo dialekta)* [Morphological dictionary of the Ket verb (on the basis of the southern Ket dialect)]. All other abbreviations follow the Leipzig Glossing Rules.

³ The relationship between the different stem-components is complicated, varies substantially from one verb to another, and necessarily falls outside the scope of this paper. See Nefedov & Vajda, 2015, pp. 35–36, 50–62 and Vajda, 2014, pp. 514–518 for a description of Ket-specific non-inflectional categories, Drossard, 2002 for an excellent typologically-oriented description of some Ket word-formation processes and Mattissen (2006) for an attempt to situate Ket word-formation within a typology of word-formation in polysynthetic languages.

- (3) kkuḡdit
k=ku-g-d-it
2SBJ=2.SG.SBJ-carry-1.SG.OBJ-STEM
 ‘you carry me’ [VZ 194]

Again, this double subject marking is maintained throughout the paradigm, cf. *da=bu-gdit* ‘she carries me’, *dbo-gbit* ‘I carry it’, and so on.

This type of variability between verbs is characteristic of the system as a whole. In Ket, different verbs mark their subjects and objects using different markers in different combinations. As previous Ketological literature has demonstrated, it is usually not possible to predict which markers a verb will use to mark its arguments (subjects for intransitive verbs, subjects and direct objects for transitive) based on any syntactic, semantic, or phonological properties of the verb (Vajda, 2015). In other words, despite the best efforts of some Ket scholars in the past to prove otherwise (Reshetnikov & Starostin, 1995; Butorin, 1995; Belimov, 1991), the complexities of Ket argument marking are purely morphological, in the sense of Aronoff (1994), with the verbal lexicon being divided into a fairly complex system of inflection classes.

Building on this established analysis from the Ketological literature, this paper adds the following observation: within Ket’s inflectional class system, the same material is very frequently reused with different functions across different classes. The same marker that indexes the subject in one verb can index the object in another, and vice-versa.⁴ Consider the marker *ba* (1.SG), which marks the object in example (2). It can also mark the subject alone:

- (4) avəriḡbayara
 avəriḡ-**ba**-y-a-ra
 spend.the.day-**1SG.SBJ-STEM-NPST-STEM**
 ‘I spend the day’ [KN 435]

Cf. avəriḡ-**aḡ**-ara ‘they spend the day’.⁵

Or it can be a co-exponent⁶ of the subject with another marker:

- (5) daḡabatsaq
d-aya-ba-t-s-aq
1.SBJ-make.a.trip.to.forest.and.return-1SG.SBJ-STEM-NPST-STEM
 ‘I make a quick trip to the forest and return’ [KN 438]

Cf. **k-aya-yu**-tsaq ‘you make a quick trip to the forest and return’ [VZ 116–177].

⁴This paper follows recent work in the Ketological tradition (Vajda, 2015; Nefedov, 2015; Nefedov & Vajda, 2015) in their usage of the terms ‘subject’ and ‘object’. It also assumes, following such work, that Ket is a nominative/accusative language. At the same time, the terms *argument-marking* and *indexation* are used throughout to avoid making claim as to whether the markers in question represent agreement markers, as opposed to bound pronouns. Such a discussion would necessarily fall outside the scope of this paper.

⁵Velar obstruents are sometimes lost after [ḡ], which is the case with [y] here.

⁶The term *co-exponent* is used throughout to refer to one of multiple markers which encode overlapping information, in other words share a feature. *X* is a co-exponent of *Y*, and vice-versa, if *X* redundantly encodes overlapping information with *Y*. These are cases of Overlapping Exponence, in the sense of Matthews (1972), if one considers the relationship between markers, and of Multiple Exponence, in the sense of Harris (2017), if one considers the information (features) which are being redundantly expressed.

In some verbs, the same material used to index subjects and objects can even serve completely unrelated functions. For example, *o* can mark 3M arguments in the past tense.⁷

- (6) endonsuk
 en-d-**o**-n-suk
 forget-STEM-3SG.M.PST-PST-STEM
 ‘he forgot’ [KN 626]

cf. end-**iru**-nsuk ‘she forgot’, en-**ba**-nsuk ‘I forgot’ [KN 489].

However, it can also mark just past tense, without having any argument-marking function. In the following example, the subject is marked by *ba*.⁸

- (7) avəriŋbaɣolda
 avəriŋ-**ba**-ɣ-**o**-l-da
 spend.the.day-1SG.SBJ-STEM-PST-PST-STEM
 ‘I spent the day’ [KN 435]

This phenomenon, where the same formal material is used systemically to encode different functions is called *polyfunctionality* (Stump, 2014, 2015).⁹ This paper understands this term broadly, referring to any instance of a many-to-one mapping between function and form. This includes both the use of the same marker to encode different function across different lexemes (as is the case in 4-7 above), and the use the same marker to encode different functions across different forms of the same lexeme (in other words, syncretism). This follows the definition given in Stump (2014, p. 73): “In the domain of inflectional morphology, polyfunctionality is the use of the same morphology in the expression of distinct morphosyntactic property sets. Inflectional polyfunctionality is observable both within and across paradigms and even within individual word forms”.

Polyfunctionality represents a type of *morphological complexity* (Baerman et al., 2017), and more specifically a phenomenon known as *complexity of exponence* (Anderson, 2015), in which the relationship between grammatical information (e.g. tense, person) and the formal units which are used to encode it (e.g. affixes, stem changes, tones) is non-isomorphic (not one-to-one) or otherwise opaque.

⁷Ket has three genders (noun classes): masculine, feminine, and neuter (=inanimate), with masculine and feminine patterning together as animate in plural forms, and the morphological encoding of feminine and neuter often overlapping. Class membership is mostly semantic, but the masculine and feminine genders also include a number of lexical groupings that are somewhat arbitrary. For example, large non-human animals like bears and moose, as well as trees, are masculine, while smaller animals like foxes and squirrels are feminine, as is the word for fire. For more, see Vajda (2004, p. 18).

⁸Despite their identical function and linear adjacency in this form, *o* and *l* are clearly separate markers. Another marker can come between them and they can occur independently. See fn. 24 for examples.

⁹Ackerman and Bonami (2017) investigate systemically polyfunctional argument markers in another Siberian language, Nenets (Samoyedic, Uralic), which is areally related to Ket, but genetically unrelated. Crucially, in Nenets, the function of the marker set in question is predictable from the word class (noun, preposition, etc.) of the construction in which it occurs, whereas in Ket argument markers are polyfunctional within a single subsystem of the grammar, finite verbs. This difference is what motivates the question in Ket as to how the function of such markers can be determined by the listener.

Recent years have seen an increased interest in morphological complexity (see Arkadiev & Gardani, 2020 for a recent overview), including *complexity of exponence*. Much of this work has focused on the communicative challenges which morphological complexity would seem logically to create. Roughly, if the same information can be encoded in different ways in different contexts, how does the language user determine how that information is or should be encoded in any particular form? The type of apparent communicative challenge presented by a system like Ket is well formulated as the INFLECTED WORD RECOGNITION PROBLEM (IWRP) in Bonami and Beniamine (2021):

- (8) Inflected Word Recognition Problem (Bonami & Beniamine, 2021, p. 82)
What allows speakers to draw inferences from a word's form to its content?

In another words, as applied specifically to Ket:

In a language like Ket, in which the same formal material is frequently associated with different functions in different wordforms, how might a listener encountering a novel form understand which of these possible functions was intended in the given form?

It is assumed that many sources of information could aid the listener in this task – syntactic context, discourse context, the semantics of a given verb – however, this paper investigates only one, a property of morphological systems known as *implicative structure* (Wurzel, 1984). This can be understood as interdependencies between elements within a morphological system.

Implicative structure has been a major theme in recent work on morphological complexity, with much of it focusing *paradigmatic implicative structure*, or interdependencies between whole words within an inflectional paradigm. Building on the established model of Ket's unique argument-marking and tense-marking morphology, (Nefedov & Vajda, 2015; Vajda, 2014, 2015; Georg, 2007; Nefedov, 2015, *inter alia*), using data drawn from both published sources¹⁰ and original fieldwork,¹¹ this paper departs from previous work on the role of implicative structure in morphology, focusing instead on *syntagmatic implicative structure* – interdependencies between subword pieces. It demonstrates that uncertainty concerning the function of a polyfunctional marker in a given verbform can be greatly reduced through implicative relations which hold between that marker and other markers in the same wordform.

¹⁰Kotorova and Nefedov (2015), *A Comprehensive Dictionary of Ket*, is a massive two-volume dictionary, compiled by the Department of Siberian Indigenous Languages at Tomsk State Pedagogical University, from 2000 and 2015, with support from the Department of Linguistics at the Max Planck Institute for Evolutionary Anthropology. It brings together materials from previous dictionaries along with the Departments archival materials. For each entry, the dictionary includes translations into Russian, English and German, information about inflection class membership and segmentation, and example sentences from texts with translations into Russian. Vajda and Zinn (2004), *Morfologičeskij slovar' ketskogo glagola (na osnove južno-ketskogo dialekta)* [Morphological dictionary of the Ket verb (on the basis of the southern Ket dialect)], is a smaller paradigm dictionary, which gives exemplary partial paradigms, with translations into Russian and English and information about segmentation and inflection class membership for several hundred verbs.

¹¹Unless otherwise indicated, all data are from Valentina Andreevna Romanenkova, a native speaker of Ket and the author's teacher and collaborator.

The Ket case makes broader typological predictions about the hypothesized limits of complexity of exponence and its relationship with implicative structure.

The structure of this paper is as follows. Section 2 lays out the background on morphological complexity. Section 3 provides necessary background on relevant aspects of the Ket verbal system. Sections 4 and 5 provides an overview of polyfunctionality in the Ket argument-marking system. Section 5 lays out examples of implicative structure between markers. Section 6 discusses the relationship between the Ket system and some other types of highly complex exponence discussed in the literature, namely Gestalt Exponence (Blevins, 2016, *inter alia*), and Distributed Exponence (Carroll, 2022). Section 7 provides further discussion and concludes.

2 Complexity of exponence and implicative structure

Before discussing polyfunctionality in Ket argument marking (in §4), and how syntagmatic structure could help to resolve the challenges associated with it (§5), it is necessary to start with some background, which this section aims to provide. It is divided into three subsections.

Section 2.1 expands upon the central notion of *complexity of exponence*, the term used in this paper to refer to non-isomorphic form-meaning mappings. It clarifies what this means, and discusses what phenomena are understood to contribute to, as well as *not* contribute to, *complexity of exponence* for a particular language.

Section 2.2 discusses several inter-related issues. It discusses further the INFLECTED WORD RECOGNITION PROBLEM as well as the related PARADIGM CELL FILLING PROBLEM, used to illustrate the way in which implicative structure can be used to resolve the communicative challenges associated with complex form-meaning mappings. This section also introduces the information theoretic notions of *entropy* and *conditional entropy*, which provide a more precise way of talking about implicative structure and the role that it plays in a particular morphological system.

Finally, §2.3 suggests that the role of implicative structure in morphology is much broader than previous work has explored, namely that it has not explored the relationship between implicative structure and other types of *complexity of exponence* which do not directly bear on the PCFP, such as polyfunctionality. It has also not explored how languages with a large number of polyfunctional markers might use implicative structure differently. An exploration of the role played by syntagmatic implicative structure in resolving polyfunctional form-meaning mappings in Ket is presented as a contribution to the framework.

2.1 Complexity of exponence

Nearly any discussion of inflectional morphology from a cross-linguistic perspective will note that it is a point of massive cross-linguistic variation. Some languages are conventionally understood to lack inflectional morphology entirely (Vietnamese, Yoruba), while others have massive inflectional paradigms (Archi). Some languages have essentially no inflection classes (Turkish), while others have hundreds (Chinantepec verbs). Such observations have prompted a sizeable literature on how to conceptualize and measure the complexity of morphological systems. For overviews, see the introductory chapters of Baerman et al. (2015) and Arkadiev and Gardani (2020).

Table 1 Examples of complexity of exponence

Allomorphy	The same information is encoded differently across different lexemes or different cells of a paradigm
Cumulative Exponence (Fusion)	A unit of form encodes several features simultaneously
Multiple Exponence (Harris, 2017)	A unit of information is encoded multiple times in the same wordform
Polyfunctionality (Stump, 2014, 2015)	The same unit of form encodes different features in different contexts

This paper focuses on one particular dimension of morphological complexity, which Anderson (2015) refers to as *complexity of exponence*. Essentially the same notion, with subtle differences, has also been referred to as *non-canonicity* of exponence (Baerman et al., 2017), *non-linearity* (Dahl, 2004, 2017), and *opacity* (Hengeveld, 2011; Trudgill, 2020). Complexity of exponence refers to those cases where there is a non-isomorphic relationship between units of meaning and units of form across the lexicon. In other words, there is not a one-to-one mapping between a unit of information (grammatical or lexical) and the formal material that is used to encode that information (i.e. affixes, stem changes, tones). Examples of non-isomorphic relationships (outlined in more detail in Table 1) include phenomena like allomorphy, cumulative exponence (i.e. fusion), Multiple Exponence (Harris, 2017), also known as Extended Exponence (Matthews, 1972) and, as has already been noted, polyfunctionality (Stump, 2015, 2014). The more a morphological system instantiates any of the above, the higher its complexity of exponence. Examples of complexity of exponence are given in Table 1.

As noted in §1, this paper understands the term *polyfunctionality* broadly (in concurrence with Stump, 2014), encompassing any type of many-to-one mapping between meaning and form. Polyfunctionality of markers can arise through the reuse of material across different cells of the same paradigm (e.g. *-en* in German verbs marks the infinitive, 1PL.PRS and 3PL.PRS, a case of classical syncretism¹²), across different inflection classes (e.g. *-u* in Russian marks accusative singular with class I nouns, but dative singular with class II nouns), and across different morphosyntactic categories (e.g. person/number suffixes in Nenets mark possessors on nouns and direct objects on verbs).

Complexity of exponence is not directly related to the number of distinct categories which a morphological system distinguishes (i.e. the number of cells in a paradigm), or to the amount of information that can be expressed in a single wordform (i.e. the degree of synthesis, cf. Bickel & Nichols, 2007). Some extremely complex systems involve very small paradigms and only a few formatives per wordform. For example, nominal inflection in Nuer (Baerman, 2012) involves only 3 distinct forms per lexeme and a small number of formatives per wordform (a stem, which may undergo alternations, plus a suffix), but the system exhibits a massive amount

¹²Syncretism is understood in this paper to be a property of paradigms, wherein material is shared between multiple cells of a paradigm, while polyfunctionality is a property of markers, wherein the same formal material encodes different functions in different contexts. Syncretism in a paradigm creates polyfunctional markers. Thank you to a reviewer for helping to clarify this point.

of allomorphy and unpredictable syncretism. Conversely, systems with very large paradigms and a very high degree of synthesis may be very simple if the relationship between formatives and the functions which they encode is very transparent and consistent across the lexicon. This would be the case with, for example, Turkic languages.¹³

When there is a non-isomorphic relationship between form and function across different lexemes, then this creates *inflection classes*, groupings of lexemes which encode information in the same way, different from other groups of lexemes (as a phenomenon, this is sometimes called *flexivity*, Bickel & Nichols, 2007).

Inflection classes have been the primary focus of much of the literature on morphological complexity, because they makes things more difficult for the language user. For a speaker of a language with inflection classes, it is not enough to know what information they want to encode (i.e. what case or what tense to use); they must also know how to encode that information for a particular lexeme. For the listener, inflection classes make things more difficult when the formal material used to encode a particular feature in one class is used to encode different information in another class (i.e. when allomorphy creates polyfunctionality). A quite sizeable literature has developed around the communicative task posed to speakers (for good reason, as the problem is much bigger than it might seem at first glance). The next sections aims to introduce the reader to this literature.

2.2 The paradigm cell filling problem and conditional entropy

As noted in the last section, the presence of inflection classes in a language forces a speaker of that language to know, for every inflected lexeme, not only what information to encode, but also how to encode it.

The speaker must be able to do this without any guarantee of having heard the form that they want to use before. This is because linguistic input follows a Zipfian distribution, where a small number of forms are very common, while all other forms may be vanishingly rare. Studies have shown that even increasingly large corpora may never contain all inflected forms for morphologically complex languages (Baayen, 2002; Blevins et al., 2017; Sims & Parker, 2016) and have suggested that the need to predict unknown forms continues throughout the lifespan (Bonami & Beniamine, 2016; Sims & Parker, 2016).

Work beginning with Ackerman et al. (2009) terms the puzzle of how speakers are able to predict all forms of all lexemes based on at best some subset of them the PARADIGM CELL FILLING PROBLEM, and as the solution it implicates the property of morphological systems known as *implicative structure* (Wurzel, 1984). A language

¹³ A reviewer asks specifically about Inuit languages. Acknowledging his very superficial knowledge about Inuit languages, the author understands them to represent a good illustration of what complexity of exponence is *not*. Complexity of exponence is only concerned with how much the relationship between units of meaning and units of form deviates from one-to-one. A language that can encode complex multi-clausal structures morphologically using long strings of affixes is certainly morphologically complex in some other sense, but it still shows very low complexity of exponence if there is a one-to-one relationship between those affixes and the information they encode: if each unit of information is encoded with only one affix, and that affix always encodes that same information in every lexeme in which it occurs.

Table 2 Russian 1st, 2nd and 3rd Nominal Inflection Classes (singular forms, without stress)

	Class I	Class II	Class III
Nominative	sobak-a	gorod-∅	krov ^j -∅
Genitive	sobak ^j -i	gorod-a	krov ^j -i
Dative	sobak ^j -e	gorod-u	krov ^j -i
Accusative	sobak-u	gorod-∅	krov ^j -∅
Instrumental	sobak- <i>oj</i>	gorod- <i>om</i>	krov ^j - <i>ju</i>
Prepositional	sobak ^j -e	gorod ^j -e	krov ^j -i
Gloss	DOG	CITY	BLOOD

exhibits implicative structure if known forms provide information which can be used to predict unknown forms.

Implicative structure is a property, at minimum, of *inflectional paradigms*. An inflectional paradigm refers to all inflected forms of the same lexeme taken collectively.¹⁴ As an example of how these exhibit implicative structure, consider a highly simplified version of Russian nominal inflection in the singular (excluding stress), given in Table 2.

Russian nouns provide an example of implicative structure at work. Case and number are encoded cumulatively using inflectional suffixes. However, the same value can be encoded with one of several different suffixes (different *allomorphs*) depending on the noun: dative singular with *-e*, *-u* or *-i*, instrumental singular with *-oj*, *-om*, or *-ju*, and so on.

The task of predicting which allomorph will go with which noun however is made much simpler by the fact that they co-vary to a great extent with one another. If a user of Russian knows e.g. that a given noun has an accusative singular in *-u*, they know that (almost always) the dative singular suffix will have *-e* and not *-u*, and the instrumental will have *-oj* and not *-om*, and so on.

The more forms that one knows, the easier this task becomes. For example, the suffix *-i* for the genitive singular is shared between declensions I and III, making it poorly predictive of the other forms. However, if the speaker knows any other form in addition to the genitive, then this is enough to predict all other forms.¹⁵

The networks of implicative relations that are learned based on frequent lexemes can then be analogized to produce new forms, allowing the speaker to accurately predict all forms of even very rare lexemes. For example, suppose that a Russian-user has never encountered the form *argišu* ‘a seasonal migration by reindeer caravan among Siberian peoples (dative singular)’. If they know that this is the dative singular from

¹⁴Since the division between inflection and derivation, between a word and a phrase, and often between morphology and syntax, are fuzzy concepts (especially in cross-linguistic work), inflectional paradigms are necessarily fuzzy as well. In other words, inflectional paradigms should not be seen as exhaustive sets; they are simply a group of forms which exist in a relationship of opposition. The point is the core relationship, not the boundaries of the set. Many of the same notions used in inflectional paradigms can be applied to derivational paradigms (Bonami & Strnadová, 2019) and to families of syntactic constructions as well.

¹⁵Compare the classical notion of *principal parts*, which is used to the same effect in works like Finkel and Stump (2009).

the syntactic context, then they can be reasonably certain that the genitive singular will be *argiṣa*, without having ever encountered it before.¹⁶

Although most previous work has focused on the communicative challenge associated with producing novel forms (with encoding), some recent work has shown a shift towards the task of comprehending novel forms (with decoding) (Bonami & Beniamine, 2021), and with its relationship with complexity of exponence (Carroll, 2022). This is the so-called INFLECTED WORD RECOGNITION PROBLEM, cited in §1. This paper represents an extension of this line of research.

The next section discusses in more detail the implications of information theoretic approaches to morphology for complexity of exponence specifically.

2.3 Implicative structure, decoding, and polyfunctionality

Information theoretic approaches to morphology, like those outlined above, have underscored how implicative structure can be used in a language to resolve complex form–function mappings.

At the same time, the full ramifications of this model for different kinds of morphological complexity has not been sufficiently explored. As work in this area largely originates from attempts to solve the PARADIGM CELL FILLING PROBLEM, it has largely focused on certain types of complexity of exponence which are directly relevant to the PCFP, while not exploring the implications of its findings for other types of complexity.¹⁷ Some kinds of complexity of exponence, like affix, stem, or tonal allomorphy are directly related to the PCFP. If the same information can be encoded in many different ways across different lexemes, this creates a challenge for the speaker in solving the PCFP. These phenomena, as such, have been a major point of study.

Other kinds of complexity of exponence, like polyfunctionality, do not directly relate to the PCFP. If the same formal material can encode different functions, this is not a major problem in solving the PCFP, since the speaker presumably knows which meaning they intended.¹⁸ It is a problem however for the listener, who must

¹⁶It is important to stress that questions about uncertainty and predictability matter regardless of what theory of morphology one assumes, e.g. whether inflected forms are stored as whole chunks, or minimal form–meaning parings (classical morphemes), or abstract bundles of syntactic features, and whether they are retrieved whole or built up incrementally. In other words, it is irrelevant to the present discussion (and an empirical question) whether the analogical processes described here are accomplished through direct storage of (probably partial) exemplary paradigms, or through decomposition of exemplary forms into smaller units with diacritics for how to recombine them, e.g. STEM+*a* [NOM.S.CLASS1]. This paper remains agnostic on this point.

¹⁷Ackerman et al. (2009) and following primarily discuss morphological complexity in terms of an opposition between what they term *enumerative complexity* (E-complexity) and *integrative complexity* (I-complexity). The former refers to the number of components in the system, i.e. the number of paradigm cells and inflection classes, while the latter refers to the predictability of forms in the system based on other forms (the conditional entropy). E-complexity can be high, but I-complexity must be low. Some kinds of E-complexity are also complexity of exponence, and vice-versa, but the two can also be orthogonal. The size of the paradigm is irrelevant to complexity of exponence, while syncretism and polyfunctionality are irrelevant to E-complexity.

¹⁸Note that allomorphy can cause polyfunctionality, in cases where different inflectional classes use the same exponent, as in the case with *-u* in Russian above. However, the fact that it creates polyfunctionality does not make such a case of allomorphy any more of a challenge for the speaker than other kinds of allomorphy.

determine which meaning was intended, which, for a language with very high exponence complexity, may be a non-trivial task. As such, most previous literature has focused primarily on communicative challenges in encoding information (which exponent is needed for this lexeme?), not in decoding information (which of the things this exponent can mean does it mean in this form?).

Owing to its origins, previous work has also focused almost exclusively on implicative structure within paradigms, and has not focused on how other types of implicative structure might be important for resolving complex form-meaning mappings in languages with diverse morphological structures.

The remainder of this paper sets out to begin filling these gaps. It makes two proposals. The first is that high predictability (low conditional entropy) is a necessary property of morphological systems when mapping from form to meaning, not only from meaning to form. In other words, uncertainty must be low for both the speaker and the listener, and in some languages implicative structure plays a crucial role in achieving this.

The second is that the importance of implicative structure in decoding is a point of cross-linguistic variation. This builds on work by Sims and Parker (2016), who demonstrate that the amount of ‘work’ done by implicative structure in predicting unknown forms (formally, the difference between entropy and conditional entropy) varies across languages. However, the present paper also goes further, proposing that the role of implicative structure across different languages varies not only quantitatively, in the amount of work done by implicative structure, but also qualitatively, in the kind of implicative structure that systems exhibit. More specifically, it is argued that for some languages, as illustrated with Ket, implicative structure is found not only within networks of related words, but also between individual morphs in wordforms. This type of structure is hypothesized to be an adaptation characteristic of languages with very high complexity of exponence, especially those which are strongly head-marking, equivalent to the notion of an *old polysynthetic language*, proposed by Fortescue (2013).

To illustrate an example of both points, consider again the Russian example of a rare form *argišu* ‘a seasonal migration by reindeer caravan among Siberia peoples’ (dative singular). The suffix *-u* in this word could hypothetically encode either the dative (II declension) or the accusative (I declension). However, encountering another form, like the nominative singular *argiš* would disambiguate the function of *-u* in the first case (since if it encoded the accusative, the nominative should be encoded with *-a*).¹⁹

In actual practice though, the need to rely on implicative structure for this purpose in a language like Russian is likely to be quite low; Russian is strongly dependent marking and has robust case/number/gender agreement on nominal modifiers, and hence the syntactic context of a given form provides ample opportunity for disambiguating its function. A quick websearch is enough to confirm this, as in the actual title for an article about an event dedicated to the memory of the Nenets writer Leonid Lapsuy, given in example (9).²⁰

¹⁹Hypothetically, *-u* could be part of the stem and the noun could be indeclinable, in which case knowledge of another form like the prepositional *argiše* could exclude that possibility.

²⁰<https://depcul.yanao.ru/presscenter/news/33235/>, accessed online February 8th, 2021.

- (9) жизнь человека, подобна аргишу
ʒɪʒnʲ čelovʲek-a, podobna arɣiʂ-u
 life[FEM.NOM.S] person-GEN.S similar[FEM.NOM.S]
 seasonal.migration-DAT.S

‘The life of a person is like an *argish*’

Here *arɣiʂu* is clearly the complement of the predicative adjective *podobna* ‘similar (to)’, which invariably takes a complement in the dative, and not the accusative.²¹ Implicative structure could help disambiguate the form, but the need to rely on it for this function is low.

This is not the case for all languages though. By contrast, consider the sentence *ditanbuyolʲ betin* ‘they migrated seasonally, made the *argish*’ in the northern dialect of Ket.²² This is a perfectly grammatical sentence by itself, with no syntactic information available outside the verbform.

Like many Ket verbs, the way that the subject is marked for this verb is fairly complex (a full discussion of the Ket verbal system is given in the next section). It exhibits Multiple Exponence, marking the subject (and the past tense) twice. Three of the markers associated with marking subjects, *d*, *bu*, and *o*, are all, on their own, partially ambiguous with regard to their function.

However, the function of each of these markers can be disambiguated by the presence of other markers in the verbform. The combination of the prefix *d* (person) and the suffix *in* (pl.number) indicates that the subject of the verb is either 1.PL or 3.PL but leaves it ambiguous as to which. The marker *bu* redundantly encodes a 3rd person subject, ambiguous as to whether it is masculine, feminine, or plural. However, *bu* is a special form that indicates it is a co-exponent with a subject marker *d*-. The presence of *bu* therefore disambiguates the person encoded by *d*. The presence of *in* in turn makes it clear that *bu* encodes a 3PL subject, and not 3M or 3F. Finally, *o* always encodes past tense, but in some verbs additionally encodes a 3rd person masculine singular argument (which can be either the subject or object). The fact that the subject-marking function is absent here is not discernible by considering the marker itself in isolation, but can be discerned by looking at other markers. Among other things (discussed in §5), its number value here is incompatible with *in*.

To make things clearer, we can represent implicative structure between morphs using a dependency graph, shown in Fig. 1, with the arrows representing information that the markers at the beginning of the arrow provide about the function associated with the marker at the end of the arrow. Note that, although in this case morphs later in the string provide information about those earlier in the string, in Ket, these dependencies can go in any direction; morphs earlier in the string can also provide information about the function of morphs later in the string as well, as will be seen later on.

²¹As a reviewer rightly points out, this example illustrates government and not agreement, but is nonetheless an illustration of the overall dependent-marking structure of Russian.

²²In the central and southern dialects of Ket, this verb belongs to a different class with a simpler argument-marking pattern (class v1, see next section), and so would not illustrate the present point.

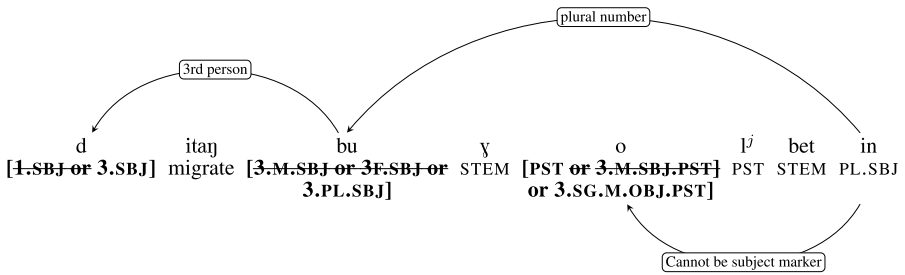


Fig. 1 Syntagmatic implicative structure in the Ket verb

Putting all of this together, the structure of nearly every component in the verb is unambiguous, and the meaning of all components in the verb is less ambiguous,²³ as shown in Fig. 1.²⁴

- (10) ditaŋbuɣolʲbetin
 d-itaŋ-bu-ɣ-o-lʲ-bet-in
 3.SBJ-migrate-3.PL.SBJ(JOINT)-STEM-PST-PST-STEM-PL.SBJ
 ‘They migrated’ [KN 536]

As this example illustrates, in Ket verbs, we see a similar dynamic playing out at the syntagmatic level as has been described for the structure of paradigms in other

²³This verb is intransitive, and *o*₄ is only a tense marker. However, there is technically nothing in the syntagmatic structure alone which excludes a transitive reading with a masculine object, perhaps ‘they took him on *argish*’, especially since many verbs with this particular subject-marking pattern are transitive verbs of motion. The stem, literally ‘migration-do’ or ‘migration-VBLZ’, also has no information about transitivity.

²⁴Note that despite their identical function and linear agency in this form *o* and *l* (*il*, when required by the phonotactics; sometimes also *lʲ*, *ilʲ*) are clearly separate markers, based on comparison with other forms. For one, the 3N argument marker or empty morph (see §4.2) *b* ([v] between sonorants) can occur between them, cf.

- (i) *kətəŋaɣovla*
 k-ətəŋ-a-ɣ-o-v-l-a
 2.SBJ-cover-3.SG.M.OBJ-STEM-PST-EM-PST-STEM
 ‘you covered him’.

Second, they can occur separately, cf.

- (ii) *sildobqon*
 sil-d-o-b-qon
 come.of.summer-STEM-PST-3.SBJ-STEM
 ‘summer came’
- (iii) *dabulgut*
 da=bu-l-gu-t
 3.F.SBJ-3.SBJ.JOINT-PST-2.SG.OBJ-carry
 ‘she carried you.’

languages. Even though many individual morphs are *polyfunctional*, any potential ambiguity is greatly reduced by implicative structure which holds between morphs.

Of course, implicative structure between morphs does not need to do all the work. Presence of a full NP (e.g. $de^2\eta$ ‘people’) or an overt pronoun (e.g. $b\bar{u}\eta$ ‘they’), or knowledge of related forms, or simply discourse context, could all accomplish some of the decoding work done by implicative structure in this example. Similarly, the fact that implicative structure exists in the syntagmatic structure is not direct evidence that it is used by Ket listeners.

Nevertheless, the fact that implicative structure allows for such a striking difference between the language’s very high complexity of exponence and actual ambiguity is highly suggestive of the idea that such structure is a communicative adaptation. This is especially the case for a head-marking language with prodrop,²⁵ like Ket, in which examples like the above, with just a verbform, may constitute a substantial portion of the input.

The remainder of this paper strives to develop of a more complete understanding of the role of implicative structure, as a formal organizational property of the system, in Ket argument marking. The next section lays out essential background information on the overall structure of the Ket verb.

3 Background on Ket verbal morphology

As noted in §1, the Ket verb is often described as polysynthetic (Vajda, 2017b), and exhibits many traits characteristic of an *old polysynthetic language* (Fortescue, 2013). Waves of chronologically distinct material have been grammaticalized on top of existing material, as older components of the verb have been metathesized, fused together, or reanalyzed (Vajda, 2010, 2017a), producing a seemingly random interdigitation of inflectional, derivational, and lexical material.

The resulting system is often described as *templatic*, meaning simply that the order of morphs in the verb reflects the grammaticalization history and is not derivable from semantic scope or abstract syntactic structure in any obvious way (see Mithun, 2011 for a defense of this assessment and parallel developments in Navajo²⁶). Specific templatic models have been used to describe the Ket verb since the 1990s (Reshetnikov & Starostin, 1995; Butorin, 1995; Werner, 1997; Vajda, 2001, 2003, 2004). Such models, similar to those used for the Na-Dene languages, divide the verb into a fixed number of morphological ‘slots’, or *position classes* (*P*), which different markers can be said to occupy. Table 3 gives a simplified version of the current predominant templatic model of the Ket verb, adapted from Nefedov and Vajda (2015).²⁷

²⁵Nefedov (2015, p. 56) “In general, core noun phrases can be freely omitted in the discourse as the presence of the cross-referencing markers makes it possible to easily recover these arguments. Therefore, any verbal predicate in the above examples can constitute a fully grammatical sentence on its own.”

²⁶For criticisms of the notion of a template in Navajo, see McDonough (2000) and Beniamine et al. (2017). The specifics of the criticisms of templatic models differ between the two analyses.

²⁷The model in Table 3 is based on the model worked out by Edward Vajda in Vajda (2001, 2003, 2004) and revised in subsequent work (Vajda, 2010, 2013, 2015; Nefedov & Vajda, 2015; Vajda, 2017a). It is

Table 3 Templatic model of the Ket verb (simplified from Nefedov & Vajda, 2015, p. 36)

P8	P7	P6	P5	P4	P3	P2	P1	P0	P-1
subject class/person	stem	subject or object features	stem	subject or object features or tense	subject or object features or empty morph	tense/mood	subject or object features or resultative	stem	plural subject (in some verbs that use P8 for subject)

As the templatic model shows, the Ket verb goes against cross-linguistic tendencies in affix ordering. The stem, whose internal structure is not considered in this paper, consists of up to three discontinuous pieces (P7-P5-P0), which are interdigitated with markers that indicate the subject and object, tense, or some limited valence-changing operations.²⁸ The templatic model also gives some sense of the complexity of the argument-marking system. Argument markers are copious, all but one of the sets of argument markers can mark either the subject or the object, and many of the argument markers in the area P4-P1 represent a fusion or reanalysis of historically distinct markers, and alternatively or simultaneously encode completely orthogonal functions (tense, valence, or seemingly nothing at all).

Templatic models like the above should not be seen as explanatory mechanisms or as something cognitively real. Rather, as (Crippen, 2012, p. 43) puts it for Tlingit, a templatic model should be understood as “merely a descriptive tool that aids in understanding the positions and interrelationships of different morphological elements within the verb.” Elements assigned to the same position class are those which are in paradigmatic opposition (complementary distribution) and share the same morphotactic behavior (occur in the same position in the string), even if they have contradictory or orthogonal functions (cf. discussion in Vajda, 2001, pp. 371–372). In this sense, a position class is equivalent to Gurevich (2006, p. 8)’s notion of a distribution/form class.

To underscore this point, this paper deviates from previous Ketological work in that it avoids the metaphor of a position class as a ‘slot’, which markers ‘occupy’ or ‘fill’. Instead, it takes the position classes as labels for sets of markers. For example,

used also in Kotorova and Nefedov (2006), Georg (2007), Nefedov (2015), Nefedov and Vajda (2015), Kotorova and Nefedov (2015).

The other model that still finds occasional usage is that proposed by Reshetnikov and Starostin (1995) and Butorin (1995). A recent version of it is presented in Butorin (2018), a simplified version of which is given below (translation mine):

- (i) [Agent]-[Incorporant]-[Lexical]-[Dative]-[Lexical]-[Tense]-[Patient]-[Lexical]-[Agent.PL]

The main differences between the two models is that the Reshetnikov & Starostin/Butorin model assigns specific (questionable) semantic roles to different sets of argument markers and glosses over many of the complexities at the right edge of the verb (P4-P1 in Vajda’s model).

²⁸Some transitive verbs have a resultative form, derived with the P1 marker a¹ (homophonous with one of the many argument markers). These forms are not discussed in this paper. Aside from this, derivational processes in Ket are often quite opaque, in that the line between derivational affixes and roots is fuzzy at best. This topic necessarily falls outside the scope of this paper.

rather than saying that “*ba* occupies P6”, it will refer to the “P6 set of markers” (= {*ba*, *ku*, *a*, *i...*}), which are in complementary distribution and share the same morphotactic behavior. This removes an unnecessary level of abstraction.²⁹

Something that is not immediately apparent from the templatic class model is the tremendous variability across Ket verbs. Of the three possible stem-components, only the one represented in Vajda’s model by P0 is present in all verbs. In the tense system, one of two different tense markers represented by P4 ($s \sim \emptyset$ and $a \sim o$, where NPST \sim PST) may or may not be present. Finally, the argument-marking system is the most variable and complex, and the remainder of this section turns to describing it in more detail.

3.1 Argument markers

The main sources of morphological complexity in the Ket verb is the way that the system marks verbal arguments. Ket indexes up to two arguments verb-internally, usually a subject and a direct object (although some trivalent verbs require agreement with the indirect object over the direct object).

Verbs make use of essentially three distinct sets of markers. This paper refers to these sets using the names of the corresponding position classes in Vajda’s templatic model. These are the P8 set, the P6 set, and the P4/P3/P1 set respectively. In older literature (e.g. Krejnovich, 1968), the equivalent labels are the *di/du*, *ba/a* and *di/a* sets (based on the 1S and 3M forms respectively).³⁰

These sets of markers are distinct both in their segmental content (with some overlap across sets) and in their ordering relative to other morphs, when such morphs are present. This section will examine each of the sets in turn. A summary of all argument markers is given in Table 4.³¹

The P8 set of markers, shown in Table 4, occur at the leftmost edge of the verb. It only indexes the person and (for the 3rd person) the gender of the argument, whereas the other sets mark person/gender and number cumulatively. Many verbs use the suffix *-in* (P-1 in Vajda’s model), in conjunction with the P8 set to indicate a plural subject.

In examples throughout, indices below a morph represent the position class to which it belongs.³²

²⁹A reviewer stresses the need to clarify this point, and asks if this approach is primarily for the purpose of avoiding “empty” slots. Yes, but not only. This paper takes where a marker occurs in the string to be a property of that marker, and hence homophonous forms with different morphotactic behavior are different markers. We can describe these directly as properties of the marker-sets, which in turn are parts of bigger constructions, without needing to rely on abstract slots that markers are thought of as being inserted into. It is more illuminating to focus on which markers are in complementary distribution and have the same morphotactic behavior, which is lost, or can only be talked about indirectly, in the slot metaphor.

³⁰Reshetnikov and Starostin (1995) and Butorin (1995)’s model, which attempts to assign distinct semantic roles to each set, labels them respectively the agentive, dative, and patient markers.

³¹Markers separated by \sim represent allomorphs. Those in parentheses are variants derived by predictable phonological processes, namely voicing assimilation and intervocalic lenition ($/t, d/ \rightarrow r, /k/ \rightarrow \gamma, /h/ \rightarrow \chi, /q/ \rightarrow \beta, /b/ \rightarrow \nu$). Lenition of $/t/$ and $/d/$ is characteristic of the southern dialect, in which most of the data in this paper are cited.

³²The present paper departs from previous Ketological work in using indices to indicate position classes instead of superscripts. This is done at the suggestion of one reviewer in order to avoid confusion with foot-

Table 4 Summary of Ket argument markers (based on Nefedov and Vajda (2015, pg. 38))

Krejnovich's system	<i>di/du</i> set	<i>ba/a</i> set	<i>di/a</i> set		
Vajda's system	P8	P6	P4	P3	P1
1SG	di~d (t, r)	ba~bə			di~d (t, r)
1PL	di~d (t, r)	dəŋ			dəŋ (taŋ, raŋ)
2SG	ku~k (g, ʎ)	ku (gu, ʎu)			ku~k (g, ʎ)
2PL	ku~k (g, ʎ)	kəŋ			kaŋ (gaŋ, ʎaŋ)
3M	du~d (t, r)	a~ɔ~bu	a~ɔ~aj		a
3F	da~dʌ~də	i~u~bu	i~dit (dir)~it~iru		a
3N	da~dʌ~də	ø~i~u~bu		b (v, p)	a
3PL	du~d (t, r)	aŋ~ɔŋ~bu	aŋ~ɔŋ~aŋa~ɔŋə		aŋ

- (11) at turə jəsla divrət
 at turə jəsla **di**₈-v-rət
 I this.FEM/NEUT paper 1.SBJ-3.N.OBJ-read

‘I’m reading this book’

- (12) ətn turə jəsla divrətɪn
 ətn turə jəsla **di**₈-v-rət-**in**₋₁
 we this.FEM/NEUT paper 1.SBJ-3.N.OBJ-read-**PL.SBJ**

‘We’re reading this book’

The P8 markers show distinct long (12) and short (13) variants. The long variants, with a vowel, appear only with certain basic verbs and only in the non-past.³³ The short forms show less phonological boundedness with the verb than do other verbal affixes, and in certain phonological environments may be pronounced separately, or appended to the preceding word (13), or elided altogether except in very careful speech. The phonological details necessarily fall outside the scope of this paper, but in relevant instances I represent the short P8 markers as clitics (=). I return to this issue in Sect. 5.6.

- (13) **bud** lovərabet
 bu=**d**₈ lovə_{r7}-a₄-bet₀
 3.SG.ANIM=3.M.SBJ₈ work₇-NPST₄-STEM₀

‘He works’ [KN 575]

- (14) əkŋ **klov**ərabetin
 əkŋ **k**₈=loβə_{r7}-a₄-bet₀-in₋₁
 you.PL 2.SBJ₈-work₇-NPST₄-STEM₀-PL.SBJ₋₁

‘You (plural) work’

notes or tone markers, and to be in more in line with general notational precedent outside the Ketological tradition.

³³I am only aware of one exception, that being the verb ‘die’, e.g. *kuno* ku_g-n-o (2.SBJ₈-PST-die) ‘you died’.

The P6 set of markers appears to the right of the P8 set and any P7 lexical morphs, if the latter are present, and to the left of all other components of the verb.³⁴

- (15) sa:lin**bayara**
 sa:lin-**ba**-y-a-ra
 spend.nights-1.S.SBJ.-STEM-PRS-STEM
 ‘I spend nights’ [VZ 136]
- (16) at arendiŋa d=**batsaq**
 at aren-di-ŋa d=**ba**-t-s-aq
 I forest-FEM-DAT 1.SBJ=1.SG.SBJ-STEM-PRES-STEM
 ‘I will make a quick trip to the forest’

The P4/P3/P1 set of markers, is the oldest, and shows many complexities which the other sets do not. For one thing, the 3rd person animate markers cumulatively mark tense (PST/NPST) alongside person, which is not true of any other argument markers.

The basis for splitting this set into three distinct position classes, an innovation introduced in Vajda (2001, 2003, 2004), is twofold. First, the third person animate markers occur to the left of the past tense markers {1, n}, while the 1st/2nd person markers occur to their right.

- (17) kildis, kild**an**s
 k-il-**di**-s, k-il-**dan**-s
 2.SBJ-PST-1.SG.OBJ-dress, 2.SBJ-PST-1.PL.OBJ-dress
 ‘you dressed me, us’ [VZ 189]
- (18) dilgus, dilg**an**s
 d-il-**gu**-s, d-il-**gan**-s
 1.SBJ-PST-2.SG.OBJ-dress, 2.SBJ-PST-2.PL.OBJ-dress
 ‘I dressed you (s, pl)’ [VZ 189]
- (19) dols, diruls
 d-**o**-l-s, d-**iru**-l-s
 1.SBJ-3.M.OBJ.PST-PST-dress, 2.SBJ-3.F.OBJ.PST-PST-dress
 ‘I dressed him, her’ [VZ 189]

This variable morphotactic behavior is not exhibited by the special ‘jointly-indexed’ 3rd person forms, *a* and *aŋ*, which occur in the same position as 1/2 markers.³⁵

³⁴The distinct rounded and unrounded variants of the first and third person markers (bo for the 1st person, u for the 3F, and so on) were taken in some treatments of the Ket verb (Vall & Kanakin, 1988; Reshetnikov & Starostin, 1995) to be a different marker set. Vajda (2004) argues convincingly that these allomorphs are the result of historical rounding before certain dorsal consonants.

³⁵The *o* in this example is only a past tense marker (see discussion beginning next paragraph). Cf. *digdondajcotn*, d⁸-igd⁷-o⁴-n²-**dan**¹-goi⁰-n⁻¹, 1.SBJ⁸-go.down.to.the.riverbank⁷-PST⁴-PST²-1.PL.SBJ¹-STEM.PL⁰-PL⁻¹ ‘we were going to the riverbank’ [KN 523]. The *o* remains the same when the person changes. This can be contrasted with pairs like the following, where *o* is present only with a masculine argument:

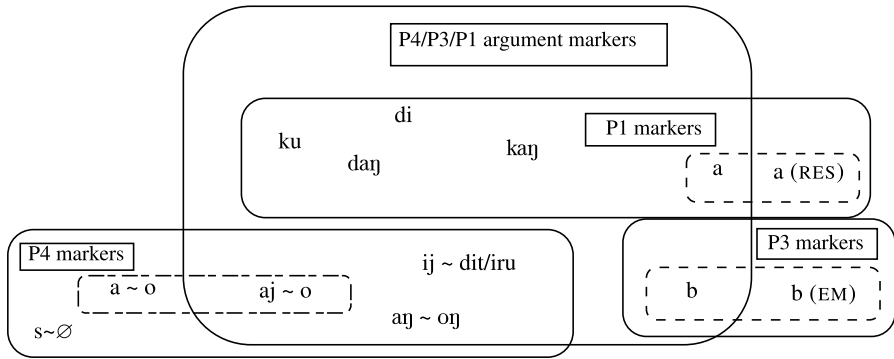


Fig. 2 Set Representation of the P4/P3/P1 Position Classes: solid lines share the same morphotactic behavior, dashed lines also share phonological material

- (20) digdonaraq
 d-igd-o-n-a-raq
 3.SBJ-go.down.to.the.riverbank-PST-PST-3.CO-IND-STEM
 ‘he was going to the riverbank’ [KN 523]

Second, the three subsets differ from one another in their relationship to other markers. Each of the subsets shows identical morphotactic behavior with, and is in complementary distribution with, at least one marker that is not in complementary distribution with the members of the larger set.

It may be helpful to visualize this using some actual modified set notation. In Fig. 2, markers which are in complementary distribution are shown as belonging to the same set.

The P4 argument markers {*aj~o*, *ij~dit/iru*, *aŋ~oŋ*}, for example, are in complementary distribution with the P1 and P3 sets of argument markers, and hence are elements in the same superset. However, they are also in complementary distribution with, and occur in the same position in the string as, the tense markers {*a~o*, *s~∅*} (NPST~PST). At the same time, the tense markers {*a~o*, *s~∅*} are not in complementary distribution with the P3 or P1 argument markers.

As an illustration, consider the forms in (21)–(23). The tense marker *a~o* (here *a* for NPST) freely co-occurs with *b* and *daŋ* (*v* and *raŋ* due to lenition, see fn. 31), but cannot co-occur with *ij*.

- | | |
|---|--|
| <p>(i) donolbet
 d⁸-o⁴-l²-bet⁰
 have.a.knife⁷-3.SG.M.PST⁴-PST²-STEM⁰
 ‘he had a knife’ [Vajda, 2004, p. 49]</p> | <p>(ii) donildəŋbet
 don⁷-il²-daŋ¹-bet⁰
 have.a.knife⁷-PST²-1.PL.SBJ¹-STEM⁰
 ‘we had a knife’ [Vajda, 2004, p. 50]</p> |
|---|--|

The [i] in *il*, as in (ii) only appears when required by the phonotactics, and is therefore likely epenthetic, although the idea that it is original, and is then deleted in cases like (i), has also been suggested. See Reshetnikov and Starostin (1995, p. 36) for an articulation of the latter view.

- (21) dusqavra
 d-us-q-**a**-v-da
 1.SG.SBJ-warm.up-STEM-NPST-**3.N.OBJ**-STEM
 ‘I warm it up’ [VZ 133]
- (22) dusqaraqda
 d-us-q-**a-raŋ**-da
 1.SG.SBJ-warm.up-STEM-NPST-**1.PL.OBJ**-STEM
 ‘I warm us up’
- (23) dusqijda
 d-us-q-**ij**-da
 1.SG.SBJ-warm.up-STEM-**3.F.OBJ**.NPST-STEM
 ‘I warm her up’

This same relationship holds between the P1 argument markers and the resultative marker *a* and between the argument marker *b* and the empty morph *b* (see §4.2).

Those sets which are encircled by dashed lines represent markers which show not only the same morphotactic behavior (they show up in the same position in the string, and it is not possible to have more than one of them in a single form), but also share the same phonological material in at least some instances. The tense marker *a~o* and the cumulative argument/tense marker *aj~o* are usually (but not always) distinct in the non-past (*a* vs *aj*), but are formally (i.e. phonologically and morphotactically) identical in the past (as *o*, seen already in *ditaybuyolʹ betin* ‘they migrated’ in §2.3).³⁶ The two *a*’s and two *b*’s, on the other hand, show complete formal identity in all cases.

All three pairs are the result of historical reanalysis, wherein the form and distribution of a marker has been altered through it being equated with another, historically unrelated, marker (Vajda, 2010). As such, their similarity clearly goes beyond accidental homophony. This leaves the difficult analytical question of whether these pairs are instantiations of *the same* marker, or not. Even if this were merely a case of accidental syncretism within the paradigm,³⁷ these pairs would still fit the definition of *polyfunctionality* used in this paper as “reuse of the same formal material with different functions”, and their formal identity clearly present a challenge for a listener, attempting to determine what each marker in a given verbform does. As such, given the narrow analytical focus of this paper, as investigating the information available to a listener in accomplishing exactly this task, this paper treats these pairs as instances

³⁶There are a few exceptions, wherein the two become completely identical. One is the verb ‘walk around’, where *aj~o* marks only tense:

- | | | | |
|-----|--|------|--|
| (i) | ktajga
k-t- aj -ga
2.SBJ-walk.around-NPST-STEM
‘you walk around’ | (ii) | ktolya
k-t- o -l-ya
2.SBJ-walk.around-PST-PST-STEM
‘you walked around’ |
|-----|--|------|--|

³⁷Thank you to a reviewer for suggesting this term in this instance and for pointing out the need to clarify this point.

Table 5 Summary of Ket intransitive argument-marking classes

		P8	P6	P4	P3	P1
v1	3.N.SBJ other subjects	SBJ			SBJ	
v2			SBJ			
v3		SBJ	SBJ			
v4	3.N.SBJ other subjects	SBJ			SBJ	SBJ
v5	3.ANIM.SBJ 3.N.SBJ 1/2 SBJ			SBJ	SBJ	SBJ

of the *the same* polyfunctional marker in those cases where they are homophonous, while acknowledging that different analytical choices might be appropriate under different analytical goals.

These markers have an important role to play in the story, and are revisited in §4 and §5.

Finally, returning to Fig. 1, there is one marginal exception to the given generalizations, that being that *b* (3N) and *a* (3.JOINT) can co-occur in class IV intransitive verbs, which is seemingly a by-product of the interaction of two separate families of constructions which these markers participate in (more on this in §3.2).

3.2 Argument-marking classes

Arguments are indexed using different combinations of these markers. The choice of which sets of markers are used to mark which arguments defines a given verb as a member of a particular inflection class. Classes are split according to transitivity, but among verbs of the same transitivity, class membership is largely arbitrary, and is usually not predictable from any semantic, syntactic, or phonological property of the given verb, although certain semantic clusterings can be observed within the classes, and certain constructions require a particular argument-marking pattern (Vajda, 2015).

This paper will refer to verbs which exhibit the same argument-marking pattern as belonging to the same *argument-marking class*, and will follow the system developed in Vajda (2015), Nefedov and Vajda (2015), Nefedov (2015), and Kotorova and Nefedov (2015), which classifies verbs into nine classes.

A summary of the intransitive argument-marking classes is given in Table 5. Note that classes v3 and v4 multiply exponence the subject.

Intransitive class 1 (v1) verbs mark their subjects with the P8 (di/du) marker set, as in (24).

- (24) ditolʁut / kutolʁut
 di₈-t₅-o₄-l₂-ʁut₀ / ku₈-t₅-o₄-l₂-ʁut₀
 1.SBJ₈-lie₅-PST₄-PST₂-STEM₀ / 2.SBJ₈-lie₅-PST₄-PST₂-STEM₀
 I lay, slept / you lay, slept [VZ 187]

Verbs of this class typically show an animacy split, where they mark neuter-class 3rd person subjects not with a member of the P8 set, but instead with the 3N marker from the P4/P3/P1 (di/a), b (here lenited to *v*, see footnote 31).

- (25) *tovilʁut*
 t₅-O₄-v₃-il₂-ʁut₀
 lie₅-PST₄-3.N₃-PST₂-STEM₀
 ‘it lay’ [VZ 187]

This is not always the case though, and some neuter subjects, particularly those that are perceived as playing a more active role in the event, are instead marked like feminine nouns.

- (26) *qoŋloq dʁvəl*
 qoŋloq dʁ-v-əl
 bell 3.F.SBJ-SING
 ‘a bell is ringing (lit. singing)’

Intransitive class II verbs (*v2*) mark their subjects with the P6 (*ba/a*) set.

- (27) *boʁatn, kuyatn*
 bo₆-ʁ₅-atn₀, ku₆-ʁ₅-atn₀
 1.S.SBJ₆-gO₅-STEM₀, 2S.SBJ₆-gO₅-STEM₀
 ‘I go, you go’ [Vajda, 2013, p. 634]

Intransitive class III (*v3*) verbs mark their subjects with both the P8 (*di/du*) set and the P6 (*ba/a*) set.

- (28) *dbatsaq*
 d₈=ba₆-t-s-aq
 1.SBJ₈-1S₆-make.a.quick.trip-N.PST-STEM
 ‘I make a quick trip and return’ [Vajda, 2013, p. 635.]
- (29) *kkutsaq*
 k₈=ku₆-t-s-aq
 2.SBJ-2S-make.a.quick.trip-N.PST-STEM
 ‘you make a quick trip and return’

As mentioned briefly in §2.3, the 3rd person members of the P6 set have a special ‘jointly-indexed’ form (this author’s term) *bu*, which is shared across classes. This marker is used only in the 3rd person by verbs which use Multiple Exponence of the subject (i.e. the P6 markers alongside another marker set, which for the P6 set is always the P8 markers).

- (30) *dbutsaq*
 d₈=bu₆-t-s-aq
 3.SBJ₈=3.CO-EX₆-make.a.quick.trip-N.PST-STEM
 ‘he makes a quick trip and returns’

Intransitive class 4 (v4) verbs mark their subjects with both the P8 (di/du) set and the P4/P3/P1 (di/a) set.

- (31) digdaddaq
 d₈-igda-d₁-daq
 1.SBJ₈-go.down.to.the.riverbank-1.SG.SBJ₁-STEM
 ‘I go to the riverbank’ [Vajda, 2013, p. 635; VZ 141]
- (32) kigdayuraq
 k₈-igda-yu₁-raq
 2.SBJ₈-go.down.to.the.riverbank-2.SG.SBJ₁-STEM
 ‘you go to the riverbank’ [Vajda, 2013, p. 635; VZ 141]

As with v3 verbs, there are special ‘jointly-indexed’ forms for the 3rd person P4/P3/P1 markers. Unlike v3, these distinguish number (*a* for S, *ay* for PL) and show different morphotactic behavior, occurring immediately before the right-most part of the stem.

- (33) digdonaraq
 d₈-igda-o-n-a₁-raq
 3.SBJ₈-go.down.to.the.riverbank-PST-PST-3.CO-IND₁-STEM
 ‘he was going to the riverbank’ [KN 523]

Like v1 verbs, this class also marks neuter subjects differently, using *b* from the P4/P3/P1 set in place of a marker from the P8 set.

- (34) oblatij
 o-b₃-l-a₁-tij
 PST-3.N₃-PST-3.CO-IND₁-grow
 ‘it grew’ [VZ 216]

There are some verbs which show a mixed v1/v4 class, following the v1 pattern with singular subjects, but the v4 pattern with plural subjects.

- (35) diroq, dirajdoβin
 di₈-roq, di₈-raj₁-doβ-in
 1.SBJ₈-fly, 1.SBJ₈-1.PL.SBJ₁-fly-PL.SBJ
 ‘I fly, we fly’ [Nefedov & Vajda, 2015, p. 44]

Finally, intransitive class 5 verbs mark their subject using the P4/P3/P1 (di/a) set.

- (36) s^jen^jdivet
 s^jen^j-di₁-vet
 have.a.reindeer-1.S₁-STEM
 ‘I have a reindeer’ [KN 596]

Table 6 Summary of Ket transitive argument-marking classes

	P8	P6	P4	P3	P1
vt1	SBJ		3.ANIM OBJ	3.N OBJ	1/2 OBJ
vt2	SBJ	OBJ			
vt3	SBJ ¹	SBJ ¹	3.ANIM OBJ	3.N OBJ	1/2 OBJ
vt4	SBJ ¹	OBJ			SBJ ¹

- (37) terijvet
 ter-ij₄-vet
 have.a.husband-3.F.S₄-STEM
 ‘She has a husband’ [KN 656]

The model adopted here distinguishes four transitive argument-marking classes, a summary of which is given in Table 6. Note that each verb marks only one object, but for those that use the P4/P3/P1 markers for the object, which of the three subsets is used depends on person and animacy (as was the case before with v5 verbs which P4/P3/P1 markers for subjects).

Note that, while the transitive classes are labelled by analogy with the intransitive classes which share the same subject-marking pattern (except for transitive class 2 and intransitive class 2, where the object of the former is marked like the subject of the latter), this convention is simply an artefact of how the analysis developed over time. The transitive classes are different inflection classes from the intransitive classes, and could just as appropriately be labelled classes 6–9, or something equivalent.

Transitive class 1 (vt1) verbs mark their subjects with the P8 (di/du) set and their objects with the P4/P3/P1 (di/a) set.

- (38) dausqolgura
 da₈=us₇-q₅-o₄-l₂-gu₁-ra₀
 3.F.SBJ₈=warm.up₇-STEM₅-PST₄-PST₂-2.SG.OBJ₁-STEM₀
 ‘she warmed you up’ [VZ 133]

Transitive class II (vt2) verbs mark their subjects using the P8 (di/du) set and their objects using the P6 (ba/a) set.

- (39) daejkugovla
 da₈=ej₇-ku₆-g₅-o₄-v₃-la₀
 3.F.SBJ₈=throw₇-2.SG.OBJ₆-STEM₅-PST₄-EM₃-STEM₀
 ‘she threw you’ [VZ 109]

Transitive class III (vt3) verbs mark their subjects using both the P8 (di/du) set and the P6 (ba/a) set, and mark their objects using the P4/P3/P1 (di/a) set. Like v3 verbs, the P6 set has a special ‘jointly-indexed’ form *bu* for 3rd person arguments.

- (40) dabulgut
 da₈=bu₆-l₂-gu₁-t₀
 3.F.SBJ₈=3.CO-EX₆-PST₂-2.SG.OBJ₁-carry₀
 ‘she carried you’ [VZ 194]

Transitive class IV (vt4) verbs mark their subjects using both the P8 (di/du) and the P4/P3/P1 set. Like v4 verbs, the P4/P3/P1 set has special jointly-indexed forms for 3rd person arguments, *a* and *aj*, occurring immediately before the right-most part of the stem. This class is very small and is limited to a few lexemes (Vajda, 2015).

- (41) dayuksaʁa
 da₈=y_u₆-k₅-s₄-a₁-ʁa₀
 3.F.SBJ₈=2.SG.OBJ₆-STEM₅-N.PST₄-3.JOINT₁-sell₀
 ‘she sold you’ [VZ 174]

Having laid out this necessary background on the Ket verb, the next section turns directly to the issue of polyfunctionality in the Ket verb.

4 Polyfunctionality in the Ket verb

The last section laid out the relevant components of the predominant model of Ket verbal morphology established in the Ketological literature, with some slight conceptual and terminological changes.³⁸ Assuming this model as a foundation, this section is now able to demonstrate how Ket argument markers instantiate polyfunctionality, expanding upon issues that have been alluded to or touched upon in §3. This is followed in §5 by a discussion of how syntagmatic implicative structure could aid in resolving complex form-function mappings.

All cases of polyfunctionality in Ket in this paper relate to the argument-marking system. They can be divided into two groupings: cases where the same marker is reused with different functions within the argument-marking system, and cases where formal material is shared between the argument-marking system and some other subsystem of the language, such that the same marker may or may not encode an argument, depending on the form.

The former is discussed in detail in section §4.1. As for the latter, one instance of this has already been discussed in §3 (pp. 27–29), that being the relationship between the P4 tense marker *a~o* (PST~NPST), which encodes only tense, and the P4 argument marker *aj~o*, which encodes both 3.SG.M arguments and tense cumulatively. The two markers are both morphotactically and phonologically identical in the past tense, and are treated for the purposes of this paper as representing a single polyfunctional marker in such cases.

The other instance concerns the relationship between the argument-marking system and the empty morph *b*, which has historically broken off from parts of the stem. This was mentioned briefly in section §3, and is discussed in full in §4.2.

4.1 Polyfunctionality within the argument-marking system

The broad observation that some markers can mark either the subject or the object was made already by Krejnovich (1968, pp. 22–23). As such, the present paper, like

³⁸Specifically, the avoidance of a spatial or *slot* metaphor for the position classes, in favor of a more set-theoretic terminology which refers directly to sets of markers and their morphotactic behavior.

Table 7 Argument marker sets by function (A/S/O) across classes

	A	S	O
P8	vt1, vt2, vt3, vt4	v1, v3, v4	
bu ₆ (3.joint)	vt3	v3	
a ₁ (3.joint)	vt4	v4	
P6 (not 3.joint)	vt3	v2, v3	vt2, vt4
P4		v5	vt1, vt3
P3		v1, v4, v5	vt1, vt3
P1 (not 3.joint)	vt4	v1, v4, v5	vt1, vt3

much of Ket linguistics, can be seen as an attempt to build on his observations using new theoretical tools.

Table 7 lays out whether a particular position class is associated with the subject or the object by inflection class. Recall that position classes here are understood as markers which have the same morphotactic behavior and are in complementary distribution. The P8 set always marks the subject, as do the special 3rd person ‘jointly-indexed’ forms of the P6 and P1 sets, but all other argument markers (P6, P4, P3, P1, except the 3rd person ‘jointly-indexed’ forms) can mark either the subject or the object.

Lest one think that some of the markers are simply tracking different types of arguments,³⁹ we can break down the notion ‘subject’ further into transitive (A) and intransitive subjects (S), which gives us the distribution in Table 7.

As Table 10 shows, the P8, P6 and P1 markers are all readily associated with both transitive and intransitive subjects, depending on the class. What is the case is that all four transitive verb classes use the P8 marker set, alone or in addition to another argument marker. Hence, the P6 and P1 markers are never the sole exponents of a transitive subject (with some possible exceptions, see §5.4). Nevertheless, in transitive verbs they can freely mark either transitive subjects (alongside P8), or objects, in a way that is clearly not reducible to tracking a single argument.

Consider the following two verbs with the P6 marker *ba*, where *ba* in the first verb marks a transitive subject (along with the P8 marker *d*), while *ba* in the second marks an object.

(42) dbatavraq
 d₈=**ba**₆-a₄-v₃-raq₀
 1.SBJ₈-**1.SG.SBJ**₆-pull.out₅-NPST₄-3.N.OBJ₃-STEM₀

‘I pull it out’

(43) dinejbatayit
 d₈-inej₇-**ba**₆-t₅-a₄-yit₀
 3.SBJ₈-scratch₇-**1.SG.OBJ**₆-STEM₅-NPST₄-STEM₀

‘he [a cat] scratches me’ [KN 530]

³⁹E.g. that some track might track absolutive arguments, as one reviewer suggests.

Table 8 Partial paradigm for ətiŋ7-q5-a4-da0, cover

NON-PAST		
3.N.OBJ	d8-ətiŋ7-(Ø6)-G5-a4-v3-ra0	‘I cover it’
1.PL.OBJ	d8-ətiŋ7-daj6-G-a4-v3-ra0	‘I cover us’
2.SG.OBJ	d8-ətiŋ7-ku6-ʙ5-a4-v3-ra0	‘I cover you’
PAST		
3.N.OBJ	d8-ətiŋ7-(Ø6)-G5-o4-v3-il2-da0	‘I covered it’
1.PL.OBJ	d8-ətiŋ7-daj6-G-o4-v3-il2-da0	‘I covered us’
2.SG.OBJ	d8-ətiŋ7-ku6-ʙ5-o4-v3-il2-da0	‘I covered you’

For another example, the following two verbs with the P1 marker *di*, where in the first verb *di* marks a transitive subject along with the P8 marker *d*, and in the second *di* marks an object.

- (44) doldiʙa
 d8-o4-l2-**di**1-ʙa0
 1.SBJ8-3.SG.M.OBJ4.PST2-1.SG.SBJ1-sell0
 ‘I sold him’ [KN 547]

- (45) daayabundiʙos
 da8=ayə7-bu6-n2-**di**1-ʙos0
 3.F.SBJ8=take.to.the.forest7-3.SBJ.JOINT6-PST2-**1.SG.OBJ**1-STEM0
 ‘she took me to the forest’ [KN 437]

The only markers associated with intransitive subjects but not transitive subjects are the P4 and P3 markers. Prefiguring the discussion in section §5, this gap has nothing to do with the type of arguments tracked by the P4 and P3 markers, but is instead because these two markers always encode the 3rd person. Since, in order to mark transitive subjects, they must be a co-exponent with a P8 marker, the special 3rd person co-referential form of the P4/P3/P1 markers, *a*₁, is used instead.

4.2 Argument marking and empty morph *b*

Work as early as Krejnovich (1968, p. 38) has noted that *b* (phonetically often *v* or *p*), usually the 3rd person neuter member of the P4/P3/P1 argument marker set, appears in many verbs which lack any neuter argument. In such cases, it occurs throughout the paradigm and lacks any obvious grammatical function. However, it is phonologically identical with the neuter argument marker, which occurs in exactly the same morphotactic position – after any P4 marker and before the P2 past tense markers *l* and *n*.

For example, consider the partial paradigm of the verb ‘cover’, given in Table 8. Here the actual object is marked with the P6 marker set.⁴⁰

⁴⁰The variation between *g* and *ʙ* (both > /q/) is regular phonology. Recall that, in Ket, stops lenite intervocally, and voice.

Table 9 Partial paradigm for *us7-q5-a4-da0* WARM.UP

NON-PAST		
3.N.OBJ	d8-us7-q5-a4-v3-ra0	'I warm it up'
1.PL.OBJ	d8-us7-q5-a4-daŋ1-da0	I warm us up
2.SG.OBJ	d8-us7-q5-a4-γu1-ra0	'I warm you up'
PAST		
3.N.OBJ	d8-us7-q5-o4-v3-il2-da0	'I warmed it up'
1.PL.OBJ	d8-us7-q5-o4-l2-daŋ1-da0	'I warmed us up'
	d8-us7-q5-o4-l2-gu1-da0	'I warmed us up'

Table 10 Possible functions for each set of potential argument markers

Marker set	Possible function
P8	SUBJECT
bu ₆ , a ₁	SUBJECT
P6	SUBJECT or OBJECT
P1	SUBJECT or OBJECT
P4 (not o ₄)	SUBJECT or OBJECT
o ₄	SUBJECT or OBJECT or TENSE
b ₃	SUBJECT or OBJECT or EM

Compare this to a verb like 'warm up', given in Table 9, where *b* is a true agreement marker, in paradigmatic opposition to other P4/P3/P1 markers indicating the object. This case is far from anomalous: Kotorova and Nefedov (2015) list 921 verbs which have *b* throughout their paradigms.

Based on comparison with Kott, Vajda (2017a, 2013) demonstrates that non-argument-marking cases of *b* derive historically from unrelated sources via metathesis from other parts of the stem, which were then reanalyzed.

Non-argument-marking *b* might be analyzed as an *empty morph*, a unit of form without any corresponding function or meaning (Hockett, 1947). Krejnovich (1968, p. 38) explicitly opts for this analysis, calling it an 'empty morpheme' (*pustaja morfema*). More recent literature has typically opted for more neutral, but also more idiosyncratic, terminology. Vajda (2013) and following work refers to *non-agreement b*, while Nefedov and Vajda (2015) talk about *thematic b*.

This paper adopts the empty morph analysis of Krejnovich (1968), and glosses *b* in relevant instances as such (EM).

4.3 Summary of polyfunctionality in Ket and discussion

A useful concept in summarizing the Ket system is that of a *potential argument marker*, a marker that indicates arguments in at least some verbs. Taken together, all of the possible functions for Ket potential argument markers can be summarized as in Table 12.

5 Implicative relations in syntagmatic structure in Ket

This section lays out examples of implicative structure internal to Ket verbforms, between potential argument markers. Although these markers are highly polyfunctional across the language as a whole, the listener can often substantially narrow down the range of possible functions for a given potential argument marker based on which other potential argument markers are present in the same wordform, and which person/number/class features those markers encode.

It should be stressed that the purpose of this section is not to claim that Ket listeners *do* pick up on these patterns and use them in online processing, which is of course an empirical psycholinguistic question.⁴¹ The point is simply to point out that these patterns exist, are available in the input, and allow for a massive formal difference between the range of possible functions for a given marker in the language as a whole, and the range of possible functions which that marker can have in context.

5.1 Single argument marker

The absence of other potential argument markers implies that the marker that is present indicates an intransitive subject.

- (46) diyaraq
 di₈-γ-a-raq
 1.SBJ₈-live-NPST-STEM
 ‘I live’
- (47) boyatn
 bo₆-γ-a-tn
 1.SG.SBJ₆-go-NPST-STEM
 ‘I go’
- (48) s^jen^jdivet
 s^jen^j-di₁-vet
 have.a.reindeer-1.SG-STEM
 ‘I have a reindeer’ [KN 596]

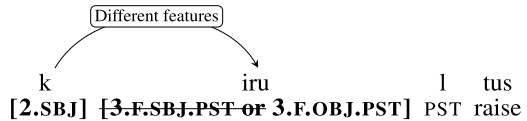
5.2 More reliable markers disambiguate less reliable ones

Ket potential arguments markers can be organized along a scale, from least to most polyfunctional, depending on how greatly their functions can vary across different lexemes. Recall that bu₆ and a₁ are the special ‘jointly-indexed’ 3rd person forms of the P6 and P4/P3/P1 sets respectively.

- (49) P8 SET, bu₆, a₁ »P6 SET, P1 SET, P4 SET (not O₄) » o₄, b₃

⁴¹As a reviewer rightfully points out, this would be an extremely difficult question to address given the state of the language and the age of the remaining fluent speakers.

Fig. 3 P8 marker and another marker have different features (ex. (51))



A frequent phenomenon is that the presence of markers further to the left along this scale has implications which help to disambiguate the intended function of the markers further to the right along the scale. This is especially true with regard to the P8 set of markers. (54) is the first of several generalizations about the P8 set of markers, which has important implications for other markers in the verb:

- (50) P8 markers only ever mark subjects, never objects or something outside of the argument system.

This generalization means that if a P8 marker occurs in a given verbform alongside markers of other sets, those other markers can only be exponents of the subject if they are co-exponents with P8. This in turn is only possible if the given marker has compatible features with the P8 marker.

For example, consider a form like *kirultus* ‘you raised her’.

- (51) kirultus
 k₈-iru₄-l-tus
 2.SG.SBJ-3.F.OBJ.PST-PST-raise
 ‘you raised her’

In this case, *iru*₄, as a P4/P3/P1 marker, is more polyfunctional than the P8 marker *k*₈. Depending on the verb, it might indicate either the subject or the object. However, the P8 marker must mark the subject. Since the two markers have incompatible features, meaning they cannot be co-exponents of the same subject, the P4 marker must mark the object in this form.

As before, we can represent implicative relations between morphs using a dependency graph, as shown in Fig. 3.

Conversely, if a form has a P8 marker and another marker which has the same features, then either the two must be co-exponents of the subject (i.e. there must be subject Multiple Exponence), or the form must be reflexive. In Ket, these two scenarios are morphologically identical. Compare a true reflexive verb in (51), with a verb with double subject marking in (52).⁴²

- (52) da₈-qo:l-q-o-l-a_{4/3/1}-ra
 3.F.SBJ8-treat-stem-PST-PST-3.SG.JOINT-STEM
 ‘she treated herself’

- (53) da₈-bo₆-k-s-a_{4/3/1}-ba
 3.F.SBJ8-1.SG.OBJ6-sell-NPST-3.JOINT-STEM
 ‘She sells me’ [KN 547]

⁴²Note that this verb lacks reflexive semantics whatsoever. Many verbs with double subject-marking might be analyzed as cases *deponency*, wherein a syntactically transitive verb exhibits intransitive morphology.

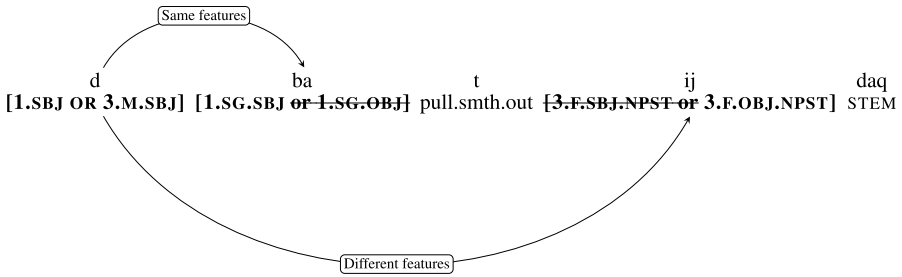
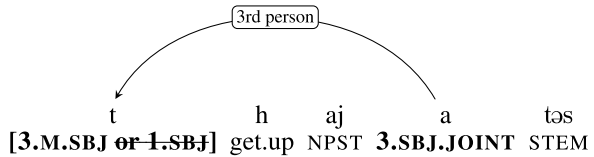


Fig. 4 Same features as P8 marker is subject, other is object (ex. (54))

Fig. 5 3rd person jointly-indexed form disambiguates P8 marker (ex. (56))



When there are three markers present, if one of them has the same features as the P8 marker, then that marker is a co-exponent of the subject, while the other non-P8 marker is the object, as shown in example (54) and Fig. 4.

- (54) dbatijdaq
 d₈=ba₆-t-ij₄-daq
 1.SBJ₈=1.SG.SBJ₆-STEM-[3.F.SBJ or 3.F.OBJ]₄-pull.out
 ‘I pull her out’

Recall that in the 3rd person the P6 and P4/P3/P1 marker sets have special forms (*bu*₆ and *a*₁ respectively) when they are jointly indexed with a P8 marker, as shown in (55) (*bu*₆) and (56) (*a*₁).

- (55) dbutsaq
 d₈=bu₆-t-s-aq
 3.SBJ-3.JOINT-make.a.quick.trip-N.PST-STEM
 ‘he makes a quick trip and returns’
- (56) thajatæs
 t₈=h-aj-a₁-tæs
 [1.SBJ or 3.SBJ]₈-get.up-NPST-3.SBJ.JOINT₁-STEM
 ‘he gets up’

These forms have important implication themselves. For one, they disambiguate the 1st person and 3rd person masculine forms of the P8 set (both *d* or *t*), which are otherwise syncretic, as shown in Fig. 5.

They also disambiguate some forms where the subject and object are of the same class and number, as shown in Ex. (57) and Fig. 6.

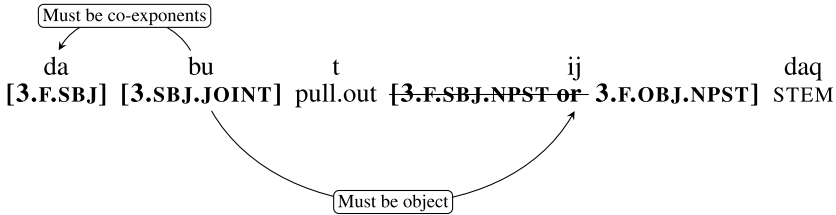


Fig. 6 3rd person jointly-indexed forms disambiguate cases where multiple markers have same features (ex. (57))

Table 11 P8/P6/P1 Generalizations

Combinations	Form-function mappings	Notes
P8 + P6	P8 is subject, P6 is object or Reflexive (if same features) or Co-exponents (if same features)	bu ₆ always jointly indexed with P8 subject
P8 + P1	P8 is subject, P1 is object or Reflexive (if same features) or Co-exponents (if same features)	a ₁ always conidexed with P8 subject

(57) dabutijdaq
 da₈=bu₆-t-ij₄-daq
 3.F.SBJ₈=3.SBJ.JOINT₆-pull.smth.out-3.F.OBJ₄-STEM
 ‘she pulls her out’

We can further summarize the relevant generalizations as in Table 11.

Things become more complicated once one considers forms with the potential argument markers *o* and *b*, since these can either mark arguments or can serve other functions. However, many of the implicative relations already discussed can also help to narrow down the range of possible functions for these markers in any given form as well.

For example, in a verb with a P8 marker and *b*, the latter can either indicate a neuter object (58) or be an empty morph (59).⁴³

⁴³A reviewer asks whether phonological context might be useful in disambiguating the function of *b* in a given form. Vajda (2017a) notes about the metathesis process that created empty morph *b* “Ket metathesis typically involved labial and non-labial segments changing places, though the phonological trigger is not fully clear. [p. 22]”. It does seem to be the case that none of the verbs in Kotorova and Nefedov (2015) with empty morph *b* have a P0 stem component beginning with a labial consonant (of which phonemically Ket just has /b/ and /m/, not counting /p/ and /ɸ/ in some recent Russian loans), so it is possible that that is a generalization that speakers could pick up on. It would only help in one direction though, since of course there are very many verbs that use *b* as an argument with P0 stem components that do not begin with a labial consonant. It is important to stress though that this paper is focused only on how syntagmatic

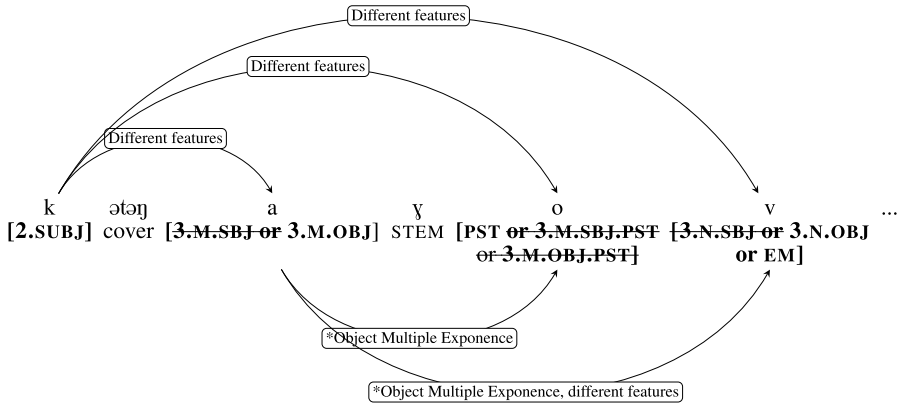


Fig. 9 Marker with different features from P8 marker cannot mark the subject; a_6 must mark an argument, multiple object marking not possible, therefore o and v cannot mark arguments here (ex. (65))

means that if the subject and object can be determined, all other potential argument-markers in the form must serve a non-argument-marking function.

For example, consider the form *kətəŋayovla* ‘you covered him’, in example (65), which has four potential argument markers, k_8 , a_6 , o_4 , v_3 .

- (65) *kətəŋayovla*
 k_8 - $ətəŋ$ - a_6 - $ʧ$ - o_4 - v_3 - l - a
2.SBJ-cover-3.M.OBJ-STEM-PST
-3.N.SBJ-EM₃-PST-STEM
 ‘you covered him’

The first marker, k_8 must represent the subject. None of the other potential argument markers have compatible features, which rules them out as co-exponents of the subject. See Fig. 9 for illustration.

This leaves all three of a_6 , o_4 and v_3 as potential object markers. Of the three, only a_6 must always represent an argument. Since it cannot mark the subject here, it must mark the object. Since Multiple Exponence of the object is impossible, a_6 being the object implies that the other two markers must not be argument markers. In this way again, the less polyfunctional marker (a) helps to disambiguate the function of the more polyfunctional markers (o and v). Again, see Fig. 9 for illustration.

‘I pull her/it, he pulls her/it), with the empty morph b_3 reinterpreted as a genuine neuter object marker. Many thanks to Edward Vajda for pointing out and discussing these data with me. In a similar vein, Olivier Bonami (p.c.) asked whether the o_4 in cases like *k-ətəŋ-a-ʧ-o-v-l-a* ‘you cover him’ could be interpreted as a second object marker, since it has compatible features with the P6 object a_6 . The data from Krejnovich suggest that this is a possible generalization speakers could make, although paradigmatically related forms suggest that they have not, cf. *k-ətəŋ-i-ʧ-o-v-l-a* ‘you cover her’, and not **k-ətəŋ-i-ʧ-iru-v-l-a*, with two feminine object markers. Forms like *d-o-gd-aj-taj* and **k-ətəŋ-i-ʧ-iru-v-l-a* would cause no problem in decoding; *aj* and *iru* both must always represent arguments, and have different features from the subject (or, for the third person, are not in their jointly-indexed forms), and so are unambiguously object markers.

5.4 Restrictions on transitive subject marking

Recall again that all of the argument-marking classes for transitive verbs involve a P8 marker as one of the exponents of the subject. This generally means that if the verb does not have a P8 marker, it is intransitive, which in turn leads to the following generalization in (66).

(66) If the verb has no P8 marker, it also has no object.

There do seem to be some marginal exceptions to this. Specifically, some class v2 verbs which have *b*, seemingly as an empty morph, appear to be able to take 3rd person neuter objects.

(67) āt jekkənnə dūy bayabda
 āt jekkən=nə dūy ba₆-y-a-b₃-da
 I thunder.PL=GEN.PL yell 1.SG.SBJ-hear-N.PST-EM-STEM
 ‘I hear thunder’, lit. ‘I hear the thunder’s yell’

However, objects of other persons do not seem to be possible. The meaning of ‘I heard you’, for example can only be expressed using a paraphrase with a 3rd person neuter object, ‘I heard your words’.

(68) āt ūk qáàn bayovilda
 āt ūk qáàn ba₆-y-o-v₃-il-da
 I your words 1.SG.SBJ-hear-NPST-EM-PST-STEM
 ‘I heard you’, lit. ‘I heard your words’

It is likely that such verbs have *b* as an empty morph, and lack true object agreement,⁴⁷ and yet the formal identity with the neuter object marker licenses a neuter object, although it could also be that *b* is a real object marker in such cases, and these verbs can only take neuter objects for some other reason.

In either case, it is not possible for any other marker to indicate an object in the absence of a P8 marker. For example, in *bayovilda* ‘I heard’, neither *ba₆* nor *o₄* can indicate the object, as shown in Fig. 10.

Now the question arises of which of the three potential argument markers marks the subject. In cases like this, the least polyfunctional marker is *ba₆*. It must indicate an argument, unlike *o₄* and *b₃*. Since it is not possible for it to mark the object here, its presence excludes the possible subject function of the other two markers. This is also illustrated in Fig. 10.

These implicative relations hold for any verbform that lacks a P8 marker but has a P6 marker, a P1 marker other than *a₁*, or a P4 marker other than *o₄*. That marker will indicate the subject, while all other potential argument markers must serve a non-argument-marking function (save for the aforementioned sort of neuter quasi-object marking). This is summarized in Table 12.

⁴⁷Some verbs in Ket lack object agreement in some or all cells of the paradigm without obvious reason, with a lack of overt agreement with neuter objects being the most frequent. Cf. *qim bə'n=da had-dit-l-a*

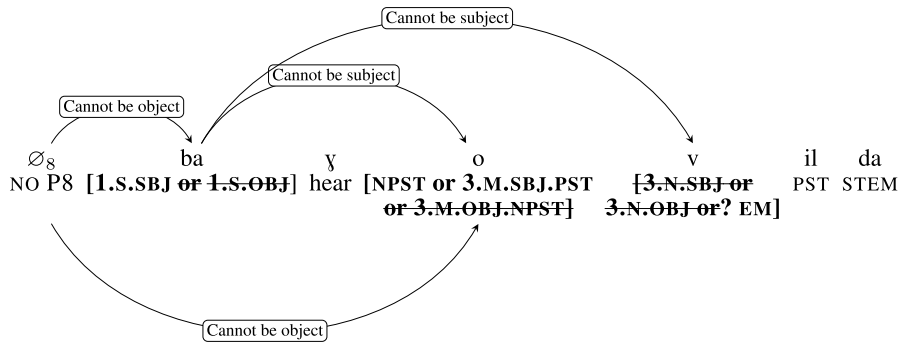


Fig. 10 Absent a P8 marker, verb usually cannot have an object; P6 markers must mark subject, others cannot mark subjects and usually cannot mark objects (cf. ex. (68))

Table 12 Summary of generalizations from Fig. 10

Combination	Form-function mappings	Notes
P6 + P4	P6 marks subject	P4 only tense
P6 + P3	P6 marks subject	P3 empty morph
P6 + P4 + P3	P6 marks subject	P4 only tense, P3 empty morph

It also means that the following combinations of markers in the same verbform should be impossible, since it would require that the verb have two subjects: P6 + P1, P6 + P4 other than o₄, and P4 other than o₄ + P1 other than a₁.

Of course, the question remains as to why (nearly?) all transitive verbs in Ket use the P8 marker set, alone or in addition to another marker set. In attempting to answer this, it should be kept in mind that the P8 set of markers is the most recent component of the verb to grammaticalize. The P1 and P6 sets are much older; they were present in Ket’s now dormant relative Kott,⁴⁸ and are reconstructed to Proto-Yeniseian. Based on Vajda’s reconstruction (Vajda, 2010, 2017a), all 1/2 subjects in Proto-Yeniseian (S or A) were marked with the P1 set, while the P6 set only ever marked objects (though not all objects were marked with it).⁴⁹ However, both sets of markers in

(woman duck=3.F.SG cut-3.F.OBJ.PST-PST-STEM, ‘the woman cut up the duck’), with an object marker, vs. *qim kilaj da=ha:l-a* (woman threads 3.F.SBJ=cut-pst-stem ‘the woman cut the threads’), without one.

⁴⁸Kott was another Yeniseian language, spoken around the territory of modern Krasnoyarsk, in southern Siberia. Its speakers had been assimilated by their Turkic, Russian and Samoyedic-speaking neighbors and the language ceased to be spoken by the second half of the 19th century (Werner, 1990). Most of what is known about the language comes from the work of the Finnish linguist Mattias Castrén (Castrén, 1858), who worked with some of its last speakers during his travels in Siberia in the 1840’s. Despite the similar names, Ket and Kott are not closely related, coming from different branches of the family.

⁴⁹3rd person subjects, and it seems also some 3rd person objects, were indicated using markers ancestral to *b* from the P3 set and *aj~o* from the P4 set. However, it is unclear from the literature what exactly the pattern was (which objects were marked with these as opposed to P6) or how it related to transitivity. Modern Ket avoids using *b* to mark semantic agents, whether they correspond to a transitive or intransitive syntactic subject, and instead uses the feminine P8 marker in such instances (see p. 32). Hence, specifically

the modern language, he argues, are heterogeneous in origin, resulting from fusion of previously distinct markers and metathesis with reanalysis of material from other components of the verb, including one another, leading to their polyfunctionality in the modern language. This means that, at minimum historically, the requirement that transitive verbs include the P8 markers has nothing to do with the inability of either the P6 or P1 sets to index transitive subjects – at least the P1 set has always been able to do that. Rather, as argued further on in §5, resolving which markers, if any, indicate transitive subjects is one of several disambiguating functions which the P8 markers serve, reflecting a way in which the language has increased its syntagmatic implicative structure in order to deal with the more complex form-meaning mappings of its older markers.

5.5 Co-occurrence restrictions within the P4/P3/P1 set

As discussed in §3.1, $aj\sim o_4$ as an argument/tense marker and $a\sim o_4$ as a tense marker show identical morphotactic behavior, and identical phonological behavior in the past. However, they are not distributionally identical. Tense-marking $a\sim o_4$ is in complementary distribution with the P4 markers: the argument-marking $aj\sim o_4$ as well as ij_4 and its allomorphs and 3PL aj and its allomorphs. In (69) and (70), $a\sim o$ is elided when another P4 marker is present.

- (69) d-us-q-**aj**₄-da
1.SBJ-warm.up-STEM-3.M.OBJ.NPST
'I warm him up'
- (70) d-us-q-**ij**₄-da
1.SBJ-warm.up-STEM-3.F.OBJ.NPST
'I warm her up'

However, $a\sim o$ freely co-occurs with other members of the P4/P3/P1 argument set, such as b (v, p, 3.N.OBJ or EM) and daj (raj , 1PL). This is shown in (71) and (72).

- (71) d-us-q-a-v-ra
1.SBJ-warm.up-STEM-NPST-3.N.OBJ-STEM
'I warm it up'
- (72) d-us-q-**a**₄-**raj**₁-da
1.SBJ-warm.up-STEM-NPST-1.PL.OBJ-STEM
'I warm us up'

Argument/tense marking $aj\sim o_4$ does not have this property. It is in complementary distribution with all other members of the P4/P3/P1 set. This means that whenever another P4/P3/P1 marker is present, this implies that o_4 can only be a tense marker. This is illustrated in Fig. 11.

The other member of the P4/P3/P1 set which can serve a non-argument marking function, b , seems to show a different distributional restriction. Namely, b as an empty

for b , the idea that it can only track arguments corresponding to a particular semantic role (patients) is a reasonable one (cf. §4.1), even though this is not the case for the P6 and P1 sets.

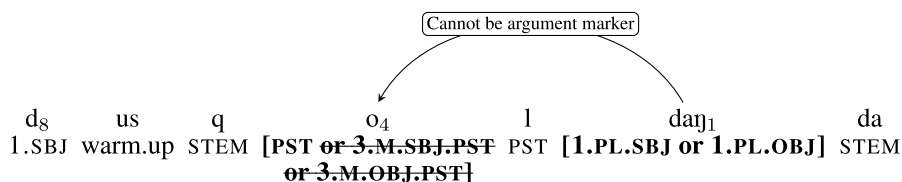


Fig. 11 Presence of other P4/P3/P1 markers disambiguates *o*

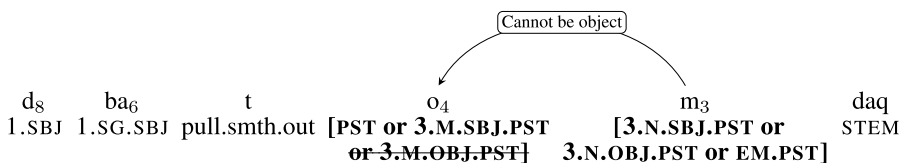


Fig. 12 With *b* present, *o* cannot mark object

morph never occurs in the same paradigm as *b* as an object marker,⁵⁰ maybe for historical reasons or maybe because the former would be likely to be reanalyzed as the latter. A search of Kotorova and Nefedov (2015) confirms this: of the 921 verbs that are listed with *b* as an empty morph, only 8 of them are listed as belonging to classes vt1 or vt3, which are the classes that use P4/P3/P1 for object marking, and all of them appear to be mislabelled vt2 or v1 verbs.

This does nothing for disambiguating the function of *b* itself, but it does mean that any time *b*, as either an argument or an empty morph, occurs alongside *o*, the latter cannot be an object. This is illustrated in Fig. 12. Note that *b*₃ in Fig. 12 has an allomorph *m*, which also marks past tense, due to historical coalescence between *b*₃ and a following plural marker *n*.

- (73) dbatomdaq
 d₈=ba₆-t-o₄-m₃-daq₀
 1.SBJ-1.SG.SBJ-pull.smth.out-PST-3.N.OBJ.PST-STEM
 ‘I pulled it out’

5.6 Elision of the P8 markers

Before moving on, there is one potential challenge to the system laid out above, which should be addressed. As noted in §3.1, the P8 markers have long and short forms. The short forms, which are used by the vast majority of verbs, show clitic behavior before consonants, and often encliticize to the end of the preceding word rather than appearing with the rest of the verb. All but the 3rd person feminine marker are single consonants. If the P8 marker is not encliticized to the end of the preceding

⁵⁰Whether *b* as an empty morph and as a subject marker can co-occur in the same paradigm is unclear. The search turned up only 3 verbs with *b* as an empty morph belonging to v5, and they all seem to be mislabelled v1 verbs. However, v1 verbs with stem-component *b* are much more common. The question then would be whether any of these verbs allow neuter subjects, and what happens in such cases. This is a topic for future research.

word, it is often elided entirely, except in careful speech, especially if the preceding word also ends with a stop (*at d=lovərəbət* ‘I work’, or *at lovərəbət*).⁵¹ The 3.F P8 short form marker *da* never elides (*bu da=lovərəbət* or *bu=da lovərəbət* ‘she works’, **bu lovərəbət*), and the short P8 markers are always retained remain with the rest of the verb when they precede a vowel (*bu d-ik-sə-bəs* 1.SBJ-come-STEM, ‘he comes’, **bu=d-ik-sə-bəs*, **bu ik-sə-bəs*).

If P8 markers are so important to the comprehension of unknown verbal forms, what can be said about cases in which the P8 marker is elided? Any amount of elision of the P8 markers greatly increases the amount of uncertainty about the function of a novel form. Large speech corpora, which unfortunately are almost impossible to collect for Ket at this stage, would be necessary to determine what percentage of novel verbs encountered by speakers occur in contexts in which the P8 marker is completely elided, but if they represent a substantial portion of the input, then this might be a serious challenge to the idea that the P8 markers are a functional adaptation meant to increase syntagmatic implicative structure.

At the same time, there is some diachronic evidence which supports a functional adaptation story for the P8 markers. It seems to be the case that the elision of the P8 markers is a recent development. Texts collected by Anuchin at the turn of the 20th century show cases of long-form P8 markers, with full vowels, in cases where the short form would be used and likely elided in modern Ket (cf. modern Ket (*t=*)*kajnam* ‘I took it’).⁵²

- (74) *daljam di kaj-n-am*
 half 1.SBJ take-PST-STEM
 ‘I took half’ (Dul’zon, 1972, pp. 200-201)

If the elision of the P8 marker is a recent development, and if the P8 markers are responsible for much of the syntagmatic implicative structure of the verb, and if there is indeed a connection between syntagmatic implicative structure and the maintenance of high complexity of exponence in Ket, then we would expect the elision of the P8 markers to be associated with a corresponding drop in complexity of exponence. And indeed there is some evidence for this in modern Ket. The author’s teacher seems to show certain signs of paradigm leveling, wherein the 3rd person form of the older, non-P8, exponent in some verbs with Multiple Exponence is extended to other, though not all, cells of the paradigm. Note (*g*)*bugbus* for the 2nd person singular, instead of the expected *u=k ku-g-b-us*, *you=2.SBJ 2.SG.SBJ-carry-3.N.OBJ-STEM*.

- (75) *ad bo-g-b-us*
 I 1.SG.SBJ-carry-3.N.OBJ-STEM
 ‘I carry it’
- (76) *u=g₈ bu₆-g-b-us*
you=2.SBJ₈ 3.SBJ.JOINT-carry-3.N.OBJ-STEM
 ‘you carry it’

⁵¹The frequency of elision is impressionistically not the same before all consonants, being more likely before stops, although the actual phonological details do not seem to have been worked out.

⁵²Many thanks to Edward Vajda, who pointed out these data and discussed them with me in depth.

Table 13 Estonian noun ‘lock’, taken from (Blevins, 2016, pp. 16–18)

Nominative	lukk
Genitive	luku
Partitive	lukku

- (77) u akws bu-g-b-us?
 you what 3.SBJ.JOINT-carry-3.N.OBJ-STEM
 ‘what are you carrying?’

In all of these forms, there is a consonant immediately following where the P8 marker would be expected. In (76), the P8 marker is saved by being encliticized to the end of the preceding, vowel-final, word. In (75) and (77), the preceding word ends with a consonant, preventing encliticization, and the P8 marker is elided. These data show that the 3rd person jointly-indexed member of the P6 marker set has been extended to the 2nd person singular, although not the 1st person singular.⁵³ Note that it is specifically *bu* which is the least polyfunctional member of the P6 set, always representing the subject, and hence the subject can be easily recovered even when the P8 marker is lost (as in (77)).⁵⁴ This seems consistent with a system which is “readjusting”, finding ways to compensate for the loss in some instances of the P8 markers, and hence the decoding cues which they provide. More research is needed on this topic.

6 Ket and the typology of exponence

Before concluding, it is worth briefly exploring how the type of complex exponence found in Ket relates to other types of highly complex exponence discussed in the literature, and specifically the notions of *Gestalt Exponence* (Blevins, 2016) and *Distributed Exponence* (Carroll, 2022).

Gestalt Exponence is not precisely defined in the literature, but is conventionally applied to those cases in which individual formatives cannot be associated with a particular meaning in a given form, but rather encode meaning only through their combination. Consider the partial paradigm for Estonian LUKK ‘lock’, given as an example of Gestalt Exponence in Blevins (2016, pp. 16–18), shown in Table 13.

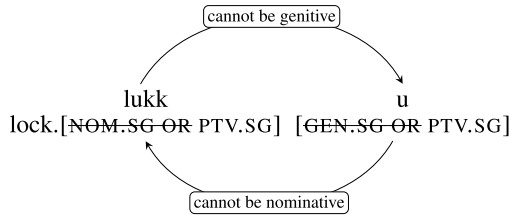
Estonian encodes case through a combination of stem alternations (strong *kk* stem vs. weak *k* stem) and inflectional suffixes. Neither the strong stem nor the suffix *u* can be associated with the meaning ‘partitive singular’, as the strong stem is shared with the nominative, and the suffix is shared with the genitive. Only in conjunction do the strong stem and suffix *u* encode the meaning ‘partitive singular’.

Ket superficially bears some resemblance to such exponence patterns in that individual elements are reused across the morphology in different combinations for different functions. The crucial difference appears to be that Ket reuses pieces for

⁵³It would seem logical that retention of the more conservative form in the 1.SG is connected with the 1.SG being a relatively more frequent form than other persons.

⁵⁴In this case, there is an overt pronoun *u*, but the sentence is full grammatical without it.

Fig. 13 Estonian Gestalt
Exponence as polyfunctionality
with syntagmatic implicative
structure



different functions only across lexemes, not within the inflectional paradigms of single lexemes. In other words, within the paradigm of a given lexeme (say *dbatsaq* ‘I make a quick trip to the forest and return’), a given marker (say *ba*) always encodes the same function (here 1.SG.SBJ). Deviations from consistent form~meaning mappings arise only when one compares across lexemes (say, *ba* in *kətəɣbayavra* ‘you cover me’, which encodes 1.SG.OBJ). This gives Ket an essentially morphemic quality which differs from those systems described as Gestalt Exponence, such as Estonian, Georgian (Blevins, 2016; Gurevich, 2006), and Yam languages (Carroll, 2022; Carroll et al., 2016). For each individual lexeme, the meaning of the whole is essentially the sum of the meaning of the parts.

At the same time, this is partly a question of analytical choice. The Estonian system could also be analyzed in similar terms to the Ket system in this paper. The strong stem and the suffix *u* could both be analyzed as polyfunctional morphs; the strong stem could be taken to encode either nominative or partitive singular, and the *u* suffix to encode either genitive or partitive singular. The form then would exhibit Multiple Exponence, and the two morphs would exhibit syntagmatic implicative structure, disambiguating one another, as shown in (78) and Fig. 13.

- (78) *lukk-u*
 lock.PTV.SG-PTV.SG
 ‘lock (partitive singular)’

The reverse does not appear to be true; it is not possible to analyze the Ket system in terms of recombinant Gestalts, which lack compositional structure. Ket verbs are not always able to be fully disambiguated based on syntagmatic structure alone (for example, verbs with empty morph b_3 are often ambiguous in their transitivity based on syntagmatics alone). Hence, the same combination of markers need not represent the same paradigm cell in all cases, unlike systems like Estonian. Hence, a Gestalt analysis for Ket not only fails to capture the unique properties of the system, but makes the wrong predictions about interpretability based on different combinations of formants.

Some recent work by Matthew Carroll has suggested an approach which allows for a unified treatment of both types of systems (Carroll, 2022). In defining his notion of *Distributed Exponence*, Carroll presents an approach which somewhat abstracts away from this difference between Ket and languages with prototypical Gestalt exponence. He defines *Distributed Exponence*⁵⁵ as the following (Carroll, 2022, 2):

⁵⁵Distributed Exponence is a type of *Verbose Exponence*, a term with Carroll uses to encapsulate both Multiple Exponence and Distributed Exponence, i.e. cases of morphological redundancy.

Distributed exponence is the co-occurrence of multiple formatives within a single word such that more than one formative is required to provide a fully specified reading of a feature or category (Caballero & Harris, 2012; Carroll et al., 2016; Harris, 2017).

The Ket argument marking system clearly exhibits this property, as outlined throughout this paper, and so does the Estonian data cited above. Carroll relies on his notion of *Informativeness*, as a more theoretically neutral alternative to exponence, defined informally in the following way:

What information does a language learner or hearer have about the (grammatical) meaning of a word given this formative? (Carroll, 2022, p. 6)

Under this approach, e.g. Estonian *u* and Ket *ba* serve the same function: they reduce the range of possible paradigm cells which the form could represent, and have in common the fact that they must be considered in conjunction with other markers in order to provide unambiguous information. Carroll's approach and the polyfunctionality analysis taken here for Ket seem to represent essentially notational variants of one another, although future research will show whether some crucial difference exists between the two approaches.

7 Discussion and conclusion

This paper has shown that Ket exhibits remarkably high complexity of exponence, owing to the abundance of polyfunctional argument markers. At the same time, such markers are organized into networks of implicative relations with one another, such that actual ambiguity about the function of any given marker is substantially lower than if the marker were considered in isolation.

At the core of these networks is the P8 set of markers, which frequently act as the starting point for chains of implications about other markers. Given that the P8 markers were the last component of the verb to grammaticalize, and are the first markers in the string (and therefore the first to be processed), this dynamic bears some resemblance to the notion of Reinforcement Multiple Exponence (Harris, 2017, pp. 61–64, 163–165; cf. also Caballero & Kapatsinski, 2015). This is where a second, often more frequent or transparent, exponent is added, due to the older exponent of the relevant feature becoming opaque or difficult to parse.⁵⁶ The difference is that rather than just reinforce one feature or set of features, the P8 markers often allow for whole cascades of implications, often as much through their differences with other markers as through their similarities, or through their absence as much as through their presence. Rather than simplify the remarkable complexity of exponence towards the right edge

⁵⁶Class v3, v4, vt3, vt4 verbs, which use the P8 markers as co-exponents of the subject alongside another marker set, would seem to be literally cases of Reinforcement Multiple Exponence. (Vajda, 2017a, pp. 7–9) explicitly connects the grammaticalization of the P8 set in Ket with the loss of the P1 subject markers in many verbs, which gives further support to this idea. On the other hand, see Vajda's review of Harris (2017), where he suggests that some cases of Multiple Exponence in Ket in fact result from semantic reinterpretation of something at the right edge of the verb, in cases where the P8 marker was already present (Vajda, 2018).

of the verb, Ket has instead added additional implicative structure, which allows for otherwise unfeasibly complex form-function mappings to be maintained.

The Ket data underscore the point made by Sims and Parker (2016) that the role and importance of implicative structure varies across languages with different morphological structures. Ket suggests 1) that implicative structure can play a role in resolving complex form-function mappings not in encoding, but also in decoding and 2) that implicative relations need not only hold between paradigmatically related words or families of constructions, but can hold also between markers.

In languages like Russian or English, both of these points might be true in some trivial sense as well. However, since complexity of exponence is comparatively very low, the number of morphs per word is small, and the syntactic structure provides ample cues for resolving what complexity does exist, the amount of *work* done by syntagmatic implicative structure for the listener is likely negligible. It is only Ket's combination of an abundance of polyfunctional markers, a large number of morphs per word, and exclusive head-marking of core syntactic arguments, that allows for syntagmatic implicative structure to do so much work.

Of course, high complexity of exponence, a high degree of synthesis, and predominant head-marking are all properties characteristic of *old polysynthetic languages* (Fortescue, 2013): those languages whose structures are the result of long periods of successive grammaticalizations and reanalysis. Hence, the particular set of historical contingencies observable in Ket might have played out in a similar way in other languages which have followed a similar historical trajectory. Only further research would say.

In terms of broader implications for an understanding of polyfunctionality and complexity of exponence more generally, Ket as a case study suggests a sort of LOW CONDITIONAL ENTROPY CONJECTURE for the listener. As a hypothesis, there is no *a priori* limit on how many functions a single marker may have across a language. It may have as many functions as history would have it, provided that the intended function can be determined in any particular instance, although how that result is achieved will vary from language to language.

Acknowledgements Enormous thanks to my teacher, Valentina Andreevna Romanenkova, for sharing her incredible language with me. Enormous thanks also to Farrell Ackerman, Olivier Bonami, Gabriela Caballero, Matt Carroll, Emily Clem, Nina Hagen Kaldhol, Elena Kryukova, Zoya Maksunova, Rob Malouf, Andrey Nefedov, Andrea Sims, Edward Vajda, Michelle Yuan, and the audiences at the 53rd Annual Meeting of the Societas Linguistica Europaea, the 5th American International Morphology Meeting, the 29th International Conference “Dulzon Readings”, and the Spring 2022 Seminar Series of the Linguistic Convergence Laboratory (Higher School of Economics, Moscow, Russia), who provided essential feedback or support at various stages of this project. Enormous thanks also to two anonymous reviewers and to the editor of Morphology for their many helpful comments. All errors are my own.

Declarations

Competing Interests The author declares that he has no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article

are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ackerman, F., & Bonami, O. (2017). Systemic polyfunctionality and morphology-syntax. In N. Gisborne & A. Hippisley (Eds.), *Defaults in morphological theory*. Oxford: Oxford University Press.
- Ackerman, F., Blevins, J. P., & Malouf, R. (2009). Parts and wholes: patterns of relatedness in complex morphological systems and why they matter. In J. Blevins & J. Blevins (Eds.), *Analogy in grammar: form and acquisition* (Vol. 54, p. 82). Oxford: Oxford University Press.
- Anderson, S. R. (2015). Dimensions of morphological complexity. In M. Baerman, D. Brown, & G. Corbett (Eds.), *Understanding and measuring morphological complexity* (pp. 11–26). Oxford: Oxford University Press.
- Arkadiev, P. & Gardani, F. (Eds.) (2020). *The complexities of morphology*. Oxford: Oxford University Press.
- Aronoff, M. (1994). *Morphology by itself: stems and inflectional classes*. Cambridge: MIT Press.
- Baayen, H. (2002). *Word frequency distributions* (Vol. 18). Berlin: Springer.
- Baerman, M. (2012). Paradigmatic chaos in Nuer. *Language*, 88, 467–494.
- Baerman, M., Brown, D., & Corbett, G. G. (Eds.) (2015). *Understanding and measuring morphological complexity*. New York: Oxford University Press.
- Baerman, M., Brown, D., & Corbett, G. G. (2017). *Morphological complexity*. Cambridge: Cambridge University Press.
- Belimov, E. I. (1991). *Ketskij sintaksis*. Novosibirsk: Nauka.
- Beniamine, S., Bonami, O., & McDonough, J. (2017). When segmentation helps: implicative structure and morph boundaries in the navajo verb. In *Proceedings of the first international symposium on morphology* (pp. 11–15).
- Bickel, B., & Nichols, J. (2007). Inflectional morphology. *Language Typology and Syntactic Description*, 3(2), 169–240.
- Blevins, J. P. (2016). *Word and paradigm morphology*. Oxford: Oxford University Press.
- Blevins, J. P., Milin, P., & Ramscar, M. (2017). The Zipfian paradigm cell filling problem. In *Perspectives on morphological organization: data and analyses* (Vol. 10, p. 141).
- Bonami, O., & Beniamine, S. (2016). Joint predictiveness in inflectional paradigms. *Word Structure*, 9(2), 156–182.
- Bonami, O., & Beniamine, S. (2021). *Leaving the stem by itself*. By Marcia Haag et al. (pp. 81–98). Amsterdam: John Benjamins.
- Bonami, O., & Strnadová, J. (2019). Paradigm structure and predictability in derivational morphology. *Morphology*, 29(2), 167–197.
- Butorin, S. S. (1995). *Opisanie morfoložičeskoj struktury finitnoj glagol'noj slovoformy ketskogo jazyka s ispol'zovanijem metodiki porjadkovogo členija: avtoreferat dissertacii*. Novosibirsk: Institut Filologii.
- Butorin, S. S. (2018). Glagol'naja derivatsija v ketskom jazyke: aktantnaja derivatsija i inkorporatsija. *Tomskij žurnal lingvističeskikh i antropologičeskikh issledovanij* (2), 9–18.
- Caballero, G., & Harris, A. C. (2012). *A working typology of multiple exponence*. *Current issues in morphological theory: (Ir) regularity, analogy and frequency* (pp. 163–188). Amsterdam: John Benjamins.
- Caballero, G., & Kapatsinski, V. (2015). Perceptual functionality of morphological redundancy in Choguita Rarámuri (Tarahumara). *Language, Cognition and Neuroscience*, 30(9), 1134–1143.
- Carroll, M. J. (2022). *Verbose exponence: integrating the typologies of multiple and distributed exponence*. *Morphology*, 32(1), 1–24.
- Carroll, M. J. (2016). *The Ngkolmpu language with special reference to distributed exponence*. PhD Dissertation, the Australian National University.
- Castrén, M. A. (1858). *Versuch einer ostjakischen Sprachlehre*. St. Petersburg: Buchdruckerei der Kaiserlichen akademie der wissenschaften.

- Crippen, J. (2012). The basics of Tlingit verbal structure. Unpublished manuscript. <http://www.drangle.com/~James/verbal-structure>.
- Dahl, Ö. (2004). *The growth and maintenance of linguistic complexity*. Amsterdam: Benjamins.
- Dahl, Ö. (2017). Polysynthesis and complexity. In M. Fortescue, M. Mithun, & N. Evans (Eds.), *The Oxford handbook of polysynthesis*. Oxford: Oxford University Press.
- Drossard, W. (2002). Ket as a polysynthetic language, with special reference to complex verbs. In *Problems of polysynthesis* (pp. 223–256). Berlin: Akademie Verlag.
- Dul'zon, A.P. (1972). *Skazki narodov sibirskogo severa* (Vol. 1). Izd-vo Tomskogo Universiteta.
- Finkel, R., & Stump, G. (2009). Principal parts and degrees of paradigmatic transparency. In J. P. Blevins & J. Blevins (Eds.), *Analogy in grammar: form and acquisition*. Oxford: Oxford University Press.
- Fortescue, M. (2013). Polysynthesis in the Arctic/Sub-Arctic. In *Language typology and historical contingency: in honor of Johanna Nichols* (pp. 241–264). Amsterdam: Benjamins.
- Georg, S. (2007). *A descriptive grammar of Ket (Yenisei-Ostyak): part 1: introduction, phonology and morphology*. Leiden: Brill.
- Gurevich, O. (2006). Constructional morphology: the Georgian version. PhD Dissertation. Stanford: Stanford University.
- Harris, A. C. (2017). *Multiple exponence*. Oxford: Oxford University Press.
- Hengeveld, K. (2011). Transparency in functional discourse grammar. *Linguistics in Amsterdam*, 4(2), 1–21.
- Hockett, C. F. (1947). Problems of morphemic analysis. *Language*, 23(4), 321–343.
- Kotorova, E., & Nefedov, A. (2006). Tipologicheskie karakteristiki ketskogo jazyka: vershinnoe ili zavisimostnoe markirovanie? *Voprosy Jazykoznanija* (5), 43–56.
- Kotorova, E., & Nefedov, A. (2015). *Comprehensive Ket dictionary*. München: Lincom Europa.
- Krejnovich, E. (1968). *Glagol ketskogo jazyka*. Leningrad: Nauka.
- Matthews, P. H. (1972). *Inflectional morphology: a theoretical study based on aspects of Latin verb conjugation* (Vol. 6). Cambridge: Cambridge University Press.
- Mattissen, J. (2006). The ontology and diachrony of polysynthesis. In D. Wunderlich (Ed.), *Advances in the theory of the lexicon* (pp. 287–354). Berlin: de Gruyter.
- McDonough, J. (2000). Athabaskan redux: against the position class as a morphological category. In W. U. Dressler, O. E. Pfeiffer, M. A. Pöchtrager, & J. R. Rennison (Eds.), *Morphological analysis in comparison*. Amsterdam: Benjamins.
- Mithun, M. (2011). Grammaticalization and explanation. In H. Diessel, H. Narrog, & B. Heine (Eds.), *The Oxford handbook of grammaticalization*. Oxford: Oxford University Press.
- Nefedov, A. (2015). Clause linkage in Ket. PhD Dissertation. University of Leiden. Utrecht: LOT.
- Nefedov, A., & Vajda, E. (2015). Grammatical sketch of Ket. In A. Nefedov & E. Kotorova (Eds.), *Comprehensive dictionary of Ket* (pp. 27–68). Munich: Lincom Europa.
- Reshetnikov, K. Y., & Starostin, G. (1995). Struktura ketskoy glagol'noy slovoformy. In *Lingvistika: Vol. 4. Ketskiy sbornik* (pp. 7–121). Moscow: Vostochnaja Literatura.
- Sims, A. D., & Parker, J. (2016). How inflection class systems work: on the informativity of implicative structure. *Word Structure*, 9(2), 215–239.
- Stump, G. (2014). Polyfunctionality and inflectional economy. In *Linguistic issues in language technology: Vol. 11. Theoretical and computational morphology: new trends and synergies*.
- Stump, G. (2015). *Inflectional paradigms: content and form at the syntax-morphology interface* (Vol. 149). Cambridge: Cambridge University Press.
- Trudgill, P. (2020). *Millennia of language change: sociolinguistic studies in deep historical linguistics*. Cambridge: Cambridge University Press.
- Vajda, E. (2010). A Siberian link with Na-Dene languages. In J. Kari & B. Potter (Eds.), *The Dene-Yeniseian connection, fairbanks, Alaskan native language center* (pp. 33–99).
- Vajda, E. J. (2001). The role of position class in Ket verb morphophonology. *Word*, 52(3), 369–436.
- Vajda, E. J. (2003). Ket verb structure in typological perspective. *STUF – Language Typology and Universals*, 56(1–2), 55–92.
- Vajda, E. J. (2004). *Ket*. Munich: Lincom Europa.
- Vajda, E. J. (2013). Metathesis and reanalysis in Ket. *Tomsk Journal of Linguistics and Anthropology*, 1(1), 14.
- Vajda, E. J. (2014). Yeniseian. In R. Lieber & P. Štekauer (Eds.), *The Oxford handbook of derivational morphology*. Oxford: Oxford University Press.
- Vajda, E. J. (2015). *Valency properties of the Ket verb clause. Valency classes in the world's languages* (pp. 630–668). Berlin: de Gruyter.

- Vajda, E. J. (2017a). Patterns of innovation and retention in templatic polysynthesis. In M. Fortescue, M. Mithun, & N. Evans (Eds.), *The Oxford handbook of polysynthesis*. Oxford: Oxford University Press.
- Vajda, E. J. (2017b). Polysynthesis in Ket. In M. Fortescue, M. Mithun, & N. Evans (Eds.), *The Oxford handbook of polysynthesis*. Oxford: Oxford University Press.
- Vajda, E. J. (2018). A review of Multiple Exponence by Alice C. Harris, Oxford: Oxford University Press, 2017. *Word*, 64(1), 43–48.
- Vajda, E. J., & Zinn, M. (2004). *Morfologičeskij slovar' ketskogo glagola (na osnove južno-ketskogo dialekta)*. Tomsk: Tomsk State Pedagogical University.
- Vall, M. N., & Kanakin, I. A. (1988). *Kategorii glagola v ketskom jazyke*. Novosibirsk: Nauka, Sibirskoe otd-nie.
- Werner, G. (1990). *Kotskij jazyk*. Rostov-na-Donu: Izdatel'stvo rostovskogo universiteta.
- Werner, H. (1997). *Die ketische Sprache* (Vol. 3). Wiesbaden: Harrassowitz.
- Wurzel, W. U. (1984). *Flexionsmorphologie und Natürlichkeit (studia grammatica xxi)*. Berlin: Akademie.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.