Is Typology Possible?

Luci Keen *

1 Introduction

Typology has long been a cornerstone in linguistic research, exploring the similarities and differences between the grammatical structures of the world's languages through systematic classification. The very concept of typology, however, relies on the assumption that languages are comparable; that a single classification system can be meaningfully applied across the board. In order to dig deep into the complications cross-linguistic compatibility faces, this essay will follow the typology of a single linguistic feature, the order of verbal A and P person markers, analysing the hiccups, pitfalls, and roadblocks we run into along the way.

The World Atlas of Language Structures (WALS) classifies 193 languages as having verbal person marking of both the A and the P arguments (Siewierska, 2013b), with Feature 104A exploring the order these markers appear in cross-linguistically (Siewierska, 2013a). There are a number of reasons why this question is of interest to linguists. For one, it relates to speculation of statistical universals in affix order such as that explored in Harmon (1994)'s research of tense and aspect marker order, Hawkins and Gilligan (1988) research into suffixation, and Trommer (2002)'s exploration of person and number marker order. Furthermore, the question is relevant to discussion around the morphology-syntax interface (Mykhaylyk, 2010) - for example, exploring if morpheme order has syntactic origins as Givón (1971)'s proposed 'Diachronic Universal' model claims. With regard to the current study, however, the feature is also of interest for the range of subfields it is influenced by and relevant to. Morphology requires phonological understanding (Bertinetto and Jetchev, 2005, p. 16; Saldanya and Vallès, 2005, p. 54; Viaplana, 2005, p. 172), whilst person marking is influenced by syntactic principles and affix order, itself subject to semantic restrictions (Ryan, 2010, p. 758; Korotkova and Lander, 2010, p. 345). As such, the order of A and P verbal person marking affixes in relation to the root is a perfect feature for exploring difficulties in linguistic typology whilst also being relevant to the linguistic theories which typology explores.

2 Defining 'A and P Verbal Person Marking'

The first step towards systematically classifying the world's languages is developing criteria to identify languages with A and P verbal person marking affixes. Languages must have person markers for both the agent of a transitive verb and the patient. Those person markers must

*Originally submitted in 2019 as part for the subject Exploring Linguistic Diversity (LING30001)

be affixes, and those affixes must be attached to verbs. For example, Tawala (Ezard, 1997) and Macushi (Abbott, 1991) provide prototypical demonstrations of this feature, with Tawala exemplifying an AVP order, whilst Macushi demonstrates a PVA order:

(1) Tawala (Ezard, 1997, p. 99)

Kedewa *Kamkam* *i-uni-hi*
Dog chicken 3sg.A-kill-3pl.P
'A dog killed the chickens.'

(2) Macushi (Abbott, 1991, p. 24)

i-koneka-'pî-i-ya
3sg.P-make-pst-3sg.A-erg
'He made it'

Apply this criterion across the breadth of the world's documented languages however, and a number of complications emerge.

Firstly, it is worth noting that the concept of 'language' itself is blurred. Dialects are generally said to be mutually intelligible, whilst different languages are not (Nomoto & Long, 1999, p. 298), however this theoretical distinction can be elusive in practise. Political boundaries are particularly salient in overruling linguistic distinctions. Bugarski (2012)'s discussion of Yugoslavian language classifications revealing how linguistic evidence is not sufficiently definitive to withstand this pressure. Furthermore, the lines of dialect and language have a tendency to be drawn differently by different linguists. Gunyan, for example is considered a dialect of Bidyara by Breen (1981), whilst WALS classifies it as a language in its own right (Dryer, 2013). Yet it is not only linguists who disagree over language/dialect boundaries. A 2010 Scottish government study revealing 64% of the population do not view Scots, an official language of Scotland, as a language (Sebba, 2018, p. 342). Non-Scots speakers were the least likely to consider it as such. There is also evidence that perception of dialectal differences differs between individuals. This is unsurprising given that language variation is a continuum, and where your own variant sits on that continuum will determine what variants are most similar to your own. However, this is not the only factor. Research into children's perception of dialects reveals five-year olds show an impaired ability compared to adult speakers to perceive regional dialects (Wagner et al., 2013, p. 1081). This is of interest as it suggests that dialectal differences have a learnt component, further complicating our understanding of the continuum between language and dialect. Typology compares languages, but 'language' itself is a subjective classification.

'Subject' is an equally tricky concept for which to provide a definitive definition, classification of A and P relying on identifying the subject of transitive constructions. There are many approaches to determining subject - Chomsky (1965), for example, labels the defining feature of grammatical subject its 'immediate dominance'. Paul (2010) argues against this criterion. Keenan (1979) also comments that it is 'problematic' for SOV, VSO and variable word order languages. Instead, Keenan promotes three other criteria: autonomy principles, case marking,

and semantic distinctions. Each, however, contains their own shortfalls. Paul's discussion of Keenan's criteria stresses that it is 'crucial' to understand there is no single measure that can be applied to all languages, autonomous features such as acting as antecedent to reflexives only a 'tendency' of subjects. Case marking also illustrates a difficulty with subject identification; it requires knowledge of how a language marks subject in the first place. Meanwhile, passive constructions highlight just one flaw with semantic distinctions, showing how patients can be subjects too. Indeed, pro-drop languages prevent ellipsis of subject being a universally meaningful distinction between passive voice in a nominative-accusative language, and active voice in an ergative (de Zarobe, 1998). Rather subject itself is a concept that can only be completely understood within the context of a specific language, reducing the comparability of A and P person marker order cross-linguistically.

Whilst the concept of subject and verb provide their own complications, such as whether auxiliary verbs will be considered, by far the greatest complication in defining the feature itself is what can be classified as an affix. Indeed, the WALS entry for Verbal Person Marking bows out of this deliberation, including both affixal and clitic marking (Siewierska, 2013a). When it comes to morpheme order, however, clitics are influenced by different factors than affixes, their preference for external positions potentially skewing the data (Terzi, 1999, p. 69; Goria, 2000, p. 144). This raises the issue of actually identifying clitics, the term having become something of a catch-all for those language segments that don't quite behave like affixes, nor demonstrate the properties of a word. Russell, 2006, p. 342 illustrates the potential this approach opens to inconsistent classification, discussing how focusing on distinguishing clitics from words yields a different definition to concentrating on their distinction from affixes. It is Mansfield's discussion of clitics that captures the difficulties with this category most clearly, his research into Murrinh-patha morphology highlighting how clitics are haunted by multiple defining features. Clitics are syntactically unselective, modify entire phrases, and lack the independent phonology of a prosodic word whilst remaining prosodically distinguishable from their host (Mansfield, 2019, pp. 167–168). Whilst in some languages these features align, Mansfield, 2019, p. 169 outlines the varied analyses of Murrinh-patha's grammatical morphemes given by focusing on each criterion in turn, coming to no clear consensus on which distinction is central. Even English has segments that rebel against easy classification. Take, for instance, the possessive marker "s'. It demonstrates syntactic independency, whilst requiring a host to be utterable. Unlike clitics however this host effects the phonological realisation of "s' - compare 'car's', /kɐːz/, to 'bus's', /bɐsəs/. Indeed, in a thorough discussion of criteria frequently applied to distinguish between affixes and word, (Haspelmath, 2011, p. 37) makes a compelling conclusion that such notions of 'affix' 'clitic' and 'word' can only be defined language specifically, vastly reducing the ability to objectively compare languages.

Finally, there is a question regarding what should be the threshold of person marking for a language to be considered as having this feature. Will only languages with compulsory A and P marking be considered, or is any degree of person marking enough? Here, the answer seems obvious considering the original interest in exploring potential universals and correlations in affix order. All that is required is enough data to determine what order A and P person markers take when they are present. This is of interest in a discussion of the nature of typology, as it demonstrates that judgement calls around parameters of specific typologies are driven

by more than just "charting linguistic diversity" (Plank, 2016, p. 2). Rather, these decisions are shaped by comparative aims. Typology is inherently targeted towards finding correlations (Comrie, 1988, p. 145), and increasing the "predictive power" (Mithun, 2016, p. 467) of the linguistic community's grammatical knowledge. That language-specific judgements are required to determine subject person markers and carve the continuum that stretches between affix and word, undermines this comparative goal.

3 Subcategorising Verbal Person Marking Affix Order

Having established that determining which languages contain A and P verbal person marking affixes can be imprecise, there is then the task of deciding how to subcategorise these languages. On the surface this seems simple enough: examine each relevant language and state the order of A person marker, P person marker, and the verbal root. As such it would be expected the scope of possible values are APV, AVP, VAP, VPA, PAV, and PVA. However, in studying the actual diversity of languages and morpholzzogical phenomenon, this reveals itself to be a large oversimplification. These categories only capture the potential for prefix and suffix person markers, barely scratching the surface of observed affix placements and patterns. Whilst affixes have traditionally been viewed as requiring a set-position, emerging evidence from languages such as Chintang (Bickel et al., 2007), Tagalog (Ryan, 2010), and Murrinh-patha (Mansfield, 2015) promotes the phenomenon of variable order. Chintang is particularly relevant, the below examples illustrating how the language demonstrates both APV order and PAV order, with no semantic difference:

(3) Chintang (Bickel et al., 2007, p. 44)

a. *u-kha-ma-cop-yokt-e.*
3NS.A-1NS.P-NEG-see-NEG-PST
'They didn't see us'

b. *kha-u-ma-cop-yokt-e.*
1NS.P-3NS.A-NEG-see-NEG-PST
'They didn't see us'

Even when languages do have set affix placements however, these do not always fit snugly into categories such as AVP and PAV. Consider the below example of infixation:

(4) Lakhota (Albright, 2000)

máni 'he walks'	ma-wá-ni 'I walk'
aphé 'he hits'	a-wá-phe 'I hit'
hoxpé 'he coughs'	ho-wá-xpe 'I cough'

Without the addition of more values, Lakhota would not be classifiable in this typology, however simply adding an infixation category is insufficient. Not only does this group all infixing languages together regardless of what person marker is being infixed, it also fails to

account for the ”analytical problem” (Albright, 2000, p. 4) of variable realisation of infixes. Infixes in Ulwa, for example, are placed after the first foot, resulting in infixes being realised as suffixes when attaching to roots with only a single foot.

Circumfixation offers up similar complications, including ambiguity around what classifies as a circumfix. Take the below paradigm of Georgian A person markers:

(5) Georgian (Hewitt, 1995, p. 128)
Affixal Agreement - Set A

	Singular	Plural
1st person	v-	V- -t
2nd person	∅(/x)-	∅(/x)- -t
3rd person	-s/a/o	-(a/e)n/es/nen

At first glance it would appear the paradigm includes circumfixes, however alternatively the ’-t’ affix could be analysed as a suffix corresponding to number, plural person markers formed through a process of parasynthesis. These two processes are usually distinguished by considering whether one of the morphemes can be used independently of the other (Klégr, 2018, p. 54). However, in the case of a language with obligatory person marking, this criterion can become meaningless. In the example of Georgian the lack of a ’-t’ component of the 3rd person plural marker leans towards a circumfix analysis.

The Georgian paradigm also showcases another complication. Third person singular A person agreement is a suffix, whilst first and second person singular is a prefix. This is a very different kind of variation to that seen in Chintang, and to group it along with no dominant order languages, or to ignore the variation and simply classify it by the most common order would vastly simplify the reality of affix placement.

Less transparent person markers must also be taken into account. One example of this phenomenon where a single segment does not translate to a single meaning is portmanteau morphology, such as in the examples from Sienna Populuca below:

(6) Sierra Popoluca (Elson, 1960, p. 211)

i-koʔc-pa
3A.3P-hit-INC
’He is hitting him’

a-koʔc-pa
3A.1P-hit-INC
’he is hitting me’

aŋ-koʔc-pa
1A.3P-hit-INC
’I am hitting him’

mi-koʔc-pa
3A.2P-hit-INC
’he is hitting you’

iŋ-koʔc-pa
2A.3P-hit-INC
’You are hitting him’

Capturing the order of person marking affixes in Sierra Popoluca would therefore require a category for A+P V that indicates whilst A and P precede the verb, it is not possible to break the order down further.

Finally, there is the issue of systematic variation caused by principles of grammatical relations. One such complicating phenomenon is split ergativity, where the animacy of the subject determines whether A and S, or P and S are marked the same (Tollan, 2014, p. 417). For languages where affix placement marks grammatical relations, this can result in different affix order of A and P depending on the nature of the subject. The category of 'split ergativity' would still not be sufficient to capture this idea however, languages such as Yimas demonstrating the relationship between affix order and grammatical relations can be extremely complicated, and highly language specific:

(7) Yimas (Foley, 1991, p. 172)

a. *pu-ka-tay*
3plO-1sgA-see
'I saw them'

b. *Mpu-ŋa-tay*
3plA-1sgO-see
'they saw me'

At first glance the examples above would appear to be explainable through the concept of split ergativity, 1st and 2nd person pronouns higher on the animacy hierarchy and therefore being treated with Nominative-Accusative alignment (Deal, 2015, p. 534). Examining Yimas's paradigm however, and the story becomes a lot more complex:

(8) Yimas (Foley, 1991, p. 170)

	A	O	S
1DL	ŋkra-	ŋkra-	kapa-
1PL	kay-	kra-	ipa-
1SG	ka-	ŋa-	ama-
2DL	ŋkran-	kul-	kapwa-
2PL	nan-	kul-	ipwa-
2SG	n-	nan-	ma-
3SG	n-	na-	na-
3PL	mpu-	pu-	pu-
3DL	mpi-	impa-	impa-

Rather, Foley describes Yimas's person marker order as being governed by two principles, first a 'person hierarchy' where 1st outranks 2nd, which in turn outranks 3rd. Secondly, there is a 'role hierarchy' where for 1st and 2nd person, O outranks A, but for third person A outranks O (Foley, 1991, p. 173). The highest ranked argument occupies the position closest the verb. As complicated as all this sounds, what it demonstrates most clearly is for all these examples, a value which represents this specific circumstance is not practical, whilst a definitive classification of affix order without such a value is simply not possible.

At this point it should be becoming clear that the degree of diversity within affix ordering is at odds with the typological aim of illuminating patterns. The number of values required, and specificity of those values to each language, would create a typology far more confusing than insightful, and as such, choices have to be made about what values to focus on, and what to group together in an 'other' category. As this feature is about affix order, the grouping of languages that use similar processes, such as infixation, but with different orders effectively excludes this data from the analysis, editing the true diversity of the world's languages.

4 Classifying Languages

Finally, there is the process of actually assigning the appropriate value to each language. Regardless of the values being used, there remains the complication of non-systematic variation itself, and how to distinguish between language inherent variation, social variation, and individual variation. The WALS entry for Order of Subject, Object and Verb discusses the concept of 'dominant order', approaching language variation by looking at what is statistically most common (Dryer, 2013). However, this opens the potential for grammatically acceptable variation to be obscured, whilst classification of no dominant order could be effected by social variation, such as that associated with age and indicative of diachronic change (Traugott & Smith, 1993, p. 272). That is to say, this method does not distinguish between variation between variants, and variation within variants. Further, the use of percentages and majorities relies on having sufficient quantities of data from a range of speakers, some languages such as Bidyara, lacking this wealth of data (Breen, 1981), risking idiolects and sociolects being recorded as representative of the language. Assigning values accurately requires diverse and representative data, with the use of one-size fits all thresholds creating potential for yet more inaccuracies and biasing against less common phenomenon. Relying instead on language-specific judgements however, once again reduces the cross-linguistic compatibility of these classifications.

5 Conclusion

Language is complex, and cross-linguistic analysis faces many challenges. Typology demands continuums are carved into categories, definitions stretched and pulled to encompass the breadth of linguistic diversity and most damaging of all, variation smoothed over in the path of definitive rulings. Ultimately, typology is a compromise between being accurate and being informative, every step of the process steeped in judgement calls, imprecision and language-specific analyses. This is not to say that typological research is worthless, however, or incapable of providing insight. Rather, in order to make best use of this research tool it is important to understand its shortfalls and acknowledge that truly objective cross-linguistic categorisation is an illusion.

References

Abbott, M. (1991). Macushi. In D. C. Derbyshire & G. K. Pullum (Eds.), *Handbook of amazonian languages* (pp. 23–160). Mouton de Gruytero.

Albright, A. (2000). The productivity of infixation in lakhota. *UCLA Working Papers in Linguistics*, *0*, 1–18.

Bertinetto, P. M., & Jetchev, G. (2005). Lexical access in bulgarian: Nouns and adjectives with and without floating vowels. *Catalan Journal of Linguistics*, *4*(1), 171. https://doi.org/10.5565/rev/catjl.118

Bickel, B., Banjade, G., Gaenszle, M., Lieven, E., Paudyal, N. P., Rai, I. P., Rai, M., Rai, N. K., & Stoll, S. (2007). Free prefix ordering in chintang. *Language*, *83*(1), 43–73.

Breen, J. (1981). Margany and gunya. In R. Dixon & B. J. Blake (Eds.), *Handbook of australian languages* (pp. 274–393). John Benjamins Publishing Company. https://doi.org/10.1075/z.hal2.08bre

Bugarski, R. (2012). Language, identity and borders in the former serbo-croatian area. *Journal of Multilingual and Multicultural Development*, *33*(3), 219–235. https://doi.org/10.1080/01434632.2012.663376

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.

Comrie, B. (1988). Linguistic typology. *Annual Review of Anthropology*, *17*(1), 145–159. https://doi.org/10.1146/annurev.an.17.100188.001045

Deal, A. R. (2015). Person-based split ergativity in nez perce is syntactic. *Journal of Linguistics*, *52*(3), 533–564. https://doi.org/10.1017/s0022226715000031

de Zarobe, Y. R. (1998). El parámetro pro-drop y la adquisición del inglés como segunda lengua. *ITL - International Journal of Applied Linguistics*, *119-120*, 49–63. https://doi.org/10.1075/itl.119-120.04dez

Dryer, M. S. (2013). Order of subject, object and verb. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology. https://wals.info/chapter/81

Elson, B. (1960). Sierra popoluca morphology. *International Journal of American Linguistics*, *26*(3), 206–223. https://doi.org/10.1086/464578

Ezard, B. (1997). *A grammar of tawala : An austronesian language of the milne bay area, papua new guinea*. Dept. of Linguistics, Research School of Pacific Studies, The Australian National University.

Foley, W. A. (1991). *The yimas language of new guinea*. Stanford University Press. https://books.google.com.au/books?id=5TCsAAAAIAAJ

Givón, T. (1971). Historical syntax & synchronic morphology: An archeologist's field trip.

Goria, C. (2000). *Syntax and morphology of subject clitics in piedmontese: Analysis based on the minimalist program and optimality theory* (Doctoral dissertation). https://search.proquest.com/docview/2178829436

Harmon, L. L. (1994). *Verb suffixing and language universals* (Doctoral dissertation) [Copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Last updated - 2020-02-21]. https://search.proquest.com/docview/304130476

Haspelmath, M. (2011). The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica*, *45*(1). https://doi.org/10.1515/flin.2011.002

Hawkins, J. A., & Gilligan, G. (1988). Prefixing and suffixing universals in relation to basic word order. *Lingua*, *74*(2-3), 219–259. https://doi.org/10.1016/0024-3841(88)90060-5

Hewitt, B. (1995). *Georgian: A structural reference grammar*. John Benjamins Publishing Company. https://books.google.com.au/books?id=MUBtUcqOF-AC

Keenan, E. (1979). Towards a universal definition of "subject". In C. N. Li (Ed.), *Subject and topic* (pp. 303–333). Academic press. https://linguistics.ucla.edu/people/keenan/papers/Keenan-universal-subject.pdf

Klégr, A. (2018). Language is embiggened by words that don't exist: The case of a circumfix. *Linguistica Pragensia*, *1*, 53–70.

Korotkova, N., & Lander, Y. (2010). Deriving affix ordering in polysynthesis: Evidence from adyghe. *Morphology*, *20*(2), 299–319. https://doi.org/10.1007/s11525-010-9185-y

Mansfield, J. (2015). Morphotactic variation, prosodic domains and the changing structure of the murrinhpatha verb. *Asia-Pacific Language Variation*, *1*(2), 163–189. https://doi.org/10.1075/aplv.1.2.03man

Mansfield, J. (2019). *Murrinhpatha morphology and phonology*. De Gruyter. https://doi.org/10.1515/9781501503306

Mithun, M. (2016). Typology, documentation, description, and typology. *Linguistic Typology*, *20*(3). https://doi.org/10.1515/lingty-2016-0019

Mykhaylyk, R. (2010). Diachronic universals and morpheme order in the ukrainian synthetic imperfective future. *Morphology*, *20*(2), 359–380. https://doi.org/10.1007/s11525-010-9182-1

Nomoto, K., & Long, D. (1999). Consciousness of linguistic boundaries and actual linguistic boundaries. *Handbook of perceptual dialectology* (pp. 63–69). John Benjamins Publishing Company. https://doi.org/10.1075/z.hpd1.12nom

Paul, I. (2010). Subjects: Grammatical relations, grammatical functions and functional categories. *Language and Linguistics Compass*, *4*(9), 890–902. https://doi.org/10.1111/j.1749-818x.2010.00234.x

Plank, F. (2016). Call for contributions to this collection on the relevance of typology.

Russell, M. A. (2006). *The syntax and placement of wolof clitics* (Doctoral dissertation) [Copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Last updated - 2020-02-21]. https://search.proquest.com/docview/305308807

Ryan, K. M. (2010). Variable affix order: Grammar and learning. *Language*, *86*(4), 758–791. https://doi.org/10.1353/lan.2010.0032

Sebba, M. (2018). Named into being? language questions and the politics of scots in the 2011 census in scotland. *Language Policy*, *18*(3), 339–362. https://doi.org/10.1007/s10993-018-9488-0

Siewierska, A. (2013a). Order of person markers on the verb. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology. https://wals.info/chapter/104

Siewierska, A. (2013b). Verbal person marking. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology. https://wals.info/chapter/102

Terzi, A. (1999). Clitic combinations, their hosts and their ordering. *Natural Language & Linguistic Theory*, *17*(1), 85–121. https://doi.org/10.1023/a:1006075626603

Tollan, R. (2014). Review of aspects of split ergativity, by jessica coon. *The Canadian Journal of Linguistics / La revue canadienne de linguistique*, *59*(3), 415–418. https://www.muse.jhu.edu/article/565773

Traugott, E. C., & Smith, H. (1993). Arguments from language change. *Journal of Linguistics*, *29*(2), 431–447. https://doi.org/10.1017/s0022226700000384

Trommer, J. (2002). The interaction of morphology and syntax in affix order. *Yearbook of morphology 2002* (pp. 283–324). Springer Netherlands. https://doi.org/10.1007/0-306-48223-1_9

Wagner, L., Clopper, C. G., & Pate, J. K. (2013). Children's perception of dialect variation. *Journal of Child Language*, *41*(5), 1062–1084. https://doi.org/10.1017/s0305000913000330