



Universiteit
Leiden
The Netherlands

A functional approach to differential indexing: combining perspectives from typology and corpus linguistics

Just, E.C.

Citation

Just, E. C. (2022, April 20). *A functional approach to differential indexing: combining perspectives from typology and corpus linguistics*. LOT dissertation series. LOT, Amsterdam. Retrieved from <https://hdl.handle.net/1887/3283627>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3283627>

Note: To cite this publication please use the final published version (if applicable).

Chapter 6

A structural and functional comparison of differential A and P indexing¹

Abstract

Indexing P arguments on bivalent predicates is often considered more restricted and less often obligatory than A indexing. However, differential A indexing, i.e. the absence vs. the presence of an index referring to the A argument role, is not uncommon either: usually present A indexes can be omitted in particular discourse settings. However, differential A indexing has been a Cinderella subject in the typological study of differential marking, as opposed to differential P indexing or differential A flagging. This paper scrutinizes various cases of both differential A and P indexing and examines structural and functional differences and similarities. It will be shown that exploring differential indexing helps to understand how indexing in general is linked to referential prominence which surfaces as factors such as identifiability, animacy or topicality. Cases where indexing is particularly sensitive to referential prominence, and where it thus is employed only if the referent fulfills certain criteria, bring out the fact that A and P indexing have a common purpose, namely tracking referents through discourse. In this context, the paper also points out that differential A indexing presents an exception from generalizations concerning the amount of material in coding asymmetries.

¹This chapter is submitted as: Just, Erika. *A structural and functional comparison of differential A and P indexing*.

6.1 Introduction

This paper provides an overview of differential indexing of both A and P referents, suggesting that the investigation of differential indexing (i.e. the variable occurrence of an index for reasons other than the referent's argument role) helps improving our understanding of the functionality of indexing in general, and why it is different from flagging. Whereas the latter can be considered as a strategy of individual role assignment, indexing is a means of reference tracking, which becomes evident through the investigation of differential indexing systems. This has been comprehensively shown for P arguments (e.g. Croft 1988, Iemmolo 2011), but A arguments have been paid less attention to in the cross-linguistic study of differential indexing. Showcasing the parallel behavior of differential A and P indexing, the paper aims at highlighting how indexing, irrespective of argument role, serves to keep track of prominent referents in discourse.

Indexes have been defined as bound markers (commonly on the predicate) expressing argument features (Haspelmath 2013). Indexing, contrary to agreement, does not presuppose any syntactic relationship between the marker and the referential noun phrase, nor whether the latter is obligatorily expressed. In other words, how a language handles overt NP referents on the one hand, and whether it indexes them on the other hand, should be considered as logically independent features (cf. Haig & Forker 2018). Languages show a lot of diversity in how and how often referents are expressed; just as there are languages like German where (S and A) referents are usually expressed by an overt NP and are also obligatorily indexed, there are also a number of languages where there is no indexing, but zero-anaphora is nevertheless an option, i.e. there can be no overt reference of an argument at all. There is no one-to-one relationship between indexing and the omission of lexical or pronominal referents (see Gilligan 1987). Between those extremes, there are various possibilities how a referent can be indicated. This diversity, coupled with the traditional understanding of agreement in some Indo-European languages (cf. Haspelmath 2013: 209) has led to the agreement vs. pronoun debate: the attempt to classify a bound person marker as either being redundant feature matching and thus coding grammatical relations, or as being the one true instantiation of a referent.

In the context of this debate, often also the morphological status of the index as a clitic or an affix comes up. It is considered irrelevant here, as the latter is often unjustifiably equated with obligatoriness of marking (cf. Haig & Forker 2018: 720), though there are clear examples of syntactically obligatory clitics, as well as of affixal indexes which are syntactically optional. An example for obligatory enclitic person marking comes from Kharia, where the subject index is enclitic to the verb in affirmative clauses, as in (44a), whereas it attaches to the negative particle in negated clauses, as in (44b) (Peterson 2011: 58):

(44) Kharia (Munda, Peterson 2011: 58)

- a. *kayom = ta = ɲ*
 speak = MID.PRS = 1S
 ‘I speak.’
- b. *um = ɲɲ kayom = ta*
 NEG = 1S speak = MID.PRS
 ‘I do not speak.’

Languages where a syntactically optional index is analyzed as an affix are, for example, Juang (also Munda, Patnaik 2008), or Uralic languages like Hungarian (Coppock & Wechsler 2012), Northern Ostyak (Nikolaeva 1999: 64–76) or Eastern Mansi (Virtanen 2014). In Juang, object indexing by means of a suffix seems to have been very productive, but has been declining in the course of the last decades (Anderson 2007: 83, Patnaik 2008: 529–530). In some Uralic languages, there are two paradigms for the suffixal indexes for an A referent, one of which also indexes the number of the P referent, depending on the information-structural status of the latter. In Eastern Mansi, for instance, the so called subjective paradigm (as used in example (45a)) indexes only person and number of the A referent, whereas the objective paradigm (used in example (45b)) also indexes the number of the P argument and is used when the P referent is the secondary topic of a proposition (Virtanen 2014: 404).

(45) Eastern Mansi (Uralic, Western Siberia, Virtanen 2015: 28–30)

- a. *äj-nø = tee-nø wöär-s-øt*
 drink-GER = eat-GER make-PST-3PL
 ‘They made something to eat and drink.’

- b. *õõw-øm öät kont-iitø*
door-ACC NEG find-3SG > SG
'He does not find the door.'

That a clitic-affix-distinction is probably not expedient in the discussion of the use of bound person marking is also reflected in cases where even for one and the same language, scholars do not agree on how to categorize bound person forms: Osada (2008) and Anderson (2007) call the same Mundari set of S/A indexes “suffixes” and “clitics” respectively. The same goes, for instance, for Siwi object indexes (Ouali 2011 vs. Souag 2014.) Settling on this morphologically and syntactically unconstrained definition of indexing opens up quite a range of phenomena which can be referred to as *differential* indexing. Differential indexing falls in the category of differential marking as defined by Witzlack-Makarevich & Seržant (2018), which includes both flagging and indexing. It refers to situations where an argument role can be coded in different ways, depending on factors other than the argument role itself. These factors can either be inherent to the referent (like animacy, for instance) or non-inherent (like information-structural status).

The macroroles (or generalized semantic roles) which are of interest in this paper are A and P as the more agent-like and the less agent-like argument of a two-place predicate, giving priority to the semantic relationship between the predicate and its argument over their structural relationship (Bickel & Nichols 2009, Bickel 2011, Witzlack-Makarevich 2019). This study only encompasses transitive predicates, and therefore only A and P arguments, leaving out S arguments of intransitive clauses, as well as T and G in ditransitive predicates. Although there is a strong cross-linguistic tendency towards A and S to align with regard to indexing (Bickel et al. 2013), in many languages S can also align with P, and obligatory S indexing might affect the obligatoriness of P indexing, as brought forward by Haig (2018: 788). Therefore, I consider the investigation of differential indexing and alignment worth to be followed up in its own right.

For the present analysis, I want to exclude clausal properties which lead to differences in indexing, like TAM distinctions, clause type, or polarity, and bring into focus differences in indexing due to referential features of the argument, as well as discourse-structural conditions on the whole clause. Therefore, example (44) from Kharia does not fall under my definition of differential indexing, as the position of the index is absolutely predictable based on po-

larity.² Example (45) constitutes my first example for differential indexing, as here the choice of indexing P is controlled by the discourse status of the referent.

In this context, I also have to briefly dwell upon the notions of *topic* and *focus*. Although there is a general consensus about the pragmatic effects associated with each of these categories, the meanings conveyed by different constructions in different languages ascribed to focality or topicality are so manifold that these information-structural categories should actually be seen rather as interpretive effect of certain constructions, and not as being at their core (see Matic & Wedgwood 2013 on focus). Additionally, speakers of languages that have comparable constructions to map information structure, such as clefts or left dislocation, might still use different structures under identical discourse conditions, so there is no one-on-one mapping of information-structural categories and particular constructions (Skopeteas & Fanselow 2010, Ozerov 2018). Differential indexing can also be considered one such construction: as outlined in Section 6.2, the presence of a P index is often connected to the topicality or topicworthiness³ of the respective referent, whereas the absence of an otherwise present A index has been attributed to the referent being in some kind of focus construction. However, the pragmatic and referential features that indexing is sensitive to have to be considered language specific, even if the differences may seem subtle.

The paper proceeds as follows: In the following Sections 6.2 and 6.3, I will provide examples of differential indexing from various languages, starting with differential indexing of the P argument, as this is probably the more familiar phenomenon in the context of differential marking, and then proceed to the A argument. It will be shown that following the definitions just provided, differential A indexing is not that uncommon and in principle exhibits the same general pattern as differential P indexing. In line with the findings from these two sections, Section 6.4, deals with differential indexing and referential prominence (Haspelmath 2021b), addressing both the role of the referents individually as well as that of co-arguments. Section 6.5 elaborates

²Admittedly, demarcating polarity from information structure can be a balancing act. In fact, languages where the position of a syntactically mobile index is determined by discourse effects also very often display an interplay between the placement of an index and polarity (see Cysouw 2003)

³Topicworthy referents display semantic features associated with topicality (such as being definite or high on the animacy scale) without necessarily being a topic.

on the functions of indexing in general thus revealed, highlighting the need to consider indexing independently from the overtness of an NP referring to the same referent, and arguing to functionally demarcate indexing more from flagging. I will conclude in Section 6.6, by reflecting on how (or rather if) the thoughts brought forward in the paper can be brought in line with the notion of obligatoriness of marking.

6.2 Differential P indexing

Differential P indexing, in whichever guise it comes along, has been given a lot of attention in the studies on individual languages, language families, and also cross-linguistically. The phenomenon can be encountered, for instance, as “clitic doubling” (see e.g. Aoun 1999 on Arabic dialects), “object reduplication” (see e.g. Friedman 2008 on the languages of the Balkans), or “optional agreement” (see e.g. Muxí 1996 on Catalan).

Many instances of differential indexing have typically involved the presence of an overt lexical NP. It is very often associated with animacy, humanness, specificity, definiteness, or topicality of P. As the P argument is generally not associated with such factors, but with new or contrastive information, a P referent which behaves rather atypically seems to call for a distinct marking pattern like differential flagging or differential indexing (Iemmolo 2011, Dalrymple & Nikolaeva 2011). However, although topicality is often the common ground on which differential P flagging as well as differential P indexing operate, the two have to be distinguished with regard to their function. As the main functions of flagging are to reflect referent features and to discriminate the different actants of an action (Bakker & Siewierska 2009), differential P flagging is often associated with effects of discourse discontinuity, such as topic-shift or topic promotion (Iemmolo 2011: 52). The primary function of indexing, on the other hand, is reference tracking (Lehmann 1982, Givón 1983, Siewierska 1997). Consequently, (differential) P indexing is also associated with tracking an (unexpectedly) topical P referent (Croft 1988, Iemmolo 2011).

However, cross-linguistically, the interpretive effects of differential indexing are very idiosyncratic, which makes it impossible to find a unitary explanation (Kallulli & Tasmowski 2008a: 10). This diversity is connected to

the fact that language internally, there is also often variation, even if a feature has been identified to be the triggering factor for the presence of a P index. For example, in Macedonian, P indexing is clearly associated with definiteness and specificity. If a referent is either inherently definite (like proper nouns) or marked by a definite article, it has to be indexed (Tomić 2008: 70). This is shown in example (46), where P ('the movie director') has to be indexed, as it is marked by a definite article, irrespective of an interpretation as being specific or non-specific.

(46) Macedonian (Balto-Slavic, Tomić 2008: 70)

Jana **(go) = bara* *režiser-ot*
 Jana 3SG.M.ACC = look.for.3SG movie.director-DEF.M

'Jana is looking for the movie-director (namely for X, who happens to be the movie-director).' or: 'Jana is looking for the movie-director (whoever that may be).'

However, depending on the context, also non-definite referents can be indexed if they are specific. But unlike definiteness, specificity does not force indexing, i.e. there can be specific referents that are not indexed. Also, based on acceptability judgements, humanness plays a role in indexing in Macedonian (Tomić 2008: 71–72). In other languages of the Balkans, P indexing is similarly associated with pragmatic and semantic features of the referent, and has grammaticalized to various extents (or not at all) in the different languages (Ivanov 2012: 350), and this phenomenon was given some attention in an areal context (e.g. Kallulli & Tasmowski 2008a).

There has also been some interest in differential P indexing from a family perspective, for example regarding the Romance languages (e.g. Jaeggli 1981, Miller & Monachesi 2003, De Cat & Demuth 2008, Fischer & Rinke 2013, or Fischer et al. 2019) or Bantu languages (see Downing & Marten 2019: 278–280 for a recent overview of contributions).

Whereas in some languages of the Bantu family, the constraints on P indexing seem to be purely formal in nature, in others, P indexing is licensed by inherent semantic properties of the referent. For instance, in Kagulu, P indexing can be considered as being differential in that it is described as "optional" for animate referents as soon as there is an overt NP referring to the same referent, exemplified in (47a) and (47b); likewise, if there is P indexing

without an NP, as in (47c), this lack of an NP is lead back to the fact that there is already indexing (Petzell 2008: 169).

(47) Kagulu (Bantu, Tanzania, Petzell 2008: 169–170)

- a. *Awafele ha-wa-koma dijoka*
C2.woman PST-C2.A-kill C5.snake
'The women killed the snake.'
- b. *Ka-mu-on-aga imukulu akwe*
PST.3SG.A-C1.P-see-IPFV C1.big 3SG.POSS
'S/he sees his/her older sibling regularly.'
- c. *Mheho i-ku-mu-ogoh-es-a*
C9.cold C9.A-PRS-3SG.P-fear-CAUS-FV
'The cold scares him/her.'

I would like to put it a bit differently, namely that P indexing and the overt-ness of the NP are logically independent, but that they are both sensitive to the discourse pragmatic status of the referent and that in certain cases (depending on the discourse pragmatic effect the speaker wants to achieve) they co-occur. Thus, the information-structural status of some referents allows for the expression by an overt NP, as well as for indexing.

In other Bantu languages, differential indexing is considered on purely syntactic grounds and indexing together with a co-referential NP is limited to cases where the latter appears in a pragmatically marked position. For instance, in Nkore-Kiga, a topicalized P is expressed by an overt NP which appears in a clause initial topic position (the pragmatically neutral position would be after the verb). This P argument then also has to be obligatorily indexed on the verb (Taylor 1985: 78, 91) and the co-occurrence of index and clause-initial NP has thus fully grammaticalized.

That P indexing has become obligatory only in combination with other structural features which can themselves be considered differential is not uncommon. It often co-varies with word order alternations, for instance, in Amuesha (Duff-Tripp 1997) or Burunge (Kießling 1994). Moreover, in some languages, differential indexing has grammaticalized to the extent that it goes hand in hand with differential flagging (see Arkadiev 2013), e.g. in Romanian (Cojocaru 2004) or Lebanese Arabic (Aoun 1999), or in that only a subset of nouns are indexed, e.g. in Makhuwa-Enhara (van der Wal 2009).

Although the co-occurrence of a lexical NP and an index might become obligatory, like in Nkore-Kiga, they should still be considered as logically independent. This becomes especially clear when looking into languages where indexing and the overtness of the referent NP do not totally correlate, but are similarly connected to the information-structural status of a referent. For the Austronesian language Larike (Laidig & Laidig 1990), for instance, the authors state quite clearly how both indexing and the overtness of the NP are independently associated with the prominence⁴ of the referent, but that their co-occurrence as well as their joint absence can enhance the discourse structural effect. Although indexes (for both A and P arguments in transitive clauses) generally co-occur with the NPs referring to the same referent, a sentence may consist of only a verb with the appropriate indexes, or there can be an overt NP only, without indexing. In addition, there can be neither, i.e. there can be no overt reference to a P argument at all. The authors state that the choice of how to code a referent is directly linked to the pragmatic status of the referent, and that each way of reference in a clause results in a different interpretation (Laidig & Laidig 1990: 93n11). Despite their insight, however, Laidig & Laidig (1990) unfortunately do not further elaborate on the different discourse effects that differential indexing and its interplay with the lexical argument can have.

In other accounts, the conditions of indexing or not-indexing in the respective language is somewhat clearer, for instance for Teiwa.⁵ In Teiwa, P indexing is differential in that firstly, it is strongly associated with animacy: animate referents are obligatorily indexed (48a) whereas inanimate referents are only rarely indexed. Secondly, P indexing is also differential in that its omission is associated with a particular discourse effect: the verbal prefix is omitted, also for animate referents, if they are in contrastive focus; then, a free pronoun is used instead (Klamer & Kratochvíl 2018: 81–82) as in (48b).

(48) Teiwa (Alor-Pantar, Indonesia, Klamer 2010: 407)

- a. *Miaag yivar ga-sii*
 yesterday dog 3SG.P-bite
 ‘Yesterday a dog bit him.’

⁴In Laidig & Laidig’s (1990) terms, which they do not specify.

⁵There are more members of the Alor-Pantar family which are noteworthy with regard to the importance of referential properties as well as lexical restrictions for indexing, see e.g. Fedden et al. (2014).

- b. *Miaag yivar ga'an sii*
yesterday dog 3SG bite
'Yesterday a dog bit HIM (not me).'

In this section it was shown how indexing P referents can be associated to various factors: discourse pragmatic effects (as in Macedonian, Nkore-Kiga or Teiwa), or referent semantics (as in Kagulu and Teiwa, where animacy has a strong effect, or Macedonian, where humanness has an impact on indexing). In each language, the relevant factors have different impacts on indexing, and their interplay can be quite complex. Also, the degree of obligatorification ranges from something like “always obligatory if” to “usually not”. Probably the only way to do justice, at least to some extent, to the complexity behind differential marking phenomena in many languages is to evaluate the impact of the relevant variables on the basis of corpus annotation as done, for instance, by Schikowski (2013) on object flagging in Nepali, Bresnan et al. (2007) on the English dative alternation, Just & Čéplö (to appear) on P indexing in Maltese, or Goldstein (2021) on the marking of passive agents in Ancient Greek.

Nevertheless, a digest of differential P indexing like the present account demonstrates that, this diversity notwithstanding, the phenomenon often has to do with certain expectations concerning the referent and, depending on whether these expectations are met, the referent is either not indexed or indexed. This also holds for A referents, but often in a reversed manner: the omission of A indexes is often associated with unexpected properties of the A referent. Although it has been observed for various languages that focused A NPs are often in a non-default position with regard to the verb and lack indexing (e.g. Lambrecht & Polinsky 1997, Mereu 1999), differential A indexing has not been addressed as extensively as differential P indexing. Section 6.3, deals with this phenomenon in more depth, ultimately showing that it is driven by the same mechanism as differential P indexing.

6.3 Differential A indexing – the reverse pattern of differential P indexing?

Differential A indexing has been claimed to be globally more restricted than differential P indexing (Haig 2018: 789). As opposed to differential P indexing, where the presence of the index often marks an unusual situation, in the

case of differential A indexing, the absence of a usually present index marks an atypical situation.⁶ Like with differential P indexing, lack of A indexing can be induced by semantic features such as non-volitionality or the referent being inanimate (Malchukov & Ogawa 2011: 32–36), but also by discourse pragmatic features. In some languages, differential A indexing is directly related to A displaying certain properties related to focus (such as newness, indefiniteness, contrastiveness), as has been noted by Siewierska (2004: 159–162). On that note, it is in some languages syntactically tied to interrogatives, quantification, or to relative clauses or complement clauses (e.g. Ouhalla 1993, Lambrecht & Polinsky 1997).

In line with this, it has been brought forward that the loss of topicality of the A referent can result in the omission of indexing (e.g. Givón 1976, Lambrecht & Polinsky 1997, Mereu 1999, Malchukov & Ogawa 2011: 29–32). For instance, the suspension of the index for an A argument (“lack of subject-verb agreement”) is one of a couple of prosodic and morphosyntactic strategies which languages make use of in order to express sentence focus, i.e. a proposition in which both the predicate as well as the subject are in focus (Lambrecht & Polinsky 1997). This lack of indexing is in some languages realized as impersonal or singular marking also for plural third person referents, like English *There’s three women in the room*. Lambrecht & Polinsky (1997) analyze the subject in such a construction as being in focus, together with the predicate, without really stating in detail what pragmatic effects focus entails in this case.

In many Bantu languages, lack of A indexing together with the co-referential noun phrase in a non-default, postverbal position is also often described as reflecting that the referent is in focus. Some Bantu languages feature indexing of a locative noun class instead of the noun class of the A argument,⁷ others feature indexing with the P argument, and some seem to have both constructions, like Kirundi (Ndayiragije 1999).

(49) Kirundi (Bantu, Burundi, Ndayiragije 1999)

⁶As for terminology, for differential A indexing one also finds, for instance, “anti-agreement” (Ouhalla 1993, Baier 2018), “impersonal agreement (Lambrecht & Polinsky 1997) or “default agreement” (Borsley et al. 2007).

⁷It has to be added that some Bantu languages which display this locative inversion only allow it with intransitive, some only with unaccusative verbs (see Buell 2005: 48–50).

- a. *Abâna ba-ára-nyôye amatá*
C2.children C2-PST-drink:PFV C1.milk
'Children drank milk.'
- b. *Amatá y-á-nyôye abâna*
C1.milk C1.PST-drink:PFV C2.children
'Children (not parents) drank milk.' [Lit.: 'Milk drank children.']
- c. *Ha-á-nyôye amatá abâna*
C10.LOC-PST-drink:PFV C1.milk C2.children
'Children (not parents) drank milk.' [Lit.: 'There drank milk children.']

In sentence (49a), the constituents are in canonical AVP-order and the A argument is indexed on the verb; in (49b), A and P have swapped their positions and P is indexed. In (49c), neither A nor P are indexed, but there is a prefix for the locative noun class instead. According to Ndayiragije (1999), (49b) and (49c) convey the same meaning and imply a contrastive focus reading of *abâna*. In fact, it has even been argued for some Bantu languages that indexing is forced by topicality rather than the syntactic status of the referent as an argument (e.g. Morimoto 2000), or that an index can be considered an antifocus marker (Zeller 2008). The relation between indexing A and its discourse status as not being the focus of the proposition becomes even more obvious in languages where A referents can only be indexed if they are also overtly marked for topicality (e.g. the Cushitic language Oromo, Malchukov & Ogawa 2011: 31). However, very often indexing is sensitive to the pragmatic status of a referent without this status being morphosyntactically indicated. Some cases reported by Siewierska (2004: 159–162) are Konjo (Austronesian), Chalcatongo Mixtec, and the Arawakan languages Bare, Yagua and Apurinã.

Provided that A indexing – like P indexing – is reserved for referents with a particular discourse status, there can be some other structural features that are involved in differential indexing, which similarly mark or even enhance this discourse status. In Welsh, for instance, there is a restriction with regard to the part of speech of the referential NP: only pronominal referents like in (50a) are indexed, whereas full NPs such as in (50b) are not indexed for number (the verb is in “default form”, Borsley et al. 2007). Thus, one could say that only pronominal referents fulfill the information-structural requirements to be indexed.

(50) Welsh (Celtic, Great Britain, Borsley 2009: 3)

- a. *Gwel-on nhw ddraig*
see-3PL.PST they dragon
'They saw a dragon.'
- b. *Gwel-odd y bechgyn ddraig*
see-3SG.PST DEF boy.PL dragon
'The boy / boys saw a dragon.'
- c. **Gwel-on y bechgyn ddraig*
see-3PL.PST DEF boy.PL dragon

Another type of structural dependency can be found in Koorete: A indexing correlates with the presence of the assertive focus marker on the predicate: only if the verb carries the focus marking morpheme *-ko* is A indexed (Mendisu 2010: 166, 172), as in (51a). If the verb does not carry a focus marker, as in (51b) and (51c), there is no indexing. Example (51d) is thus ungrammatical. So the index here is omitted not only if the A argument is in focus, as in (51c), but if there is any deviation from predicate focus, which is considered a universally unmarked discourse configuration (cf. Lambrecht 1994: 296).⁸

(51) Koorete (Omoti, Ethiopia, Mendisu 2010: 172, 180-181)

- a. *nun-i doro woon-d-uu-ns'i-ko*
we-NOM sheep buy-PFV-PST-1PL-FOC
'We BOUGHT sheep.'
- b. *nun-i doro-ko woon-d-o*
we-NOM sheep-FOC buy-PFV-PST
'We bought SHEEP.'
- c. *tamba-ko doro woon-d-a*
me-FOC sheep buy-PFV-REL
'I bought sheep.'
- d. **nun-i doro-ko woon-d-uu-ns'i*
we-NOM sheep-FOC buy-PFV-PST-1PL

In another Omotic language, Zargulla, the presence of the focus marker on the verb is a prerequisite for indexing as well. But even if there is focus marking by means of the respective marker, and indexing can therefore occur,

⁸Another example where indexing interacts with the presence as well as the position of overt focus marking, and which has received some attention in the literature, is Somali (e.g. Saeed 1984, Mereu 1999, Tosco 2002): A is only indexed if constituents other than A are focused (Mereu 1999: 231–232).

it is still variable, i.e. not obligatory, but sensitive to identifiability as well as animacy (Amha 2007: 200–201). The association of animacy and A indexing is also known, for instance, from Standard Persian (Sedighi 2010: 35), or Georgian (Harris 2009: 21).

As with differential P indexing, also with A indexing there can be a correlation between indexing and word order. In Anuak, for instance, various constituent orders in transitive clauses are possible, depending on information management. Reh (1996) argues that only a sentence like (52a) with a clause-final verb plus an A index can be considered pragmatically unmarked, (1996: 350–351) and the A referent as topical (Reh 1996: 339, 347–357). In all other cases, i.e. with a topical P in clause (52b), a focalised P in (52c) and a focalised A in (52d) there cannot be an index (Reh 1996: 348–350).⁹

(52) Anuak (Nilotic, Ethiopia, Reh 1996: 348)

- a. *jìlàal kwǎn ā-cám-ē*
child porridge PST-eat-3SG.A
'A child ate the porridge.'
- b. *kwǎn ā-cám jìlàal(-lì)*
porridge PST-eat child(-DEF)
'The child ate the porridge.'
- c. *jìlàal cám-á kwǎn*
child eat-FOC porridge
'The child eats (the) porridge.'
- d. *kwǎn cám-á jìlàal(-lì)*
porridge eat-FOC child(-DEF)
'A (The) child has eaten the porridge.'

In another language of Ethiopia, Sheko, there is differential A indexing without a correlating differential structure (like word order deviation) or overt pragmatic marking: A referents are usually indexed in main clauses, as in (53a), and only in those clauses where the referent NP is in focus is A indexing omitted, as in (53b):¹⁰ Additionally, similarly to Koorete, indexing in Sheko is not only sensitive to the discourse status of the A referent itself, but also to

⁹A similar case can be found, for instance, in Trentino, where focalized, post-verbal 3rd person referents are not indexed (Mereu 1999: 238).

¹⁰Similarly, in Jamsay (Dogon, Heath 2008: 455–456) a focalized preverbal A NP is not indexed, and can be focus-marked or not. In the latter case, the lack of an index alone is an indicator of focus status of the referent.

the pragmatic configuration of the whole proposition in that the position of the index is not fixed: it procliticizes to the verb stem in the case of predicate focus, corresponding to an unmarked topic-comment structure. But the index can also be enclitic to the verb, as in example (53a) in the case of verb polarity focus andthetic sentences. What is more, it can leave the verbal position altogether and can encliticize to any constituent apart from the A NP if this constituent is focused (Hellenthal 2010: 429-432). It is not uncommon cross-linguistically that index placement can be determined by discourse effects, irrespective of the argument role (see Cysouw 2003).

(53) Sheko (Omotic, Ethiopia, Hellenthal 2010: 430–436)

- a. *gébèn bây dādù nyààs = í-k*
 Geben female child give.birth = 3SG.F.A-REAL
 ‘Geben has given birth to a daughter!’
- b. *m-bāyñ nata gasku-k-ə*
 1SG.POSS-wife 1SG insult-REAL-IND
 ‘MY WIFE insulted me.’

This section has shown that differential A indexing – although probably not as common as differential P indexing – is not a rare phenomenon, as its lack of attention from a typological side (compared to other kinds of differential marking) might suggest. In languages which display overt A indexing, the lack of the index can mark deviations of the referent from being high in in some language specific factors usually ascribed to the A role, like topicality or animacy. So structurally, differential A indexing somehow looks like the mirror image of differential P indexing: instead of an additional index showing that something is amiss (such as the P role being occupied by a referent high in topicality), it is the lack of the A index which is often used to indicate something is out of order. However, both phenomena actually simply boil down to the fact that indexing is linked to referential prominence, which will be discussed more in-depth in the following section.

6.4 Differential indexing and referential prominence

6.4.1 Indexing and referential properties

The last two sections served to illustrate that differential indexing should not be considered particularly exceptional, either for P or for A. If there is a P which is high in identifiability (defined by definiteness and specificity), animacy or topicality, it can be indexed, or if an A misses a certain mark with regard to factors such as these, an index can be omitted. In order to account for the connection between various kinds of differential marking and a referent's features (inherent as well as non-inherent), different scales have been proposed, e.g. the *potentiality of agency scale* by Dixon (1979: 85), the *empathy hierarchy* by DeLancey (1981), the *prominence scale* by Aissen (1999), the *D-hierarchy* by Kiparsky et al. (2008) or the *referential hierarchy* (e.g. Bickel 2008). I will not go into these scales or hierarchies (see Witzlack-Makarevich & Seržant 2018: 5-10 for an overview) but use the more convenient term *referential prominence*¹¹ (Haspelmath 2021b) to refer more generally to a referent's status with regard to identifiability, animacy, or person ranking, etc., whichever factors are relevant in a given language.¹²

Although the precise nature as well as the impact of referential prominence always has to be considered as being language specific, its universal character lies at the heart of what Haspelmath (2021b) refers to as role-reference associations. These role-reference associations imply the ranking of roles with respect to each other (e.g. A ranked higher than P), as well as the role's characteristics regarding prominence. Deviations from these associations, like a scenario in which the role ranking is reversed, or an argument role showing an unexpected degree of prominence, can lead to a number of coding splits: Asymmetries in marking, such as differential object flagging or split ergativity, can ultimately be reduced to deviations from referential features associated with a particular role.

Crucially, Haspelmath also argues that such deviations from role-reference associations tend to be coded by longer forms as they are less predictable and

¹¹A term which has similarly been used is *salience* (Croft 1988).

¹²For an extended definition of prominence as a structure-building principle accounting for phenomena on different levels of grammar, and a discussion of its relation to other concepts of referential management see von Heusinger & Schumacher (2019).

less frequent (also see Haspelmath 2021a, as well as Croft 2003 for similar observations). The proposed explanation is that it is more efficient to explicitly mark less frequent meanings, and to not or to a lesser extent mark the more frequent ones. And this generalization works perfectly well for the coding asymmetries he presents, as in differential object (P, but also R and T) flagging, where an NP receives additional or differential case marking if the referent deviates from their role association. It not only works for splits with regard to flagging, but also for the encoding of voice on verbs, like inverse marking, passives, or antipassives, as with these categories, the verbs receive special or longer marking in situations where the arguments deviate from default associations.

As for splits in argument coding, indexing is explicitly exempted from the generalizations concerning the amount of coding material (Haspelmath 2021b: 131n6). The reason for this is that referential prominence is connected to indexing in general, not just with objects, but also with the higher ranked A arguments. Thus, whereas the economical idea to explicitly mark less frequent meanings seemingly fits with differential P indexing (where there is an index for a referent violating their role-reference association), it is not compliant with cases of differential A indexing: the deviating, unexpected, and therefore supposedly less frequent and less predictable construction receives *less* marking in that the index is omitted. Moreover, also with differential P indexing, it is not always the case that the presence of an index is the exceptional pattern, but there are also languages where there is differential P indexing surfacing by the omission of an index. Consider, for instance, the situation in the Austronesian language Makasar, where in transitive clauses both A (proclitic) and P (enclitic) are usually indexed, as in (54a). The exception to this normal transitive pattern occurs if either A or P are in focus (in 54b and 54c), in which case the respective argument NP is fronted¹³ and there is no indexing of that argument (Jukes 2015: 55–58):

(54) Makasar (Austronesian, Indonesia, Jukes 2015: 55–58)

- a. *Na = cini' = i tedong-ku i Ali*
 3SG.A = see = 3SG.P buffalo-1SG.POSS PERS Ali
 'Ali sees my buffalo.'

¹³without prosodic break, so the construction is different from left dislocation, which has different discourse pragmatic effects and would entail indexing (Jukes 2015: 60)

- b. *Kongkong* = *a* *a-buno* = *i* *miong* = *a*
 dog = DEF AF-kill = 3SG.P cat = DEF
 ‘THE DOG killed the cat’
- c. *Miong* = *a* *na* = *buno* *kongkong* = *a*
 cat = DEF 3SG.A = kill dog = DEF
 ‘The dog killed THE CAT.’

This is similar to the Teiwa example shown in (48) above, where P indexing for animate referents has to be omitted if they are in contrastive focus. Thus, the data support this view that indexing is primarily connected to prominence, irrespective of the argument role, and this is where it differs from flagging or other devices which encode the who-does-what (which I will return to in Section 6.5). Indexing indicates a certain level of language specific prominence that a referent has in discourse. If a referent does not have this particular level of prominence, or loses it (e.g. becomes focalized), it is not indexed. Thus, differential indexing should not be considered as a priori marking deviations from role-reference associations (Haspelmath 2021b), but as being only indirectly linked to roles, as an A is typically more prominent (i.e. more index-worthy) than a P, which leads to A indexing becoming grammaticalized more readily. This, in turn, might have distorted our view a bit towards the idea that A arguments (or subjects) are more prone to indexing than other roles (or relations), while this is just a side effect of the prominence level associated to that role.

6.4.2 Properties of the co-argument

Some phenomena which also fall under coding splits are scenario induced, i.e. cases where it is not referential features of the affected argument itself that cause some kind of differential marking, but the nature of the co-argument as well, i.e. “the whole configuration of who is acting on whom” (Witzlack-Makarevich & Seržant 2018: 12). Whereas differential flagging depending on co-argument features is not very common in the world’s languages (also see Haspelmath 2021b: 143–151), they play a role in indexing in a number of languages.

Such indexing systems have been referred to as hierarchical alignment (see e.g. Creissels 2009 or Witzlack-Makarevich 2011: 181–194 for discussing whether the combination of the terms ‘hierarchical’ and ‘alignment’ is appro-

priate). A distinction has been made between two systems: languages where two roles are considered to compete for a given slot and only the referent outranking the other in terms of a language specific hierarchy is indexed, and languages where indexing a referent is permitted or blocked depending on what kind of referent takes the co-argument role. However, Witzlack-Makarevich et al. (2016) clearly show that both types can be explained in terms of the latter type, namely in terms of co-argument sensitivity: hierarchies are not needed to describe, explain, or compare such systems.

The following examples provided in (55) come from Reyesano, a Tacanan language spoken in Bolivia. Whether A or P are indexed¹⁴ depends on the person as well as on the role of the co-argument: in scenarios involving a locuphoric (i.e. 1st and 2nd person) referent together with a 3rd person, as in (55a) and (55b), the locuphoric referent is indexed, whether it is A or P, while a 3rd person co-argument is indexed only when it is the A argument. In scenarios involving only 3rd person referents in both the A and the P role, exemplified in (55c), only A is indexed (the index for the third person thus encodes the A role). Lastly, in scenarios involving locuphoric referents in both roles (55d) and 55e), A or P is indexed if it is second person (Guillaume 2009: 35–40).¹⁵ For better clarity, Table 6.1 illustrates whether A or P or both are indexed in Reyesano transitive clauses, depending on the co-argument.

(55) Reyesano (Tacanan, Bolivia, Guillaume 2009: 37–40)

- a. *K-a-maneme-a awadza*
1PL-PST-kill-PST tapir
'We killed a tapir.' (*'A tapir killed us.')
- b. *K-e-dai-ta-da chenu te tue*
1PL-IPFV-cure-3SG.A-IPFV EMP BM 3SG
'She cures us.' (*'We cure her.')
- c. *A-kachi-ta-a te iba te awadza*
PST-bite-3SG.A-PST BM jaguar BM tapir
'The tapir bit the jaguar.' OR: 'The jaguar bit the tapir.'

¹⁴For 1st and 2nd person, indexes do not encode role, i.e. the same set of indexes is used

¹⁵There is no role distinction in independent pronouns (Guillaume 2009: 31). Therefore, isolated from context, the examples in (55d) and (55e) could be ambiguous in two ways: (55d) could be translated as 'You saw me' with the 2nd person as the A referent, or also be interpreted as having a third person P, i.e. 'You saw him/her'. Similarly, (55e) could also mean 'We won't forget you' in a different context.

- d. *Mi-a-b-a te miwe*
2SG-PST-see-PST BM 2SG
'I saw you (crossing the plaza yesterday afternoon).'
- e. *Ma te mi-e-deta te ekama*
NEG BM 2SG-FUT-forget BM 1PL
'You won't forget us.' OR: 'We won't forget you.'

Arguments	Indexing
A1st with P2nd	P
A1st with P3rd	A
A2nd with P1st	A
A2nd with P3rd	A
A3rd with P1st	A and P
A3rd with P2nd	A and P
A3rd with P3rd	A

Table 6.1: A and P indexing in Reyesano transitive clauses

In a system like this, referential prominence is primarily defined in terms of empathy (cf. DeLancey 1981), and how referents relate to one another with regard to it. That is, the lineup of the referents determines which role is indexed. Thus, indexing based on co-argument sensitivity can be referred to as differential in that it is not the argument role itself which triggers whether a referent is indexed. It is not essentially different from systems like those found in Koorete (examples in 51) or Anuak (examples in 52) or other languages where it is not or not only some inherent referential features playing into prominence and thus trigger indexing, but where the whole configuration of a clause is relevant.

6.4.3 Prominence and lexical NPs

What can also be considered in the light of prominence-level and is thus ultimately linked to role associations, is whether an argument is expressed by an overt NP. Lexical NPs are used for new information, contrastive information, topic shifts or for referents at a long lexical distance; non-lexical forms, on

the other hand, are used for more accessible information (Givón 1983, Ariel 1990), i.e. for referents higher in prominence.

As the A role is usually occupied by referents which are high in identifiability, animacy and topicality, A arguments are less commonly occupied by lexical NPs than P arguments (see DuBois 1987, as well as Haig et al. 2020 for confirming the observation in corpus data from a sample of typologically diverse languages). P, on the other hand, is more commonly used to introduce new or non-prominent referents than A (e.g. Givón 1976, DuBois 1987, Comrie 1988, Schnell et al. 2020). Therefore, to be expressed by a lexical NP is more of an exception for the A role than for the P role (also see Lambrecht 1994: 189-190).

And this is the point where we come back to the issue of indexing and the co-occurrence of a full lexical NP, one of the parameters often used in the agreement-vs-pronoun-debate. Although I do not want to overgeneralize too much here – prominence obviously has a common denominator cross-linguistically, but it also has a specific character in every language – indexing and the overtness and also the position of a lexical NP in the clause are subject to referential prominence.

6.5 Some more thoughts on the function of indexing

Regarding indexing and overt NPs, their co-occurrence can in some cases grammaticalize, either with an additional prerequisite (word order, differential flagging etc., like in the case of Nkore-Kiga preverbal P NPs), or unconditionally, in languages with grammatical agreement (Bresnan & Mchombo 1987, Siewierska 1999). Although grammatical agreement, where both the index and the overt NP (aka the controller) are confined to the same clause, is very rare cross-linguistically (Siewierska 1999), it has been considered as the logical endpoint on the basis of which deviating scenarios can be described (Corbett 2006).

It seems to be hard to give up on the assumption that indexing has to be investigated with regard to the co-occurrence of a referential NP. Even Haspelmath (2013) who argues for treating indexes as “phenomena sui generis” (2013: 213) in order to detach them from the futile agreement-vs-pronoun

discussion, eventually classifies them on the basis of whether they can, may, or must co-occur with a co-nominal in the same clause (cross-indexes, gramm-indexes and pro-indexes, respectively, Haspelmath 2013: 218–221).

In order to uncouple indexing from overt NPs and to truly consider them as logically independent (as also advocated by Haig & Forker 2018), one has to bear in mind that indexes, in the course of their emergence, become less referential than they used to be in their time as an anaphoric pronoun (Kibrik 2011). Anaphoric pronouns, which are by definition used for prominent referents, can be considered as the source material for indexes (e.g. Givón 1976, Lehmann 1982). When grammaticalizing into indexes, they gradually lose referential potential (Siewierska 1999: 225) and it is the transmission of the prominence level of the referent which remains. This would explain how indexing is employed to facilitate tracking of prominent referents, irrespective of the argument role (e.g. Lehmann 1982, Givón 1983, Siewierska 1997) which can best be demonstrated by looking into cases of differential indexing as presented in Sections 6.2 and 6.3. An index is omitted if the referent is not prominent enough (however prominence might be spelled out in the respective language) even though this index might have grammaticalized in other contexts. Conversely, there can be an index which is syntactically optional in order to assign the appropriate level of prominence to the respective referent.

Therefore, when considering indexes as a role identifier like case marking (e.g. Dixon 2010, Haspelmath 2019), one probably runs into danger of confusing cause and effect. What I mean by this is the following: as indexes serve to keep track of a prominent referent, indexing of a particular role depends on which kind of referent (semantically and/or pragmatically speaking) occupies this role. Following this assumption, an index indexing referential features of a particular argument role and thus marking this role on the predicate should be considered as a side effect of following this particular referent through the discourse.¹⁶

So I think that what has been stated by (Iemmolo 2011: 47–60), that differential P flagging and differential P indexing, albeit often sensitive to similar referential features, do not serve the same purpose, can be expanded

¹⁶ Arguably, there are languages (like German) where over the course of time indexing has become more of an automatism and dissociated from prominence, which might have led to statements that grammatical agreement is a purely redundant expression of features (Bresnan & Mchombo 1987: 741).

to flagging and indexing in general. True role assignment can be achieved through flagging which serves to distinguish the referents involved in an action (Bakker & Siewierska 2009), whereas indexing a referent, occupying a particular role, marks this one as prominent and as to be tracked. That flagging is immediately tied to roles in contrast to indexing would also explain why flagging splits can be explained by deviations from role-reference associations which ultimately lead to longer coding for unexpected and/or less frequent combinations (Haspelmath 2021b), while this does not really work for indexing: in cases of differential A indexing, the deviating and supposedly less frequent and less predictable construction receives less marking instead of longer coding (see Section 6.4.1).

6.6 Conclusion

In a nutshell, investigating differential indexing opens a window to looking at the core of indexing in general, and to understanding its link to referential prominence. Prominent arguments, be it A or P, tend to be indexed more readily than arguments which are low in identifiability, animacy or topicality. That both A and P indexing have a common purpose, namely tracking referents through discourse, becomes especially evident in languages where indexing is particularly sensitive to referential prominence, and where it thus is employed only if the referent fulfills certain prominence criteria.

The A role is more commonly associated with prominent referents, so indexing it is in many languages more prevalent than P indexing, and has more often grammaticalized. But indexing can also be associated with the P argument, if the referent fulfills certain requirements.¹⁷ (Haig 2018) discusses whether the grammaticalization paths of A and P indexes are the same. He argues that differential P indexing constitutes an attractor state, i.e. a pattern constituting the endpoint towards which languages as complex dynamic systems tend to settle during the course of change. This would mean that differential indexing is more preferable for P referents than “fully obligatory” (Haig 2018: 788) indexing.

¹⁷Categorical splits due to information structural effects can be considered rare (Schultze-Berndt 2018), and phenomena which are not hard-coded (i.e. not syntactically required) might escape grammatical description (Fauconnier & Verstraete 2014: 10n1). So not only differential P indexing, but also differential A indexing might be more wide spread than it appears.

On that note, I would like to point out again that even with languages where indexing is described as obligatory (like in Bantu languages), it can be differential to satisfy certain discourse pragmatic effects. And even if such a pragmatic effect is seemingly identified as focus or emphasis, it still remains to be discerned what this actually means for the language in question (cf. Matić & Wedgwood 2013, Ozerov 2018). In addition, in cases where there is an interplay of various factors involved, the question remains whether these factors can be translated into obligatorily choosing one pattern over another. It is probably more fruitful to conceive of many differential marking patterns as tendencies rather than rules (cf. Witzlack-Makarevich & Seržant 2018: 28), and the often cited optionality can be well motivated, although the motivation for the choice of a particular marking pattern might not be fully understood.