

The grammar of engagement II: typology and diachrony

NICHOLAS EVANS

Australian National University & ARC Centre of Excellence for the Dynamics of Language

HENRIK BERGQVIST

Stockholm University

AND

LILA SAN ROQUE

Radboud Universiteit Nijmegen & Max Planck Institute for Psycholinguistics

(Received 08 March 2017 – Revised 29 August 2017 – Accepted 31 August 2017)

ABSTRACT

Engagement systems encode the relative accessibility of an entity or state of affairs to the speaker and addressee, and are thus underpinned by our social cognitive capacities. In our first foray into engagement (Part 1), we focused on specialised semantic contrasts as found in entity-level deictic systems, tailored to the primal scenario for establishing joint attention. This second paper broadens out to an exploration of engagement at the level of events and even metapositions, and comments on how such systems may evolve. The languages Andoke and Kogi demonstrate what a canonical system of engagement with clausal scope looks like, symmetrically assigning ‘knowing’ and ‘unknowing’ values to speaker and addressee. Engagement is also found cross-cutting other epistemic categories such as evidentiality, for example where a complex assessment of relative speaker and addressee awareness concerns the source of information rather than the proposition itself. Data from the language Abui reveal that one way in which engagement systems can develop is by upscoping demonstratives, which normally denote entities, to apply at the level of events. We conclude by stressing the need for studies that focus on what difference it makes, in terms of

Address for correspondence: Nicholas Evans. e-mail: nicholas.evans@anu.edu.au

communicative behaviour, for intersubjective coordination to be managed by engagement systems as opposed to other, non-grammaticalised means.

KEYWORDS: engagement, accessibility, epistemic, evidential, perspective, intersubjectivity.

Bien des notions en linguistique ... apparaîtront sous un jour différent si on les rétablit dans le cadre du discours, qui est la langue en tant qu'assumée par l'homme qui parle, et dans la condition d'intersubjectivité, qui seule rend possible la communication linguistique.

(Benveniste, 1966, p. 266)

1. Introduction

Engagement refers to *a grammatical system for encoding the relative accessibility of an entity or state of affairs to the speaker and addressee*. While many linguistic elements can be deployed to express intersubjective meanings of this kind (e.g., asserting that I know something and you don't), the possibility that grammatical systems can be built around such values – themselves fundamental to social cognition – has barely been explored and remains an open question. In Part I we introduced the notion of engagement with an initial example from Andoke, where a four-way auxiliary choice, which is a core part of the grammar and has clause-level scope, encodes the speaker's assumptions about the accessibility of the represented proposition to speaker and/or hearer across all four logical permutations (speaker only, hearer only, both, or neither). From there we passed to a discussion of the broader question of intersubjectivity in language – not necessarily grammaticalised – and then back to the 'primal scene' of attentional coordination as it is played out through the use of deictics to coordinate attention to objects. We placed special emphasis on systems like Turkish or Jahai, in which attentional coordination appears to be the primary function of at least one demonstrative.

In this second part of the paper, we return to systems where the scope is the proposition or clause rather than the entity or NP. We also broaden our typological base to show that systems of engagement with clausal scope are found in several geographical hotspots – particularly the Colombian Andes and Western Amazonia, and several parts of New Guinea. In §2 we examine two systems from South America, Kogi and Kakataibo, which resemble Andoke in taking the event as a whole, rather than an individual object, as the level at which grammaticalised engagement coordinates mutual attention. In §3, we examine how engagement can interact with other knowledge-related categories, for example, by taking not just the proposition itself but the evidence for it within its scope. Having worked our way upwards in terms of

scope levels, from demonstratives (entities) through basic propositions (states of affairs) to meta-propositions (certain evidentials), we show the interconnections between them in §4, by examining a language (Abui) which has coerced the rich set of speaker- vs. addressee-based contrasts in its demonstrative system into use at different grammatical levels (interclausal marker, clause-final marker); in the process it has developed a set of engagement markers from more basic deictic contrasts. We conclude in §5 by drawing together the threads of these various systems, suggesting some directions in which a more comprehensive typology of engagement can be developed in future research.

2. Engagement and states of affairs

We have already introduced one example of a language of Colombia, Andoke, with grammatical marking indicating the presumed degree of speaker and/or addressee knowledge or attention (broadly, accessibility) regarding an event, drawing on the seminal study by Landaburu (2007). We now examine in detail two further languages where engagement has scope over clauses / states of affairs. In §2.1, we turn to Kogi, an unrelated Colombian language, which organises the four-way choice of engagement values into two sets of two, defined by a contrast between speaker-perspective and addressee-perspective. In §2.2 we look at Kakataibo, which also clearly manifests contrasts between speaker-focused and addressee-focused evidence, but in a way that is structurally less neat than either Andoke or Kogi.

2.1. EPISTEMIC MARKING IN KOGI

Kogi (Arwako-Chibchan) has a tightly structured, paradigmatic set of epistemic markers, prefixed to an auxiliary verb, whose function is to signal the speaker's assumptions regarding epistemic (a)symmetries between the speech participants with respect to an event (see Bergqvist, 2011, 2016). 'Symmetry' denotes a situation where speech participants have shared access to an event, whereas 'asymmetry' indicates that access is exclusive to one party. Accessibility is subject to epistemic authority, which may reside with the speaker, or the addressee (see directly below). The set of epistemic markers consists of five prefixes: *na-*, *ni-*, *sha-*, *shi-*, and *ska(n)-*.¹

Na- and *ni-* both signal that the epistemic authority rests with the speaker. *Na-* denotes the speaker's exclusive access to an event, while *ni-* denotes shared access between the speaker and the addressee. Consider the examples in (1):

[1] While *ska(n)-* forms part of the set of epistemic prefixes for both paradigmatic and functional reasons, it will be left out of the present account in order to allow for a clearer focus on the speaker/addressee-authority contrast found with *na-/ni-/sha-/shi-*.

- (1) a. *kwisa-té* *na-nuk-kú*
 dance-IMPF SPKR.ASYM-be.LOC-1SG
 ‘I am/was dancing.’ {I am informing you} (JM_130613)
- b. *kwisa-té* *ni-nuk-kú*
 dance-IMPF SPKR.SYM-be.LOC-1SG
 ‘I am/was dancing.’ {as you know / are aware} (BUN_090824)

The verb form *namukkú* in (1a) is appropriate in a situation where the speaker claims epistemic authority (in this case related to performing the action in question) without assuming that the addressee is aware of or knows the event referred to. For example, it could be uttered in a situation where the addressee has just asked the speaker what they are doing in another room. Access in (1a) is thus asymmetrical. The form *ninukkú*, on the other hand, is appropriate when the speaker claims epistemic authority while at the same time assuming that the addressee already knows, or is aware of, the event. Thus (1b) could be uttered in a situation where the speaker is asked to do something else and replies that they can’t do this right now because of their current activity, namely dancing. Access is in this case presented as symmetrical.

The forms, *shi-* and *sha-*, in contrast, pass the epistemic authority to the addressee. *Sha-* denotes the addressee’s exclusive access (2a), while *shi-* denotes shared access between the addressee and the speaker (2b):

- (2) a. *nas* *hanchibé* *sha-kwisa=tuk-(k)u*
 1SG.IND good ADR.ASYM-dance=be.LOC-1SG
 ‘I am dancing well.’ {don’t you think?} (BUN_090824)
- b. *kwisa-té* *shi-ba-lox*
 dance-IMPF ADR.ASYM-2SG-be.LOC
 ‘You are/were dancing.’ {right?} (BUN_090824)

As would be expected from ‘territory of knowledge’ considerations, vesting of the epistemic authority with the addressee frequently correlates with second person subject markers, as shown in (2b), but the distribution of the addressee-authority forms *sha-* and *shi-* is by no means restricted by the person of the subject, as shown in (2a) where the event concerns the actions of the speaker. Example (2a) could be uttered in a situation where someone learning how to dance seeks an evaluation from the instructor. By uttering the sentence in (2a), the speaker indicates that they think they are dancing well, but leaves it up to the addressee to agree or disagree. Example (2b) could be uttered in a situation where the speaker comments on the obvious activity of the addressee, but invites agreement from the addressee, who is offered the ultimate authority for the assertion. The paradigm of forms is shown in Table 1.

There is a functional overlap between the notions of speaker- vs. addressee-authority and of sentence-type. While *na-/ni* clearly occur in declarative

TABLE 1. *Meaning dimensions of epistemic marking prefixes in Kogi (after Bergqvist, 2016)*

	Speaker-authority	Addressee-authority
Asymmetric	<i>na-</i>	<i>sha-</i>
Symmetric	<i>ni-</i>	<i>shi-</i>
Non-Speech Participant		<i>ska(n)-</i>

clauses, the addressee-authority forms *shi-* and *sha-* might appear *prima facie* to be interrogative markers, as is suggested by the paraphrases in curly brackets (i.e., *don't you think?* / *right?*). However, there are both grammatical and distributional reasons to analyse these as occurring in declarative clauses as well.

First, interrogative constructions can be formed without *sha-/shi-*, for example with a content interrogative (3a) or the interrogative marker *-é* (3b):

- (3) a. *sakí mi-k-zéi-shi²*
 what 2O-DAT-feel-PTCP
 'How are you?' (DAM_090819)
- b. *néi ma-gu-ngu-é*
 go 2SG-do-PST-INT
 'Did you go?' (DAM_090820)

Second, the interrogative marker *-e* and the engagement prefix *sha-* are in complementary distribution (4): it is ungrammatical to combine the *shi-/sha-* prefixes with the interrogative *-e*. The semantic difference between *-e* and *sha-* is suggested by the translation of example (4) where 'thinking about something' (e.g., what to eat, or where to go) differs from 'having an opinion about something' (cf. (2a) above). The key difference in meaning is whether the speaker expresses his/her assumptions regarding the addressee's thoughts and opinions, or not. In (4a), the speaker avoids making such assumptions by using *-e*. In (4b), on the other hand, the speaker assumes that the addressee has an opinion/thought about something and signals, at the same time, that the addressee has epistemic authority concerning what this opinion consists of. Given an otherwise identical construction, this difference in meaning must be attributed to the semantics of the individual forms, which in the case of *sha-* aligns with its proposed exclusive meaning (asymmetry).

- (4) a. *sakí hangwa-ba-lóx-e*
 what think-2SG-PROG-INT
 'What are you thinking about?'

[2] The *shi-*suffix that is glossed PRTC (participle) is not related to the epistemic prefix *shi-*.

- b. *sakí sha-hangwa-ba-lóx*
 what ADR.ASYM-think-2SG-PROG
 ‘What do you think (about something)?’
 (BUN_090826)

The presence of the speaker’s assertion in the *shi-/sha-* forms is also apparent from their use in narratives. Depending on the specific setting for a narrative, an addressee-oriented stance may be adopted by marking monologic stretches of speech with either *shi-* or *sha-*. Consider the extract in (5), taken from a first person account of what life was like in the region of the Sierra Nevada de Santa Marta before the colonisers came and claimed much of the Kogi’s traditional lands.

- (5) *hate-kwe-ha~ Ø-izhi-hĩ dzaldzí-chi hixa aró hixa*
 father-PL-AGT 3SG-bring-PRTC non.indigenous-ABL nor rice nor
aka-té Ø-to-a-kí hei-ni zeldázã
 eat-PROG 3SG-see-PERF-NEG this-LOC food
 ‘The elders were not bringing (food) from the outsiders; not rice, nor
 had they seen eating (of this kind), only traditional food.’
 [...]
hei-kí hei-kí shi-tu-lo-ku-ã
 this-FOC this-FOC ADR.SYM-see-PROG-1SG-PERF
 ‘This, this is what I saw.’
 (JM_130613)

The use of *shi-* in the final utterance of a longer stretch of speech serves to invite the (potentially) overlapping points of view of the speaker’s peers, who are present during the performance of the narrative. Notably, in other parts of the narrative, *sha-* is used interchangeably with *shi-* (see Bergqvist, 2016). Comparable narratives that are told to foreigners, or persons unfamiliar with the Kogi way of life, do not feature the *shi-/sha-* forms. Instead, they usually feature the *na-/ni-* forms, which, as stated, focus on the epistemic authority of the speaker.

While Kogi epistemic prefixes are frequent in discourse, they are not obligatory. Their grammatical status is also restricted in that the *na-/ni-/sha-/shi-* forms are mainly found in auxiliary constructions where they attach to the auxiliary head. Non-auxiliary (synthetic) verb phrases cannot directly take the epistemic prefixes. A way around this restriction is available, however, by using periphrastic auxiliaries (6):

- (6) *nas kwisa-nuk-ku-gé na-kla*
 1SG.IND dance-PROG-1SG-HAB SPKR.ASYM-be
 ‘(Can’t you see) I am dancing!’ (ARR_120520)

Nakla is arguably not part of the verbal core, which is limited to the synthetic verb phrase (*kwisanukkugé*). Exactly what the functional and/or

semantic difference between examples (1a) and (6) consists of remains to be explained.

The semantic scope of the prefixes includes tense, aspect, mood, and polarity. An example of how epistemic asymmetry scopes over modality is in (7a, b). In these examples the impossibility of sleeping is modified by the *ni-/na-* contrast, which target differences in epistemic symmetry:

- (7) a. *kaba-gasā* *ni-ba-kú*
 sleep-NEG.POT SPKR.SYM-2SG-do
 ‘(Now) you can’t sleep anymore.’ (e.g., because it’s morning)
- b. *kaba-gasā* *na-ba-kú*
 sleep-NEG.POT SPKR.ASYM-2SG-do
 ‘You can’t sleep anymore.’ (e.g., because I say so, or for reasons unknown to you)
 (ARR_120520)

Pragmatic interpretation effects that cannot be attributed to the encoded meaning of the forms, but which may result from their combination with certain contextual cues, include temporal displacement and attitudinal shades of meaning, such as ‘familiarity’ and ‘affection’. These are both forms which interact with time reference (see Bergqvist, 2011, 2016).

Given the non-obligatory status of the discussed forms, what motivates the use of *ni-/na-/shi-/sha-* and when are they omitted? While the pragmatic considerations relevant to predicting the use of these prefixes have not yet been exhaustively explored, there are some initial indications. An important determinant of the (a)symmetry marker’s distribution is purely interactional: if there is an opposing claim to the one held by the speaker, then this may be contradicted by asserting (asymmetric) epistemic authority (cf. *I do like the Eagles’ first album!*). Conversely, the speaker may be forced to defer authority to the addressee in order to be able to talk about certain topics at all, such as the opinions of the addressee. Drawing on a model for stance-taking that aligns the speaker’s evaluation/positioning of an event with the addressee’s evaluation/positioning of the same event (Du Bois, 2007), we see that the notion of epistemic asymmetry in Kogi is most likely to be used when an event has direct relevance for the speaker and/or the addressee. This pertains especially to events within the speech participant’s presumed ‘territory of information’ (Kamio, 1997), including ones that involve family members, expert knowledge, and personal experience. In contrast, engagement prefixes will be omitted where the speaker judges an event as inconsequential to him/herself and the addressee, for example, events involving third persons that do not require an evaluation.

2.2. SPEAKER- VS. ADDRESSEE-PERSPECTIVE IN KAKATAIBO

While there is a solid tradition for the study of speaker's perspective (in modality and evidentiality systems, for instance), the cross-linguistic apparatus for the study of the encoding of the perspective of the addressee is currently being built.

(Zariquiey, 2015, p. 161)

Kakataibo is a Panoan language of Peru that is of special interest for the number of markers it devotes to encoding "the expectations of the speaker about the perspective of the addressee in relation to the information presented in an utterance" (Zariquiey, 2015, p. 143).³ These markers are found both in the final affix slot on verbs, and in special slots at the end of clitic strings in clause-second position.

A primary category distinction that affects the set of addressee-sensitive grammatical choices in Kakataibo is the difference between narrative and conversational genres, reflecting differences in the differential accessibility of information between recounted events vs. the here-and-now.

In the narrative genre, verbal suffix morphology opposes *-a* 'unmarked' to *-ín* '(unexpectedly) proximal / accessible to the addressee'. The default is to use the unmarked form, since normally one talks about things not known to the addressee, but Zariquiey discusses some revealing cases where the narrative passes from information (correctly) assumed by the speaker not to be known to the addressee, to information with which the addressee is familiar. For example, in (8a) the speaker begins a text with clan information unknown to the addressee, and uses the unmarked suffix *-a*, but somewhat later in the text (8b) he passes to the mention of a particular man (the son of one of the three brothers referred to in (8a)). This man was a close friend of the addressee, triggering a shift to *-ín*. Note that, while the key addressee-accessible information is the NP *este Nicolás Aguilar* 'this Nicolás Aguilar', the addressee-proximity is marked on the head of the clause as a whole, namely the verb. This resembles the location of engagement marking in Andoke and Kogi that we discussed above.⁴

- (8) a. *A kimisha uni i-akë-x-a tres hermanos*
 That three man.ABS be-REM.PST-3-UNM three brothers
 'Those three men were three brothers.'

[3] For interesting discussion of another language, the Tibetic language Denjongke, see Yliniemi (2016). Denjongke possesses a special clitic, =*co*, which Yliniemi shows is used to indicate the preceding material is "particularly attention-worthy, ... because it is unexpected, surprising, counter-expectational, newsworthy, important to know, a counter-claim, or the main point of a story or teaching" (p. 106).

[4] Though of course the use of Spanish *este* is also indexing addressee-familiarity, something that could be rendered in English through the use of 'your' in the alternative translation offered here. See also Manning (2001) for an equivalent method in Welsh.

- b. *Este Nicolás Aguilar a-x i-akë-x-ín*
 This Nicolás Aguilar 3 PL-S be-REM.PST-3 PL-PROX
 ‘that (man) ... His son was this Nicolás Aguilar’ (perhaps better
 rendered in English as ‘was your Nicolás Aguilar’)

As well as *-ín*, there is what Zariquiey (2015, pp. 154–155) calls a special second person malefactive suffix *-ié*.⁵ This is used when reporting an event that will impact negatively on the addressee, but only when “the event is assumed by the speaker to be non-proximal from the perspective of the addressee in the sense that the information is not perceptually accessible for him or her”. An example:

- (9) *Goliath=n kamënë’ mi=n*
 Goliath=ERG NAR.3 PL.MIR you=GEN
kuriki mēkamat-ié:
 money.ABS steal-3 PL.2 MAL.NON.PROX
 ‘Goliath took your money.’

Within the conversational genre, addressee perspective is manifested in a different grammatical site – at the end of a string of second position clitics. As with *-ié*: but in opposition to *-ín*, the assumption of addressee ignorance attaches to these clitics, but in contradistinction to both cases there is a focus on the speaker’s (cognitive integration of) knowledge: certainty, previously established, in the case of the ‘certitudinal’, and surprise in the case of the mirative. More specifically, the *=pa* ‘certitudinal’ clitic is used in recounting events which the addressee wasn’t present to witness, while the *=pënë* ‘mirative’ “indicates that the addressee and the speaker have different perspectives or are in different places at the moment of the speech act” (Zariquiey, 2015, p. 158). For example, if the speaker discovers something about the addressee’s son, and reports it, he would use one of two forms depending on the time of the discovery. He would use *pa* (in the sequence *riapa*) if he discovered it earlier and then went to tell the addressee it is true, but *pënë*’ (in the sequence

[5] Zariquiey uses the term ‘accusatory’, but we use the more standard cross-linguistic term ‘second person malefactive’. Note in passing that there exist languages with pairs of distinct benefactive forms that can distinguish between whether the effect is known or not to the beneficiary (not necessarily second person, though). An example is Lakhota (Boas & Deloria, 1941): cf. the following pair of examples courtesy of the late Regina Pustet (p.c.), illustrating the difference between the two benefactive prefixes *wakí-* and *wéci-*:

- (i) *mázaska ki wakí-yuha*
 money DEF 1 SG.AGT.3 SG.BENA-keep
 ‘I keep the money for her, and she doesn’t know’
 (ii) *mázaska ki wéci-yuha*
 money DEF 1 SG.AGT.3 SG.BENB-keep
 ‘I keep the money for her, and she knows it’

riapëñë) if he is seeing it at the moment of reporting, but the addressee can't, e.g., because he is too far from where the event takes place.

While it is clear on the one hand that there are a number of categories in Kakataibo relevant to the monitoring of addressee's presumed knowledge or access to information, the organisation of the grammar differs from Andoke or Kogi in not presenting a single organised paradigm detached from other categories. There are different grammatical strategies depending on whether the genre is narrative or conversational, leading to different locations for the addressee-oriented marker (verbal suffix vs. second position clitic). The encoding of presumed addressee non-knowledge gets melded in with second person malefactive in the case of *-ié:*, and within the conversational genre it is mixed up with degrees of speaker integration and ratification of knowledge. Finally, there are differences in whether the relevant markers emphasise accessibility to the addressee (*-in*), against the presumed background of inaccessibility in narratives, or inaccessibility (*=pa* and *=pëñë*), against the background of presumed accessibility in face to face interaction.

3. Engagement, evidence, and other epistemic categories

In the preceding section we have focused on the expression of accessibility and knowledge as either present or absent across speech act participants, with this mental directedness portrayed as either particular to speaker or hearer, or shared between them. However, we cannot stop there, as additional qualities of knowledge (for example, source and certainty) may also be incorporated with engagement-type values. Here we discuss some examples of how the more classic knowledge-related category of evidentiality, and to some extent those of epistemic modality and mirativity, can combine with the grammaticalised marking of engagement. In certain cases we can view these systems as *metapositional* operators, where attention is coordinated not necessarily towards an event itself, but rather to the evidence for it. This represents a similar shift in level as that from entity (typically, the province of demonstratives in the noun phrase) to state of affairs (typically, the province of verbal operators in the clause), as discussed previously.

Evidentiality is conservatively defined as 'grammaticised information source' (Aikhenvald, 2004). Typically, evidential morphemes specify the kind of evidence that an assertion is based on, for example, whether the event was seen to happen, or is being reported from hearsay. More rarely, evidentials may take scope over a referent (e.g., stating that an entity is known about through hearsay) rather than a state of affairs (see, e.g., Aikhenvald, 2015; Gutiérrez & Matthewson, 2012; Hanks, 1990; Jacques & Lahaussais, 2014; San Roque, 2008).

For some constructions in a language that marks an event like ‘peccaries crossed the path here’ with a perceptual evidential, there is a metapositional operator, representing the epistemic commitment of perception, which takes the basic proposition in its scope. Exactly how this epistemic commitment is best represented for individual morphemes and languages is an interesting problem – at one extreme (e.g., Fleck, 2007; Speas, 2004) are analyses that treat the epistemic commitment as a (fully tensed) proposition with an identifiable perceiver-argument (I saw that ...), at the other extreme (see, e.g., San Roque, 2015) are underspecified representations that do not anchor the information source to any particular deictic centre (e.g., ‘through visual evidence’). For present purposes, our main goal is to show that these metapositions of evidence can themselves be modulated according to the same categories of engagement that apply to propositions.

Studies of evidentiality have usually focused on the speaker as an experiencer of evidence, and it certainly seems to be the case that evidential markers tend to be used to express the speaker’s perspective. As a general rule, we make claims about our *own* evidence for the things we say. However, it has long been known that evidential morphology can also represent non-speaker perspectives. For example, questions typically take the evidential perspective of the addressee (Aikhenvald, 2004; San Roque, Floyd, & Norcliffe, 2017), while third person narratives may be at least partly told from the evidential perspective of a central protagonist (see examples in Brugman & Macaulay, 2015). Certain languages appear to have taken this ability to represent the evidence of others a step further, and encode not one but two evidential perspectives simultaneously: that of both the speaker and the hearer. While such systems have been described (or at least sketched) for several different languages, our understanding of them is still in its infancy, and, with some exceptions, little material is available on how such distinctions are operationalised in discourse. We limit ourselves here to outlining a few of the known contrasts, looking first at several languages that appear to make specific claims about the nature of an addressee’s evidence.

Several languages of New Guinea, including Foe (Rule, 1977), Wola (Sillitoe, 2010), and Pole (Rule, 1977), are described as encoding whether an information source is shared between speaker and hearer, or exclusive to one of them (see also San Roque & Loughnane, 2012a, 2012b). Foe (Rule, 1977) has a rich evidential system in independent clauses that distinguishes up to five information source categories (participatory, visual, non-visual sensory, inferred, assumed) across four tenses (present, near past, far past, future), three moods (indicative, customary, abilitative) and two sentence types (declarative and interrogative). These evidentials reflect a single perspective, typically that of the speaker in statements and the addressee in questions.

However, in nominalised clauses, an additional distinction is introduced into the participatory/visual evidential paradigm: whether or not the *addressee* witnessed the event or situation in question. Thus, Rule (1977, p. 97) describes a set of nominalisers used for a “fact known to speaker but unseen by person spoken to” as opposed to events “seen by both speaker and person spoken to” (see Table 2). Nominalisers also have special forms to indicate non-visual sensory and inferred evidence, but for these suffixes the addressee’s (presumed) perspective is not specified.

Examples of the contrastive far past nominalisers *-ira* and *-bo’owa* (as used in the formation of a relative clause) are shown in (10a) and (10b), respectively. While Rule does not provide details of context, we can assume that in (10a) only the speaker witnessed or was otherwise involved in the killing of the men long ago, whereas in (10b), both speaker and addressee saw the pig being killed:

- (10) a. *amena gahaye hü-ira bi hüyoga-bi’ae*
 ?men previously hit/strike-FP.KTS.NMZ ?here bury-FP.PTCY
 ‘The men who were killed a long time ago, we buried here.’
 (Rule, 1977, p. 97, gloss added)
- b. *nami davi hü-bo’owa to’ae*
 pig 2.days.away hit/strike-FP.SSA.NMZ ?this
 ‘This is the pig which was killed two days ago.’
 (Rule, 1977, p. 97, gloss added)

The Engan language Pole uses a special marker on main verbs when referring to past events that both the speaker and addressee saw (Rule, 1977). Another Engan language, Wola, has a more complex system of evidential contrasts in independent clauses. According to Sillitoe’s (2010) analysis, in the near and far past tenses Wola contrasts five kinds of speaker/addressee evidence:

- i. both speaker and hearer witness [or participate in]⁶
- ii. either speaker or hearer witnesses [or participates in]
- iii. hearer did not witness but heard of previously
- iv. speaker did not witness
- v. neither speaker nor hearer witness

[6] From Sillitoe’s (2010) discussion it appears that ‘witnessing’ refers to either participation and/or observation, that is, potentially covering both participatory (egophoric) and visual evidence. From paradigms provided by Sillitoe and information on related languages (e.g., Madden, n.d.; H. Reithofer p.c.), it appears that verb forms encoding mutual knowledge (that is, categories (i) and (iii) are compositional, at least diachronically, with a morpheme that specifies addressee knowledge (‘you know this too!’) being added following an inflection that specifies individual evidence (typically understood as that of the speaker in such contexts). However, more data are needed to gain a fuller understanding of this fascinating aspect of Wola and of other Angal language varieties, as well as of their evidential systems as a whole (see also Reithofer, 2011; Tipton, 1982).

TABLE 2. *Selected evidential nominalisers in Foe (from Rule, 1977, p. 97)^a*

	Present	Near past	Far past	Future
Known to speaker, unseen by addressee	<i>-bora</i>	<i>-ra</i>	<i>-ira</i>	<i>-'abora</i>
Seen by both speaker and addressee	<i>-boba</i>	<i>-ba'a</i>	<i>-bo'owa</i>	(none)

[^a] Note the recurrent formal opposition between *-ra* in the first row and *-bo/ba* in the second. It is tempting to relate the *-ba* formative to the distal demonstrative free word *ba* in Foe; a *-ba* formative also occurs in other nominalisations, namely those making statements determined on grounds of present evidence. The corresponding proximal demonstrative is *-to* (Rule 1977, p. 19), and the only way to relate this to *-ra* would be by means of some change like *-to > -ro > -ra*.

Sillitoe (2010) outlines how persuasion in Wola society is only regarded as effective if the status of propositions can be epistemically upgraded, through conversation and praxis, from the 'witnessed by speaker' to the 'witnessed by speaker and addressee' categories. Understanding how this epistemic distribution interacts with evidentials, he argues, is crucial for development workers in countries like Papua New Guinea: only by understanding the operation of grammatical markers of who knows what can they establish plausibility and trust in the message they wish to convey:

[I]n parts of the Papua New Guinea highlands where the authority of the nation-state is weak to non-existent ... participation (featuring *bisumindis* 'we do, both parties witness' knowledge) will be necessary if development initiatives are to have any hope. Agencies will have not only to involve people but also to demonstrate the effectiveness of their views and proposals. People will not heed what others direct as best unless they can 'see' – i.e. think or know – that it will work for them. They are suspicious of experts (with, at best, their *biso*, 's/he did, speaker only witnessed', knowledge) given a propensity to question the necessary validity of others' experience and only fully to trust in their own, paying heed to what they 'see' themselves. (Sillitoe, 2010, p. 26)

In the evidential systems found in the New Guinea Highlands, markers that indicate awareness of the addressee's *visual* experience, or lack of it, thus appear to be especially prominent. This suggests the comparative ease of assessing whether or not an addressee was an eyewitness of some event, as opposed to more 'hidden' mental processes such as inference and assumption (see also San Roque et al., 2017). The Papuan language Duna (which neighbours the Engan language family) shows a spin on this tendency by including an inflection (*-noko ~ -naoko*) that does not make a definitive claim about the addressee's visual experience, but suggests that he or she *could* have seen something that the speaker already knows about. An example is shown in (11). The (hypothetical) context is that Speaker A has asked B if they went

to the market, and B has said they did, in company with another person (Mary). Speaker A finds this surprising, as she saw Mary but not Speaker B. Speaker B asserts that nevertheless they were there in plain view.

- (11) A: *ko no na-ke-ya, Mari no ke-o.*
 2SG 1SG NEG-see-NEG PSN 1SG see-PFV
 ‘You I didn’t see, Mary I saw.’
 B: *neya=nia, no ngo-naoko.*
 not=ASSERT 1SG go-POT.OBS
 ‘No, I went (you could have seen me).’

Utterances marked with this inflection are often functionally interpreted as questions concerning what the addressee has seen (San Roque, 2008). In (12) this implicit question (‘did you see?’) is made explicit. In this hypothetical context, the speaker is relatively certain that the addressee would have walked past the burned school building in order to reach the place where they are now talking.

- (12) *skul-anda khira-noko, ke-o=pe.*
 school-ENCL burn-POT.OBS see-PFV=Q
 ‘The school burned, did you see it?’

In some instances the Duna ‘potential observation’ inflection thus appears to instruct the addressee to reflect on and perhaps to talk about their visual experience (see also San Roque, 2015). It may be that this is one of the important pragmatic functions of addressee-oriented visual evidentials more generally.

Outside of New Guinea, evidential systems that include a contrast between exclusive speaker knowledge as opposed to inclusive, shared knowledge have been briefly described for several languages of South America, such as Jaqaru (Hardman, 1986) and Southern Nambikuara (Kroeker, 2001). For example, according to existing analyses Southern Nambikuara distinguishes between ‘individual’ (speaker-based) and ‘collective’ (speaker + hearer-based) observation, using the suffix *-na²* in the first case and *-ti²tu³* in the second (subject to different tense distinctions). Compare: *wa³kon³na²la²* ‘He worked today (I saw it, but you didn’t)’ (Kroeker, 2001, p. 63) vs. *wa³kon³tait²ti²tu³wa²* ‘He worked (we both saw it)’ (Lowe, 1999, p. 276). More recently, a related contrast has been discussed for the Tibeto-Burman language Kurtöp (Hyslop, 2014).

The evidential markers discussed so far are described as encoding a specific kind of information source (e.g., direct observation) that (the speaker claims) an addressee has for an event. Contrasts relevant to engagement can also be embedded within what have been analysed as evidential systems in other ways, without identifying the exact nature of the addresser’s evidence. For example, according to Willett (1991, pp. 162–165), evidentials of Southeastern Tepehuan mark (i) the information source of the speaker and, in the reported category only, (ii) whether (the speaker claims that) the proposition is old or

new knowledge for the addressee. The particle *sap* is used for “reported evidence previously unknown to the hearer” (13), whereas *sac* is used for reported evidence where “the speaker reminds the hearer of information he already knows the hearer is aware of” (14). Willett notes that *sac* is much less frequent than *sap* in both conversation and folklore narratives, suggesting that it may be a situationally and pragmatically marked choice.

- (13) *Oidya*-’-*ap* *gu-m* *tat.* *Jimi-a*’ *sap* *para*
 go.with-FUT-2SG ART-2SG father, go-FUT REU to
Vódamtam cavuimuc.
 Mezquital tomorrow
 ‘(You should) accompany your father. He says he’s going to Mezquital tomorrow.’
 (Willett, 1991, example (465))
- (14) *Va-jípir* *gu-m* *bí* *na-p* *sac* *tu-jugui-a*’.
 RLZ-get.cold ART-2SG food SUB-2SG REK EXP-eat-FUT
 ‘Your food is already cold. (You said) you were going to eat.’
 (Willett, 1991, example (471))

An important thing to note in the Tepahuan case is that, whereas the epistemic channel by which the speaker gained their knowledge is explicitly identified as reported, that of the addressee is unspecified. In this respect, the assessment of evidential source as between speaker and addressee is less clearly symmetric than in such examples as Foe. Rather, the assessment of addressee knowledge seems to be straying into the (embattled) territory of mirativity, the marking of knowledge as new or unexpected, as already mentioned in relation to Kakataibo, above. We are yet to note a fully-fledged grammatical system that paradigmatically distinguishes (a)symmetric combinations of mirativity and engagement (e.g., with such specifications as ‘this is news to both of us’ versus ‘this is old news for you, but news to me’). However, the potential for a language to have dedicated addressee-oriented mirative markers (‘this is news for you!’) has received more attention of late (e.g., Hengeveld & Olbertz, 2012; Mexas, 2016; see also Gossner, 1994), and an interest in the general problem goes back at least to discussions of the ‘hot news’ use of the English perfect by McCoard (1978) and McCawley (1981), of the type *Malcolm X has just been assassinated*. This suggests that the newness of knowledge of some state of affairs may be coded independently for speaker and hearer in some grammars.

To take a different approach again, Hintz and Hintz (2017) describe how in South Conchucos Quechua the category of ‘mutual knowledge’ between speaker and addressee actually has a dedicated marker (the morpheme *-cha:*) within the evidential system. The exact nature of the source for this mutual knowledge can be quite varied, so there is a focus on the end state of shared awareness, rather than on the way this knowledge was acquired. (‘This could

even be interpreted as a *non*-mirative marker in relation to speaker and addressee.) They also describe the evidential system of another variety, Sihuas Quechua, where an ‘individual’ vs. ‘mutual’ contrast is available for all evidential contrasts, symmetrically organised so that *-i* indicates ‘individual’ and *-a* ‘mutual’. Summarising the interaction of evidence type and its epistemic distribution, they conclude:

[I]nformation sources for the evidential category of mutual knowledge include the contributions of conversational participants, the beliefs and assumptions of the participants when interpreting shared experiences, and what members of the speech community can be expected to know about the world. Speakers use individual knowledge evidentials to introduce information and then use mutual knowledge evidentials once the fact has been established by consensus. (Hintz & Hintz, 2017, p. 107)

The South Conchucos Quechua case shows similarities to Kogi, but in the Quechua variety this category is marked in contrast to evidential values such as ‘reported’, rather than being part of a paradigm that deals primarily with epistemic (a)symmetry.

Like Andoke and Kogi, all of the languages discussed above have developed morphemes that encode a range of epistemic configurations between speaker and hearer, but intertwined with the evidence for a proposition rather than simply for the proposition itself. Communicatively they can be used for such functions as to remind the addressee of shared knowledge and experience, to highlight the speaker’s more exclusive access to a particular event, to acknowledge or direct the addressee’s attention to relevant evidence, or to confirm the status of information as mutually known and agreed upon.

As has been extensively discussed and disputed in the literature, there is a close relationship between the semantic domains of information source and *certainty*, and thus, the grammatical categories of evidentiality and epistemic modality (e.g., Aikhenvald, 2004; Chafe & Nichols, 1986; Palmer 2001). Similarly as for evidentials in Foe and other languages, languages may offer options for a speaker to encode whether an epistemic modal value (e.g., certain, probable) is assumed to be shared by the addressee.

One example of this is found in the language Yurakaré (Gipper 2011, 2015). Yurakaré has two morphemes, =*ya* and =*laba*, both of which indicate that “the speaker considers the proposition to be possibly or probably true” (Gipper, 2015). The difference between them is that the ‘intersubjective’ =*ya* is used with assertions where the speaker expects the addressee to share his or her belief, whereas the ‘subjective’ =*laba* does not express any assumptions concerning the addressee’s state of mind. Gipper (2015) describes how this difference in meaning has consequences for the distribution of the two markers: intersubjective =*ya* is typically found in situations of ‘symmetric’

knowledge, where both speaker and addressee have equal access to the information upon which the judgement is based. Her findings are based on quantitative and qualitative analyses of a video corpus of approximately 5.25 hours of (mostly dyadic) conversation. An example is shown in (15), where two speakers discuss the state of the lagoon in their village.

(15) Yurakaré [160906_conv]

M: ((turns his head, chin-points to the lagoon outside))

ujmanaj tishi kadyimta (.) tajudawa=
ujwa-ma=naja tishilë

look-IMP.SG=NEW.SITUATION NOW

ka-dyimta-ø ta-kudawa

3SG.OBJ-subside-3SG.SBJ 1PL.POSS-lagoon

‘Look, the water in our lagoon has subsided.’

P: =*të* *bij*: [*binta dyimta kompadre yosse*]

të bij~binta dyimta-ø kompadre yosse

INTJ INTS~strong subside-3SG.SBJ compadre(SP) again

‘Yes, it has subsided very much again.’

M: [*të::j*] (0.7)

të

INTJ

‘Yes.’

P: *namashtay tajudawa yosse*

nama-shta-ø=ya ta-kudawa yosse

dry-FUT-3SG.SBJ=INTSUBJ 1PL.POSS=lagoon again

‘Probably our lagoon will dry out again.’

By contrast, the subjective marker =*laba* is commonly used in both symmetric and asymmetric contexts, as the addressee’s knowledge state is not at issue. An example with an asymmetric context is shown in (16), where the addressee has superior access to the information in question: the epistemic perspectives of speaker and addressee are disparate, not shared, and the intersubjective marker =*ya* would not be appropriate:

(16) Yurakaré [290906_convI]

A: *batanlab tishil na loma alta (chi)* ((gaze to addressee)) (.)

bata-m=laba tishilë naa loma alta=chi

go.FUT-2SG.SBJ=SUBJ now DEM Loma Alta=DIR

‘You are going to Loma Alta today, I think?’

E: *nijtala*

nijta=la

NEG=COMM

‘No.’

Gipper (2015) further notes (among other findings) that =*ya* is used comparatively more frequently than =*laba* in ‘agreeing responses’, where the speaker agrees with what has just been said, and (unlike =*laba*) is never used in disagreeing responses. She argues that, as an intersubjective marker, =*ya* is highly compatible with agreeing responses because these are situations where “a shared epistemic perspective is explicitly expressed”. By the same token, =*ya* is not appropriate to disagreeing responses, where the epistemic perspectives of speaker and addressee are explicitly at odds.

A further example of engagement combined with epistemic modality is found in the Tibeto-Burman language Kinnauri (Saxena, 2000). In this case, the copula *ni* expresses *contrastive* values of speaker and addressee certainty, being used where the speaker is confident about what they are asserting, against the addressee’s perceived doubts. In (17), *to* would be used “when Sonam is either a family member of the speaker, or is presently with the speaker. *Du* is used when Sonam is not a family member of the speaker, nor is ... in physical proximity to the speaker. *Ni* is used if the hearer has some doubts about Sonam being a good person and the speaker knows that she is a good person” (Saxena, 2000, p. 473). While the first two copula forms contrast different degrees of authority / epistemic access on the part of the speaker, the third form combines an authoritative positive modal assessment by the speaker with an assumption that the addressee does not share this assessment.

- (17) *Sonam* *dam to* / *du* / *ni*
 [proper.name] good be1:PRES / be2:PRES / be3:PRES
 ‘Sonam is good.’

Overall, then, various additional qualities of knowledge (evidence, oldness/newness, certainty) can be expressed not only in regard to a single perspective, but also in regard to both speaker and hearer, and/or as a relation between them. There is no reason to assume that the expression of engagement is limited to these specific qualities, but we can rather expect that many other aspects of the mental directedness of interlocutors can be grammaticalised (§5). At the same time, however, we note that it is very unusual to find a comprehensive grammatical system of engagement and evidential (etc.) contrasts. That is, the full range of logical possibilities (e.g., speaker saw the event, hearer saw the event, both saw it, neither saw it; speaker inferred the event, hearer inferred ... etc.) is rarely, if ever, morphologically differentiated within a single paradigm. This rarity of bidimensional systems may reflect the regular correlation, in most situations, between accessibility and evidence: direct access allows direct evidential reading, lack of direct access means that an assertion is founded on some form of evidence other than current mutual accessibility.

4. Engagement, level-shifting, and diachrony

Part of our rationale in progressing from demonstratives through engagement markers operating at clausal level, and on to markers with evidence/certainty in their scope, has been that the same processes of mutual coordination are at work, whatever the level in terms of syntactic or semantic structures. Up to this point, however, we have not examined languages where this connection is made clear. But we now pass to a Papuan language, Abui (Kratochvil, 2011a, 2011b), which illustrates the connections remarkably clearly thanks to the way it deploys its demonstratives with various levels of syntactic scope – a way somewhat reminiscent of how some Australian languages deploy case-suffixes at various syntactic levels (embedded NP, NP at clause level, embedded clause) with differential semantic effects; see also Schapper and San Roque (2011) concerning clause-level demonstratives in other Timor-Alor-Pantar languages.

We have already surveyed, in §5 of Part I, an interesting system of basic demonstratives in Abui, which recombines the proximal vs. medial distinction with both speaker and addressee anchor-points. In doing so, the language draws on two sets, a ‘basic’ set which most commonly functions adnominally and situates individual entities, and an ‘adverbial’ set which situates states of affairs more generally and has meanings like ‘be here’, ‘be there near you’, etc. (though they are not true verbs in the sense of being able to be used alone).

We will now see that, by applying these demonstratives with sentential scope, a range of engagement-type meanings can be coerced. Note that the engagement-related meanings are only a subset of the very rich range of metaphorical extensions found with the Abui demonstrative system – others, which we do not discuss, include their use to indicate tense and various kinds of modality.

Both basic and adverbial Abui demonstratives can be used in ways that are relevant to engagement. From the adverbial set (shown in the right half of Table 3), “the addressee-based forms are used when the speaker wants to evaluate or interact with addressee’s perspective” (Kratochvil, 2011a, p. 8). For example, say the addressee and the speaker are sitting in a traditional house with a leaking thatched roof. The speaker inquires whether the addressee is affected by the rain (there are no windows and it’s dark inside). Since they are together, he may simply say (18a). However, it is also possible to say (18b), using the addressee-proximal form *ta* to specifically invite the addressee’s assessment of the quality of the thatched roof above where the addressee is seated.

- (18) a. *anui ma* *o-pa=ng* *sei?*
 rain be.SP.PRX 2SG.RECIPI-touch.IPF=see come.down.CONT
 ‘Is it raining on you here?’
- b. *anui ta* *o-pa=ng* *sei?*
 rain be.AD.PRX 2SG.RECIPI-touch.IPF=see come.down.CONT
 ‘Is it raining on you here (where you are)?’

TABLE 3. *Basic and adverbial demonstratives in Abui (omitting elevation-based forms for adnominal demonstratives)*

Distance	Speaker-viewpoint	Addressee-viewpoint	Speaker-viewpoint	Addressee-viewpoint
Proximal	<i>do</i>	<i>to</i>	<i>ma</i>	<i>ta</i>
Medial	<i>o, lo</i>	<i>yo</i>	<i>la</i>	<i>fa</i>
Distal		<i>oro</i>		<i>ya</i>
Demonstrative type		Basic		Adverbial

The addressee-based medial *fa* is used to indicate non-proximate location with respect to the addressee. Typically, this occurs when “the speaker wants to stress that the addressee is in another place or not aware of the location of an event or participant” (Kratochvil, 2011a, p. 9). For example, in performing a ‘matching task’, the speaker may be describing a picture to the addressee, so that he can match the description to a picture in the set he was given. Here “the speaker uses *fa* to locate two balls on the picture that the addressee is unable to see”:

- (19) *kaan-r-i, bal do fa ayoku*
 good.CPL=reach-PFV ball SP.PRX be.AD.MED be.two
 ‘right, there are two balls there’
 [perhaps a closer translation would be ‘right, these balls (i.e., “these” on my picture) there’s (a picture) there (on your side) (where) there are two (of them)’]

At a higher syntactic and semantic level, members from the basic set can be placed in sentence-final position to index the distribution and extent of knowledge among speech-act participants. Speaker-proximal *do* can stress the speaker’s foundation for his assertion in immediate experience (20):

- (20) *na nala nee-ti beeka do*
 1SG.A something.eat-PHSL bad cannot SP.PRX
 ‘I couldn’t eat up (swallow) anything.’

In questions, the addressee-based medial form can be used to appeal to what they may know of a situation, while the addressee-proximate form, if used with exclamatory force, can indicate that the question is redundant and that the information should be available to the addressee, thus functioning as a reproach – invoking both a type of evidence (perception) and a judgment about what the addressee could vs. did perceive. This is reminiscent of the Duna *-noko* suffix discussed above.

- (21) A: *mangmat*, *ma* *e-ya* **yo?**
 foster.child SP.PRX 2SG-mother AD.MED
 ‘Child, what about your mother?’
 B: *ni-ya* *ha-rik* **to!**
 1PL.EXC-mother 3PATIENT-hurt AD.PRX
 ‘My mother is sick (as you could see).’

The addressee-medial form, likewise, may be used in sentence-final position in a reproachful way – in this context, “the speaker stresses that the addressee knew about the funeral and yet failed to attend” (Kratochvil, 2011b, p. 773).

- (22) *pi* *yaar-i* *ni-ya* *do* *nabuk* **yo**
 1PL.INC go-PFV 1PL.EXC-mother SP.PRX bury AD.MED
 ‘We went to bury our mother (as you could have known).’

The essence of the Abui system of recycling demonstratives is thus to shift their function upward, from coordinating attention to objects, in their basic use, to coordinating attention to states of affairs and their epistemic status, in the extended uses we have discussed (examples (20)–(22)). It is not unreasonable to see the unusual starting point of the basic system – which separates the proximal vs. medial contrast from that between speaker and addressee anchor-point – as providing an ideal semantic affordance for the extension into the more general management of epistemic gradients between speaker and hearer.⁷

In the case of Abui, the demonstratives remain as separate words even as their function and syntactic position shifts to higher scopes. However, an interesting case where original demonstratives turn into verbal prefixes encoding semantic values of engagement is found in Marind, a language of Southern New Guinea (Olsson, 2016). In the present tense, Marind features two sets of verbal subject prefix complexes, encoding person, number, gender, and a category Olsson terms ‘absconditive’ (< Latin *absconditus* ‘hidden, concealed’), which are “used to establish joint attention, by instructing the ADR to ‘align’ her attention with Spr’s, and thereby get access to previously unavailable information” (p. 3). Summarising Olsson (2016), the two main circumstances in which absconditive-series prefixes are used are when the speaker:

- (i) “wants to draw attention to something outside ADR’s visual focus”
 (p. 1), e.g. when a speaker tells a child’s mother that the child’s nose is snotty, something the mother cannot see because the child is sitting on her lap, and

[7] Kratochvil also mentions the use of the remote demonstrative, in sentence-final position, to mark “a reliable and recognised source, such as the tradition of ancestral stories” (2011a, p. 18). This then gives examples of four out of the five demonstratives. He does not give an example with the speaker-medial form used in this syntactic position.

- (ii) to “‘update common ground’ by denying Adr’s presuppositions” (p. 1), e.g. when someone tells an old woman that she should be talking Marind to the linguist so that he can learn, and the woman retorts that she is indeed doing that, using the absconditive in a way that would be translated into English as ‘I AM talking to him’ or ‘BUT I AM SO talking to him’.

What is relevant to our argument here is that the forms of absconditive prefixes can be broken down into two parts: an initial gender element, and a second deictic element identical in form to demonstratives. Interestingly, the use of the absconditive can be triggered either by the addressee’s (non-)attention to an entity, or to a state of affairs; it appears that the proximate vs. distal semantics of the deictic element is primarily exploited when the location of an entity is involved. Where the focus is on a state of affairs, the one example given by Olsson employs the form derived from the distal form.

The Marind absconditive is thus intriguingly parallel to the level-shifting trajectory we saw for Abui, but in a way takes it further by grammaticalising the deictic elements into actual prefixes on the verb. In doing this, it illustrates one grammaticalisation path by which verbs can evolve engagement morphology. What these two languages clearly demonstrate is the logical link between achieving mutual attention to objects in the here-and-now, and the more abstract job of producing convergence of epistemic positioning between speaker and addressee.

5. Conclusion

We have tried to shatter the illusion that definite reference is simple and self-evident by demonstrating how it requires mutual knowledge, which complicates matters enormously. But virtually every other aspect of meaning and reference also requires mutual knowledge, which also is at the very heart of the notion of linguistic convention and speaker meaning. Mutual knowledge is an issue we cannot avoid. It is likely to complicate matters for some time to come.

(Clark & Marshall, 1981, p. 58)

The languages we have surveyed illustrate the proposition with which we began this paper: that it is possible for languages to place epistemic coordination systems right in the heart of their grammars. Languages like Andoke and Kogi have paradigmatically structured categories that show the speaker’s epistemic access, and their assessment of that of the addressee, as potentially independent variables to be monitored and appealed to as conversation unfolds. Such languages thus place, at the core of the grammatical system, the central role of

dialogue as an ongoing transaction in which mutual attention and knowledge is closely monitored and repeatedly recalibrated.

The grammaticalisation of epistemic assessment is not virgin territory to linguistic investigation. There are long-standing traditions for investigating the modelling and updating of mutual knowledge that is needed to successfully use a system of definite articles (see, e.g., Clark & Marshall, 1981; Epstein, 1997; Verhagen, 1986) and discourse particles of an epistemic nature (e.g., Enfield, Brown, & de Ruiter, 2013; Hayano, 2012; Simon-Vandenbergen & Aijmer, 2007; Verhagen, 2005, ch. 4). There have also been a growing number of studies illustrating the ways in which speaker assessment of addressee attention can be built into demonstrative systems, as we illustrated in §5 of Part I. What has remained unclear, however, has been the way that comparable intersubjective assessment can be integrated into grammatical paradigms with scope over clauses or propositions, or even potentially evidential qualification (§3), depending on whether we characterise scope syntactically or semantically.

As in many other areas of typology, it is useful to set up canonical cases as clear conceptual reference points (cf. Brown, Chumakina, & Corbett, 2013). The systems found in Andoke (Part I, §2) and Kogi (this Part, §2.1) demonstrate with particular clarity what a canonical system of engagement with clausal scope looks like, because of the symmetry with which they independently assign positive and negative epistemic values to speaker and addressee.

On the other hand, we also find languages that exhibit only some of the characteristics of canonical engagement paradigms—just as we find departures from semantic purity in virtually every grammatical category, e.g., the much better-known dimension of tense, with its cross-linguistically variable differences in degree of structuration, from neat paradigms to relatively unintegrated free words, strung out along a grammaticalisation trajectory including more heterogeneous options such as systems that mix in periphrasis. Kakataibo (§2.2) was presented as an example in which engagement is grammaticalised in a less canonical way: it includes a number of values, on verbal inflections and second position clitics, that correspond to key values in canonical engagement systems, but compared to Andoke and Kogi they are less integrated into a single, symmetric paradigm.

The same point about variability in canonicity may be made in terms of grammaticalisation, since the emergence of one category (here: engagement) from another (e.g., demonstratives) is typically marked by phenomena exhibiting transitional or mixed status. An interesting example of this is the grammaticalisation of engagement examined for Abui in §4, which lifts the speaker vs. addressee x proximal vs. medial contrast found in its basic demonstratives and reapplies it at clausal level to produce an engagement system with propositional scope, though one in which the

relevant markers (demonstratives in sentence-final position) remain transparently multifunctional without becoming a specialised grammatical system as they are in Andoke and Kogi. While the Abui case provides a good example of engagement categories appearing to have been recruited from demonstratives, it is unlikely that this is the only diachronic source: other candidates include time adverbials in Lakandon Maya⁸ (Bergqvist, 2008, in press), pronominal clitics in Jaminjung/Ngaliwuru (Schultze-Berndt, 2017),⁹ and “ethical datives”, also called ‘non-selected arguments’ (Bergqvist & Kittilä, 2017; cf. Bosse, Bruening, & Masahiro, 2012).

We have taken the canonical case of engagement as a grammatical system for encoding the relative mental directedness of speaker and addressee towards an entity or state of affairs – thus allowing knowledge and attention (etc.) to be tracked and dynamically updated in discourse. This leads naturally to the question of what is the full set of typological dimensions involved? In this paper we have focused on two – the set of permutations of epistemic authority across the speaker and addressee, and the semantic and syntactic level at which this applies – (i) deictically indicated entity (demonstratives), (ii) state-of-affairs/proposition/clause, and (iii) metaproposition in the case of certain evidentials. But other syntactic levels and semantic dimensions may also prove relevant.

One promising dimension for future investigation concerns the interaction of engagement values with tense/time. In other words, is the monitoring of relative epistemic authority/directedness confined to the here-and-now, or can it be displaced? For example, work by Fleck (2007) on the Peruvian language Matses has shown that the psychological event of inferring an action from evidence can be located in time independently of the speech event or the reported event (e.g., recently or long ago, I may have seen the tracks of a peccary that crossed the path; and that path-crossing may have been immediately prior to or a long time before I saw the tracks, generating a four-way system of tensed evidentials in Matses).

[8] Southern Lakandon has two time adverbials, *uich* and *kuich*, which originally featured a semantic contrast between ‘long ago’ and ‘recently’. These have developed into a semantic contrast between ‘a past event that is unknown to the addressee’ (*uich*) and ‘a past event that is known to the addressee’ (*kuich*). Bergqvist (in press) details this development as an instance of “intersubjectification” (Traugott & Dasher, 2002).

[9] The cliticisation of absolutive pronoun enclitics to inflected verbs, in addition to a prefixal layer encoding actual arguments, is reported by Schultze-Berndt (2017) for Jaminjung/Ngaliwuru: the absolutive first person singular *ngarndi* signals the speaker’s exclusive epistemic authority, while the first person inclusive, *mirndi*, marks shared epistemic authority between the speaker and the addressee. In Jaminjung/Ngaliwuru, the function of absolutive markers as P arguments in transitive clauses aligns with their subsequent epistemic function, namely to signal epistemic authority over an event that involves the speaker and/or the addressee as observers and experiencers. This specific development is conceptually comparable to evidential forms that feature engagement semantics, albeit with a focus on claim of epistemic authority rather than source of information.

A priori, we may expect the object of presumed attention or knowledge to be likewise locatable in time. We have already seen a hint of this in the contrast between Kakataibo *-riapa* and *-riapëñë*, where in both cases the addressee is presumed to be unaware of what the speaker reports, but the speaker has discovered the fact at different times – earlier in the case of *-riapa*, vs. at the moment of speech in the case of *-riapëñë*.

A second dimension for future investigation concerns the type of mental disposition involved. In our discussion throughout this paper we have focused on attention and knowledge. But the very rich literature on epistemic clitics and discourse markers focuses on other cognitive dispositions – particularly belief and expectation – and there is a long tradition of investigating their use as key argumentative resources to manage and overcome divergences in the belief states of speaker and addressee in the unfolding discourse, such as Foolen (2003) on Dutch *toch*, Hayano (2011, 2012) for Japanese *yo*, Schwenter (1996) on Spanish independent *si* clauses, Wilkins (1986) for several Mparntwe Arrernte particles, Sekiguchi (1977 [1939]) for German *doch* and Leiss (2012) for the German particle *ja*, and Matthews and Yip (2011) on a variety of Cantonese particles. Evaluative attitude – like and dislike with respect to the event – as well as emotional disposition such as fear (in categories like the apprehensive) are also relevant parameters worth exploring.

In many well-known cases, such as German *doch* and Italian *mica* (Cinque, 1991; Visconti, 2009), there is a statistical bias towards an interpretation where the speaker asserts a state of affairs to hold, against a contrary belief imputed to the addressee (*Er ist doch hier!* ‘But he IS here!’, *Non è mica freddo!* ‘But it’s not cold at all’). But this alignment is not a necessary one, and it is also possible to use these particles in circumstances where the particle signals that it was the prior expectations of the speaker him or herself which turn out to be incorrect. It will now be interesting to revisit the study of particles from the perspective of more tightly structured systems of grammaticalised engagement marking, focusing on the extent to which they form tightly integrated systems patterning with the dimensions we have presented here.

Determiners of noun phrases are a third obvious dimension for the investigation of engagement, and indeed, as seen in our quote from Hawkins in Part I, §4,¹⁰ to use the determiner system in English or similar languages the speaker “must constantly take into consideration knowledge of various kinds which he assumes his hearer to have”. We also know that determiners can “escape from the noun phrase” (Epps, 2009, p. 87) to take scope over a clause as stance markers, like the Abui demonstratives (see also Yap, Grunow-Hårsta, &

[10] Hawkins (1978, p. 97): “the speaker when referring [and choosing between definite and indefinite articles – authors] must constantly take into consideration knowledge of various kinds which he assumes his hearer to have” (see Part I, §4).

Wrona, 2011). Now we know about the sort of four-way set of epistemic contrasts we saw earlier for Andoke and Kogi, we can ask whether such rich systems are also found in determiner systems. We can see that the English indefinite article is ambiguous between readings where the referent is not identifiable to the speaker or the addressee, vs. identifiable to the speaker but not the addressee – as with the ambiguity of ‘A man was waiting outside your door at 6 am’ between a situation where I know who the man is (following with ‘It was your brother.’), and one where I don’t know either. Some of this ambiguity can be removed with less grammaticalised means, such as ‘a certain’, as in ‘a certain colleague of mine always reacts that way’. In languages like Russian one can draw on words like *nekto*, to give expressions like *nekto čelovek* ‘a certain man, whose identity I know, but who I assume you don’t’ (cf. explication in Wierzbicka, 1980, p. 326).

We close this paper with a final unanswered question. The sorts of epistemic management mechanisms we have illustrated, in the pointedly grammaticalised forms we have been calling engagement, have been widely investigated in the conversation analysis literature, but in the languages examined there the formal coding is much more diffuse – involving prosody, gesture, tactical restatement, or the use of epistemic particles or adverbials like *well* or *actually*. What difference does this semiotic investment make? For example, are speakers of languages with engagement markers dragged into monitoring relative epistemic state much more frequently, even obligatorily? (A related issue, which current descriptions don’t fully resolve, is how far the marking of engagement is obligatory as opposed to strongly encouraged.) Alternatively, could the effects go the other way, with the smaller palette of a grammaticalised system offering fewer alternate ways of organising the task of epistemic management? Or could it simply make no difference – the epistemic management tasks go on being handled just the same, whether there is a grammaticalised system of engagement or not? As a next step in the research, we need studies of naturalistic conversation, closely analysed for the attentional states of both parties, across a sample of languages that includes some with canonical engagement systems. Only then can we understand the full import of these fascinating linguistic systems for the interface between grammar, intersubjectivity, and the management of interaction.

Abbreviations

1: first person, 2: second person, 3: third person, A: Actor, ABL: ablative, ABS: absolutive, AD.PRX: addressee proximal, ADR: addressee, AGT: agent, ART: article, ASSERT: assertion, ASYM: asymmetric, BENA: benefactive type A = unknown to beneficiary, BENB: benefactive type B = known to

beneficiary, COMM: commitment, CONT: continuative, DAT: dative, DEF: definitive, DEM: demonstrative, DIR: directional, ENCL: enclosure, ERG: ergative, EXC: exclusive, EXP: expected knowledge, FOC: focus, FP: far past, FUT: future, GEN: genitive, HAB: habitual, IMP: imperative, IMPF: imperfective, INC: inclusive, IND: independent, INT: interrogative, INTJ: interjection, INTS: intensifier, INTSUBJ: intersubjective, IPF: imperfective, KTS, known to speaker, LOC: locative, MAL: malefactive, MIR: mirative, NAR: negative for addressee, NEG: negative, NMZ: nominaliser, O: object, OBJ: object, PERF: perfective, PFV: perfective, PHSL: phasal (roughly: 'after'), PL: plural, POSS: possessor, POT: potential, POT.OBS, potential observation, PRES: present, PROG: progressive, PROX: proximal/proximate, PSN: personal name; PST: past, PTCP: participial, PTCY, participatory evidential, Q: question, RECIPI: recipient, REK: reported evidence known, REM: remote, REU: reported evidence previously unknown, RLZ: realization, S: intransitive subject, SBJ: subject, SG: singular, SPKR: speaker, SP.PRX: speaker proximal, SSA: seen by speaker and addressee, SUB: subjunctive, SUBJ: subjective, SYM: symmetric, UNM: unmarked.

REFERENCES

- Aikhenvald, A. Y. (2004). *Evidentiality*. Oxford: Oxford University Press.
- Aikhenvald, A. Y. (2015). Evidentials: their links with other grammatical categories. *Linguistic Typology*, **19**(2), 239–277.
- Benveniste, E. (1966). *Problèmes de linguistique générale* (Vol. 1). Paris: Gallimard.
- Bergqvist, H. (2008). *Temporal reference in Lakandon Maya: speaker and event perspectives*. Unpublished doctoral dissertation, University of London.
- Bergqvist, H. (2011). Complex perspectives in Arwako languages: comparing epistemic marking in Kogi and Ika. In P. K. Austin, O. Bond, D. Nathan, & L. Marten (Eds.), *Proceedings of conference on language documentation & linguistic theory 3* (pp. 49–57). London: SOAS.
- Bergqvist, H. (2016). Complex epistemic perspective in Kogi (Arwako). *International Journal of American Linguistics*, **82**(1), 1–34.
- Bergqvist, H. (in press). Time and commitment: the grammaticalization of uúch in Lakandon Maya. In V. Vapnarsky & F. de-Pierrebourg (Eds.), Special Issue: Mesures et Textures du Temps chez les Mayas: le dit, l'écrit et le vécu, *Journal de la Société des Américanistes*.
- Bergqvist, H., & Kittilä, S. (Eds.). (2017). Person and knowledge: introduction. *Open Linguistics*, **3**, 18–30.
- Boas, F., & Deloria, E. (1941). *Dakota grammar* (Memoirs of the National Academy of Sciences 23). Washington: US Government Printing Office.
- Bosse, S., Bruening, B., & Masahiro, Y. (2012). Affected experiencers. *Natural Language & Linguistic Theory*, **30**(4), 1185–1230.
- Brown, D., Chumakina, M., & Corbett, G. (Eds.). (2013). *Canonical morphology and syntax*. Oxford: Oxford University Press.
- Brugman, C. M., & Macaulay, M. (2015). Characterizing evidentiality. *Linguistic Typology*, **19**(2), 201–237.
- Chafe, W., & Nichols, J. (1986). *Evidentiality: the linguistic coding of epistemology*. Norwood, NJ: Ablex.
- Cinque, G. (1991). *Teoria linguistica e sintassi italiana*. Bologna: Mulino.

- Clark, H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge: Cambridge University Press.
- Du Bois, J. (2007). The stance triangle. In R. Englebretson (Ed.), *Stancetaking in discourse* (pp. 139–182). Amsterdam: Benjamins.
- Enfield, N., Brown, P., & de Ruiter, J. (2013). Epistemic dimensions of polar questions: sentence final particles in comparative perspective. In J. P. de Ruiter (Ed.), *Questions: formal, functional and interactional perspectives* (pp. 193–221). New York: Cambridge University Press.
- Epps, P. (2009). Escape from the noun phrase: from relative clause to converb and beyond in an Amazonian language. *Diachronica*, **26**(3): 87–318.
- Epstein, R. (1997). Viewpoint and the definite article. In A. E. Goldberg (Ed.), *Conceptual structure, discourse and language* (pp. 99–112). Stanford: CSLI.
- Fleck, D. (2007). Evidentiality and double tense in Matses. *Language*, **83**(3), 589–614.
- Foolen, A. (2003). Niederländisch *toch* und Deutsch *doch*: gleich oder doch nicht ganz. *Linguistik online* 13, 1/03:85–101 <http://www.linguistik-online.de/13_01/foolen.html>.
- Gipper, S. (2011). *Evidentiality and intersubjectivity in Yurakaré: an interactional account*. Unpublished doctoral dissertation, Radboud Universiteit, Nijmegen.
- Gipper, S. (2015). (Inter)subjectivity in interaction: investigating (inter)subjective meanings in Yurakaré conversational data. *Sprachtypologie und Universalienforschung (STUF)*, **68**(2), 211–232.
- Gossner, J. (1994). *Aspects of Edolo grammar*. Retrieved from UMI Dissertation services, Ann Arbor, MI: University Microfilms.
- Gutiérrez, A., & Matthewson, L. (2012). Evidential determiners: best (sensory) evidence. In E. Bogal-Allbritten (Ed.), *Proceedings of the sixth meeting on the semantics of under-represented languages in the Americas and SULA-Bar* (pp. 63–79). Amherst, MA: GLSA.
- Hanks, W. (1990). *Referential practice, language and lived space among the Maya*. Chicago: University of Chicago Press.
- Hardman, M. (1986). Data-source marking in the Jaqi languages. In W. Chafe & J. Nichols (Eds.), *Evidentiality: the linguistic coding of epistemology* (pp. 113–136). Norwood, NJ: Ablex.
- Hawkins, J. (1978). *Definiteness and indefiniteness: a study in reference and grammaticality prediction*. London: Croom Helm.
- Hayano, K. (2011). Claiming epistemic primacy: yo-marked assessments in Japanese. In T. Stivers, L. Mondada, & J. Steensig (Eds.), *The morality of knowledge in conversation* (pp. 58–81). Cambridge: Cambridge University Press.
- Hayano, K. (2012). *Territories of knowledge in Japanese interaction*. Unpublished doctoral dissertation, Radboud University.
- Hengeveld, K., & Olbertz, H. (2012). Didn't you know? Mirativity does exist! *Linguistic Typology*, **16**, 487–503.
- Hintz, D. J., & Hintz, D. M. (2017). The evidential category of mutual knowledge in Quechua. *Lingua*, **186/187**, 88–109.
- Hyslop, G. (2014). On the category of speaker expectation of interlocutor knowledge in Kurtöp. *Berkeley Linguistics Society (BLS)* **40**, 201–214.
- Jacques, G., & Lahaussais, A. (2014). The auditory demonstrative in Khaling. *Studies in Language*, **38**(2), 393–404.
- Kamio, A. (1997). *Territory of information* (Pragmatics and beyond new series 48). Amsterdam: Benjamins.
- Kratochvil, F. (2011a). *Demonstratives as markers of stance: evidence from Abui*. Unpublished manuscript.
- Kratochvil, F. (2011b). Discourse-structuring functions of Abui demonstratives. In F. H. Yap, K. Grunow-Härsta, & J. Wrona (Eds.), *Nominalization in Asian languages: diachronic and typological perspectives* (pp. 757–788). Amsterdam: Benjamins.
- Kroeker, M. (2001). A descriptive grammar of Nambikuara. *International Journal of American Linguistics*, **67**(1), 1–87.

- Landaburu, J. (2007). La modalisation du savoir en langue andoke (Amazonie colombienne). In Z. Guentchéva & J. Landaburu (Eds.), *L'énonciation médiatisée II: Le traitement épistémologique de l'information; Illustrations amérindiennes et caucasiennes* (pp. 23–47). Leuven: Peeters.
- Leiss, E. (2012). Epistemicity, evidentiality, and theory of mind (ToM). In W. Abraham & E. Leiss (Eds.), *Modality and theory of mind elements across languages* (pp. 39–66). Berlin: Mouton de Gruyter.
- Lowe, I. (1999). Nambiquara. In R. Dixon & A. Aikhenvald (Eds.), *Amazonian languages* (pp. 268–291). Cambridge: Cambridge University Press.
- Madden, B. (n.d.). [c. 1960]. *An introduction to Mendi grammar*. Unpublished manuscript.
- Manning, H. P. (2001). On social deixis. *Anthropological Linguistics*, 43(1), 54–100.
- Matthews, S., & Yip, V. (2011). *Cantonese: a comprehensive grammar*. London: Routledge.
- McCawley, J. D. (1981). Notes on the English perfect. *Australian Journal of Linguistics*, 1, 81–90.
- McCoard, R. W. (1978). *The English perfect: tense-choice and pragmatic inferences*. Amsterdam: North-Holland.
- Mexas, H. (2016). *Mirativity as realization marking: a cross-linguistic study*. Unpublished Master's thesis, University of Leiden.
- Olsson, B. (2016). *The Marind absolutive: attention alignment and knowledge asymmetries in a language of south New Guinea*. Paper presented at the 8th Austronesian and Papuan Languages and Linguistics conference, May 2016, SOAS, University of London.
- Palmer, F. R. (2001). *Mood and modality*, 2nd ed. Cambridge: Cambridge University Press.
- Reithofer, H. (2011). Skywalkers and cannibals: chanted tales in the Angal (Mendi) language area. In A. Rumsey & D. Niles (Eds.), *Sung tales from the Papua New Guinea Highlands: studies in form, meaning and sociocultural context* (pp. 207–245). Canberra: ANU EPress.
- Rule, W. (1977). *A comparative study of the Foe, Huli and Pole languages of Papua New Guinea*. Sydney: Oceania Linguistic Monographs, University of Sydney.
- San Roque, L. (2008). *An introduction to Duna grammar*. Unpublished doctoral dissertation, Australian National University.
- San Roque, L. (2015). Using you to get to me: addressee perspective and speaker stance in Duna evidential marking. *Sprachtypologie und Universalienforschung (STUF)*, 68(2), 187–210.
- San Roque, L., Floyd, S., & Norcliffe, E. (2017). Evidentiality and interrogativity. *Lingua*, 186/187, 120–143.
- San Roque, L., & Loughnane, R. (2012a). The New Guinea Highlands evidentiality area. *Linguistic Typology*, 16, 111–167.
- San Roque, L., & Loughnane, R. (2012b). Inheritance, contact and change in the New Guinea Highlands evidentiality area. *Language and Linguistics in Melanesia: Special Issue 2012, Part II*, 397–427.
- Saxena, A. (2000). Evidentiality in Kinnauri. In L. Johanson & B. Utas (Eds.), *Evidentials in Turkic, Iranian and neighboring languages* (Empirical Approaches to Language Typology) (pp. 471–482). Berlin: Mouton de Gruyter.
- Schapper, A., & San Roque, L. (2011). Demonstratives and non-embedded nominalisations in three Papuan languages of the Timor-Alor-Pantar family. *Studies in Language*, 35, 380–408.
- Schultze-Berndt, E. (2017). Shared vs. primary epistemic authority in Jaminjung/Ngaliwurru. *Open Linguistics*, 3, 178–218.
- Schwenter, S. (1996). The pragmatics of independent *si*-clauses in Spanish. *Hispanic Linguistics*, 8(2), 316–351.
- Sekiguchi, T. (1977 [1939]). Was heißt *doch*? In H. Weydt (Ed.), *Aspekte der Modalpartikeln* (pp. 3–9). Tübingen: Niemeyer. (Originally published 1939.)
- Sillitoe, P. (2010). Trust in development: some implications of knowing in indigenous knowledge. *Journal of the Royal Anthropological Institute (N.S.)*, 16, 12–30.
- Simon-Vandenberg, A.-M., & Aijmer, K. (2007). *The semantic field of modal certainty: a corpus-based study of English adverbs*. Berlin: Mouton de Gruyter.

- Speas, M. (2004). Evidentiality, logophoricity and the syntactic representation of pragmatic features. *Lingua*, **114**, 255–276.
- Tipton, R. (1982). *Nembi procedural and narrative discourse*. Canberra: Pacific Linguistics.
- Traugott, E., & Dasher, R. (2002). *Regularity in semantic change*. Cambridge: Cambridge University Press.
- Verhagen, A. (1986). *Linguistic theory and the function of word order in Dutch*. Dordrecht: Foris.
- Verhagen, A. (2005). *Constructions of intersubjectivity: discourse, syntax and cognition*. Oxford: Oxford University Press.
- Visconti, J. (2009). From textual to interpersonal: on the diachrony of the Italian particle *mica*. *Journal of Pragmatics*, **41**, 937–950.
- Wierzbicka, A. (1980). *Lingua mentalis: the semantics of natural language*. Sydney: Academic Press.
- Willett, T. (1991). *A reference grammar of Southeastern Tepehuan*. Dallas, TX: SIL. <<http://www-01.sil.org/acpub/repository/29375.pdf>>.
- Wilkins, D. (1986). Particle/clitics for criticism and complaint in Mparntwe Arrente (Aranda). *Journal of Pragmatics*, **10**(5), 575–596.
- Yap, F. H., Grunow-Härsta, K., & Wrona, J. (Eds.) (2011). *Nominalisation in Asian languages: diachronic and typological perspectives*. Amsterdam: Benjamins.
- Yliniemi, J. (2016). Attention marker =co in Denjongke (Sikkimese Bhutia). *Linguistics of the Tibeto-Burman Area*, **39**(1), 106–161.
- Zariquiey, R. (2015). The encoding of addressee's perspective in Kakataibo (Panoan, Peru). *Sprachtypologie und Universalienforschung (STUF)*, **68**(2), 143–164.