Chapter 5

A nascent definiteness marker in Yokot'an Maya

Maurice Pico

Leiden University

This paper examines the characteristics of a nascent definiteness marker in the Yokot'an language from the Mayan family from both a synchronic and diachronic perspective. The paper examines the contemporary distribution of the determiner *ni*, comparing it to that of the enclitic *ba*, which roughly corresponds to a topic marker. It employs Centering Theory to analyze oral materials, concluding that the use of the two particles is partially motivated by the processing cost of attentional shifts. Given that the determiner *ni* has been argued to develop from the distal demonstrative *jini* through grammaticalization, a diachronic perspective is also considered. The different synchronic uses of the determiner illustrated in this paper are then compared to the grammaticalization stages proposed for the development of definite articles. Both approaches ultimately suggest that *ni* conveys definiteness based on discourse-salience, not identifiability. The diachronic analysis further suggests that *ni* has started to bear some contrastive meaning related to reference in restricted contexts (reference to kinds in generic statements and specific reference in negative existential statements), indicating that the use of *ni* has spread beyond a pure topicality marker. Furthermore, the synchronic textual analysis in terms of Centering Theory clarifies some of the claims in Grammaticalization Theory regarding the early stages of definite articles by linking their emergence to the need of flagging attentional shifts in utterance-by-utterance processing of discourse.

1 Introduction

Yokot'an, a Mayan language from the Ch'olan branch spoken in the state of Tabasco, Mexico, makes use of demonstratives and deictic enclitics as NP mod-

ifiers but has also developed a reduced form *ni* which no longer seems to have deictic value (1).[1]

(1) *A-x-e   tä   num-e   t-u-pat   **ni bojte'.***
     ERG2-go-IPFV PREP pass-INF PREP-ERG3-back DET fence
     (Car-f): 'You are going to pass behind **the fence**.'
     [chf_MG_CAR_28-30_(1:32-1:35), Delgado-Galván 2018]

In her 1984 dissertation on Yokot'an morphosyntax, Knowles-Berry (1984: 209) proposes *ni* as a "definite determiner", but does not attempt to illustrate any further such characterization of its behavior. The goal of this paper is precisely to explore the main functionality of the determiner *ni* of Yokot'an. I will show that this determiner does not easily fit the usual characterization of definite articles as items with high textual frequency conveying familiarity, uniqueness of reference or identifiability via general knowledge (Himmelmann 2001: 832). Instead, I will show through textual analysis of oral materials that the distribution of *ni* exhibits a discourse-salience related role. I will unravel the main function of *ni* on the basis of two axes. The first axis is the *synchronic* perspective whereby *ni* overlaps in function with a topic marker, the enclitic *ba*. This overlapping relation will emerge through textual analysis performed with the help of a theory developed within computational linguistics: Centering Theory (Grosz et al. 1995). The second axis is the *diachronic* perspective whereby the form *ni* is a reduction of the distal demonstrative *jini*, a form that has been reconstructed all the way up to Proto-Mayan **ha+in*, through intermediary reconstructions **hin+i* for Western Ch'olan and **ha'in+i* for Proto-Ch'olan.[2] While the diachronic relation *ni* – *jini* has been proposed and argued for elsewhere (Mora-Marín 2009), it will be my contribution to try and relate synchronic uses of *ni* with different stages attested in the grammaticalization theory of articles. Furthermore, I suggest that the textual analysis in terms of Centering Theory links together two

---

[1]The abbreviations used in the examples can be found at the end of this paper. I have replaced the labels A and B used for pronominal indexes in traditional Mayan linguistics by the more standard ERG and ABS, respectively. A disadvantage, however, is that such glosses misleadingly suggest that the corresponding forms always convey ergative or absolutive grammatical relations, which is not accurate. Firstly, the same set of pronouns are also used in the nominal domain for possession and predication (respectively), and secondly, if seen as an "ergative" language, one must concede that Yokot'an presents a split on imperfective clauses.

[2]Throughout this paper, I will make use of a practical alphabet for the transcription of examples from Yokot'an, which conforms to the extent possibble to current practice in Mayan languages with a standardized alphabet. The values of the orthographic symbols are as expected but for *ä*=[ə], *ch*=[tʃ], *x*=[ʃ], *j*=[h], and *'*=[ʔ]. The only exception will be in the context of Mayan historical linguistics where, following its tradition, I will write h=[h].

independent observations on the grammaticalization of definite articles and illustrates how they fit together, thereby providing a better understanding of the early stages of grammaticalization of articles and their initial parallelism with the development of topic markers from demonstratives. No attempt whatsoever is made to put forward a semantic characterization of the meaning of *ni*, but I hope that this first text-oriented and functional analysis will lay out the ground that will make possible such undertaking.

This paper will be organized as follows. In §2, I review the standard conception of definiteness as rooted in uniqueness or familiarity of reference. I then show that neither seems a natural choice to represent the main motivation behind the use of *ni*. Moreover, I discuss the relative optionality of *ni* to argue that its function is likely sensitive to discourse-management motivations. In §3, I turn to an utterance-by-utterance discourse analysis of the texts to justify a discourse-salience definiteness for *ni*, or, as Walker & Prince (1996) would put it, a view of *ni* as a marker of the "Discourse-status" of the entity evoked by an NP, as opposed to its "Hearer-status" (its availability in the background knowledge of speaker and/or hearer). To this end, in §3.3, I illustrate attentional transition types in Yokot'an discourse within the Centering Theory framework, with preliminary concepts given in §3.1–§3.2. In §3.4, I show the association of *ni* occurrences with attentional transitions of some type, where its functional overlap with the topic marker *ba* will be apparent. In §4, I incorporate a diachronic perspective by looking at the current distributional properties of *ni* through the glass of the well-attested path of grammaticalization from demonstratives to articles. In §4.1, I assess whether the determiner *ni* has departed from being a demonstrative and I do so by following two criteria: a quantificational one (§4.1.1) and a qualitative one (§4.1.2). Once we have seen that *ni* has undergone progress along the grammaticalization path towards a definite article, away from its demonstrative source, the stages proposed in the literature of grammaticalization become relevant and I proceed, in §4.2, to pinpoint the stages at which *ni* currently stands with its several uses. The textual distribution that the Centering Theory analysis revealed in §3.4 now comes to clarify how two independent observations on the early stages of definite article grammaticalization fit together. Finally, given that topic markers can also develop from demonstratives, I point out in §4.3 a specialized use of *ni* as a marker of specific reference, which happens in the restricted context of negative existential constructions. In this way, I argue that nevertheless its main function as a discourse marker of topicality shifts, *ni* is better seen globally as a nascent definite article based upon salience-management, rather than as a pure topic-marker. In §5, I summarize the conclusions of this first study on the determiner *ni*.

We are now ready to initiate §2, where I will show that the standard cognitive correlates associated with definiteness are not sufficient to explain the distribution of *ni*. Furthermore the reader is faced with the scarcity and seeming optionality of the form *ni*. This will motivate the view of *ni* as a discourse-oriented particle.

2 Which sort of definiteness for *ni*?

In this section I illustrate some of the difficulties that can be encountered when trying to understand the contribution of a previously undescribed determiner. I briefly compare the distribution of *ni* with what would be expected from the standard treatment of definiteness which is informed by the historical debates on definite descriptions in more familiar languages. This will make apparent the need to move on to discourse motivations behind the use of *ni*, which then will be seen as a marker of NP discourse status (or of transitions between them) at the end of §3. Reasons to maintain *ni* as a *nascent* definite determiner rather than as a purely pragmatic particle will be apparent in §4 with the insights from Grammaticalization Theory.

The treatment of definiteness in linguistics emerged from an originally philosophical debate around the contribution of the so-called "definite descriptions" to the meaning of the utterances in which they appear. Most accounts of definiteness take definite descriptions to denote identifiable referents and are built around three main ideas:

- The definite article indicates the *identifiability* of the NP's referent.

- Identifiability stems from the *uniqueness* of the referent that satisfies the descriptive content of the NP (within a given situation), or

- Identifiability stems from such referent being already *familiar* to both the speaker and the addressee – in particular through previous discourse mention – regardless of its descriptive uniqueness.

These ideas have been exploited independently or in a combined fashion.[3] The intuition that definiteness involves the uniqueness of the referent is motivated by

---

[3]The initial philosophical discussion can be found in Frege (1892); Russell (1905) and Strawson (1950). For modern accounts of definiteness as uniqueness I refer the reader to Hawkins (1978; 1991) and Abbott (1999). The familiarity perspective is embodied by a dynamic semantic analysis of anaphora resolution. This kind of analysis embeds utterance interpretations into their discourse context to allow for inter-sentential anaphora resolution, including anaphoric def-

cases where the referent is picked out through an immediate and unambiguous situational availability without the need of any previous linguistic co-text (see Hawkins 1978: 103, 110). Example (2) illustrates such cases:

(2)   Context: In a carpentry workshop after some time silently working together.
       *Could you please hand me **the smoothing plane** on the workbench?*

Familiarity, on the other hand, aims to reflect cases like (3), where no visual/situational input is needed for the hearer to properly interpret the utterance, rather relying on a previous mention:

(3)   *While I was fixing my bike yesterday, **a man** and a woman approached me and asked for directions. **The man** had a strange accent. I couldn't guess where he was from.*

When a definite article is known to have developed diachronically from a demonstrative, uniqueness and familiarity can both be seen as an outcome of a specialized use of deixis. Uniqueness within a situation would then develop from spatial exophoric uses of a demonstrative while familiarity would develop from anaphoric uses (there is some discussion about whether one use is more fundamental, see Lyons 1999: 160). Some languages even develop two different articles, each specialized in one of the uses, an article for expressing uniqueness-based definiteness (which would correspond to a *weak* article in Schwarz 2013) and another for expressing familiarity-based definiteness (corresponding to a *strong* article in Schwarz 2013). Given that *ni* likely originates from the distal demonstrative *jini* one may be led to expect it to fit the previous picture. However a first difficulty arises already with its rather scarce presence in texts, as compared to the rather common situation in which an entity has already been mentioned or in which it is ostensibly unique or perceptually salient in the context.

Interestingly, the oldest texts that I could consult of modern Yokot'an (two texts collected by Keller & Harris in 1946 or earlier) do not contain a single occurrence of the determiner *ni*. Its absence in a given corpus, especially a corpus not exceeding two pages, cannot be taken as evidence of non-existence, however. If we assume that the determiner *ni* was already in the language, I find it significant

---

inite NPs. Discourse Representation Theory (Kamp 1981) and File Change Semantics (Heim 1982) are the main starting points in the formalization of this idea. Both characterizations of definiteness have also been combined in other accounts either to jointly provide a treatment of a given definite article (Farkas 2002; Roberts 2003) or to account for different articles with their own specialized meaning contribution (Schwarz 2009; 2013).

that "indefinite" NPs are always retaken as bare nouns without any determiner, as I show in the textual sequence (4a–c). Both *ajyäx* 'crab', and *ixmuch* 'frog', are introduced with the numeral 'one' (+ classifier), but their respective references are resumed later with bare nouns, rather than with a sequence *ni*+N.[4]

(4)   Yokot'an (Tapotzingo, Nacajuca; Keller & Harris 1946: 138)

    a. *Ajn-i*   ***um-p'e***   ***aj-yäx.***   *Näts'ä ti'*   *pa'*
         be.located-PFV[ABS3] one-NUM.CLF CLF.M-crab close   mouth lake
         y-otot.   [...]
         ERG3-house
         'There was **a crab**. Near the bank of the river was his house. [...]'

    b. *Bix-i*   *tä*   *wa'wa'n-e*   *pan ji'.*   *I*
         go-PFV[ABS3] PREP drift.around-INF PREP sand and
         u-nuk't-an   ***un-tu***   ***ix-much.***
         ERG3-find-IPFV[ABS3] one-NUM.CLF CLF.F-frog
         'He went for a walk on the sand. And met **a frog**.'

    c. *U-pek-än*   *tä*   *ts'aji.* ***Ix-much*** *uy-äl-e'*   *tan*
         ERG3-call-IPFV[ABS3] PREP chat CLF.F-frog ERG3-say-IPFV[ABS3] PREP
         u-k'ajalin: ya' kä-x-e   *kä-xik'-e'*   ***aj-yäx.***
         ERG3-mind SD ERG1-go-IPFV ERG1-fool-IPFV[ABS3] CLF.M-crab
         'He [the crab] spoke to it [the frog]. **The frog** talked in his mind: "I'm going to make a fool of **the crab**".'

This could suggest that the determiner *ni* is not associated with anaphorically-based familiarity, i.e. it is not an article of the strong type, in terms of Schwarz (2013). Thus, *ni* would not be required for an NP to be interpreted as referring to the same entity than previously introduced in the discourse. The example (5) shows a mention (mid-text) of one of the main characters of a story. Thus, both speaker and hearer know, and are assumed to know, the referent. Clearly, a bare noun is enough.

(5)   *Y-äl-i*   ***balum****: "kä-x-e*   *tä*   *och-e*   *tan noj*
      ERG3-say-PFV[ABS3] jaguar ERG1-go-IPFV PREP enter-INF PREP big
      bujchach".
      basket
      (Alb-m): '**The jaguar** said: "I am going to enter into the big basket".'
      [chf_HT_ALB_624_(24:10-24:12), Delgado-Galván 2018]

---

[4]The reader may notice that the nouns *yäx* and *much* are preceded by gender classifiers *aj-* and *ix-*. These are not crucial for the current discussion, as we will observe later in example (5) that their absence does not hinder the capability of a noun to be interpreted definitely.

As an anonymous reviewer kindly noted, one may wonder whether the NP *balum* has received special treatment or has turned into a proper name in view of the mythological character of its referent and its cultural prominence. However we can see in example (6) that this is rather the standard treatment of NPs. The monkey, *ajpum*, gets introduced with the numeral 'one' (+ classifier), *un-tu*, and then mentioned again later. Once more, a bare noun is enough.

(6)   *I*   *ya'-i*   *u-nuk't-i*   ***un-tu***   ***aj-pum.***   […]
     and SD-DIST ERG3-find-PFV[ABS3] one-NUM.CLF CLF.M-monkey […]
     u-k'ech-i   ***aj-pum.***   […] *i*   *u-bis-an*
     ERG3-grab-PFV[ABS3] CLF.M-monkey […] and ERG3-bring-IPFV[ABS3]
     aj-pum   *t-u-chejpa.*
     CLF.M-monkey PREP-ERG3-rib

     (Bla-m): 'He found **a monkey**. […] he took **the monkey**. […] and brings **the monkey** with him on his side.' [chf_HS_BLA_27-30_(03:41-04:03), Delgado-Galván 2018]

In this case, the extracted example comes from an elicited picture-story and thus none of its characters can be assumed to be culturally prominent. Given that anaphoric familiarity doesn't seem to trigger the use of *ni*, one may try to verify whether it behaves akin to a weak-type article (Schwarz 2013), with definiteness based upon uniqueness. Starting with example (7) we see that the determiner *ni* is not used – and in fact is unnatural to use – in cases of global uniqueness like *the sun*, *the moon*, etc.[5]

---

[5]I say "unnatural" rather than "ungrammatical" since Knowles-Berry (1984) provides a counter-example, with *ni* introducing global entities like *the sun* (i).

    (i)   *'A*   *tik-i*   *ni*   *nok'*   *k'a*   ***ni*** ***k'in.***
        AUX.PFV dry-PFV[ABS3] DET clothes PREP DET sun
        'The clothes dried because of **the sun**.' (Knowles-Berry 1984: 309)

    Still, the more acceptable strategy is to avoid using the determiner *ni* in this cases, as can be seen in another example from the literature below (and is also confirmed by my collaborators in the field):

    (ii)   *Jik'in a*   *tuts'-i*   ***k'in*** *a*   *bix-on tä*   *patan.*
        when AUX.PFV appear-PFV[ABS3] sun   AUX.PFV go-ABS1 PREP work
        'When **the sun** appeared, I went to work.' (Schumann Gálvez 2012: 113)

(7) *tä   ke   t'äb-o   (??**ni**) **k'in**
     PREP COMP ascend-INF   DET sun

(Luc-m): 'until **the sun** rises'

[chf_TwoFishingmen_178_(10:10-10:11), Delgado-Galván 2018]

Example (8) illustrates the case of uniqueness within a restricted situation, in which the determiner *ni* is not used either. The context of the utterance is one in which only one dog (the family dog) is known to be behind the house and it is recognized by its barking.

(8) *Ya' an   tä   woj   **wichu'** nanti.*
     SD EXIST[ABS3] PREP bark dog   over.there

(Mar-m): '**The (family) dog** is barking over there (behind the house).'
[My_elicitation, elic_deif_marc_08]

However, in contrast to (7), the determiner *ni* is perfectly fine in this context and can appear used in such examples, as can be seen in (9). The same translation is kept to signal a lack of meaning difference.

(9) *Ya' an   tä   woj   **ni**   **wichu'** nanti.*
     SD EXIST[ABS3] PREP bark DET dog   over.there

(Mar-m): '**The (family) dog** is barking over there (behind the house).'
[My_elicitation, elic_deif_marc_08b]

Thus, at least in some contexts of use, there is some freedom as to marking the NP with *ni.* As a matter of fact, a narrative sequence similar to the sequence in (4) above, nowadays, would still allow a near absolute absence of *ni.*[6] Although Yokot'an has been considered to be a language with a "definite word distinct from demonstrative" by Dryer (2005) – probably based upon examination of Knowles-Berry's (1984: 209) proposal of *ni* as a "definite determiner" – a large portion of NP instances in Yokot'an which would be translated by a definite noun phrase in English fail to have any determiner at all, i.e. they are bare nouns. This points to an aspect that complicates the cross-linguistic picture of definiteness. It is the non-negligible number of "languages where there is an article that is restricted to but not obligatory in definite contexts" (Dryer 2014: e234), i.e. languages which

---

[6]As an example, not a single instance of *ni* appears in the sample text provided in appendix by Knowles-Berry (1984: 371–382). An exception to this scarcity of *ni* is written Yokot'an, where Spanish as a model of literacy exerts an enormous influence and tends to impose the art+N nominal template.

do have definiteness markers of some kind but whose definite NPs, somewhat paradoxically, do not seem to require them in the first place.

The need for motivation is twofold. Diachronically, the optionality – to varying degrees – raises the question of why would a language develop a seemingly dispensable marker. Synchronically, the optionality of definite articles raises the question about the reason a speaker might use them. Both aspects are linked. According to Hawkins (2004: 84), the *compelling* motivation for the diachronic emergence of a definite article from a demonstrative "to express meanings that are perfectly expressible in languages without definite articles", originates from synchronic *processing* needs of grammar rather than from semantics or pragmatics. Interestingly, Givón (2001: 474) points out that "Grammaticalized definite markers [...] arise first to mark *topical* definites.", which implies that nascent definite markers do not systematically accompany every NP interpreted as identifiable, but rather seem to come associated with a change of discourse-status regarding the NP concerned.

In the next section, I will introduce two notions to capture these two aspects of an NP: the Hearer-status (related to identifiability and to the common-ground) and the Discourse-status (related to processing and to the referent's status in the short-term memory). Under this view, nascent definite markers are better seen as some sort of Discourse-status markers which are concerned with the optimization of both discourse and utterance processing. It is precisely in this way that topicality gets modeled by Centering Theory. In §3, I will introduce this theory and use it as a heuristic device to guide our quest for the functionality of *ni* in oral texts. To this end I will apply the theory to a selection of samples from oral materials to better understand how attentional shifts in utterance sequences affect the likelihood of an NP to be introduced by *ni*.

3 Centering Theory and the discourse-management use of *ni*

3.1 Framework

Centering Theory, which is a component of a less well-known discourse theory from computational linguistics, could be perceived as one more approach to address pronominalization/anaphoric resolution and, in that way, as a competitor to other theories addressing the anaphoric properties of NPs. More established theories of discourse-oriented analysis of sentences exist, like DRT, but these were not originally proposed in order to model attention management (or infor-

mation structure) and its interaction with the *shape* of NPs and their *structural position* in sentences.[7] This difference stems from a different approach to the dual nature of referring expressions, which can be seen from a semantic or from a syntactic viewpoint. From the syntactic viewpoint, referring expressions have an impact on sentence linking and processing. From the semantic perspective they have an impact, via evoked entities, on the common ground of speaker and hearer.[8]

Let me explain. NPs can uncontroversially be taken to evoke discourse entities. These entities may bear information statuses of different nature. This has been noticed – among others – by Walker & Prince (1996: 291–294) which propose to distinguish the *Hearer-status* of a discourse entity from the *Discourse-status* triggered by its evoking formal device. I summarize my interpretation of their views in Figure 1, below. The Hearer-status is the belief, by the speaker, as to whether a discourse entity is known or inferable for the intended audience and thus can be assumed to be in the common ground (or not). If it is believed to be known or inferable, the NP will tend to be marked as definite, otherwise, as indefinite. Under this point of view, definiteness is nothing else than identifiability via general knowledge. But the discourse entities are evoked through formal devices, and these formal devices – which can range from full NPs to referential indexes in the verb – have formal discourse-properties of their own, regardless of the identifiability of the evoked entity. A referring formal device has a potential for *salience* which emerges from its overall structural role in the sentence. Moreover, in a sequence of utterances, the same discourse entity might have been evoked by devices with *different* salience. A given level of salience of an NP may affect the *activatedness* of the evoked discourse entity in the *next* utterance.[9]

---

[7]In particular, the concepts of topic and focus were not included in the standard format of DRT (see Kamp & Reyle 1993: 360, 639).

[8]The complementarity of Centering Theory, which emphasizes the first perspective, with other approaches that emphasize the second perspective has been noticed by many, with suggestions towards integration in Walker & Prince (1996) and Gundel (1998) for the Givenness Hierarchy and in Roberts (1998; 2012) for DRT.

[9]The term *activation* is usually preferred within linguistics literature and it is often associated with a single Familiarity/Givenness/Accessibility scale for NP classification (cf. Ariel 1990; Gundel et al. 1993; Kibrik 2011), but Walker & Prince (1996: 294) use the term *activatedness* "or Discourse-status" to make clear that they consider *givenness* and *activation* as independent, orthogonal, scales to be treated separately. Thus, *activation* usually involves an amalgamated scale with *givenness*, while *activatedness* is roughly *activation* considered separately. I stick to the latter term since I have based my framework on Walker & Prince (1996). Kantor (1977) introduced the term *activatedness* within computational linguistics covering a loosely similar idea. The discussion of similarities and differences in the use of these terms from author to author should not concern us here. Since I use Centering Theory to model *activatedness*, just as Walker and Prince (1996) propose, there is no risk of vagueness or confusion in the use of this term.

Two types of *Information status* for a discourse entity

1. HEARER-STATUS (related to givenness and inferability)

    - entity known or inferable (NP coded as definite)
    - entity not known and not inferable (NP coded as indefinite)

2. DISCOURSE-STATUS (related to activatedness and salience)[10]

    a) SALIENCE (*upcoming* activatedness): The formal SALIENCE of the evoking NP (or referential index) in the utterance U_i *currently* being processed. It affects the ACTIVATEDNESS of the discourse entity for the *next* utterance U_{i+1}.

    b) ACTIVATEDNESS (*former* salience): The formal SALIENCE of the evoking NP (or referential index) in the utterance U_{i-1} that has been processed *before* the current one. This affects the ACTIVATEDNESS of the evoked discourse entity in the utterance U_i *currently* being processed.
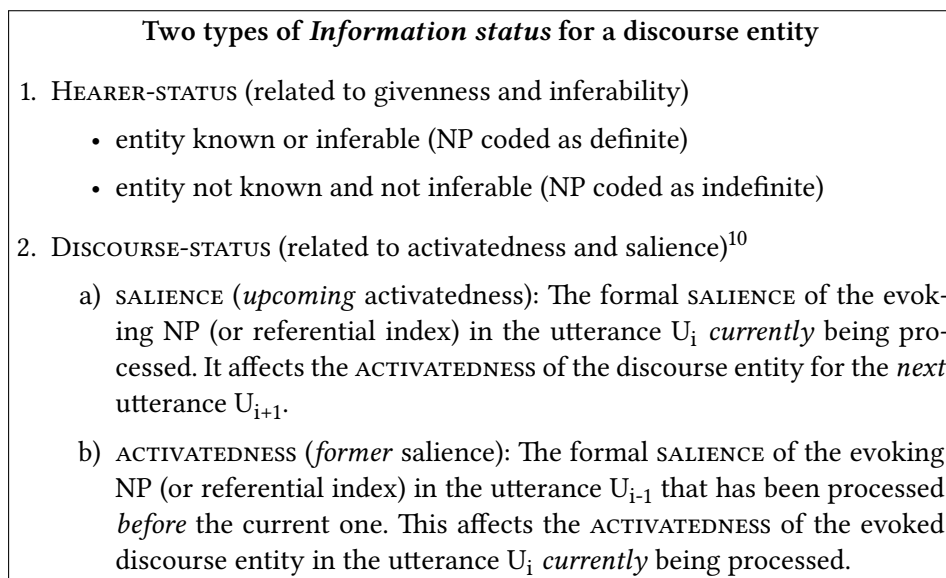

Figure 1: My interpretation of Walker & Prince (1996)

In other words, an entity evoked by the discourse has two orthogonal, but logically independent statuses: a *Hearer-status* (Is it familiar to the hearer or inferable?) and a *Discourse-status* (Is the evoking device formally salient in the utterance currently being processed? Was it formally salient in the precedent utterance (thus promoting an *activated* referent in the current one)?).

In §2 I have shown that the Hearer-status cannot by itself account for insertions of *ni*, reason for which I now turn to Centering Theory to inspect how the Discourse-status of NPs or, rather, their changes of such status (attentional transitions) relate to the presence of the determiner *ni.* Centering Theory is well suited to this aim, since it is precisely an attempt to model the way in which the changing salience of referring expressions in an utterance helps to manage attention and attention shifts throughout a discourse progression. As such, it is also intended to be a component of a larger theory of discourse coherence.

Discourse typically involves utterances organized in smaller discourse segments. Thus, the coherence of a discourse emerges at two levels: between the utterances within a single discourse segment (local coherence), and between that segment and other discourse segments (global coherence) (Grosz et al. 1983: 44). Each level of discourse structuring and coherence is associated with a corre-

---

[10]In Centering Theory, the notions (2a) and (2b) are *locally* modeled, respectively, by the concept $Cp(U_i)$ and by the preference for $Cb(U_i) = Cp(U_{i-1})$, these will be presented in §3.2, below.

sponding level of attention or *focusing*:[11] local attention (or centering) and global attention. Centering Theory is devoted to the study of local coherence and the attentional transitions from one utterance to the next, that is, it is a theory of local discourse structure (Grosz et al. 1995; Grosz & Sidner 1998; Walker et al. 1998).

3.2 Centers of an utterance

Centering Theory models the contribution of NPs (or, more generally, referential indexes) to the coherence of a local discourse segment by recognizing two ways in which an utterance affects the structure of a coherent discourse. Both ways involve the fact that any utterance U evokes a set of discourse entities which can then be used as a cohesive link with adjacent utterances. The first way is by establishing a link with the previous utterance through topic continuity. The second way is by establishing a discourse entity evoked in the current utterance as the *default choice* for being picked-up as topic by the next utterance. This prospective suggestion regarding topicality crucially involves the structural salience of a referring device and exploits the relation between the salience and the activatedness illustrated in Figure 1 above. When considered in this way, as links between adjacent utterances, the discourse entities evoked in U are named the *centers* of U (Grosz et al. 1995: 208). Since all entities in this set can potentially be talked about in the next utterance, its members are called *forward-looking centers* (Cf). Among these, an utterance often has a *center of attention*, a privileged center which constitutes the main link to the previous utterance, *i.e.* the *backward-looking center*. It roughly can be seen as a special kind of topic: a strictly *local* topic (as opposed to a *global* topic, which encompasses the entire discourse or discourse segment).

As I anticipated above, one of the main claims of Centering Theory is that each utterance has not only a current center of attention (the Cb), but also a proposed *anticipation* of what the center might be in the next utterance (the default choice for its Cb), which depends on a ranking of the Cfs according to their salience, mostly determined by grammatical structure. For the present discussion I take grammatical relations as the main ranking factor, as follows: SUBJ > OBJ > ADJUNCT. That is, the entities evoked by arguments rank higher up than

---

[11]Grosz et al. (1983: 44) use the term *focusing*, but to avoid confusions with the more specialized use of the term in information structure studies, I will rather speak of *attention*. Hence I will speak of global attention and local attention for what is termed *global focusing* and *local focusing* in the original paper. For a discussion of the relation between the concepts of *focusing* from CT and *focus* and *topic* from Information Structure, I refer the reader to Gundel et al. (1993: 279, footnote 10).

those evoked by non-arguments and, for transitive clauses, the ergative argument ranks higher than the absolutive as well.[12] The highest ranked Cf is singled out as the *preferred center* (Cp) which is the default candidate to be the *backward-looking center* (Cb) of the next utterance. To summarize:

Forward-looking centers (Cf): $Cf(U_i)$ = the set of discourse entities evoked by an utterance U_i

Preferred center (Cp): $Cp(U_i)$ = the highest ranked element of $Cf(U_i)$ in terms of salience.

The Cp constitutes a prediction about the Cb of the following utterance.

Backward-looking center (Cb): $Cb(U_i)$ = the highest ranked element of $Cf(U_{i-1})$ realized in U_i

Observe that the $Cb(U_i)$ does not coincide with the preferred center Cp of U_{i-1} when the latter is not evoked in U_i (in such case, the next highest ranking entity of $Cf(U_{i-1})$ will be taken as Cb, if evoked). Depending on the continuity or disruption between the *local topic* $Cb(U_i)$ and the *anticipated topic* $Cp(U_i)$ of an utterance U_i or between the *local topic* $Cb(U_{i-1})$ of a previous utterance U_{i-1} and the one from the current utterance, $Cb(U_i)$, we can have several types of center attention transitions, which are displayed in Table 1 of the next section.

3.3 Transitions between utterances

Since every utterance evokes entities (and therefore has centers), there can be continuity of centers from utterance to utterance or there can be shifts of centers. Two main parameters govern the quality of a transition from one utterance to the next. One parameter is whether both utterances maintain the same local topic (Cb) or not (first and second columns in Table 1 below). The second parameter is whether the local topic (Cb) of the second utterance corresponds to its anticipated or suggested topic Cp (upper row in Table 1) or not (bottom row).

A CONTINUE transition type is the least disruptive one, as the center of attention (or roughly the "local topic") in the current utterance does not replace a previous one and is additionally set up as the preferred center (Cp), the "anticipated

---

[12]For this ranking, which has proven to be accurate enough for my textual analysis, I follow Hedberg (2010: 1837-1838). The segmentation of utterances based upon the logic of clausal units rather than pure intonation follows Prince (1999) and Kibrik (2011).

Table 1: Center Transitions (Walker et al. 1998)

	$Cb(U_{i-1}) = Cb(U_i)$ (or $Cb(U_{i-1}) = ?$)	$Cb(U_{i-1}) \neq Cb(U_i)$
$Cb(U_i) = Cp(U_i)$	CONTINUE	SMOOTH-SHIFT
$Cb(U_i) \neq Cp(U_i)$	RETAIN	ROUGH-SHIFT

or suggested topic" for the next utterance.[13] According to this model, a maximally gradual change of attention ideally would involve a sequence of two transitions (so: minimally two utterances), one RETAIN transition which anticipates a shift in topic and one SMOOTH-SHIFT transition which executes it. However, more abrupt shifts can involve both transitions compressed and collapsed into a single transition executed within a single utterance: the ROUGH-SHIFT transition, which would naturally be expected to invite the use of the most marked structures. Centering Theory has the ordering rule in Figure 2 reflecting the intuition that speakers try to maximize coherence and that these transitions are increasingly less coherent (or, equivalently, coherent at a higher processing cost).

> **Transition states are ordered**:
> CONTINUE > RETAIN > SMOOTH-SHIFT > ROUGH-SHIFT

Figure 2: Ordering rule (Walker et al. 1998)

One limitation of the basic format of Centering Theory presented above is that it deals with transitions *within* topical chains (conceived as chains of utterances where pairwise sharing of at least one center is maintained and thus $Cb(U_i)$ is always available). Not much is said about utterances lacking a Cb (Cb = none or, equivalently, Cb = ?) which are the utterances that *start* a topical chain, either because they are absolute discourse-initial, or because they don't share any of its centers with the previous utterance (Hedberg 2010: 1831).[14]

For the present discussion, all I need is to complement Table 1 with Table 2, below.

---

[13]The informal expressions *local topic* and *anticipated or suggested topic* are mine. They are just intended to guide the intuition of a reader who has no prior contact with Centering Theory.

[14]Walker et al. (1998) label these transitions simply as NO CB transition. But then both kinds of topical chain starts would be collapsed. Intuitively, it is a more drastic shift to ignore all the centers introduced by a previous utterance than to start a discourse with no previously specified information in the background. Some further refinements and a classification of the transitions with the parameter (Cb = ?) have been proposed, see Poesio et al. (2004) and references therein.

Table 2: Center Transitions for chain-initial U_i (Poesio et al. 2004; Hedberg 2010)

	$Cb(U_{i-1}) = ?$	$Cb(U_{i-1}) = c$
$Cb(U_i) = ?$	NULL	ZERO (≈ ROUGH-SHIFT)

The row represents the chain-initial utterance U_i, where *chain-initial* is taken as the fact of not having a backward-looking center Cb (Cb = ?). The first column represents the situation in which the previous utterance is also "chain-initial". The special case where there is no previous utterance is not of importance here. The second column represents the case where the previous utterance had a Cb (Cb = c, for some entity c), and it was ignored by the current utterance. This case, the ZERO transition, is really some kind of shift so I will treat it as a special case of ROUGH-SHIFT transition, see example (20) further down. Observe that when Cb = ?, neither (Cp = Cb) nor (Cb = previous Cb) are true (which for all practical matters, almost boils down to Cp ≠ Cb and Cb ≠ previous Cb). Furthermore, in the case of the ZERO transition, it is known for sure that the Cb and the Cp from the previous utterance exist and have been ignored. So I will consider this as a degenerate case of ROUGH-SHIFT transition, reason for which I added this consideration in Table 2. It is more disruptive than ROUGH-SHIFT proper, since it entirely dismisses the centers from the previous utterance. I expect it to invite even more the use of non-neutral constructions.

I will now illustrate centers and center transition types with a sequence of contiguous utterances in Yokot'an from the Frog Story elicitation task (examples (11–17) below). The utterance (11) has the kid (*yokajlo'*) as backward-looking center (Cb), given the previous context – omitted – which offers *yokajlo'* as antecedent of the absolute person mark of all verbs in (11).[15] Moreover, the kid (*yokajlo'*) is also the highest ranked forward-looking center, given its status as subject-argument (it is thus the preferred center Cp from all centers in the set Cf). The centers of the utterance (11) can thus be represented as follows:[16]

---

[15]Centers are in **boldface** to remind the reader that these are not the linguistic expressions but the entities realized by them.

[16]I first display the backward-looking center (Cb). Then I display the set of Cfs by ranking order, with its first member being the preferred center (Cp). Finally, I display the two parameters that determine the transition type. For the sake of simplicity and to avoid overloading the exposition, I will disregard many details which do not affect my analysis in crucial ways. For example, I disregard the fact that some examples like (11), include in fact two utterances of which the second is a re-elaboration, and I will also skip the details about how backgrounded clauses like e.g. the temporal clause *k'echi' ak'äb* in (11) are treated.

(10) Utterance: (11);
      [Cb(*yokajlo'*), Cf(Cp(*yokajlo'*) > *ak'äb*)];
      Cp = Cb;
      Cb = previous Cb;
      Ct: CONTINUE

(11) *Ke   ya' a   bix-i   tä   wäy-e.   K'ech-i*
      COMP SD AUX.PFV go-PFV[ABS3] PREP sleep-INF ERG3-grab-PFV[ABS3]
      ak'äb, bix-i   tä   wäy-e.
      night go-PFV[ABS3] PREP sleep-INF
      (Esm-f): 'Then he [the kid] went to sleep. When the night reached him,
      he [the kid] went to sleep.'
      [chf_FrogStory_ESM_006_(00:43-00:47), Delgado-Galván 2018]

The example (11) is, in fact, a CONTINUE transition with respect to the previous (not presented) context. Consequently the backward-looking center is evoked through the most reduced referential form: a personal index in the verb, which in this case is actually an implicit ABS3 index. This is a reminder that centers are often realized by reduced referential devices. The following utterance in the narrative, (13), has the following centers and center transition (Ct):

(12) Utterance (13):
      [Cb(*yokajlo'*), Cf(Cp(*yokajlo'*) > *wichu'* > *ts'en*)];
      Cp = Cb;
      Cb = previous Cb;
      Ct: CONTINUE

(13) *De   y-o=ba   wäy-e,   de   bo'o [u-]jin   bix-i*
      PREP ERG3-want=TOP sleep-INF PREP tired ERG3-feeling go-PFV[ABS3]
      tä   wäy-e,   une='a   pan u-ts'en,   dok yok wichu'.
      PREP sleep-INF PRO3=TOP PREP ERG3-bed COM little dog
      (Esm-f): 'He wanted to sleep, of tiredness he went to sleep on his bed,
      with his dog.' [chf_FrogStory_ESM_007_(00:48-00:54), Delgado-Galván
      2018]

Since the backward-looking center (Cb) of (11) and (13) is the same, and the preferred center is also shared, there is full continuity respect to which center

gets most attention and will preferentially get attended to on the next utterance. This illustrates the CONTINUE type of transition between utterances. However, the next utterance (15) in the sequence starts to introduce a shift. While (15) and (13) keep sharing the same Cb, (15) introduces a new Cp, the frog (***much***), which announces a future shift in center of attention (a shift in "local topic"). (15) has the following centers and center transition type:

(14)   Utterance (15):
       [Cb(***yokajlo'***), Cf(Cp(***much***) > ***yokajlo'***)];
       Cp ≠ Cb; (this announces a *future* shift of "local topic")
       Cb = previous Cb;
       Ct: RETAIN

(15)   ***Ix-much='a***   *u-chän-i*   *ke*   *a*   *wäy-i*
       CLF.F-frog=TOP ERG3-see-PFV[ABS3] COMP AUX.PFV sleep-PFV[ABS3]
       yokajlo'=ba.
       kid=TOP

       (Esm-f): '**The frog** saw that the kid was asleep.'
       [chf_FrogStory_ESM_009_(00:57-01:00), Delgado-Galván 2018]

    The frog (***much***) is highest in salience ranking than the kid (***yokajlo'***) due to the fact that it is evoked by an NP associated to the structural role of subject of the transitive main clause, while ***yokajlo'*** is evoked as intransitive subject of an embedded clause. The fact that ***much*** is evoked by the ERG3 index of the main verb makes it the Cp, but the fact that it is also mentioned with a full NP with a topic marker *ba* can be blamed on the fact that Cp ≠ Cb. As we will see later (§3.4), at this point we could have had the determiner *ni* introducing the NP *ixmuch* either redundantly with *ba* or without it.[17] This illustrates the RETAIN type of transition between utterances, which *retains* the local topic (***yokajlo'***), but announces its demise. With the next utterance (17) in the sequence, I illustrate the SMOOTH SHIFT type of transition, which executes the Cb-shift that was prepared in (15). The utterance (17) has the following centers, and center transition type:

(16)   Utterance (17):
       [Cb(***much***), Cf(Cp(***much***) > ***traste***)];
       Cb = Cp;
       Cb ≠ previous Cb; (this executes the "local topic" shift)
       Ct: SMOOTH SHIFT

---

[17]This has been also confirmed by my collaborators in the field.

(17) *U-ch-i*   *aprobecha une, a   pas-i   tan*
     ERG3-make-PFV[ABS3] advantage PRO3, AUX.PFV exit-PFV[ABS3] PREP
     traste bajka   ya' an='a.
     jar   where SD   EXIST[ABS3]=TOP
     '(Esm-f): She [the frog] took advantage, she [the frog] went out of the
     bottle where she was.'
     [chf_FrogStory_ESM_010-(01:00-01:05), Delgado-Galván 2018]

Since the backward-looking center (Cb) of (15) and (17) are different, the shift in center of attention that was anticipated with RETAIN in (15) is now completed in (17). It is interesting to note that the frog (***much***) is evoked by a highly salient formal device in both (15) and (17), namely the indexes for transitive and intransitive subject, but only in (15) is a full NP used in preverbal position and with a topicality marker. The first observation is linked to the fact that Cp(15)=***much*** and Cp(17)=***much***. The second fact (the use of a *ba*-marked NP) is linked to the switch represented by ***much***=Cp(15) ≠ Cb(15)=***yokajlo'***. This should draw our attention to the fact that under this model and analysis, *ba*-marked NPs as the one above do not flag topicality of a discourse entity as such, but the *switch* of topicality, i.e. the transitions characterized by a rupture Cp ≠ Cb (and, perhaps, the fact that the Cp has just been introduced into the Cf set of the utterance without having been present in the Cf of the previous utterance).

After such shift of center in two steps, namely a RETAIN (15) plus a SMOOTH SHIFT (17) transition, the flow of local attention proceeds with minimal disturbance. The example (19) preserves the same centers than (17) and moreover does not anticipate or announce any future shift. We do have again a CONTINUE transition type, the least "marked" of all.

(18) Utterance (19):
     [Cb(***much***), Cf(Cp(***much***))];
     Cp = Cb;
     Cb = previous Cb;
     Ct: CONTINUE

(19) *A   pas-i   ya'-i, puts'-i   pues, a*
     AUX.PFV exit-PFV[ABS3] SD-DIST escape-PFV[ABS3] so   AUX.PFV
     bix-i.
     go-PFV[ABS3]
     (Esm-f): 'She [the frog] escaped from there, she ran out, she left.'
     [chf_FrogStory_ESM_011_(01:05-01:08), Delgado-Galván 2018]

Now I illustrate the most complex transition type called ROUGH SHIFT, which, as its name suggests, introduces an unannounced shift of center, in this case, in favor of the discourse entity *yokajlo'*, 'kid'. Observe how the evoking device, the NP *yokajlo'*, is bearing a topic marker *ba*.

(20) *Ya'-i   a   ch'oy-i   isapan   yokajlo'=ba.*
     SD-DIST AUX.PFV wake-PFV[ABS3] morning kid=TOP

     (Esm-f): 'Then **the kid** woke up in the morning.'
     [chf_FrogStory_ESM_012_(01:09-01:11), Delgado-Galván 2018]

No center from the previous utterance is evoked, thus there is no backward-looking center in the current utterance and the new preferred center, the kid (*yokajlo'*), is introduced without any anticipation whatsoever in the previous utterance.

(21)   Utterance (20):
        [Cb(?), Cf(Cp(*yokajlo'*) > *isapan*)]
        Cb = ?;
        previous Cb = ***much***;
        Ct: ZERO;
        ⇓ A special case of
        Cp ≠ Cb;
        Cb ≠ previous Cb;
        Ct: ROUGH SHIFT

A few notes are in order to draw attention to something that the reader might have already deduced. First, the utterance-topic conceptualized as Cb (the "local topic") is dependent on the centers of the previous utterance. The very same sentence may or may not have such a "topic" depending on which entities have just been evoked in the previous utterance (a case in point is the jump from 19 to 20). As such, Cb is clearly a relational-discourse dependent notion (as opposed to Cp which is more closely dependent on the shape of the utterance). Second, this notion of topic and center of attention is strictly local: it concerns the immediately preceding utterance within a given discourse segment. Thus, a given entity evoked by an NP might be globally topical (in the sense that the global discourse attention is directed to it) without being locally topical, i.e. without being the Cb of an utterance. To resume a reference across a local transition (of different sorts) within the same discourse segment and to resume the same reference across different discourse segments, at the level of global discourse, is likely to involve

different linguistic resources, but might also involve a great deal of overlap as to which resources are used.

The utterances (15) and (20) represent transitions that are obviously increasingly less neutral than the default CONTINUE, they involve an anticipation of a shift and a sudden shift, respectively, in the center of attention (Cb). It is no coincidence that more complex constructions are used at this point: in the utterance (15), the NP anticipating the center shift (*ixmuch'a*) is bearing the topic marker *ba* (with allomorphic *'a*) and is occupying initial position, while the NP whose topical demotion is anticipated is also bearing the topic marker *ba*. In the utterance (20) the NP realizing the center which is promoted to default preference is again bearing the topic marker *ba*.[18] In §3.4 below, I will show that *ni* and *ba* share this discourse management functionality in the domain of attention transitions.

3.4 The overlap of *ni* and *ba* as NP marking devices

Let us now see in (22–24) what happens when an entity is introduced, not as local topic, but as global discourse topic. These utterances show the beginning of an interview with a traditional drum-maker, Alberto (Alb-m). At the beginning of the interview, Bernardino (Bern-m) directs his attention to the camera to explain what the video recording session will be about, in example (22).

(22) *Une=ba*   *u-ch-en*   *joben* *i*   *bada=ba* *kä-x-e*
      PRO3=TOP ERG3-make-IPFV[ABS3] drum and now=TOP ERG1-go-IPFV
      kä-k'at-b-en-la.
      ERG1-ask-BEN-IPFV[ABS3]-PL.INCL
      (Bern-m): 'He makes **drums** and now we are going to ask him.'
      [chf_HT_ALB_7_(00:09-00:12), Delgado-Galván 2018]

Notice that *joben*, 'drum', appears as a bare noun object. After asking for the full name of the drum-maker and his professional activity, the next utterance (23) is now directed to start the main interview on drums. Now the NP *joben* bears the determiner *ni* and is being set up as the main topic of the global discourse.

(23) *Kachka=da* *u*   *y-ut-e*   *ni* *joben?* *Kachka*
      Q=PROX   AUX.IPFV ERG3-build-IPFV[ABS3] DET drum   Q
      u-täk'-an?   *Kachka u-xup-o?*
      ERG3-start-IPFV[ABS3] Q   ERG3-finish-IPFV[ABS3]
      (Bern-m): 'How is **the drum** made? How is it started? How is it finished?'
      [chf_HT_ALB_27-28_(00:35-00:39), Delgado-Galván 2018]

---

[18]Note that the particle *ba* is quite multi-functional and flagging topicality-shifts would be only one of its possible contributions in the language.

The reply of the drum-maker validates *joben* as the main topic of the global discourse, in (24).

(24) **Ni joben** *kä-täk'-e'*   *kä-jok'-än*   *dok formon.*
     DET drum ERG1-begin-IPFV[ABS3] ERG1-dig-IPFV[ABS3] COM chisel

(Alb-m): '**The drum**, I start to dig it out with the chisel.'
[chf_HT_ALB_29_(00:40-00:42), Delgado-Galván 2018]

It would then seem that a function of *ni* is to label an NP as evoking a center that constitutes a main global topic rather than just a local topic. But in fact, *ni* can serve the same purpose that the topic marker *ba* fulfilled in the examples (15) and (20) above as a facilitator of center shifts. I illustrate this with an extract from the Frog Story narrative task. After narrating how the kid of the Frog Story arrived at the tree and climbed on it, the storyteller announces a center shift as follows (notice that *ni* contracts to *n* before vowels in fast speech).

(25) *kani aw-äl-e*   *a*   *pas-i*   *tänxin te'=ba?*
     Q   ERG2-say-IPFV[ABS3] AUX.PFV exit-PFV[ABS3] middle tree=TOP
     A   *pas-i*   **n-aj-xoch'.**
     AUX.PFV exit-PFV[ABS3] DET-CLF.M-owl

(Esm-f): 'and who do you think came out from the middle of the tree? **The owl** came out!'
[chf_FrogStory_ESM_61-63_(03:42-03:49), Delgado-Galván 2018]

While in (15–17) we had a RETAIN transition followed by a SMOOTH SHIFT, here it is a rhetoric question, rather than a RETAIN transition that prepares the center shift in the utterance that answers the question. The type of transition of the rhetoric question of (25) is a ROUGH SHIFT transition which is then followed by a RETAIN transition in the answer to the question. The ROUGH SHIFT is caused by the abrupt replacement of the kid as local topic of the previous context by an entity-variable evoked by the interrogative pronoun. Then the RETAIN transition reflects the fact that the attention is upon the subject interrogative pronoun and upon its value in the answer, **the owl**. Since the ROUGH-SHIFT introduces an interrogative pronoun as a dummy topic, in the sense that it is a variable, the real topic introduction happens when the value of this dummy topic is revealed, in the answer. The question pronoun simply removes the currently activated discourse-entity (the kid) from the center of attention while the subject NP in the answer fills in the corresponding empty spot with the help of a RETAIN transition. So the topic-shift is somehow delayed until the second utterance of (25), it is there

where the discourse entity ***ajxoch'*** is evoked. What matters for the discussion is that it is precisely at this point where the determiner *ni* appears decorating the NP, flagging a shift of center aimed at the owl. The next utterance (27) indeed has the owl as Cb and Cp, evoked by the indexes in the verbs, thus displaying a CONTINUE transition.

(26) Utterance (27):
      [Cb(***ajxoch'***), Cf(Cp(***ajxoch'***) > ***yokajlo'***)]
      Cp = Cb;
      Cb = previous Cb;
      Ct: CONTINUE

(27) *A   pas-i='a   u-bwejtes-i   yokajlo'.*
      AUX.PFV exit-PFV[ABS3]=TOP ERG3-scare-PFV[ABS3] kid
      (Esm-f): 'He [the owl] came out and scared the kid.'
      [chf_FrogStory_ESM_64_(03:50-03:53), Delgado-Galván 2018]

On the next utterance (29), however, the attention is directed to the least high-ranked forward-center of (27): the kid, ***yokajlo'***, without any allusion to the owl. This SMOOTH-SHIFT transition prompts the use of a non-neutral construction, with a preposed subject NP bearing a topic marker *ba*.

(28) Utterance (29):
      [Cb(***yokajlo'***), Cf(Cp(***yokajlo'***) > ***iski***)]
      Cp = Cb;
      Cb ≠ previous Cb;
      Ct: SMOOTH SHIFT

(29) *De ya'-i   **yokajlo'=ba** de iski a   yäl-i   une.*
      PREP SD-DIST kid=TOP   PREP above AUX.PFV fall-PFV[ABS3] PRO3
      (Esm-f): 'Afterward the kid fell from above.'
      [chf_FrogStory_ESM_65a_(03:53-03:56), Delgado-Galván 2018]

In the next few lines, the narrative describes how the dog passes nearby running away from a swarm of wasps. Because the dog (***wichu'***) is the main player in the immediately previous context to example (31) below, and it is evoked there again by an ERG3 index in the locative relative clause, it constitutes the backward-looking center of (31). However, it is evoked in an embedded position and the kid

yokajlo', being the subject of the main clause, is set up as the preferred center. There is an attention shift in progress.

(30) Utterance (31):
[Cb(**wichu'**), Cf(Cp(**yokajlo'**) > **wichu'**)]
Cp ≠ Cb;
Cb ≠ previous Cb;
Ct: ROUGH-SHIFT

(31) *De   ya'-i   käda   an   [u-]bix-e   tä   puts'-e,   yokajlo'*
PREP SD-DIST where EXIST[ABS3] ERG3-go-IPFV PREP escape-INF kid
täkä pas-i   tä   puts'-e.
also exit-PFV[ABS3] PREP escape-INF
(Esm-f): 'Afterward, where he [the dog] passed escaping, **the kid** also passed escaping.' [chf_FrogStory_ESM_67_(04:02-04:06), Delgado-Galván 2018]

While *yokajlo'* is not decorated in any way, it is in preverbal position, which adds to its salience-related position. However, the narrator immediately re-elaborates the main clause of (31) as the utterance in (33) and frames *yokajlo'* with both particles *ni* and *ba*.

(32) Utterance (33):
[Cb(**yokajlo'**), Cf(Cp(**yokajlo'**))]
Cp = Cb;
Cb ≠ previous Cb;
Ct: SMOOTH-SHIFT

(33) *A   k'ot-i   tä   puts'-e   ni yokajlo'=ba.*
AUX.PFV arrive-PFV[ABS3] PREP escape-INF DET kid=TOP
(Esm-f): 'The kid arrived escaping.'
[chf_FrogStory_ESM_68a_(04:07-04:08.5), Delgado-Galván 2018]

A new shift comes with the utterance (35) where the kid has been reduced in salience, being evoked by the ABS3 index of the transitive verb form *ubwät'esi*, while the owl (**najxoch'**) is evoked twice, by two NPs in the salient position of subject of the verb. First the narrator evokes 'the owl' with an NP composed of the general term for 'bird', modified by a relative clause (*ni mut jini kä ubwät'esiba*,

'the bird that scared him'), and then the speaker zooms in on the word she was looking for: *ajxoch'* the owl. To signal such transition from **the kid** to **the owl** as the preferred center, both NPs evoking the owl are accordingly introduced by *ni.*

(34) Utterance (35):
      [Cb(***yokajlo'***), Cf(Cp(***mut=ajxoch'***), *yokajlo'*)]
      Cp ≠ Cb;
      Cb = previous Cb;
      Ct: ʀᴇᴛᴀɪɴ

(35) *Porke [u-]num-e   u-ch-en   segui **ni   mut jin-i***
      ᴄᴏɴᴊ ᴇʀɢ3-pass-ɪᴘꜰᴠ ᴇʀɢ3-do-ɪᴘꜰᴠ[ᴀʙs3] follow ᴅᴇᴛ bird ᴅᴇᴍ-ᴅɪsᴛ

      *kä   u-bwät'es-i='a   **n-aj-xoch'=ba.***
      ᴄᴏᴍᴘ ᴇʀɢ3-scare-ᴘꜰᴠ[ᴀʙs3]=ᴛᴏᴘ ᴅᴇᴛ-ᴄʟꜰ.ᴍ-owl=ᴛᴏᴘ

      (Esm-f): 'Because he was following him, the bird that scared him, the owl.' [chf_FrogStory_ESM_68b_(04:06-04:14), Delgado-Galván 2018]

On several occasions the functional overlap of the determiner *ni* and the enclitic *ba* is apparent. Either because one appears instead of the other (25–29) or because they co-appear on the same NP, as in example (33–35). The overlap and competition between *ba* and *ni* to mark transitions in NP salience can be nicely observed in the following self-correction.

(36) *Mach kumpale peru täkä **ni   bit   anima-jo'**, täkä **bit   buch'-jo'=ba**,*
      ɴᴇɢ buddy   but also ᴅᴇᴛ small animal-ᴘʟ also small fish-ᴘʟ=ᴛᴏᴘ

      ejte jits'-o'   täkä.
      ꜰɪʟ be.hungry[ᴀʙs3]-ᴘʟ also

      (Luc-m): 'No, buddy, but also **the little animals**, also **the little fish** are hungry as well.'
      [chf_TwoFishingMen_LUC_038-39-40_(02:02-02:09), Delgado-Galván 2018]

In example (36) we can observe two mentions of the same referent (the fishes being fed by one of the participants) with alternate NPs and parallel discourse statuses. Interestingly, one of the alternatives is introduced by *ni* while the other alternative is bearing the topic marker *ba* instead. The reason of the rephrasing is evidently a rectification of the description, replacing the vague *bit animajob* ('little animals') with the more precise *bit buch'jo'* ('little fish'), but along the

correction the speaker inadvertently switches from using *ni* to using *ba* for an identical discourse status of the NP.

I now turn to a sample extracted from an interview to illustrate the association of *ni* with center transitions outside a narrative monologue. In this interview, the chapel's president, Felipe (Fel-m), explains many details of the festivities related to the agricultural cycle and to *Santiago Apóstol*. To properly understand the exchange that follows, the reader should be aware of the following cultural facts: in Yokot'an festivities, three different types of musical ensembles can be encountered with different roles. In the interview selection shown below, the attention switches from one to another type of musical ensemble regarding the question whether they get any payment for their performance.[19] After explaining how the main festivity will take place, Felipe (Fel-m) adds a final comment on how in former times the musicians, *musiku*, would get paid, and how eventually drummers, *ajjobeno'*, would show up (37).

(37) *De   ke   ajn-i=ba   u-toj-e'-o'   **musiku***
      PREP COMP be.located-PFV[ABS3]=TOP ERG3-pay-IPFV[ABS3]-PL musician
      *i   abeses   y-ajn-e   **aj-joben-o'.***
      and sometimes ERG3-be.located-IPFV[ABS3] CLF.M-drum-PL

      (Fel-m): 'As it was before, they would pay **the musicians** and sometimes **the drummers** would attend.'
      [chf_CONV_FEL_219-(08:21-08:24), Delgado-Galván 2018]

These are not kept as topics since, immediately after this comment, the conversation goes on to explain other aspects of the festivity. Nevertheless, further ahead – more than twenty lines later – the interviewer, Argelia (Arg-f), brings back the theme of the musicians and drummers and asks about whether they are paid nowadays – example (39) –, reintroducing them with *ni* in subject position of a passive sentence, i.e. as preferred centers for the next utterance while they were not evoked in previous sentences (we have a ROUGH-SHIFT transition). Observe that, since the interviewer completely changes the subject matter, there is no backward-looking center: no entity from the previous utterance is retaken in the current one.

---

[19]The loanword *musiku* from the Spanish word for musician *músico* refers to musicians playing European instruments (e.g. the snare drum, the bass drum and the saxophone). Besides this ensemble, two types of native ensembles perform. The terms *joben* and *ämay* (and their derivatives, *ajjoben* and *ajämay* which refer to the corresponding musicians) refer to double-sided drums and a cane flute respectively. Finally, the terms *tunkul* and *pochó* refer to a special slit log drum and a wax-headed flute, respectively, and which also form a special ensemble.

(38) Utterance (39):
      [Cb(?), Cf(Cp(<***ajmusiku***, ***ajjoben***>))];
      Cp ≠ Cb;
      Cb ≠ previous Cb;
      Ct: ZERO (≈ROUGH-SHIFT)

(39) *I   u-toj-k-an   kära ke   jin-i   **n-aj-joben-o'***
      and ERG3-pay-PASS-IPFV Q   COMP DEM-DIST DET-CLF.M-drum-PL
      n-aj-musiku?
      DET-CLF.M-musician

      (Arg-f): 'And do **the drummers**, **the musicians**, get paid?'
      [chf_CONV_FEL_247_(09:17-09:20), Delgado-Galván 2018]

Felipe (Fel-m) selects a subtopic as backward-looking center, the musicians, and answers about them that they are paid, in example (41).

(40) Utterance (41):[20]
      [Cb(***ajmusiku***), Cf(Cp(***ajmusiku***))];
      Cp = Cb;
      Cb ≠ previous Cb;
      Cb ⊂ previous Cb;
      Ct: CONTINUE or SMOOTH SHIFT

(41) ***Aj-musiku   u-toj-k-an   une.***
      CLF.M-musician ERG3-pay-PASS-IPFV PRO3

      (Fel-m): '**The musicians**, they get paid.'
      [chf_CONV_FEL_248_(09:19-09:21), Delgado-Galván 2018]

The interviewer (Arg-f) now reselects in example (43) the drummers as center of attention, provoking a ROUGH-SHIFT and as expected the NP is decorated with *ni*, and in fact also with *ba*. Observe that due to the lack of competition with any other referential device, the sole referential device of the utterance gets maximal

---

[20]Since the interviewer puts forward a question about two discourse entities in (39), the transition type of (41) is not exactly represented by the available types, but would rather be an intermediate case between SMOOTH SHIFT and CONTINUE, since the Cb is not identical to, but it is included in the previous Cb. It is possible to classify the transition as a CONTINUE, if the inclusion (Cb ⊂ previous Cb) gets emphasized, or as a SMOOTH-SHIFT if the inequality (Cb ≠ previous Cb) gets emphasized. These details are not important for the aim of our discussion.

salience and thus its evoked entity turns into the preferred center Cp of the utterance. No Cb exists since no entity from the previous utterance (41) is evoked again in (43).

(42) Utterance (43):
[Cb(?), Cf(Cp(***ajjoben***))]
Cp ≠ Cb;
Cb ≠ previous Cb;
Ct: ᴢᴇʀᴏ (≈ʀᴏᴜɢʜ-sʜɪꜰᴛ)

(43) *i   **ni   aj-joben-ob=ba?***
and ᴅᴇᴛ ᴄʟꜰ.ᴍ-drum-ᴘʟ=ᴛᴏᴘ
(Arg-f): 'and **the drummers**?'
[chf_CONV_FEL_249_(09:21-09:23), Delgado-Galván 2018]

Felipe (Fel-m) accordingly accepts the local-topic switch and answers about the drummers. Observe how on both examples (43) and (45) the NP is decorated in an identical way, first as flagging of a ʀᴏᴜɢʜ-sʜɪꜰᴛ and then as an acceptance of it.

(44) Utterance (45):
[Cb(***ajjoben***), Cf(Cp(***ajjoben***) > **payment**)][21]
Cp = Cb;
Cb = previous Cb;
Ct: ᴄᴏɴᴛɪɴᴜᴇ

(45) ***N-aj-joben-ob=ba***   *igual täkä u-toj-k-an   peru une   mach*
ᴅᴇᴛ-ᴄʟꜰ.ᴍ-drum-ᴘʟ=ᴛᴏᴘ same also ᴇʀɢ3-pay-ᴘᴀss-ɪᴘꜰᴠ but   ᴘʀᴏ3 ɴᴇɢ
y-o   u-ch'-e-jo'.
ᴇʀɢ3-want ᴇʀɢ3-take-ɪᴘꜰᴠ[ᴀʙs3]-ᴘʟ
(Fel-m): 'The drummers, are also paid, but they don't want to take it.'
[chf_CONV_FEL_250_(09:23-09:26), Delgado-Galván 2018]

The following sequence of utterances (separated by commas under the same example 47 and with transitions labeled as 47a, 47b and 47c) maintains the same local topic (Cb) and the same anticipated topic (Cp). Accordingly, the drummers are evoked as minimally as usual in these cases: with the person markers on the verb only, without using an NP introduced by *ni.*

---

[21]A payment of some kind is evoked by the ᴀʙs3 index from *uch'ejo'.*

(46) a. Utterance (47a), corresponding to *si uk'atäno' chich*:
[Cb(***ajjoben***), Cf(Cp(***ajjoben***) > **payment**)]²²
Cp = Cb;
Cb = previous Cb;
Ct: CONTINUE

  b. Utterance (47b), corresponding to *peru pekenia koperasion ubintejo'ne*:
[Cb(***ajjoben***), Cf(Cp(***ajjoben***) > ***koperasion***)]
Cp = Cb;
Cb = previous Cb;
Ct: CONTINUE

  c. Utterance (47c), corresponding to *mach uk'atänjo' pwej una kantidad*:
[Cb(***ajjoben***), Cf(Cp(***ajjoben***) > ***kantidad***)]
Cp = Cb;
Cb = previous Cb;
Ct: CONTINUE

(47) *Si u-k'at-än-o'   chich, peru pekenia koperasion*
yes ERG3-ask-IPFV[ABS3]-PL true but little   contribution
u-b-int-e-jo'=ne,   mach u-k'at-än-jo'   pwej
ERG3-give-PASS-IPFV[ABS3]-PL=PRO3 NEG ERG3-ask-IPFV[ABS3]-PL thus
una kantidad.
one amount

(Fel-m): 'They ask, yes, but they are given a small contribution, they don't request a (fixed) amount.'
[chf_CONV_FEL_251-252_(09:26-09:32), Delgado-Galván 2018]

But then the interviewer (Arg-f) switches once more the center of attention, now to request information on the last type of musical assembly (***tunkul-pocho*** **musicians**), performing a ROUGH-SHIFT transition on example (49).

(48) Utterance (49):
[Cb(?), Cf(Cp(***tunkul-pocho*** musicians))]
Cp ≠ Cb;
Cb ≠ previous Cb;
Ct: ZERO (≈ROUGH-SHIFT)

---

²²Again, a payment of some kind is evoked by the ABS3 index from *uk'atäno'*.

(49)   *i   ni   jin   u-jäts'-e'   este ni   tunkul   i*
       and DET DEM ERG3-hit-IPFV[ABS3] FIL DET slit.log.drum and
       pocho=ba?
       wax.headed.flute=TOP

       (Arg-f): 'and those who play the tunkul and the pocho?'
       [chf_CONV_FEL_253_(09:33-09:36), Delgado-Galván 2018]

Observe how this is a more complicated NP than the one used in (43), above. The determiner *ni* is marking a relative clause 'those who...' (*jin ujäts'e' ni tunkul i pochoba*), but it is also introducing the NP *tunkul* inside the clause. Again we find *ba* seemingly playing a similar or complementary role to *ni* in the context of a ROUGH-SHIFT transition.

From the kind of data presented in this section, I conclude that *ni* has a function related to topicality-shifting. In particular, it seems to flag mostly ROUGH-SHIFT transitions (including ZERO transitions), and occasionally RETAIN transitions. The fact that attentional shifts can be performed by a sequence of two transitions, with the first preparing the second, complicates the assessment of these results. For example, in the case of *ni* both cases of RETAIN are teamed-up with a previous transition. Also (33) is technically a SMOOTH-SHIFT transition, but it could be counted here as a ROUGH-SHIFT transition, because it is a rephrasing of a previous utterance whose transition belongs to this category. Thus I interpret the rephrasing in (33) as a correction or reinforcement rather than a genuine new transition.[23] Therefore I assign to (33) the same transition category than the previous utterance. In the case of repetitions of an NP as acceptance of a ROUGH-SHIFT transition proposed by another speaker, *ni* can appear in CONTINUE transitions (see the sequence 43–45). I do not display such repetition-cases in Figure 3.

Most of these *ni* and *ba* insertions in NPs seem to involve a transition in which Cp ≠ Cb. Regarding the overlap of function between *ni* and *ba*, it is beyond the scope of this study to establish whether there are differences between them (if any) in these contexts. The main point of these distributional analogies is to make a stronger case for *ni* to be a transition discourse-marker. Now that I have established a discourse-management basis for the use of *ni*, I will link, in §4, the synchronic array of its uses to the diachronic picture of *ni* as a development of the demonstrative *jini*. This will not only clarify the status of *ni* as a nascent definite marker but will also throw light on two apparently disparate observations in the grammaticalization literature of articles. The section begins with a very brief display of the diachronic evolution of *ni* as proposed in the literature.

---

[23]Adopting such a view would imply that sequences of utterances of which the second is a correction, re-elaboration or rephrasing of the first should be treated differently than regular sequences.
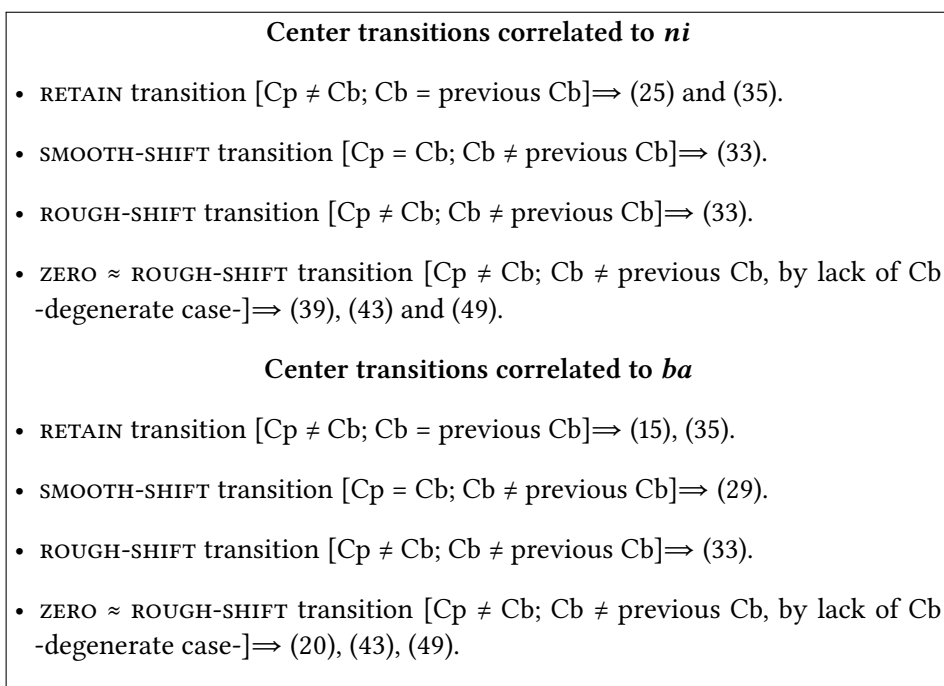

Center transitions correlated to *ni*

- RETAIN transition [Cp ≠ Cb; Cb = previous Cb]⟹ (25) and (35).

- SMOOTH-SHIFT transition [Cp = Cb; Cb ≠ previous Cb]⟹ (33).

- ROUGH-SHIFT transition [Cp ≠ Cb; Cb ≠ previous Cb]⟹ (33).

- ZERO ≈ ROUGH-SHIFT transition [Cp ≠ Cb; Cb ≠ previous Cb, by lack of Cb -degenerate case-]⟹ (39), (43) and (49).

Center transitions correlated to *ba*

- RETAIN transition [Cp ≠ Cb; Cb = previous Cb]⟹ (15), (35).

- SMOOTH-SHIFT transition [Cp = Cb; Cb ≠ previous Cb]⟹ (29).

- ROUGH-SHIFT transition [Cp ≠ Cb; Cb ≠ previous Cb]⟹ (33).

- ZERO ≈ ROUGH-SHIFT transition [Cp ≠ Cb; Cb ≠ previous Cb, by lack of Cb -degenerate case-]⟹ (20), (43), (49).

Figure 3: Transition-motivated framing of NPs with *ni* and *ba*

4 *Ni* from demonstrative to article

I mentioned earlier that *ni* is likely a recent innovation. While variants of the distal demonstrative *jini* are attested in late epigraphic writing on pottery (Mora-Marín 2009: 114, 120–121) and in the only known colonial text of Yokot'an, dating from 1610-1612, the *Maldonado-Paxbolon-Papers* (Smailus 1975), the determiner *ni* is not found on historical records.[24] Mora-Marín (2009: 120-121) makes the explicit claim that *ni* grammaticalized from *jini* (Figure 4).
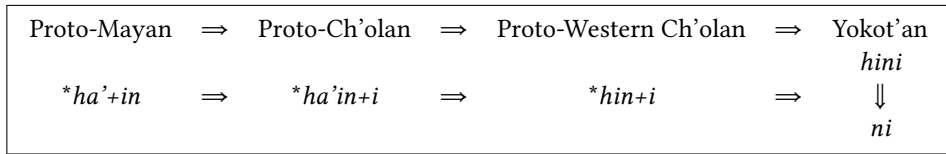
Proto-Mayan	$\Rightarrow$	Proto-Ch'olan	$\Rightarrow$	Proto-Western Ch'olan	$\Rightarrow$	Yokot'an
						hini
*ha'+in	$\Rightarrow$	**ha'in+i*	$\Rightarrow$	**hin+i*	$\Rightarrow$	$\Downarrow$
						ni

Figure 4: Reconstruction of the sources of *jini* and *ni*

This diachronic axis linking *ni* to the distal demonstrative *jini* allows me to exploit grammaticalization theory.[25] The grammaticalization approach and the development paths it suggests provide a detailed typological grid to classify and understand the functioning of article-like forms in under-described languages (Himmelmann 2001: 832). For this reason I provide in §4.2 a brief overview of the grammaticalization paths of articles from demonstratives proposed in the literature, as these developments are directly relevant to the forms available in Yokot'an. Each stage or transition between stages also helps to crystallize particular sets of uses of a form in a given language. Since *ni* presumably originates in the demonstrative *jini*, and given its main discourse function as center-attention management device, rather than as bearer of special denotational semantics, one

---

[24] A candidate for one instance of the form *ni* in the document would be the written sequence ⟨*hainniçutthan*⟩ which appears at line 13 of page 163 in the manuscript. The interlinearized version can be consulted in Smailus (1975: 71, 158) who suggests a reading of the sequence as *hain-i çut than*, rather than *hain ni çut than*. This analysis would settle the sequence ⟨*hainni*⟩ as demonstrative plus (deictic?) enclitic (*haini*) rather than demonstrative plus determiner (*hain ni*).

[25] The complex interaction of deictic enclitics, focus markers and pronominal/demonstrative roots gives some room for slightly different proposals on diachronic developments. For example, Mora-Marín (2009: 121) claims that *ha'in* was used as an article in Proto-Ch'olan and that both the Proto-Western-Ch'olan and the Proto-Eastern-Ch'olan branches developed it further as definite article. A somewhat different proposal, which shows in more detail the complexity of the process, can be consulted in Becquey (2014: 392–422). The overview of such different proposals is beyond the scope of this study but suffices to say that in either case *hini* and *ni* are linked, either by both being directly derived from a common ancestor demonstrative/focus marker *haini* or by *ni* being a further reduction of *hini*.

may ask how advanced it is in the various grammaticalization paths from demonstratives to article. I start by pointing characteristics that set *ni* apart from a demonstrative.

4.1 Telling apart articles from demonstratives

4.1.1 Frequency criteria

Faced with a puzzle similar to mine, namely, how to assess the function of a certain determiner in an under-described language, Cyr (1993) takes a small sample of languages to count the frequency of use of demonstratives and articles. She does so to propose the following frequency criterion as an auxiliary tool to assess the likelihood of a given particle of being an article in an undescribed language:

> [...] all the languages that have a definite article use it with *more than 39%* but with fewer than 55% of the NPs. Moreover, in any language, the frequency in the use of a demonstrative determiner does not exceed 7.07% of the NPs. (Cyr 1993: 222) (Sample: Finnish, French, Italian, Cree, Swedish, Montagnais, German)

I show in the Table 3 and Table 4 a similar count for Yokot'an, as established in the Frog Story narrative and the Two Fishingmen story:

Table 3: Frequency of determiners in the Two Fishingmen story

Lexical NP	Bare Noun	Indef.	*ni*	Poss.	Dem.
277	86	12	82	91	6
100%	31%	4.3%	29.6%	33%	2.1%

Table 4: Frequency of determiners in the Frog story

Lexical NP	Bare Noun	Indef.	*ni*	Poss.	Dem.
106	69	11	13	12	1
100%	65.1%	10.4%	12.26%	11.3%	0.94%

Quite clearly, on frequency figures and taking as guide the numbers from Cyr (1993), the determiner *ni* runs well below the expected article use frequency, but above the expected demonstrative use.

One may interpret this in two ways. In one of them the element counted is not really a completely developed article in the sense that its range of uses is still limited and leaves out many uses of more prototypical articles.[26] In a different perspective one may consider the possibility that the element in question can be used in every way a prototypical article can, but competes with other formal resources in many of these contexts. Both alternatives would account for a lower frequency than expected regarding Cyr (1993)'s criteria. What should be noted, however, is that in a language where such article is optional in most contexts, the frequency figures can be subjected to great variation.

4.1.2 Qualitative criteria: Anti-demonstrative contexts

Since at any stage of its grammaticalization a definite article can preserve some distributions and functions from previous stages, it can share domains of use with demonstratives. However some of the new extended uses are less well suited for demonstratives and this is one of the clues that differentiates a definite article from its ancestor. One such use is the so-called larger situation use in which the article accompanies first mentions of entities that are considered to be identifiable by general knowledge of the world and culture (Himmelmann 2001). We have seen in (7) above that with globally unique entities as *the sun*, the use of *ni* is avoided. However, *ni* becomes more readily available with institutional roles. This is shown in examples (50) and (51) in a conversation where Alfonso (Alf-m) explains the role played by some of the specialists in the village. Thus, some concrete cases are discussed, but many general statements are made which do not concern any particular individual but rather the role itself. Both (50) and (51) are generic statements not involving particular individuals.

(50)   *Dos año [u-]num-e*   **ni patron**.
          two year ERG3-pass-IPFV DET patron
          (Alf-m): '**The patron** lasts 2 years (in charge).'
          [chf_HPatron_ALF_34_(01:20-01:22), Delgado-Galván 2018]

The utterance (50) is part of a general characterization of the patron role – in fact (50) is a characterizing statement itself – and (51) is part of a general account

---

[26]Interestingly, Greenberg (1978: 62) considers an example from Bwamu (Niger-Congo family) of a "nascent article which is [...] at a point between a zero stage demonstrative and a Stage I definite article", but ultimately rejects it as a candidate for his Stage I article (definite article). One main factor that pushes him to exclude it from a Stage I status is Manessy's (1960: 93) report on the low discourse frequency and the optionality of its use. The exact same comment could be directed to the determiner *ni* of Yokot'an.

of the diseases provoked by the *yumka'ob* spirits, the "owners of the earth", but it is not a characterizing statement.

(51) *Ora aj-t'äbäla   mach une uy-äk'-e'   u-ba=une,   peru duro*
      now CLF.M-adult   NEG   PRO3 ERG3-heal-IPFV ERG3-REFL=PRO3 but   hard
      *chita   tuba u-ts'äkäl-in   **ni   yerbateru.***
      also(?) PREP ERG3-cure-IPFV[ABS3] DET healer
      (Alf-m): 'Now the adults don't heal, but it is hard as well for them to get cured by **the healer**.'
      [chf_HPatron_ALF_630-631_(28:44-28:50), Delgado-Galván 2018]

Reference to kinds is also a use where an article is better suited than a demonstrative. We can see two examples of kind-denotation to deer below.

(52) *Ida=ba   pue ajn-i   k'en   chimay. **Ni chimay***
      here=TOP so be.located-PFV[ABS3] many deer   DET deer
      u-ts'on-e-o'   ni   gente.
      ERG3-hunt-IPFV[ABS3]-PL DET people
      (Alb-m): 'Here there was a lot of deer. The people hunt **the deer**.'
      [chf_HT_ALB_72-73_(02:18-02:24), Delgado-Galván 2018]

(53) *I   che' che'chich xup-i   **ni chimay**. Ma' ni' an*
      and so   of.course finish-PFV[ABS3] DET chimay   NEG ADV EXIST[ABS3]
      bada=ba.
      now=TOP
      (Alb-m): 'So, of course **the deer** is finished. There is no more now.'
      [chf_HT_ALB_123_(04:29-04:31), Delgado-Galván 2018]

Finally, the lack of deictic contrast of *ni* can be observed in (54), which is the closing line of the Pear Story narrative. A co-occurrence within the same NP of *ni* and the proximal demonstrative *jinda* is suggestive that *ni* no longer introduces a deictic contrast. For if *ni* still held the (distal) deictic value of its diachronic source *jini*, it should be incompatible with the proximal deictic value contributed by *jinda*.[27] Such loss of deictic contrast is one of the functional criteria to identify that a former demonstrative has undergone grammaticalization (Diessel 1999: 118).

---

[27] Knowles-Berry (1984: 208, 236) provides a sample of an NP in which a distal demonstrative *jini* shares a noun with a proximal deictic enclitic (*da*): *jin-i winik-da*. I have not found such NP types in my corpus and since no context is provided – not even *sentential* context – it is hard to assess this sample.

(54) *Kama jin-i*   ***ni*** *ts'aji jin-da.*
      Q   DEM-DIST DET chat   DEM-PROX
      (Esm-f): 'That is how **this story** is.' [chf_PS_ESM_068_(03:44-03:46),
      Delgado-Galván 2018]

Notice, additionally, that *ni* can no longer inflect for deictic distance, as *jini* can: *jin-i/jin-da*, which is also a (morphological) criterion in Diessel (1999: 118). Clearly, then, the form *ni* is not just a phonological reduction of *jini*, it constitutes a new element which is located somewhere in the grammaticalization path to turn into a different marker. It is time now to compare the different uses of *ni* against the background of the paths proposed for the development of articles.

4.2 Grammaticalization path and stages

I will now assess the determiner *ni* against the grammaticalization stages of a definite article as presented by Greenberg (1978: 61–74) and Hawkins (2004: 84–86), which are presented schematically in Table 5 and Table 6. These illustrate the paths of development from a demonstrative source, other sources are not of interest here. Greenberg (1978) proposes a grammaticalization scheme in three steps for the definite article, Stage 0, Stage I early and Stage I late. Hawkins (2004) goes more into detail and proposes four logical steps of development for definite articles, but on the other hand he will not consider as definite article any determiner that still conveys deictic contrast. Thus, Stage 0 of Hawkins encompasses Greenberg's stages 0 and I early (since deictic contrast still operates), while Greenberg's stage I late is split into stages 1-2-3 of Hawkins (2004).

Table 5: Article grammaticalization stages (Greenberg 1978)

Stage 0	$\Rightarrow$	Stage I (early + late)	$\Rightarrow$	Stage II	$\Rightarrow$	Stage III
demonstrative	$\Rightarrow$	definite article	$\Rightarrow$	specific article	$\Rightarrow$	nominality marker / gender marker
pure exophoric deixis	$\Rightarrow$	identified in general	$\Rightarrow$	specific but unidentified	$\Rightarrow$	sign of nominality

In the coarsest scheme (Greenberg 1978), the main functionality of the determiner *ni* can be located in between Stage 0 and Stage I. Greenberg's Stage II (corresponding to Hawkins Stage 4) and Stage III (not represented in Hawkins 2004) have marginal relevance here as uses of *ni* related to specificity or nominality may only appear in restricted contexts (negative existential constructions and some syntactically nominalized clauses (Becquey 2014: 397, 408), respectively).

Table 6: Article grammaticalization stages (Hawkins 2004)

Stage 0	Stage 1	Stage 2	Stage 3	Stage 4
demonstrative	definite article (anchored to immediate situation)	definite article (anchored to larger situation)	definite article (unanchored)	specific article (drops inclusiveness)
exophoric via uniqueness in pointed area	exophoric via uniqueness in situation	exophoric via uniqueness within knowledge	exophoric via uniqueness within knowledge, expands to maximality	drops inclusiveness (uniqueness+maximality)
endophoric anaphoric via deixis	endophoric anaphoric via mention familiarity	anaphora via general knowledge inference	anaphora via general knowledge inference	
anchored uniqueness	anchored uniqueness	anchored uniqueness	unanchored uniqueness -generalizes to maximality (plurals) -generic reference suitable	uniqueness/maximality dropped in some context maintains existential claim
ANCHOR: immediate situation	ANCHOR: immediate situation	ANCHOR: larger situation	ANCHOR: none	
DOMAIN: deictically delimited -to subsituation -to subtext	DOMAIN: expands -to the whole visible situation -the whole text in memory	DOMAIN: expands to larger non-visible situation	DOMAIN: -restriction almost inexistent -abandons pragmatic delimitation	

Grammaticalization paths as presented in Table 5 and Table 6 are not to be taken as linear developments, but rather as logical steps that can be taken at different times or simultaneously in different pragmatic and constructional contexts. This means that the same form easily assumes different uses according to individual constructions. A case in point, to be presented in §4.3, is the negative existential construction which shelters a specialized use of *ni* which has more in common with situational uses as in the example (9) in which the NP concerned is not necessarily involved in the evolution of an anaphoric/topical chain. This lack of linearity is what leads to fragmented uses of a definite marker (see Lyons 1999: 159) which is also seen in the fact that a definite article in an early stage can already show characteristics of even the latest stages, but in restricted contexts.

The initial stage (Stage 0 in all authors) corresponds to a demonstrative, whose function is to perform situational or exophoric reference and introduces a deictic contrast with other deictic forms. Generally, it is the third-person/distal proximity deictic element from the paradigm that gives rise to the grammaticalization of an article. This is no exception in Yokot'an, as it is indeed the distal demonstrative *jini* that provides the base for *ni*. The exophoric/contrastive nature of the initial demonstrative base makes it incompatible with a generic interpretation. Clearly, then, as shown in the example (53) above, the form *ni* is beyond the initial stage (Stage 0).

The initial step of development towards an article extends the use of the demonstrative to also encompass endophoric reference, as an anaphoric (or cataphoric) device. This secondary use of the demonstrative as anaphoric device is shown, for Yokot'an *jini*, in example (55) from the Pear Story narrative. After a digression describing how a boy passed with a goat near the baskets of pears, the narrative once more returns to what the pear-collecting man is doing. The reference to him is then resumed with an anaphoric definite NP, with *jini*.

(55) *De ya'-i* **yok winik jin-i='a** *t'äb-i* *cha'-num*
PREP SD-DIST little man DEM-DIST=TOP ascend-PFV[ABS3] two-NUM.CLF
tan te'.
PREP tree

(Esm-f): 'Then **the man** (that has been mentioned) climbed again in the tree. ' [chf_PS_ESM_016_(01:16.5-01:18.5), Delgado-Galván 2018]

Such endophoric function may turn into the main or sole use of the demonstrative in its way towards developing into an article (Stage I early in Greenberg, but still Stage 0 in Hawkins as long as deixis is not dropped). At the next stage (Stage I late in Greenberg, Stage 1 in Hawkins), the identifiability of the referent is

assessed with respect to the whole visible situation or the whole previous text in memory, not just the recent text or some deictically selected subsituation. Identifiability is expanded to both textual and situational assessment and therefore the article use is restricted to anaphoric reference or to the immediate situation (for an immediate situation use of *ni*, consider that its insertion is indeed possible in an example as 9 above).

A further development is the expansion of the contexts (or "pragmatic set") within which uniqueness is assessed to also consider non-visible and/or larger situations (Stage 2 in Hawkins 2004, Stage I Late in Greenberg 1978). The association of reference gets extended from anaphoric to general-knowledge inferences, and stereotypic frames. We have seen that although *ni* has not extended to be naturally accepted with entities like *the sun*, *the moon*, etc. (see 7), it is common with institutionalized roles (50) or in relation to some stereotypic frame. Finally, a definite article reaches Hawkins's (2004: 85) Stage 3 when its use expands to unanchored uniqueness and generalizes to inclusiveness (i.e. a sort of plural uniqueness, the maximality of a group). At this point, generic reference is a suitable context for the article.

With such development path as a background, it can be observed that the determiner *ni* exhibits compatibility with some of the uses in Hawkins' Stages 1-2-3 (immediate situation-use, institutionalized roles, kind denoting). However, I wish to argue that the main function characteristic of *ni* is still at the transition between Stage 0 and Greenberg's Stage I or Hawkins' Stage 1. To see this, consider the following quote from Heine & Kuteva who, based upon Diessel (1999: 96, 128-129), explain:

> Since the adnominal anaphoric demonstrative serves a discourse internal function – to refer to the same referent as its antecedent and thus track participants of the preceding discourse – it serves as a common strategy to establish major participants in the universe of discourse. Its use involves *non-topical antecedents that tend to be somewhat unexpected*, contrastive, or emphatic. At a next stage of development, the adnominal anaphoric demonstrative becomes a definite article, whereby its use is gradually extended from non-topical antecedents to all kinds of referents in the preceding discourse. (Heine & Kuteva 2006: 101–102, emphasis mine)

It is interesting to contrast this report with the one pictured by Givón (2001: 474), which I quoted earlier: "Grammaticalized definite markers [...] arise first to mark *topical* definites."

At first, there seems to be a contradiction. Yet there isn't. By joining the observations in both quotes we can see that an attentional transition underlies the reported facts: non-topical antecedent and topical resumptive NP. Think of the antecedent as a FORWARD-LOOKING CENTER. Think of its later "unexpectedness" as reflecting the fact that it is not currently set as a PREFERRED CENTER (or, perhaps, not even as a Cb). Think of the "topical" resumptive NP as a BACKWARD-LOOKING CENTER and/or as a PREFERRED CENTER. Now we see that what seemed a contradiction hints at the specialization of an early definite of the kind found in Yokot'an. The rationale of its use is not to flag anaphoric NPs with non-topical antecedents or to flag topical anaphoric NPs, rather it is to mark the attentional transition itself.[28] A topic shift can be decomposed in two steps or components, according to the Centering Theory model. One step is to announce or prepare an incoming shift by setting Cp ≠ Cb (RETAIN transition). The second step is to execute such shift by setting Cb ≠ previous Cb (SMOOTH-SHIFT transition). Both moves can be collapsed into a single move (ROUGH-SHIFT, and ZERO as special case). From Figure 3 above, it seems that *ni* can flag both types of transition (and the one containing both moves). Given the preference of more cohesive transitions over increasingly less cohesive ones (Figure 2), however, one can expect *ni* to be more systematically used to flag the least cohesive transitions: ZERO and ROUGH-SHIFT. In fact, the condition Cp ≠ Cb across transitions covers most of the discourse-related cases I have illustrated in the present paper.[29]

Heine & Kuteva (2006) associate this particular function of flagging NPs which anaphorically evoke unexpected/non-topical entities with a stage *previous* to the demonstrative being a definite article. Givón (2001), on the other hand, associates the function of flagging topical NPs with an early definite article. Under this view, the determiner *ni* is better characterized as an early definite article, one that has not even reached Hawkins's (2004) Stage 1. Given Diessel's (1999) scheme of definite article grammaticalization (Figure 5) *ni* would be an anaphoric demonstrative specialized in anaphorically picking up non-topical referents and turn them topical (expectedly or not, i.e. with or without warning), while the original demonstrative source *jini* still holds a purely distal-anaphoric function.

Since the main focus of definiteness studies has been the Hearer-status and how it may grammaticalize, the other possible functions of an article, related to Discourse-status, have received less attention. In the above descriptions of

---

[28]The reader should be aware, however, that I am here jumping from informal notions of topical/non-topical to technical and very particular notions of "topical" vs "non-topical", as embodied by the notions of centers within Centering Theory. Yet I think the jump is enlightening for languages like Yokot'an.

[29]Further investigation would be needed, but these results are already very suggestive.

exophoric demonstrative ⟹ anaphoric demonstrative ⟹ definite article

Figure 5: Diessel's (1999) scheme of definite article grammaticalization

grammaticalization paths of the article, the Discourse-status role appears as incidental, more as an introduction context than as a main function that can be fulfilled by the article. It does not explicitly appear in Table 5 or Table 6. In many languages this particular path of evolution of articles via the Discourse-status might be more relevant to understand their synchronic use. Not only it provides a starting point to understand the distribution of an otherwise unsystematic article, but it also explains its lower frequency and its relative optionality. While *ni* has extended to being definite (in terms of the generality of contexts in which it can appear), it is still the initial specialized function of discourse-management of transitions which prompts its minimal occurrences. Since *ni* developed from a demonstrative and its main function is related to topicality while having lost any deictic value, one may wonder if it should not be regarded as a purely pragmatic marker of topicality issued from a demonstrative, in similar fashion to topic-markers in a selection of Papuan languages (de Vries 1995). Firstly, *ni* is restricted to the noun phrase, while its competitor, the topic-marker *ba* is not restricted in this way (neither are the corresponding Papuan examples of topical markers in de Vries 1995). Furthermore, in some specialized contexts, one can see *ni* inserted to convey features like specificity/referentiality, akin to more prototypical definite articles. This is what I will illustrate in the following section.

4.3 From topicality to specific referentiality marker: Special contexts

Articles often have, as the most abstract function, the function to guarantee the syntactic nominality of the expression they modify. In the most syntacticized way, this means literally creating an argument from what otherwise would be interpreted as a predicate and unable to occupy argument positions (Gillon 2015: 176). Such syntactic contrast may evolve initially from a more semantic contrast that opposes noun phrases interpreted as referring to specific entities against other noun phrases interpreted as not referring. Examples in (56) illustrate how special contexts can trigger a use of *ni* where its discourse-salience function is exploited to force (specific) referentiality. Matilde (Mat-f) is telling the story of how she got married and moved with her mother-in-law. While she was happy as to how her mother-in-law treated her, she points out an unpleasant surprise in line (56b): while the kitchen has an electrical grinder now, such grinder was not there when she moved in, she had to grind manually with a grinding stone.

(56) a. *Kol-on   kä-nojna'.   Kol-on   chich dok une, ti'i*
          leave-ABS1 ERG1-mother.in.law leave-ABS1 always COM PRO3 well
          u-k'ajalin   täkä.
          ERG3-thought also
          (Mat-f): 'I stayed with my mother in law. I stayed always with her,
          she treats me very well.'
          [chf_CONV_MAT_501-502_(19:30-19:36), Delgado-Galván 2018]
      b. *Peru mach ajn-i   **ni   molino** une.*
          but   NEG   be.located-PFV[ABS3] DET grinder PRO3
          (Mat-f): 'But **the grinder** was not [there].'
          [chf_CONV_MAT_504_(19:36-19:38.5), Delgado-Galván 2018]

The crucial point is that while on positive polarity a bare noun is generally
enough to be referential, the negative polarity in existential context forces the
speaker to call in the assistance of *ni* for the noun to unambiguously refer (56b).
The contrast is displayed below for more clarity: with *ni*, in the example (57a)
the negative context translates as negating the *location* of some referred object.
But without *ni*, in example (57b), the negative context is readily interpreted as
negating the *existence* of an object, especially when – as in this case – there was
no previous mention of the object in the conversation.

(57) a. *Mach ajn-i   ni   molino une.*
          NEG   be.located-PFV[ABS3] DET grinder PRO3
          (Mat-f): 'The grinder was not [there].'
          [chf_CONV_MAT_504_(19:36-19:38.5), Delgado-Galván 2018]
      b. *Mach ajn-i   molino une.*
          NEG   be.located-PFV[ABS3] grinder PRO3
          (Esm-f): 'There was no grinder.' [My_elicitation]
      c. *Ma' ajn-i   baile une.*
          NEG be.located-PFV[ABS3] dance PRO3
          (Mat-f): 'There was no dance.'
          [chf_CONV_MAT_474_(18:31.5-18:32.5), Delgado-Galván 2018]

Obviously, when what matters is the type of object rather than some specific
instance, no *ni* is likely to be found, like in (58), where Matilde is explaining
that you would feed small chicken with maize dough when no (industrialized)
animal-food is available:[30]

---

[30]A detail that the reader might observe is that Yokot'an has two existential verbs (in the sense
   of being used in such constructions): *an*, glossed EXIST, which does not inflect for TAM and
   ajne which inflects for TAM. Since *an* is a non-verbal predicate unable to take TAM inflection,
   ajne is used instead in all "tensed" existential constructions.

(58) I   *une, xix*   *a-b-en*   *une,   mach an*
      and PRO3 maize.dough ERG2-give-IPFV[ABS3] PRO3, NEG   EXIST[ABS3]
      alimento,   *yok   xix*   *a-b-en*   *une.*
      industrial.animal.food, small maize.dough ERG2-give-IPFV[ABS3] PRO3
      (Mat-f): 'And those, you give maize dough, there is no food, (so) you give
      them small maize dough.'
      [chf_CONV_MAT_150-151_(04:41.7-04:48), Delgado-Galván 2018]

The presence of *ni* in this negative context would again be interpreted as sug-
gesting an interpretation of *mach'an* as the negation of a location rather than of
existence (like: 'the food is not there'). It is precisely in these specialized contexts,
negative existential constructions, in which *ni* gets associated to specific refer-
ence interpretation, since in most other contexts, specific referentiality is tied to
nouns themselves as default, *ni* simply flagging a switch in attention regarding
the flow of discourse. However this marginal use, along with its inability to ap-
pear outside NPs, helps to consider the determiner *ni* in the category of definite
articles rather than in the category of topic markers.

5 Concluding remarks

I have examined Yokot'an's candidate for a definite determiner, the marker *ni*.
In trying to unravel the basis of its use, I attacked the problem from two sides. I
started with a synchronic textual-analysis perspective. In those texts with mini-
mal occurrences, I isolated a discourse pattern for the presence of *ni* using Cen-
tering Theory as a heuristic tool. On the other hand, I also used a diachronic
perspective in which I projected some of the attested possible uses of *ni* into the
grammaticalization paths proposed in the literature for the development of def-
inite articles from demonstratives. A general observation that guided this study
is the relatively low frequency and relative optionality of this particle. In this
sense, I used counting/distributional criteria regarding its frequency and its op-
tionality as compared to cross-linguistic expectations in order to determine that
a pragmatic/discourse-based explanation was called for and to show, with help
of more qualitative clues, that *ni* was beyond the grammaticalization Stage 0 as-
sociated to the demonstrative source.

I conclude that *ni* is more a discourse salience-oriented than a reference-orient-
ed resource in the sense that its likelihood to be used has more to do with atten-
tional transition types than with identifiability properties of the NP involved.
Such orientation and the overlap in function of many different linguistic re-
sources also allows more stylistic variation among speakers' use. Such variation

accounts for the fact that the low frequency does not necessarily correlate to an article with a span of uses that are limited to early stages of grammaticalization. The optionality of an article and lower frequency are in principle independent of the degree of development regarding the span of possible uses an article bears (as already suggested in general by Dryer 2014).

Both lower frequencies and relative optionality of the definite determiner in Yokot'an are a more direct reflection of multiple resources in the languages overlapping on similar functional domains than a reflection of only the intrinsic development of the article. For example, in certain contexts (as negative existential statements) it can be used to indicate referentiality/specificity, but in the overall system, (bare) noun phrases do so by themselves as default. Similarly, while the main function of *ni* is to flag attentional transitions (local topicality switches), in some contexts it can be complemented or replaced in this role by the topic-marker *ba*, which has different distributional restrictions. In other words, low-frequency of use in early definite determiners can have several independent explanations: nouns have not yet lost their capacity to be interpreted definitely (bareness is not interpreted as indefiniteness) and the nascent definite determiner might be competing with other discourse-salience markers.

Finally, the discourse-management basis of the "definiteness" underlying the use of *ni* explains why a language that does not need definite markers (since its bare nouns are generally self-sufficient in this respect) would still have them.

The logical orthogonality of two different notions, Hearer-status (identifiability) and Discourse-status (discourse-salience), and the possibility for the speakers of a language to articulate the use of a determiner around one rather than the other notion shows that to have different theories of definiteness is more interesting empirically than to reduce definiteness to a single notion that attempts to cover by generalization all the instances.

The higher frequency of *ni* in written texts and in some idiolects has undeniable relation to contact with Spanish, and at this point it is relevant to note that contact-induced change has also been blamed for the generalized spread of article systems in European languages (Schroeder 2006), which makes Mesoamerican languages a good opportunity for the study of a similar but ongoing contact-induced change.

Acknowledgements

which for reasons of time and space I could not fully implement. I am very grateful as well with the DDL research center at *Centre Berthelot* in Lyon which so kindly offered me a working space and the warm environment that made this paper possible. Thanks to Felix Ameka for some discussion on terminology. At last, but not least, I also would like to kindly acknowledge the invaluable collaboration of Esmeralda López, Bernardino Montero, Griselda Luciano, the López Méndez family and the people in San Isidro, Tapotzingo, Tecoluta and Tucta who made this research possible and the fieldwork pleasant.

Sources

Unless stated otherwise (e.g. with a label like "my elicitation" or with a reference to relevant literature), all the materials used in this study are from the *Yokot'an Space Grammar/The Oral Literature of the Endangered Cultural Practices of Yokot'an Pilgrimages Project*, lead by Amanda Alejandra Delgado Galván and to which I contributed as data collector assistant. I hereby acknowledge her kindness for granting me permission to use them. These materials are archived in the *donated archives* section of the *Language Archive* of the MPI. The "Yokot'an / Chontal de Tabasco" section from *The Language Archive* can be found at the following url address: https://hdl.handle.net/1839/00-0000-0000-001E-8B97-0. As I consider important to endorse *The Austin Principles of Data Citation in Linguistics*, I have also directly referred to this archive and to its main collector/curator in the examples and in the References section. Mainly eight texts from the area of Nacajuca municipality were consulted, with varying depths of (re-)analysis, for the present study (Table 7).

Table 7: Texts consulted

File name	Settlement	App. recording length	Type of speech event
chf_HT_ALB	Tucta	29 min	Interview/conversation
chf_CONV_FEL	Tucta	13 min	Interview/conversation
chf_TwoFishingMen_LUC	Mazateupa	21 min	Story narration
chf_MG_CAR	Tapotzingo	15 min	Match-path task
chf_HP_ALF	San Isidro	44 min	Interview/conversation
chf_CONV_MAT	San Isidro	23 min	Interview/conversation
chf_HS_BLA	San Isidro	5 min	Hunting Story task for co-motion events
chf_FS_ESM	San Isidro	6 min	Frog Story task
chf_PS_ESM	San Isidro	4 min	Pear Story task

Abbreviations

In every example from the archive in Delgado-Galván (2018) a code for the speaker identity and gender is indicated in the translation line. For example: (Fel-m) refers to a man (m := masculine gender) and (Arg-f) to a woman (f := feminine gender). The first set of numbers after the filename of the recording refer to the line numbers in the ELAN-Flex file, the second set of numbers refer to the time interval. The list of abbreviations used in the examples is the following:

EXIST	existential	PRO	pronoun
FIL	filler in conversation	SD	source deictic
NUM	numeral		
PREP	preposition		

References

Abbott, Barbara. 1999. Support for a unique theory of definiteness. *Proceedings of SALT* 9. 1–15.

Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London: Routledge.

Becquey, Cédric. 2014. *Diasystème, diachronie: études comparées dans les langues cholanes*. Vol. 1 (LOT Dissertation Series 377). Utrecht: LOT.

Cyr, Danielle. 1993. Cross-linguistic quantification: Definite articles vs demonstratives. *Language Sciences* 15(3). 195–229.

de Vries, Lourens. 1995. Demonstratives, referent identification and topicality in Wambon and some other Papuan languages. *Journal of Pragmatics* 24(5). 513–533.

Delgado-Galván, Amanda Alejandra. 2018. *Yokot'an-chontal de Tabasco*. Nijmegen: The Language Archive, MPI.

Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization* (Typological Studies in Language 42). Amsterdam: John Benjamins.

Dryer, Matthew S. 2005. Definite articles. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 154–157. Oxford: Oxford University Press.

Dryer, Matthew S. 2014. Competing methods for uncovering linguistic diversity: The case of definite and indefinite articles (Commentary on Davis, Gillon, and Matthewson). *Language* 90(4). e232–e249.

Farkas, Donka F. 2002. Specificity distinctions. *Journal of Semantics* 19(3). 213–243.

Frege, Gottlob. 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und Philosophische Kritik* 100. 25–50.

Gillon, Carrie. 2015. Investigating D in languages with and without articles. In M. Ryan Bochnak & Lisa Matthewson (eds.), *Methodologies in semantic fieldwork*, 175–203. Oxford: Oxford University Press.

Givón, Talmy. 2001. *Syntax: An introduction*. 2nd edn. Vol. 1. Amsterdam: John Benjamins.

Greenberg, Joseph H. 1978. How does a language acquire gender markers? In Joseph H. Greenberg, Charles Albert Ferguson & Edith A. Moravcsik (eds.), *Universals of human language*, vol. 3, 47–82. Stanford: Stanford University Press.

Grosz, Barbara J., Aravind K. Joshi & Scott Weinstein. 1983. Providing a unified account of definite noun phrases in discourse. In *ACL '83 Proceedings of the 21st annual meeting on Association for Computational Linguistics*, 44–50. Stroudsburg: Association for Computational Linguistics.

Grosz, Barbara J. & Candace L. Sidner. 1998. Lost intuitions and forgotten intentions. In Marilyn A. Walker, Aravind K. Joshi & Ellen Friedman Prince (eds.), *Centering theory in discourse*, 39–51. Oxford: Clarendon Press.

Grosz, Barbara J., Scott Weinstein & Aravind K. Joshi. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* 21(2). 203–225.

Gundel, Jeanette K. 1998. Centering Theory and the givenness hierarchy: Towards a synthesis. In Marilyn A. Walker, Aravind K. Joshi & Ellen Friedman Prince (eds.), *Centering theory in discourse*, 183–198. Oxford: Clarendon Press.

Gundel, Jeanette K., Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2). 274–307.

Hawkins, John A. 1978. *Definiteness and indefiniteness: A study in reference and grammaticality prediction*. London: Croon Helm.

Hawkins, John A. 1991. On (in)definite articles: Implicatures and (un)grammaticality prediction. *Journal of Linguistics* 27(2). 405–442.

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Hedberg, Nancy. 2010. Centering and noun phrase realization in Kaqchikel Mayan. *Journal of Pragmatics* 42(7). 1829–1841.

Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases*. Amherst: University of Massachusetts. (Doctoral dissertation).

Heine, Bernd & Tania Kuteva. 2006. *The changing languages of Europe*. Oxford: Oxford University Press.

Himmelmann, Nikolaus P. 2001. Articles. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language*

universals: An international handbook, vol. 1 (Handbooks of Linguistics and Communication Science 20/1), 831–841. Berlin: Walter de Gruyter.

Kamp, Hans. 1981. A theory of truth and semantic representation. In Jeroen Groenendijk, Theo Janssen & Martin Stokhof (eds.), *Formal methods in the study of language*, vol. 1, 277–322. Amsterdam: Mathematical Centre Tracts.

Kamp, Hans & Uwe Reyle. 1993. *From discourse to logic: Introduction to model-theoretic semantics of natural language, formal logic and Discourse Representation Theory*. Dordrecht: Kluwer.

Kantor, Robert. 1977. *The management and comprehension of discourse connection by pronouns in English*. Columbus: Ohio State University. (Doctoral dissertation).

Keller, Kathryn & Margaret Harris. 1946. Masculine Crab and Mosquitoes, two chontal texts. *Tlalocan* 2. 138–140.

Kibrik, Andrej A. 2011. *Reference in discourse*. Oxford: Oxford University Press.

Knowles-Berry, Susan Marie. 1984. *A descriptive grammar of Chontal Maya (San Carlos dialect) (Mesoamerica, Tabasco, Mexico)*. New Orleans: Tulane University.

Lyons, Christopher. 1999. *Definiteness* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Manessy, Gabriel. 1960. *La morphologie du nom en bwamu (bobo-oulé), dialecte de Bondoukuy*. Dakar: Université de Dakar.

Mora-Marín, David F. 2009. Reconstruction of the proto-Cholan demonstrative pronouns, deictic enclitics, and definite articles. *Transactions of the Philological Society* 107(1). 98–129.

Poesio, Massimo, Rosemary Stevenson, Barbara Di Eugenio & Janet Hitzeman. 2004. Centering: A parametric theory and its instantiations. *Computational Linguistics* 30(3). 309–363.

Prince, Ellen F. 1999. Subject pro-drop in Yiddish. In Peter Bosch & Rob van der Sandt (eds.), *Focus: Linguistic, cognitive, and computational perspectives*, 82–104. Cambridge: Cambridge University Press.

Roberts, Craige. 1998. The place of centering in a general theory of anaphora resolution. In Marilyn A. Walker, Aravind K. Joshi & Ellen Friedman Prince (eds.), *Centering theory in discourse*, 359–400. Oxford: Oxford University Press.

Roberts, Craige. 2003. Uniqueness in definite noun phrases. *Linguistics and Philosophy* 26(3). 287–350.

Roberts, Craige. 2012. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics & Pragmatics* 5. 1–69.

Russell, Bertrand. 1905. On denoting. *Mind* 14(56). 479–493.

Schroeder, Christoph. 2006. Articles and article systems in some areas of Europe. In Giuliano Bernini & Marcia L. Schwartz (eds.), *Pragmatic organization of discourse in the languages of Europe* (Empirical Approaches to Language Typology 20-8), 545–611. Berlin: Mouton de Gruyter.

Schumann Gálvez, Otto. 2012. *Introducción al chontal de Tabasco.* Mexico City: Universidad Nacional Autónoma de México, Instituto de Investigaciones Antropológicas, Programa de Investigaciones Multidisciplinarias sobre Mesoamérica y el Sureste.

Schwarz, Florian. 2009. *Two types of definites in natural language.* Amherst: University of Massachusetts. (Doctoral dissertation).

Schwarz, Florian. 2013. Two kinds of definites cross-linguistically. *Language and Linguistics Compass* 7(10). 534–559.

Smailus, Ortwin. 1975. *El maya-chontal de Acalan: Análisis lingüístico de un documento de los años 1610-1612* (Cuadernos del Centro de Estudios Mayas 9). Mexico City: Centro de Estudios Mayas, Instituto de Investigaciones Filológicas; Coordinación de Humanidades, Universidad Nacional Autónoma de México.

Strawson, Peter F. 1950. On referring. *Mind* 59(235). 320–344.

Walker, Marilyn A., Aravind K. Joshi & Ellen Friedman Prince. 1998. Centering in naturally occurring discourse: An overview. In Marilyn A. Walker, Aravind K. Joshi & Ellen Friedman Prince (eds.), *Centering theory in discourse*, 1–28. Oxford: Clarendon Press.

Walker, Marilyn A. & Ellen F. Prince. 1996. A bilateral approach to givenness: A hearer-status algorithm and a centering algorithm. In Thorstein Fretheim & Jeanette K. Gundel (eds.), *Reference and referent accessibility* (Pragmatics & Beyond New Series 38), 291–306. Amsterdam: John Benjamins.