Word order optionality in the Nata double object construction

Emily Sadlier-Brown
University of British Columbia

Abstract: This paper investigates the pragmatic motivation for the use of a non-canonical word order (NCWO) variant of the Nata (Bantu) applied double object construction. The information structural notions *focus*, *givenness*$_S$ (salience), *definiteness/givenness*$_K$ (shared knowledge) and *topichood* are examined in turn and each is eliminated as a sufficient motivator for NCWO. Givenness$_K$ is then shown to be a necessary condition because NCWO marks givenness$_K$ of the theme object in the few contexts in which the construction appears. To determine the relevant context, I propose and compare two "addressee knowledge states" that differ with respect to whether or not a speaker believes an addressee will be able to determine that a theme referent is given$_K$ based on contextual information alone. Based on the outcome of this comparison, I conclude that NCWO is a disambiguation device used to signal givenness$_K$ when a speaker believes an addressee might erroneously consider non-given$_K$ referents. I then determine that NCWO is not required in this context if ambiguity is desired on the part of the speaker. From this I conclude that the use of NCWO is not determined by context alone but by speaker choice.

Keywords: Nata, optionality, information structure, givenness$_K$, disambiguation, gradience

1 Introduction

In this paper, I investigate the pragmatic motivation for a word order alternation in the applied double-object construction (DOC) of Nata, a Bantu language spoken in Tanzania[1]. In Nata, a DOC containing a benefactive and a theme can be realized by either of two grammatical forms, one in which the order of objects is benefactive – theme (1), and one in which the order is theme – benefactive (2). The two forms below are truth-conditionally equivalent. Aside from word order, the surface form of the two constructions is identical. [2]

(1)	Masáto	a-ka-ɣór-er-a	u-mu-aarimú	e-ɣí-taβo
	Masato	SM1-PST-buy-APPL-FV	PPF-C1-teacher	PPF-C7-book
			BENEFACTIVE	THEME

'Masato bought the teacher the book.'

[1] All DOCs in this paper are "applied", containing an applicative (APPL) morpheme that increases a verb's valency by 1. In this paper, the added argument is always the benefactive. Nata has a small class of verbs that are inherently ditransitive. Since this class may behave differently, it is left for future work.

[2] Abbreviations used: 1plSM = 1st person plural subject marker; 1sgSM = 1st person singular subject marker; 2sgOM = 2nd person singular object marker; 2sgSM = 2nd person singular subject marker; APPL = applicative; C*n* = noun class marker; CAUS = causative; COMP = complementizer; COND = conditional; COP = copula; DEM = demonstrative; FV = final vowel; INF = infinitive; NEG = negative; OM1 = 3rd person singular object marker; PPF = pre-prefix; PFV = perfective; PROG = progressive; PST = past; REC = reciprocal; SM1 = 3rd person singular subject marker; SUBJ = subjunctive.

In UBC Qualifying Papers 3 (2013–2015),
University of British Columbia Working Papers in Linguistics 44,
Andrei Anghelescu, Joel Dunham, and Natalie Weber (eds.), 2016.

(2) Masáto a-ka-ɣór-er-a e-ɣí-taβo u-mu-aarimú
Masato SM1-PST-buy-APPL-FV PPF-C7-book PPF-C1-teacher
THEME BENEFACTIVE
'Masato bought the teacher the book.'

Because of this apparent optionality, the linear position of an object in a Nata DOC does not determine its semantic role. We might therefore ask on what basis an addressee interprets the semantic roles of the two objects, and equally, on what basis a speaker decides to use a particular word order. This paper focuses on the latter question. Nata is not unique among the Bantu languages in permitting an alternation in the order of semantic roles in its DOC. A large literature on Bantu DOCs refers to this possibility in some of the "symmetrical" languages. These are languages that allow both objects in a DOC to display "primary object" characteristics such as passivization and object marking (e.g. Bresnan and Moshi 1990, Baker, Safir and Sikuku 2012). Although the same question of the relationship between linear order and semantic role is raised by the symmetrical languages, it appears that Nata is not among this group. Nata does not allow passivization of the theme nor object-marking of the theme alone (Sadlier-Brown 2013).[3] Furthermore, in Nata, theme – benefactive order is highly contextually restricted. As will be shown in Section 3, this order is infelicitous in most contexts. The first goal of this paper is to discover in what context(s) theme – benefactive word order is used. This will help explain why a Nata speaker would choose to use theme – benefactive order when benefactive – theme order, at least at first glance, appears to convey the same meaning. A second goal of this paper is to discover whether theme – benefactive order *must* be used when the context licenses it, or whether optionality remains even when the correct contextual conditions are met (Section 4). That is, if context is relevant to the choice of word order, can it predict a speaker's choice of word order, or is it merely a conditioning factor?[4] The answer to this question will have ramifications for an eventual theory of the pragmatics-syntax interface; in particular, it informs the issue of whether or not contexts and syntax can be mapped 1:1.

Word order alternations such as this one are usually explained in terms of information structure (IS), and this is the approach taken in this work. The three "basic notions" of IS are *focus*, *givenness* and *topic* (Krifka 2008), to which I add a fourth notion, *definiteness* (Chafe 1976, see also Gundel, Hedberg and Zacharski 1993). These are defined in (3)–(6). In this paper, I investigate the influence of these four factors on the DOC word order alternation and discuss what the evidence suggests about the cause of the alternation in Nata. In particular, I use the results of the investigation to develop an analysis that draws both on IS and a more detailed model of addressee knowledge state.

(3) *Focus*: an assertion of the correct alternative among a (limited) set of possible candidates (Chafe 1976)

(4) *Givenness* (= *Givenness*$_S$[5]): that knowledge which the speaker assumes to be in the consciousness of the addressee at the time of utterance (established on the basis of either "extralinguistic or linguistic context") (Chafe 1976:30–31)

[3] However, it behaves "symmetrically" on some of the other tests; for example, theme deletion is permitted (Sadlier-Brown 2013).

[4] The issue is, of course, more complicated. There could be several conditioning factors, e.g. particular animacy combinations, phonological weight, lexical restrictions, and aspects of "context" more fine-grained than those explored here. I have attempted to control for as many of these other factors as possible, but in principle, the question remains the same: if it were possible to account for *every* possible motivator of word order choice, would there still be some "true" variability left over?

[5] The term *givenness*$_S$, comes from Prince (1981). The "S" stands for "salient".

(5) *Topic*: the expression whose referent the sentence is about (Reinhart 1981:57)

(6) *Definiteness* (= *Givenness*$_K$[6]): a referent the speaker has in mind, that s/he "assumes the hearer can pick out, from all the referents that might be categorized in this way" (Chafe 1976:39)

2 Methodology

Nata DOCs were elicited in contexts constructed so as to render one or both objects focused, given$_S$, given$_K$ or topicalized. The approach adopted here was to avoid confounds by investigating each notion independently, as much as possible. This decision means that focus and givenness$_S$ were investigated entirely with non-given$_K$ (indefinite) entities. (The investigation of topichood conflates givenness$_K$, for reasons given in Section 3.5.1) Using this method, it can be determined whether any of the conditions in isolation are sufficient – or not – to prompt theme – benefactive word order. I control for verb type (only applied verbs are used), sentence length, and semantic role and animacy, eliciting only the combination benefactive=human, theme=animal/inanimate. The results here are not necessarily expected to extend to other semantic roles, animacy pairs or combinations thereof.

Elicitations were conducted with one consultant, a native speaker of Nata. Elicitations proceeded as follows. I described a detailed context in English, then provided a sentence in English for translation into Nata. The consultant then provided the translation. However, in the (frequent) instances in which the Nata translation was expected to be a DOC, a slightly different method was used. Instead of providing one English sentence, I provided the two word order variants in succession (using English words). This was done to prevent the unintentional elicitation (or priming) of English word order. For example, to elicit a Nata DOC in a context containing the benefactive *omutémi* 'chief' and the theme *aŋɔ́ɔ́mbɛ* 'cow', I would offer a sentence such as 'I slaughtered 'chief-cow' or 'cow-chief''. The consultant was briefed so as to understand he should choose the Nata word order that was appropriate in the context (or a different construction if neither word order was appropriate). After I recorded the form deemed most felicitous, I asked the consultant whether the other word order variant(s) would be felicitous in the same context.

3 Felicity conditions

3.1 Newness

In keeping with the methodological decision to investigate the IS notions separately, I use *new* to refer to completely novel referents (i.e. "discourse-new, hearer-new" [Prince 1992]). However, in the literature, "new" sometimes refers to "discourse-new, hearer-old" (Prince 1992). This describes definiteness/givenness$_K$ as defined in (6). Therefore, in this paper, the latter scenario is not seen as involving a "new" referent and is instead investigated as givenness$_K$ (Section 3.4).

In the Nata DOC, "all new" word order is benefactive – theme (7bi). I will call this word order "canonical" as it is the most common word order and is the least contextually-restricted, as the examples in the following sub-sections will show. Theme – benefactive order (or non-canonical word order [NCWO]) is infelicitous in the all new context (7bii).[7]

[6] The term *givenness*$_K$ comes from Prince (1981). The "K" stands for "shared knowledge".

[7] In the examples in this paper, each turn has an alphanumeric label (a., b., c.,…). Alternatives for each turn are given roman numerals (i., ii., iii.,…). Contexts are provided in curly brackets.

(7) {You have taken a homeless child into your home and have cooked him a plantain. You run into an overly friendly stranger in the street who does not know about these events.}

Stranger:

a. ne-ke u-a-kɔr-iré rɛɛró?
COP-WH 2sgSM-PST-make-PFV today
'What did you do today?'

You:

b. i. ni-aa-tɛrɛ́k-ir-ire u-mu-aaná e-ɣi-tɔɔké
1sgSM-PST-cook-APPL-PFV PPF-C1-child PPF-C7-plantain
'I cooked a child a plantain.'

ii. #ni-aa-tɛrɛ́k-ir-ire e-ɣi-tɔɔké u-mu-aaná
1sgSM-PST-cook-APPL-PFV PPF-C7-plantain PPF-C1-child
(intended: 'I cooked a child a plantain.')

3.2 Focus

Chafe's (1976) definition of focus (3) specifically refers to a selection among a limited set of candidates. This is often called *contrastive focus* and my use of the term "focus" refers to this notion. Contrastive focus can be distinguished from another concept of focus, *presentational* or *information* focus, which in the literature refers to new information (e.g. Rochemont 1986, 2013, Kiss 1998). Here, I follow Rochemont (2013) in re-classifying constituents of this type as merely new (and not a type of focus). Therefore, constituents that others might analyze as instances of presentational/information focus are covered in other sections: Section 3.1 (Newness) or Section 3.4 (Givenness$_K$). The two sentential positions that concern us here – immediately post-verbal and clause-final – have both been described as focus positions in existing Bantu research. Aghem is said to have an immediately post-verbal focus position (Watters 1979).[8] If Nata is like this language, then theme – benefactive word order will be the result of theme focus (if, of course, focus is relevant to Nata NCWO). Kirundi has a clause-final focus position (Ndayiragije 1999). If Nata is like this language, then NCWO will be the result of benefactive focus. Therefore, both the benefactive and the theme will be investigated here.

The most common test for focus is to elicit an answer to a *wh*-question. The questioned information is said to be focused in the answer. This test for focus, however, is problematic because it usually contains a confound: when a *wh*-question is asked, any non-questioned material is rendered given$_S$. (For example, in the pair 'Who kicked the ball?' –'John kicked the ball.', *John* is focused in the answer and *kicked the ball* is given$_S$, by the definition in (4).) Because of this, an answer to a *wh*-question simultaneously elicits focus of the questioned material and givenness$_S$ of the non-questioned material; it is not possible to determine in retrospect which condition was the cause of any observed effects (or indeed whether both conditions were required). Nata illustrates this confound particularly clearly. In an answer to a DOC *wh*-question, a given$_S$, non-given$_K$, object (that is, the non-questioned object) appears at the left edge, while only the questioned object appears post-verbally. This occurs whether the questioned object is the theme or the benefactive. The pattern leaves us wondering whether NCWO would have been felicitous if givenness$_S$ had not been

[8] It is possible that Watters' use of the term "focus" includes presentational/information focus.

present to, perhaps, encourage object fronting.[9] To remedy this problem, (8) and (9) remove givenness_S from the target sentence. This is achieved by asking a transitive *wh*-question and answering it with a ditransitive such that the two objects in the answer are *focused* and *new*, respectively. (8bii) and (9bii) demonstrate that, even with givenness_S removed, NCWO is not a possible answer. Canonical word order arises.

(8) {You and your friend both love reading but you have no time to read because you are so busy with work. You are hanging out in the village during a short break.}

Your friend:

a. u-a-ŋga-βeer-é nú-u-mu-ɛja, ne-we
2sgSM-PST-COND-have-FV with-PPF-C3-time, COP-WH
u-a-ŋga-sɔm-i-iré?
2sgSM-PST-COND-read-APPL-PFV
'If you had time, who would you read for?'

You:

b. i. ɲ-a-ŋga-sɔm-i-iré u-mu-aná e-ɣí-taβo
1sgSM-PST-COND-read-APPL-PFV PPF-C1-child PPF-C7-book
'I would read a child a book'

ii. #ɲ-a-ŋga-sɔm-i-iré e-ɣí-taβo u-mu-aná
1sgSM-PST-COND-read-APPL-PFV PPF-C7-book PPF-C1-child
'I would read a child a book.'

(9) {You and your friend have no money. You are hanging out in the village.}

Your friend:

a. u-a-ŋga-βeer-é na-ʧa-Ø-heerá, ne-ke
2sgSM-PST-COND-have-FV with-PPF-C10-money, COP-WH
u-a-ŋga-ɣor-ire?
2sgSM-PST-COND-buy-PFV
'If you had money, what would you buy?'

[9] It should be noted at this point that simultaneously eliciting the *three* notions focus, givenness_S and givenness_K can sometimes produce NCWO. Specifically, NCWO is used in answers to *wh*-benefactive questions in which the theme is both given_S and given_K (answers to *wh*-theme questions in which the benefactive is given_K and given_S instead demonstrate canonical order.) Because of the three-way confound involved in these situations, we cannot tell which of the conditions (or combination thereof) was responsible for NCWO. Therefore, I continue to investigate all factors independently. Furthermore, these results cannot be teased apart from the possibility of an independent structural parallelism requirement dictating, in this case, that the constituent order of the question (e.g. [[V cook] [DO chicken]]) be replicated in the answer ([[[V cook] [[DO chicken] [IO man]]]).

You:

b. i. ɲ-a-ŋga-ɣor-i-ire u-mu-aná-umwe
1sgSM-PST-COND-buy-APPL-PFV PPF-C1-child-one
e-ke-ɣuriɣurí
PPF-C7-traditional.carriage
'I would buy some child a traditional carriage.'

ii. #ɲ-a-ŋga-ɣor-i-ire e-ke-ɣuriɣurí
1sgSM-PST-COND-buy-APPL-PFV PPF-C7-traditional.carriage
u-mu-aná-umwe
PPF-C1-child-one
'I would buy some child a traditional carriage.'

Another potential issue with eliciting focus via answers to *wh*-questions is that, according to some authors, such answers represent presentational/information focus – not contrastive focus as I have been assuming (e.g. Zubizaretta 1998). This potential objection can be countered by eliciting focus using contrastive contexts. In the contexts I used, a speaker made a statement consisting of a DOC and the addressee responded by disagreeing with the part of the statement concerning the target object(s). Like the *wh*-test, the contrastive context test renders non-contrasted material given$_S$. Therefore, I also included contrastive contexts in which the target DOC contained a focused and a new object using the same strategy as (8) and (9) above. In conditions containing the givenness$_S$ confound, when the benefactive was focused and the theme was given$_S$, the theme surfaced at the left edge, mirroring the results of the analogous *wh*-test. When the theme was focused and the benefactive was given$_S$, the left-edge effect did not occur; instead, the consultant preferred to add a clause to host the non-focused material. (The theme occurred after the first verb and the benefactive after the second.) It is not known why this form differs from the results of the analogous *wh*-test. In conditions where, like (8) and (9), givenness$_S$ was *not* present, canonical word order arose, again mirroring the results obtained in the analogous *wh*-tests. In none of these contrastive contexts was NCWO felicitous. The combined results of the *wh*-tests and contrastive context tests allow us to conclude that NCWO does not encode focus of one of the objects.

3.3 Givenness$_S$

Many authors have found that given constituents tend to appear earlier in a sentence, often resulting in a non-canonical word order (Louwrens 1979, Bock and Irwin 1980, Birner and Ward 1998, Ferreira and Yoshita 2003, Clifton and Frazier 2004, Bresnan, Cueni, Nikitina and Baayen 2007, Skopeteas and Fanselow 2009). In fact, this pattern is attested in the DOC of Shona, making givenness a likely candidate for the motivation of NCWO in Nata (Hyman and Duranti 1982, citing Hawkinson and Hyman 1974). Thus, NCWO might be encoding the givenness of its leftmost object, the theme. The given-before-new principle has been remarked of many notions of givenness, including (but not limited to) the two that will be discussed here. However, authors often do not specify what is meant by "given" when they invoke "given-before-new", making it difficult to assess the principle's actual coverage. Furthermore, since givenness does not necessarily guarantee fronting, the "principle" is best viewed as a tendency, not an absolute requirement. Keeping these

caveats in mind, I examine the predictions of given-before-new as applied to givenness_S (below) and givenness_K (Section 3.4).[10]

As stated in the introduction, I make a clear distinction between two types of givenness: givenness_S and givenness_K. (I use the subscript terminology of Prince [1981].) Givenness_S (4) refers to knowledge that is in speakers' "consciousness"; in most cases, a given_S entity has been mentioned earlier and as a result has become salient in some sense. This notion is often simply called "givenness" and its usual reflex in English is de-accenting (Chafe 1976, Rochemont 2014).

To render the theme given_S, I created contexts in which the relevant entity is mentioned earlier in the discourse, thereby providing an antecedent for the target constituent in the DOC. As shown in (10ci), this situation produces canonical word order in Nata. NCWO is infelicitous (10cii). For completeness, I also elicited a context in which the benefactive was given_S and a context in which both objects were given_S. As expected, these produced the same, canonical, results. Note that there is no benefactive in the sentential antecedent to (10ci). This ensures that the benefactive is new (not given_S and not focused) in the DOC, avoiding a confound like the one described in Section 3.2. However, a transitive sentence is arguably an unnatural antecedent for a ditransitive consequent sentence. If this is a concern, recall that in Section 3.2, I discussed the results of focus tests that used sentence pairs in which both the antecedent and consequent sentences were ditransitive. In that section, the target sentences contained one object that was focused and the other that was given_S, so givenness_S has already been elicited with ditransitive antecedent sentences. In none of these cases did NCWO arise. Therefore, givenness_S is not sufficient for the appearance of NCWO.

(10) {You and your friend are members of an organization whose aim is to deliver meals to elders in the community. You yourself have taken in a homeless child and have cooked him a hearty meal. One evening, you run into your friend in the street. He does not know about the child. You exchange greetings.}

Your friend:

a. n-ni-hɔ-ire βwahέɛnɛ. ɲ-ɲ-a-hir-i-íre
COMP-1sgSM-spend.a.day-PFV good. COMP-1sgSM-PST-deliver-APPL-PFV
i-βj-aakurí
PPF-C8-meal
'I've had a good day. I delivered a meal.'

b. ne-ke u-a-kɔr-ire rɛɛrɔ́
COP-WH 2sgSM-PST-do-PFV today
'What did you do today?'

You:

c. i. ɲ-a-kɔr-í-ire u-mu-aaná i-βi-aakurí
1sgSM-PST-make-APPL-PFV PPF-C1-child PPF-C8-meal
'I made a child a meal.'

[10] There are further possibilities which recruit the givenness_S notion. For example, it is possible that both objects must be given_S but not equally so, i.e. one must be more recently mentioned than the other, or one must be more "salient". I leave these additional possibilities to future work.

ii. #ɲ-a-kɔr-í-ire i-βi-aakurí u-mu-aaná
1sgSM-PST-make-APPL-PFV PPF-C8-meal PPF-C1-child
(Intended: 'I made a child a meal.')

3.4 Givenness$_K$

Givenness$_K$ (6) is a concept referring to mutual knowledge: "the speaker assumes that the hearer "knows," assumes, or can infer a particular thing" (Prince 1981:230). This is equivalent to Chafe's definition of "definiteness" and I will consider the two terms equivalent. Unlike a given$_S$ object, a given$_K$ object can be uttered without an immediate antecedent, as will be shown below. Thus, givenness$_K$ and givenness$_S$ are non-overlapping concepts (for further discussion of the difference, see Chafe 1976, Prince 1981 or Rochemont 2014). One of the possible reflexes of givenness$_K$ in English is the use of the definite determiner *the*.

In this section, I provide a complete givenness$_K$ paradigm: in (11) the benefactive is given$_K$, in (12) the theme is given$_K$, and in (13) both objects are given$_K$. (Recall that (7bi) already showed that two non-given$_K$ objects results in canonical word order.) In these contexts, entities are rendered given$_K$ by virtue of a shared prior experience involving the referent. This experience is explained in the context paragraph. (11), (12) and (13) show that all combinations of givenness$_K$ result in canonical word order (11bi, 12bi, 13bi). NCWO is not felicitous in any of the givenness$_K$ contexts (11bii, 12bii, 13bii). This leads to the conclusion that givenness$_K$, like givenness$_S$, is not sufficient to produce NCWO. Because all combinations of givenness$_K$ led to canonical word order, we can also conclude that canonical word order is ambiguous with regard to the givenness$_K$ status of the objects. Nata has no determiners equivalent to English *a* and *the*. However, given the detailed contexts, the consultant was able to derive the appropriate interpretation. In the constructions studied here, then, a givenness$_K$ interpretation is normally facilitated by context.[11]

(11) {You have taken a homeless child into your home and cooked him a plantain. Your friend knows that you took in the child but he doesn't know that you have cooked him something. You run into your friend in the street and exchange greetings.}

Friend:

a. ne-ke u-a-kɔr-iré rɛɛró?
COP-WH 2sgSM-PST-do-PFV today?
'What did you do today?'

You:

b. i. ni-aa-tɛrɛ́k-ir-ire u-mu-aaná e-ɣi-tɔɔké
1sgSM-PST-cook-APPL-PFV PPF-C1-child PPF-C7-plantain
'I cooked the child a plantain.'

[11] This is not an uncommon observation across the Bantu family (e.g. see Zerbian [2006] for Northern Sotho). However, it should be noted that Nata, like many Bantu languages, has an object-marking (OM) system in which an OM located within the verb complex can co-refer with an overt object NP. In these cases of "object doubling", the overt NP seems to receive a given$_K$ interpretation. The forms in this paper do not contain OMs, and it is not known what situations would require this givenness$_K$-marking strategy in place of facilitation by context. For now, I leave this important question to future research.

ii. #ni-aa-tɛrɛ́k-ir-ire e-ɣi-tɔɔké u-mu-aaná
1sgSM-PST-cook-APPL-PFV PPF-C7-plantain PPF-C1-child
(Intended: I cooked the child a plantain)

(12) {You are at your friend's house, and there are some plantains on the table. He hands you one to take home. You have taken in a homeless child, and when you arrive home you cook the plantain for this child. Your friend does not know about the child. Later, you run into your friend in the street and exchange greetings.}

Friend:

a. ne-ke u-a-kɔr-iré rɛɛró?
COP-WH 2sgSM-PST-do-PFV today?
'What did you do today?'

You:

b. i. ni-aa-tɛrɛ́k-ir-ire u-mu-aaná e-ɣi-tɔɔké
1sgSM-PST-cook-APPL-PFV PPF-C1-child PPF-C7-plantain
'I cooked a child the plantain.'

ii. #ni-aa-tɛrɛ́k-ir-ire e-ɣi-tɔɔké u-mu-aaná
1sgSM-PST-cook-APPL-PFV PPF-C7-plantain PPF-C1-child
(Intended: 'I cooked a child the plantain.')

(13) {You have taken in a homeless child. Your friend knows this. Your friend comes over to visit and brings you a plantain as a gift, then he goes home. Later, you run into your friend in the street. You exchange greetings.}

Friend:

a. ne-ke u-a-kɔr-iré rɛɛró?
COP-WH 2sgSM-PST-do-PFV today?
'What did you do today?'

You:

b. i. ni-aa-tɛrɛ́k-ir-ire u-mu-aaná e-ɣi-tɔɔké
1sgSM-PST-cook-APPL-PFV PPF-C1-child PPF-C7-plantain
'I cooked the child the plantain.'

ii. #ni-aa-tɛrɛ́k-ir-ire e-ɣi-tɔɔké u-mu-aaná
1sgSM-PST-cook-APPL-PFV PPF-C7-plantain PPF-C1-child
(Intended: 'I cooked the child the plantain.')

3.5 "Aboutness" topichood[12]

3.5.1 Given$_S$ topichood

Of all the basic IS notions, the definition of topichood is perhaps the most elusive. There are many reasons for this elusiveness, among them the fact that the term is used differently by different traditions and even by different authors within the same tradition. Most recent authors in the generative tradition agree that a topic constituent marks what a sentence is "about" (Kuno 1972, Reinhart 1981, Bresnan and Mchombo 1987, Gundel 1988, Frey 2005, Roberts 2011, cf. Chafe 1976, Jacobs 2001, Krifka 2008) but as Roberts (2011:1928) admits, the definition of "aboutness" is left vague in the literature. Outside of this central point of agreement, however, there is little concensus on what conditions must hold in order for a constituent to qualify as a topic in the information-structural sense.[13] Some authors insist that a topic must be given$_K$/definite (Kuno 1972, Gundel 1988, Zerbian 2006) or refer to "old information", while others argue that specific indefinites are allowed (Reinhart 1981, Frey 2005). In general, however, some degree of familiarity seems to be required in order for a referent to be considered a topic (Reinhart 1981, Gundel 1988, Zerbian 2006, Roberts 2011). For present purposes, I take the givenness$_K$ requirement as the most conservative point of departure. All of the topics in this paper are given$_K$. Although IS topics frequently find their syntactic reflex in left-dislocated or "topicalized" constituents, this positioning is neither necessary nor sufficient for topichood (e.g. Gundel 1988 ["not necessary"]; Frey 2005 ["not sufficient"]). In this section, I investigate whether NCWO marks a clause-internal topic.

The disagreement over the felicity conditions for topichood make it difficult to construct contexts that are guaranteed to produce a reflex that can be unequivocally labeled a topic. However, here, I will put faith in a common method for eliciting topichood, the "aboutness" test (e.g. Gundel 1974, Reinhart 1981, Frey 2005). In the aboutness test, the target (topic-containing) sentence is prefaced by an antecedent sentence or phrase referring to what the target sentence will be "about". For example, the antecedent sentence "I need to talk to you about (x)" prefaces a second sentence that, in most natural continuations, will be "about (x)". In the continuation sentence, (x) should demonstrate the language's reflex for topic. Such tests, however, are problematic for two reasons. First, most tests of this type render the topic constituent given$_S$ by mentioning it in the antecedent sentence. Therefore, results obtained from such tests cannot be reliably attributed to the notion of topic. This confound will be addressed in the next section. Second, an antecedent sentence itself cannot be spoken out of the blue. The "about (x)" test seems to prefer a context in which (x) is given$_K$ but other topic tests impose their own unique felicity conditions, as Roberts (2011) shows using the "as for (x)" test and "speaking of (x)" test. Because of this, an antecedent sentence itself seems to be more of a *reflection* of a specific context than a "test" for a uniform notion. (In fact, I suggest that the job of the antecedent sentence is merely to help a consultant accommodate the fact that a topic context appropriate to that particular topic test is in place.) Therefore, if we are to rely on the aboutness test to produce the reflex of topic in Nata, it is not the words "about (x)" that should concern us, but rather the construction of a context in which these words can be spoken felicitously.

[12] This section will not discuss contrastive topic, which arguably involves focus (e.g. Chafe 1976, Buring 2003).

[13] A notable point of disagreement among authors is the extent to which a topic should be seen as a cross-linguistically uniform notion with an (ultimately) definable conditioning context, or simply an umbrella term describing a family of cross-linguistically variable phenomena. See Chafe (1976), Jacobs (2001), Krifka (2008) or Roberts (2011) for overviews of this and other issues surrounding the notion of topic.

The contexts in (14) and (15) describe a situation in which two interlocutors are concerned about a particular, mutually familiar, entity. In (14), this entity is the benefactive and in (15) this entity is the theme. Uttering the test sentence "I need to talk to you about (x)" appears to be felicitous in these contexts as it readily elicits the continuation sentence without any hesitation or qualifying comments from the consultant. As (14) and (15) show, this classic test for topichood permits neither canonical word order nor NCWO. Instead, the test causes the given$_S$ topic to appear at the left-edge ("topicalization") (14ci, 15ci). A resumptive pronoun appears in the verb complex when the theme is in this position but not when the benefactive is, though here I will refer to both constructions as topicalization. (It is not known why the demonstrative *jiirí* is preferred in (14c). This may have to do with ambiguity in the context.)

(14) {You and your brother own one chicken. There is a medicine man who lives in your village. He is poor and hungry, and every day you and your brother talk about how to help him out. One day, while your brother is out, you make the decision to kill the chicken so that the medicine man can have a few meals. Upon receiving the gift, the medicine man is very happy and says he will find some way to thank you. This makes you very happy. Later that day, you run into your brother. He is in a rush to catch a bus.}

You:

a. ni-kwɛnd-á tu-ɣaamβ-án-ɛ a-ma-ŋána
1sgSM-want-FV 1sgSM-talk-REC-SUBJ PPF-C6-matters
ɣó-o-mu-ɣaβo
of-PPF-C1-medicine.man
'I have to talk to you about the medicine man.'

Brother:

b. ʔaʔa, n-tee-ɣo-tór-a ni-ku-aŋɣóh-i
No, 1sgSM-NEG-INF-can-FV 1sgSM-PROG-hurry-CAUS
'No, I can't; I'm in a hurry.'

You:

c. i. o-mú-ɣaβo, n-né-mo-sintʃ-ir-ire
PPF-C1-medicine.man, COMP-1sgSM-OM1-slaughter-APPL-PFV
a-ŋ-gɔkɔ́ jiirí
PPF-C9-chicken DEM
'The medicine man, I slaughtered that chicken for him.'
Consultant comment: ambiguous without *jiirí*

ii. #n-né-sintʃa-ir-ire o-mú-ɣaβo
COMP-1sgSM-slaughter-APPL-PFV PPF-C1-medicine.man
a-ŋ-gɔkɔ́ jiirí
PPF-C9-chicken DEM
(Intended: 'I slaughtered that chicken for the medicine man.')

iii.#n-né-sintʃa-ir-ire a-ŋ-gɔkɔ́ o-mú-ɣaβo[14]
COMP-1sgSM-slaughter-APPL-PFV PPF-C9-chicken PPF-C1-medicine.man
(Intended: 'I slaughtered the chicken for the medicine man.')

(15) {All of you and your brother's chickens have recently died except for one. You and your brother have been trying everything to save the last surviving chicken, but its health is declining. One day, while your brother is out, you make the decision to kill the chicken before it dies because it will make a nice gift to the village's medicine man. Later that day, you run into your brother. He is in a rush to catch a bus.}

You:

a. ni-kwɛnd-á ni-kɔ-βɔ́ɔ́rɛɛr-ɛ a-ma-ŋána ɣa-a-ŋ-gɔkɔ́
1sgSM-want-FV 1sgSM-2sgOM-tell-FV PPF-C6-matters of-PPF-C9-chicken
'I have to tell you about the chicken.'

Brother:

b. ʔaʔa, ni-tee-ɣo-tór-a ni-ku-aŋɣóh-i
No, 1sgSM-NEG-INF-can-FV 1sgSM-PROG-hurry-CAUS
'I can't; I'm in a hurry.'

You:

c. i. a-ŋ-gɔkɔ́ n-ni-ka-jé-sintʃe-er-a
PPF-C9-chicken COMP-1sgSM-PST-OM9-slaughter-APPL-FV
o-mú-ɣaβo
PPF-C1-medicine.man
'The chicken, I slaughtered it for the medicine man.'

ii. #n-ni-ka-síntʃe-er-a o-mú-ɣaβo a-ŋ-gɔkɔ́[15]
COMP-1sgSM-PST-slaughter-APPL-FV PPF-C1-medicine.man PPF-C9-chicken
(Intended: 'I slaughtered the chicken for the medicine man.')

iii.#n-ni-ka-síntʃe-er-a a-ŋ-gɔkɔ́ o-mú-ɣaβo
COMP-1sgSM-PST-slaughter-APPL-FV PPF-C9-chicken PPF-C1-medicine.man
(Intended: 'I slaughtered the chicken for the medicine man.')

3.5.2 Non-given$_S$ topichood

In this section, I investigate the effects of removing the confound introduced by the "about (x)" test itself. Mentioning (x) in the antecedent sentence rendered the target constituent given$_S$. It is possible

[14] It is not clear whether this form would improve with the addition of the demonstrative *jiirí* to *angokó* 'chicken'. However, consultant comments suggest that the lack of topicalization is an independent source of infelicity. The infelicity of (14cii) despite the disambiguation of *angokó* provides further evidence for this (also see (15cii) and (15ciii) which are infelicitous despite the unambiguity of *angokó* in this context).

[15] Unlike in context (14), *aŋgokó* 'the chicken' is not ambiguous in context (15), probably because *aŋgokó* is mentioned in the context sentence (15a). Instead, the source of infelicity in (15cii) and (15ciii) is probably the lack of topicalization of a referent that has been set up as a topic.

that givenness$_S$ caused mandatory topicalization (as it did in the *wh*-answers discussed in Section 3.2). However, NCWO might arise in topic contexts in the absence of givenness$_S$. To investigate this possibility, the contexts in (16) and (17) make one simple change: instead of prefacing the target sentence with "about (x)", I use an antecedent that can be translated as 'about something'. Otherwise the contexts in (16) and (17) are identical to (14) and (15), respectively.

(16) {You and your brother own one chicken. There is a medicine man who lives in your village. He is poor and hungry, and every day you and your brother talk about how to help him out. One day, while your brother is out, you make the decision to kill the chicken so that the medicine man can have a few meals. Upon receiving the gift, the medicine man is very happy and says he will find some way to thank you. This makes you very happy. Later that day, you run into your brother. He is in a rush to catch a bus.}

You:

a. ni-kwɛnd-á ni-kɔ-βɔ́ɔ́rɛɛr-ɛ e-ki-ɣérɔ
1sgSM-want-FV 1sgSM-2sgOM-tell-FV PPF-C7-thing
'I have to tell you something.'

Brother:

b. ʔaʔa, ni-tee-ɣo-tór-a ni-ku-aŋɣóh-i
No, 1sgSM-NEG-INF-can-FV 1sgSM-PROG-hurry-CAUS
'I can't; I'm in a hurry.'

You:

c. i. o-mú-ɣaβo, m-m-mó-sintʃa-ir-ire
PPF-C1-medicine.man, COMP-1sgSM-OM1-slaughter-APPL-PFV
a-ŋ-gɔkɔ́ jiirí
PPF-C9-chicken DEM
'The medicine man, I slaughtered that chicken for him.'
Consultant comment: if you want to talk about the same chicken, then need *jiirí*

ii. n-né-sintʃa-ir-ire o-mú-ɣaβo
COMP-1sgSM-slaughter-APPL-PFV PPF-C1-medicine.man
a-ŋ-gɔkɔ́ jiirí
PPF-C9-chicken DEM
'I killed that chicken for the medicine man.'
Consultant comment: sentence is ambiguous without *jiirí*

iii. n-né-sintʃa-ir-ire a-ŋ-gɔkɔ́ o-mú-ɣaβo
COMP-1sgSM-slaughter-APPL-PFV PPF-C9-chicken PPF-C1-medicine.man
(Intended: 'I killed the chicken for the medicine man.')
Consultant comment: could say this if there was only one chicken in the village

iv.? n-né-sintʃa-ir-ire a-ŋ-gɔkɔ́ jiirí
COMP-1sgSM-slaughter-APPL-PFV PPF-C9-chicken DEM
o-mú-ɣaβo
PPF-C1-medicine.man
(Intended: 'I killed the chicken for the medicine man.')

(17) {All of you and your brother's chickens have recently died except for one. You and your brother have been trying everything to save the last surviving chicken, but its health is declining. One day, while your brother is out, you make the decision to kill the chicken before it dies because it will make a nice gift to the village's medicine man. Later that day, you run into your brother. He is in a rush to catch a bus.}

You:

a. ni-kwɛnd-á ni-kɔ-βɔ́ɔrɛɛr-ɛ e-ki-ɣérɔ
1sgSM-want-FV 1sgSM-2sgOM-tell-FV PPF-C7-thing
'I have to tell you something.'

Brother:

b. ʔaʔa, ni-tee-ɣo-tór-a ni-ku-aŋɣóh-i
No, 1sgSM-NEG-INF-can-FV 1sgSM-PROG-hurry-CAUS
'I can't; I'm in a hurry.'

You:

c. i. a-ŋ-gɔkɔ́ (jiirí), n-né-je-sintʃa-ir-ire
PPF-C9-chicken (DEM), COMP-1sgSM-OM9-slaughter-PFV
o-mú-ɣaβo
PPF-C1-medicine.man
'The chicken, I killed it for the medicine man.'

ii. n-ni-ga-sintʃ-er-a o-mú-ɣaβo
COMP-1sgSM-PST-slaughter-APPL-FV PPF-C1-medicine.man
a-ŋ-gɔkɔ́ jiirí
PPF-C9-chicken DEM
'I killed that chicken for the medicine man.'
Consultant comment: jiirí is required

iii.? n-ni-ka-sintʃ-er-a a-ŋ-gɔkɔ́
COMP-1sgSM-PST-slaughter-APPL-FV PPF-C9-chicken
o-mú-ɣaβo
PPF-C1-medicine.man
(Intended: 'I killed that chicken for the medicine man.')
Consultant comment: could say this if it's not possible to get another chicken.

iv.ʔ n-ni-ka-sintʃa-er-a a-ŋ-gɔkɔ́ jiirí
COMP-1sgSM-PST-slaughter-APPL-FV PPF-C9-chicken DEM
o-mú-ɣaβo
PPF-C1-medicine.man
(Intended: 'I killed that chicken for the medicine man.')

For current purposes, the most important consequence of removing givenness_S from the context is that NCWO finally arises as a possible form (provided the theme is "totally unambiguous" as per consultant comments). This is seen in (16ciii), (16civ), (17ciii) and (17civ), though only (16ciii) was judged felicitous on all occasions (and note that (16civ) and (17civ) contain the demonstrative *jiirí*). Another consequence of removing givenness_S is that more forms become possible. Topicalization, canonical word order and NCWO are all possible continuations. The appearance of multiple possibilities suggests that removing the mention of an entity (x) from the aboutness test not only removes givenness_S, but also uncouples "aboutness" from entity (x), nullifying any requirement for the speaker to talk "about" entity (x) should s/he choose to mention it. If a speaker states s/he needs only to talk "about something", almost any declarative sentence can form a felicitous continuation. Thus, removing givenness_S from the aboutness test seems to have voided it as a designated test for topic. The contribution of the notion of topic to NCWO remains in question, and I will not pursue it further here. (Note, however, that topicalization remains possible (16ci, 17ci).)

In the remainder of this paper I focus on a third consequence of changing "about entity (x)" to "about something": the potential heightening of theme ambiguity. In Nata, removing the "about entity (x)" preface seems to introduce the potential that the intended referent will not be clear to the addressee when spoken in the continuation sentence. When the theme referent is mentioned in the continuation sentence, the addressee is hearing it for the first time and lacks the expectation that *it* is what the sentence is "about". In some situations, this might leave the addressee with too little contextual information to correctly interpret whether the theme is intended to be given_K or not. (Recall from Section 3.4 that givenness_K status is normally resolved with the aid of context.) Contexts (16) and (17) indeed seem to demonstrate a situation in which the theme referent has become potentially ambiguous. This is reflected by the required addition of the demonstrative *jiirí* to the canonical forms (16cii) and (17cii). Consultant comments indicate that this modification guarantees a given_K interpretation for *aŋgokó* 'chicken', repairing the ambiguity. In fact, it is this heightening of ambiguity that might play a role in the (albeit variable) appearance of NCWO in (16ciii, iv) and (17ciii, iv). Here, consultant comments are especially enlightening: in (16ciii), for example, the comment is that NCWO is felicitous "if there is only one chicken in the village". In this regard, the canonical forms modified with *jiirí* and the NCWO forms qualified by consultant comments might be providing similar contributions to interpretation. By adding the lexical item *jiirí* 'that' to *angokó* 'chicken', the speaker picks out a particular chicken from the set of all chickens. In the NCWO cases, the consultant comments suggest that the relevant context contains only a single chicken, which equally guarantees the intended chicken is picked out. Therefore, we might hypothesize that NCWO is used to pick out a given_K theme referent when there is only one such referent in the relevant context (i.e. the village). This hypothesis (NCWO Hypothesis 1) will be examined in the next section. A final note is that the consultant comments and additions of *jiirí*

apply exclusively to the theme. I will take this as evidence that it is the theme's status that is of most importance in NCWO, and not the benefactive's.[16]

3.6 Givenness$_K$ and reference ambiguity

NCWO Hypothesis 1 predicted that NCWO is used to pick out a given$_K$ theme referent when there is only one such referent in the relevant context. In this scenario, the "only chicken in the village" is, effectively, the only chicken in the "world" (or, alternatively stated, the relevant context [the village] is the largest set under consideration). In other words, the speaker believes the addressee is able to pick out the intended referent in the relevant setting and, furthermore, the speaker believes the addressee will *not* consider other referents. Figure 1 represents this model of addressee knowledge state. If Hypothesis 1 is correct, then when a speaker believes Addressee Knowledge State 1 (AKS1) holds of an addressee, NCWO will be used. However, this hypothesis encounters an immediate problem in light of the results of Section 3.4, Givenness$_K$. In this section, I showed that canonical order suffices for all combinations of given$_K$ and non-given$_K$ objects because the relevant interpretation is "facilitated by context". If this is true, then contexts where the referent is exceptionally clear (i.e. because there is only one such referent) are exactly the contexts in which we would most expect *canonical* order. Because Hypothesis 1 makes contradictory predictions about which word order will arise, I also consider an alternative scenario containing "the only chicken in the village". In the alternative scenario, there is only one chicken in the village itself, but there are many more chickens in the world. In other words, the speaker believes the addressee is able to pick out the intended referent in the relevant setting but, furthermore, the speaker believes the addressee *might* consider other referents in the superset world. The intended referent is not clear. Figure 2 represents this alternative model of addressee knowledge state. NCWO Hypothesis 2 predicts that NCWO will be used when a speaker believes an addressee's knowledge state is characterized by AKS2. Both hypotheses predict theme givenness$_K$ is a necessary but not sufficient condition for NCWO[17]. The hypotheses differ with respect to what addressee knowledge state (AKS1 or AKS2) they predict is necessary in addition to theme givenness$_K$. Because Hypothesis 2 does not suffer from the same problem as Hypothesis 1, it is considered the more plausible.

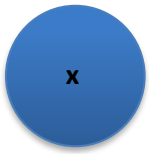

Figure 1 Addressee knowledge state 1

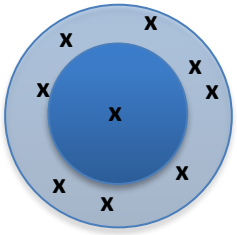

Figure 2 Addressee knowledge state 2

NCWO Hypothesis 1 can be tested by creating a context in which the intended referent is the only such referent in existence (in normal human experience). In (18), *omuɛɛrí* 'the moon' is used

[16] The noun *omúɣaβo* 'medicine man' seems to have no problem being picked out as the intended given$_K$ referent (*jiirí* is not required for this human referent). The role of real-world knowledge and the role of animacy in the interpretation of givenness$_K$ could be explored in future research.

[17] The necessity of theme givenness$_K$ is in line with the predictions of given-before-new discussed in Sections 3.3 and 3.4.

for this purpose. In this case, an addressee can be assumed to be in AKS1 because s/he will be able to pick out the intended referent, the moon, and is not expected to consider other possible moons. Examples (18ai) and (18aii) show that, as suspected, NCWO Hypothesis 1 is not borne out: when an addressee is in AKS1, canonical word order is used. Examples (19aii) and (19aiii) show the same pattern in a non-embedded context.

(18) {A beautiful nurse has moved to town. No one knows her name yet. Witare and Sabiti have been admiring her from afar, all day.}

At night, Sabiti to Witare:

a. i. ni-a-nga-tor-ire, ni-a-ŋga-ɣor-ir-ire
1sgSM-PST-COND-be.able-PFV 1sgSM-PST-COND-buy-APPL-PFV
o-mú-nesi u-mu-ɛɛrí! m-mu-tʃóm-u!
PPF-C1-nurse PPF-C1-moon COP-C1-beautiful-FV
'If I could, I would buy the nurse the moon! She is so beautiful!'

ii. #ni-a-nga-tor-ire, ni-a-ŋga-ɣor-ir-ire
1sgSM-PST-COND-be.able-PFV 1sgSM-PST-COND-buy-APPL-PFV
u-mu-ɛɛrí o-mú-nesi! m-mu-tʃóm-u!
PPF-C1-moon PPF-C1-nurse COP-SM1-beautiful-FV
(Intended: 'If I could, I would buy the nurse the moon! She is so beautiful!')

(19) {The next day, Sabiti buys some paper and makes a card in the shape of the moon and gives it to the nurse.}

Later, Sabiti to Witare (joking):

a. i. o-mu-nesi m-mo-ɣor-ir-ire u-mu-ɛɛrí
PPF-C1-nurse 1sgSM-OM1-buy-APPL-PFV PPF-C1-moon
'The nurse, I bought her the moon!'

ii. ni-ɣor-ir-ire o-mu-nesi u-mu-ɛɛrí
1sgSM-buy-APPL-PFV PPF-C1-nurse PPF-C1-moon
'I bought the nurse the moon!'

iii.#ni-ɣor-ir-ire u-mu-ɛɛrí o-mu-nesi
1sgSM-buy-APPL-PFV PPF-C1-moon PPF-C1-nurse
Intended: 'I bought the nurse the moon!'

To test Hypothesis 2, it is necessary to guarantee that the addressee will be in AKS2, the knowledge state in which s/he is likely to consider other referents besides the intended one. How can this be guaranteed? One reason an addressee might consider unintended referents is if they lack information about the immediate context of discussion. For example, perhaps there has been a lag in time or change of location since the last discussion of the intended referent, as in (16) and (17). In (20), the conversation participants are a mother and two sons, one of whom has been absent for a long time prior to the conversation. The intended referent is the family's only cow, a referent well-known to all participants. AKS2 is induced in the absentee son as he arrives home and is faced

with unexpected news about the death of the family's cow. Although he is able to pick out the relevant cow in his home setting, he has no reason to expect that the news applies to the family cow and not to some other cow. That is, he has no reason to reduce all possible cow referents down to the one he knows in the context of the family home. Hypothesis 2 predicts that NCWO will be used by the mother when delivering the news to the absentee son. This son's knowledge state is contrasted with AKS1 of the live-at-home son, who has seen evidence of the family cow's demise and therefore has no reason to consider other possible cows when delivered the news. Because this is the same knowledge state targeted in (18), canonical word order is predicted for the live-at-home son.

(20) {Your family owns one healthy cow, which your brother Masato has helped raise. Your brother lives far away in Dar es Salaam. One morning you hear that a neighbouring chief is coming into town for a feast the following night. You go to work for the day. You come home in the afternoon and are surprised to find a lot of blood in the backyard and the family's cow gone.}

You to your parents:

a. ni-hɛ a-ma-saahɛ́ ɣá-ru-ire?
COP-WH PPF-C6-blood C6-come.from-PFV?
'Where did the blood come from?'

Your mother:

b. i. n-tó-sintʃ-ir-ire o-mu-témi a-ŋ-ɔ́ɔ́mbɛ
1plSM-PST-slaughter-APPL-PFV PPF-C1-chief PPF-C9-cow
'We slaughtered the cow for the chief.'

ii. #n-tó-sintʃ-ir-ire a-ŋ-ɔ́ɔ́mbɛ o-mu-témi
1plSM-PST-slaughter-APPL-PFV PPF-C9-cow PPF-C1-chief
(Intended: 'We slaughtered the cow for the chief.')

(20, first continuation)

{You say 'Oh no, Masato will be very sad.' The next day, the blood is still in the backyard. Having heard of the chief's visit and the accompanying feast, your brother arrives home in the evening. He enters and exchanges greetings with the family. He has not yet noticed the cow is missing.}

Your mother to Masato:

c. i. #n-tó-sintʃ-ir-ire o-mu-témi a-ŋ-ɔ́ɔ́mbɛ[18]
1plSM-PST-slaughter-APPL-PFV PPF-C1-chief PPF-C9-cow
(Intended: 'We slaughtered the cow for the chief.')
Consultant comment: ambiguous

[18] But see form (21ai) for a slightly altered situation in which this sentence is felicitous.

ii. n-tó-sintʃ-ir-ire a-ŋ-ɔ́ɔ́mbɛ o-mu-témi
1plSM-PST-slaughter-APPL-PFV PPF-C9-cow PPF-C1-chief
'We slaughtered the cow for the chief.'

(20, second continuation)

{After hearing this news, Masato becomes very angry and begins shouting and your parents for killing the cow without checking with him. There is a knock at the door. It is a visitor from a neighbouring village, whom you've never met, who is in town for the chief's feast and is wondering what all the yelling is about.}

Visitor:

d. ne-ke ki-β-eere ha-nɔ
COP-WH C7-be-PFV C16-here
'What is happening here?'

Your mother:

e. i. n-tó-sintʃ-ir-ire o-mu-témi a-ŋ-ɔ́ɔ́mbɛ
1plSM-PST-slaughter-APPL-PFV PPF-C1-chief PPF-C9-cow
'We slaughtered a cow for the chief.'

ii. #n-tó-sintʃ-ir-ire a-ŋ-ɔ́ɔ́mbɛ o-mu-témi
1plSM-PST-slaughter-APPL-PFV PPF-C9-cow PPF-C1-chief
(Intended: 'We slaughtered the cow for the chief.')
(Cannot mean: 'We slaughtered a cow for the chief.')

Context (20) shows that Hypothesis 2 is borne out. NCWO is used when delivering news to an addressee who knows the referent but does not know it is this referent that is being referred to (20cii). Canonical order is considered too ambiguous for this addressee, since canonical order can be used to refer to non-given$_K$ theme referents (Section 3.4) and these are possible referents for an addressee like the absentee brother. Perhaps, then, one of the "purposes" of NCWO is to disambiguate a given$_K$ referent from a non-given$_K$ referent *when disambiguation is necessary*.[19] On the other hand, when a referent's status is clear to both speakers and disambiguation is not necessary, NCWO is infelicitous. This is illustrated by the infelicity of NCWO when spoken to an addressee who knows the referent and knows it is this referent that is being referred to (20bii) and an addressee who does not know the intended referent at all (20eii).[20] In the former case, canonical order is used to deliver the news and context facilitates a given$_K$ interpretation for *aŋɔ́ɔ́mbɛ* 'cow' (20bi). But when the same news is delivered to a stranger, this same word order is interpreted as non-given$_K$ (20ei).

If the current explanation is correct, the function of NCWO must be particularly important in a language with no equivalent to the English determiners *the* and *a*. NCWO seems to be used when a speaker thinks an addressee does not know whether to pick out a non-given$_K$ or a given$_K$ referent.

[19]It is interesting to note that the ability of NCWO to disambiguate a referent is consistent with claims that movement (which would otherwise violate economy) must have an effect on interpretation (Fox 2000). Disambiguation is often cited as such an effect (e.g. Matyiku 2013).

[20] The present analysis predicts that this infelicity is due to a presupposition failure as the stranger will not be able to pick out the given$_K$ cow that NCWO signals.

In this situation (or indeed any situation), English speakers have *the* or *a* to assert which one is intended. But a Nata speaker has no such recourse. Instead, NCWO is a syntactic strategy that can serve this purpose. Languages appear to differ in how they "carve up" this aspect of the (non)givenness$_K$ space: Nata does not mark either givenness$_K$ or non-givennness$_K$ when the correct interpretation is obvious, but the language does mark givenness$_K$ when the correct interpretation is not obvious; that is, when an addressee's knowledge state is one that requires the disambiguation.[21] On the other hand, English obligatorily marks the givenness$_K$ - non-givenness$_K$ contrast but does not distinguish between situations where an interpretation is obvious (as in "the moon") and situations where an interpretation is – prior to the utterance – ambiguous.[22] (The difference can be seen in how an English speaker would react in Context (20). *The cow* would be used both for the absentee brother and the live-at-home brother, while *a cow* would be used for the stranger.)

4 Gradience

In experimental work on the syntactic reflexes of givenness$_{(S\&K)}$, Skopeteas and Fanselow (2009) found that "given material was only optionally fronted" in the languages they looked at. When a theme was given, these languages always allowed their canonical word order as well as a language-specific strategy for marking givenness, such as object preposing.[23] Skopeteas and Fanselow described this as "gradient" behaviour. Following this use of the term, I define "gradience" as the availability, in a certain context, of two or more constructions that are identical in meaning.[24] In this section, I investigate whether or not NCWO is guaranteed in the licensing context described in Section 3.6. Is NCWO predictable or is this a case of free variation? The answer to this will have important consequences for a theoretical account of NCWO and other pragmatically-conditioned phenomena. True gradience will require a theory to justify the intermittence of whatever device is said to theoretically motivate one construction over another (e.g. a feature).[25] To prove the gradience of NCWO, it will be necessary to discover a single context that licenses both a NCWO sentence and another construction identical in meaning (a canonical word order sentence being the obvious candidate). Example (21) provides such a context. This context is identical to (20) except that in the new context, the absentee brother, Masato, stands out as a particularly angry and aggressive person.

[21] I again leave aside the issue of object-marking. It is not known if and how this strategy of marking givenness$_K$ interacts with NCWO. However, in the elicitations for this project, object-marked forms were not offered as an alternate givenness$_K$ strategy in ambiguous situations, so it seems likely that OMs are sensitive to aspects of meaning and/or context outside of those explored in this paper.

[22] It might be possible to view situations where an interpretation is obvious as reflecting the "presuppositional" use of English "the", and situations where an interpretation is not obvious as reflecting the "assertive" use of "the". I thank Patrick Littell for suggesting this terminology.

[23] It is not clear if this variability occurred within individual speakers or only within languages.

[24] I have adopted a more stringent definition than Skopeteas and Fanselow (2009), who use the term informally. S&F suggest their gradience might have resulted from givenness in combination with "further factors not controlled in the experiment", i.e. if all conditioning factors were known, it might have been possible to predict which construction would arise. This would not qualify as gradience under the definition I have laid out.

[25] Skopeteas and Fanselow touch briefly on this, concluding: "Our findings "favor syntactic models in which a 'gradient' conflict resolution is not exceptional or models in which the actual choice between syntactic constructions is not part of the theory of syntax" (Skopeteas and Fanselow 2009:25).

(21) *(= (20, first continuation) but with Mean Masato)*

{You say 'Oh no, Masato will be very angry!' The next day, the blood is still in the backyard. Having heard of the chief's visit and the accompanying feast, your brother arrives home in the evening. He enters and exchanges greetings with the family. He has not yet noticed the cow is missing.}

Your mother to Masato:

a. i. n-tó-sintʃ-ir-ire o-mu-témi a-ŋ-ɔ́ɔ́mbɛ
1plSM-PST-slaughter-APPL-PFV PPF-C1-chief PPF-C9-cow
'We slaughtered the cow for the chief.'
Consultant comment: ambiguous but ok in this case (cf. (20ci))

ii. n-tó-sintʃ-ir-ire a-ŋ-ɔ́ɔ́mbɛ o-mu-témi
1plSM-PST-slaughter-APPL-PFV PPF-C9-cow PPF-C1-chief
'We slaughtered the cow for the chief.'
Consultant comment: ok but will provoke Masato

Context (21) allows both NCWO and canonical word order, providing, at first glance, evidence for gradience. NCWO unambiguously signals a given$_K$ referent and is felicitous just as it was in (20cii). However, consultant comments warn that this sentence is likely to upset Mean Masato, who was not expecting *his* family's cow to be slaughtered. In (21ai), we see the main difference between (20) and (21). Canonical word order was considered "too ambiguous" for the absentee brother in Context (20); however, it is accepted for Mean Masato in Context (21). This is because, in (21), the speaker actually desires ambiguity of the theme referent in order to avoid the fight that will ensue if Masato discovers it is *his* family's cow that has been slaughtered. In the licensing context for NCWO, fear of reprisal can be a motivator for maintaining canonical word order. When faced with canonical word order, Masato might interpret *aŋóomβe* 'cow' as non-given$_K$ (or ask "Which cow?") but the mother is not tied to this interpretation. For example, if accused of lying, the mother could say, 'I thought you knew we'd had to slaughter your cow!' and claim she had intended to refer to Masato's cow when speaking (21ai). The use of canonical word order here represents the mother's attempt to conceal the fact that Masato's cow has been slaughtered – or at least delay revealing this fact – without actually lying.

It appears that context alone cannot predict whether NCWO will be used. In (21) at least, canonical word order is also felicitous. Because of this unpredictabilty, it is tempting to label NCWO as a gradient phenomenon. However, this is not the case. As the preceding discussion illustrates, the appearance of one construction over the other has no element of randomness (again, at least in (21)). In this situation, the meanings of the two word orders, and the outcomes each would produce, are quite different. (Because the two constructions can differ in meaning in some situations, they should not be considered truth-conditionally equivalent *a priori*, even though they become so when context values a canonical order theme as given$_K$.) Therefore, the availability of both canonical order and NCWO in (21) does not represent gradience but instead shows that it is the speaker's choice of meaning that determines which construction will be used. It is ultimately up to the speaker, not the context, to choose a construction, and this choice depends partly on the speaker's particular motivations (e.g. avoiding a fight). A future formal analysis of NCWO will have to take such motivations into account.

5 Conclusion

In this paper, I have demonstrated that Nata NCWO cannot be solely attributed to any of the four most-commonly cited IS notions: focus, givenness_S, givenness_K or topichood. Instead, NCWO signals givenness_K of a theme when a speaker believes an addressee lacks the information to arrive at this interpretation independently. However, NCWO cannot be guaranteed even when the correct conditions exist: a speaker is not forced to use this word order if s/he has reason to maintain ambiguity of the theme referent. This shows that NCWO owes its appearance to a speaker's decision to convey the meaning it carries, a choice which takes into account both the addressee's knowledge state and the speaker's own motivations. Nata NCWO illustrates how syntax can be sensitive to an interaction between context, meaning, speaker intention and a speaker's beliefs about the knowledge of others that has so far barely been explored.

One potential application of the results of this investigation would be to examine the extent to which DOC word order alternations in other ("symmetrical" and "asymmetrical") Bantu languages might be amenable to an explanation similar to the one provided here. Another avenue for future work will be to discover if and how other languages (both those with and without English-like determiners) are sensitive to the same division in addressee knowledge states. At minimum, continuing this line of research in other Nata constructions might reveal the motivation for other alternations in the language. In Bantu and beyond, many explanations for word order variation appeal straightforwardly to one of the IS notions, and Nata NCWO raises the possibility that these appeals may be too simplistic. Important areas of future research will involve developing a more articulated view of the IS notions, exploring seriously the role of speaker intention in "gradient" situations, and discovering what other variables find their encoding in word order variation.

References

Baker, M. C., Safir, K. & Sikuku, J. (2012). Sources of (A)symmetry in Bantu Double Object Constructions. In N. Arnett & R. Bennett. *Proceedings of the 30th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Proceedings.

Birner, B. J. & Ward, G. (1998). *Information status and noncanonical word order in English*. Amsterdam: John Benjamins Publishing.

Bock, J. K. & Irwin, D. E. (1980). Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Verbal Behavior, 19,* 467–484.

Bresnan, J., Cueni A., Nikitina, T. & Baayen, H. (2007). Predicting the Dative Alternation. In G. Boume, I. Kraemer, & J. Zwarts (Eds.), *Cognitive Foundations of Interpretation* (69–94). Amsterdam: Royal Netherlands Academy of Science.

Bresnan, J. & Mchombo, S. A. (1987). Topic, pronoun, and agreement in Chichewa. *Language, 63*(4), 741–782.

Bresnan, J. & Moshi, L. (1990). Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry, 21*, 147–185.

Büring, D. (2003). On D-trees, beans and B-accents. *Linguistics and Philosophy, 26,* 511-545.

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. Li (Ed.), *Subject and Topic* (25–55). New York: Academic Press.

Clifton, C. & Frazier, L. (2004). Should given information come before new? Yes and no. *Memory & Cognition*, *32*(6), 886–895.

Ferreira, V. S. & Yoshita, H. (2003). Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of Psycholinguistic Research*, *32*, 669–692.

Fox, D. (2000). *Economy and Semantic Interpretation*. Cambridge, Massachusetts: MIT Press.

Frey, W. (2005). Pragmatic properties of certain German and English left peripheral constructions. *Linguistics*, *43*(1), 89–129.

Gundel, J. K. (1988). Universals of topic-comment structure. *Studies in Syntactic Yypology, 17,* 209–239.

Gundel, J. K.. (1974). *The role of topic and comment in linguistic theory*. Ph.D. thesis, University of Texas at Austin.

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

Hawkinson, A. K., & Hyman, L. M. (1974). Hierarchies of natural topic in Shona. *Studies in African Linguistics*, *5*(2).

Hyman, L. M. & Duranti, A. (1982.) On the Object Relation in Bantu. *Syntax and Semantics, 15*, 217–239.

Jacobs, J. (2001). The dimensions of topic–comment. *Linguistics*, *39*(4), 641–681.

Kiss, E. (1998). Identificational focus versus information focus. *Language, 74*, 245–273.

Krifka, M. (2008.) Basic notions of information structure. *Acta Linguistica Hungarica 55*, 243–276.

Kuno, S. (1972). Functional sentence perspective. *Linguistic Inquiry, 3*, 269–320.

Louwrens, L. (1979). *Naamwoord-funksies in Noord-Sotho.* PhD thesis, University of Pretoria.

Matyiku, S. (2013, May). *Motivating movement: The case of negative inversion in West Texas English.* Paper presented at the 25th Scandinavian Conference in Linguistics, Reykavik, Iceland.

Ndayiragije, J. (1999). Checking Economy. *Linguistic Inquiry, 30*(3), 399–444.

Prince, E. F. (1992). The ZPG letter: Subjects, definiteness, and information status. In S. A. Thompson & W. C. Mann (Eds.), *Discourse Description: Diverse Analyses of a Fund-raising Text* (295–325). Amsterdam: Benjamins.

Prince, E. F. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.) *Radical Pragmatics* (223–56). New York: Academic Press.

Reinhart, T. (1981). Pragmatics and Linguistics: An analysis of Sentence Topics. *Philosophica, 27*(1), 53–94.

Roberts, C. (2011). Topics. In C. Maienborn, K. von Heusinger & P. Portner (Eds.) *Semantics: An International Handbook of Natural Language Meaning*. Boston: Mouton de Gruyter.

Rochemont, M. (2014). Givenness. In C. Féry & S. Ishihara (Eds.), *The Oxford University Press Handbook of Information Structure*. Oxford: Oxford University Press.

Rochemont, M. (2013). Discourse new, F-marking, and normal stress. *Lingua, 136*, 38–62.

Rochemont, M. (1986). *Focus in Generative Grammar*. Amsterdam: J. Benjamins.

Sadlier-Brown, E. (2013, June). The Nata applied double object construction: Implications for applicative and (a)symmetrical typologies. In S. Luo. *Proceedings of the 2013 Annual Conference of the Canadian Linguistic Association.* Paper presented at Canadian Linguistic Association. Victoria, Canada.

Skopeteas, S. & Fanselow, G. (2009). Effects of givenness and constraints on free word order. In M. Zimmermann & C. Féry (Eds.), *Information Structure: Theoretical, typological, and experimental perspectives* (307–331). Oxford: Oxford University Press.

Watters, J. R. (1979). Focus in Aghem. A study of its formal correlates and typology. In L. Hyman (Ed.), *Aghem Grammatical Structure, Southern California Occasional Papers in Linguistics 7* (137–197). Los Angeles: University of Southern California.

Zerbian, S. (2006). *Expression of Information Structure in the Bantu Language Northern Sotho.* Ph.D. thesis, Humboldt-Universität zu Berlin.

Zubizaretta, M. (1998). *Prosody, Focus and Word Order*. Cambridge, MA: The MIT Press.