

Antitopic and afterthought Givenness and grammar in Ponca and Omaha

Bryan James Gordon
University of Minnesota

0. Introduction

Ponca and Omaha (Siouan: Mississippi Valley: Dhegiha; henceforth OP¹) are two mutually intelligible, highly endangered languages of the Missouri River Valley. OP are head-final, but frequent postverbal referring expressions have led to OP's being described as "free-word-order". On the contrary, I show that postverbal referring expressions occupy a grammatical position, whose use is restricted to *antitopics*, an information-structural type distinct from topics.

The theoretical point of departure is described in §1 and linked to the literature; and I propose a necessary pragmatic condition upon the postverbal position. In §2 I present the qualitative and in §3 the quantitative evidence in support of this condition. I argue in §4 for considering it both necessary *and* sufficient. An extrasentential analysis of postverbal expressions is considered and rejected in §5. In §6 I compare my findings with findings from geno- and phenotypically similar languages. Finally, in §7 I propose a grammaticisation explanation, and then discuss (§8) and conclude (§9).

1. Theoretical background

Received distinctions such as *given/old/new* or *topic/focus* are ambiguous in the literature. I offer a set of terminology, and devote time to relating it to terms from the literature. In the following model and the rest of this work I adopt the practice of rendering operationalised terms in SMALL CAPS, and more general terms from the literature in *italics*.

Binary distinctions such as *topic/comment* and *focus/presupposition* gloss over both the surface complexity and the deep simplicity of information structure. Although rigid definitions of *topic* and *focus* empirically necessary, they are rooted in a less rigidly defined psychological structure. *Focus* is, quite simply, the most FIGURE-like component of a given discourse unit, and *topic* functions as the GROUND upon which that FIGURE stands.

- (1) a. **FIGURE:** a component of an information unit which is psychologically new and interesting to the hearer relative to other information in the same unit
- b. **GROUND:** information not included in FIGURE but necessarily interpreted in the same unit

At least two types of information can function as FIGURE or GROUND. Akin to É. Kiss's (1998) split between *information* and *identificational focus*, in the case of referring expressions I distinguish between propositional content (the semantic type) and referential content (the real-world referent) – henceforth P- and R-information. There are additionally at least two ways information can function as FIGURE in its context, as Gundel (1999) notes with her distinction between *semantic* and *contrastive focus*. Discourse-new P- and R-information tends to be associated with the former type, information in scalar relation with other information with the

1 OP are often called a single language, "Omaha-Ponca". The abbreviation OP is my concession to this practice; but I do not refer to them as a single language, because this offends and perplexes speakers and non-speakers.

latter. It will be necessary to operationalise these types for the present purpose:

- (2) a. **GIVEN:** an attribute of information highly accessible in the cognitive context
- a'. **NEW:** an attribute of information less highly accessible in the cognitive context
- a". **CENTRE OF ATTENTION:** (henceforth COA) the highest GIVENNESS status; c.f. *highly accessible* (Ariel 1990), *psychological focus* (Gundel 1999), *center of attention* (Chafe 1976)
- b. **CONTRASTIVE:** an attribute of information which stands in a scalar or partially-ordered-set relation (c.f. Hirschberg 1985, Ward 1985) to an antecedent or immediate succedent
- b'. **NON-CONTRASTIVE:** information that stands in a non-scalar relation to an antecedent; or information with no antecedent at all

Unlike the terms in (1), those in (2) are defined as *attributes* of information. FIGURE and GROUND are active categories. They describe the status a speaker *intends* to be mapped to a piece of information. [\pm COA] and [\pm CONTRAST] on the other hand are stative categories which describe the status information *already has* at the moment of its deployment.

A focus can be focal by virtue of being NEW, or by virtue of being CONTRASTIVE. But there may be other, non-focussed NEW or CONTRASTIVE information in the same discourse unit. For this reason Ward's (1985:87) notion of salient and relevant *open propositions* is crucial to extracting the focus or foci from discourse units. (Refer to Ward for robust discussion and further definition.)

- (3) **FOCUS:** the most FIGURE-like information in its unit. A salient and relevant open proposition is retrievable from GROUND and cognitive context, and is filled by the FOCUS.

Non-focussed NEW or CONTRASTIVE information, then, functions to alter or specify the open proposition. But the open proposition includes GIVEN, NON-CONTRASTIVE information as well, information retrievable without explicit reference, information devoid of FIGURE.

- (4) **ANTITOPIC:** information entirely in GROUND and devoid of FIGURE, i.e. [$+$ COA, $-$ CONTRAST].

Including [\pm COA], a GIVENNESS category, in a syntactic construction is more controversial than including [\pm CONTRAST], or syntactic/prosodic prominence. Such an implication has seldom been argued for. Chafe's (1976) is to my knowledge the first use of the term *antitopic* for a pragmatic category which functions to "confirm established information", and already he recognised the syntactic implications, noting that they frequently occur at the right periphery. Mithun (1999:199-200) similarly locates *topic shift* at the left and *antitopic* at the right periphery for Tuscarora (Iroquoian: Northern: Tuscarora-Nottaway). Erkü (1983:184) claims for Turkish postverbal material that "two aspects of discourse interact with word order: topic-comment structure, and activated/unactivated status." But the syntactic implications of these claims are not expounded upon by Erkü, Chafe or Mithun.

The categories defined in (3) and (4) may be seen as extreme cases of information status, between which the intermediate case, which I will call TOPIC, can be negatively defined.

- (5) **TOPIC:** information, part in GROUND and part in FIGURE, which alters the open proposition associated with the discourse unit's FOCUS, but is irretrievable without explicit reference:
 - a. TOPIC is distinguished from FOCUS by affecting rather than filling the open proposition.
 - b. TOPIC is distinguished from ANTITOPIC either by being irretrievable [$-$ COA] and augmenting the open proposition or by functioning to alter the open proposition through CONTRAST.

Few analyses of pragmatically marked constructions make a distinction like that between (4) and (5). *Focus/presupposition* contrasts (c.f. Chomsky 1971) conflate the two, and most current work in the Minimalist Program – two influential examples are Zimmermann (1999) and Rizzi (2000) – does as well. The status of ANTITOPIC vis-à-vis *topic/comment* contrasts varies. "Aboutness" approaches, which more or less paraphrase the FIGURE/GROUND distinction, tend not to distinguish antitopic from topic, since both are what the sentence is "about":

- (6) An entity E is the topic of a sentence, S, iff in using S the speaker intends to increase the addressee's knowledge about, request information about, or otherwise get an addressee to act with respect to E. (Gundel 1988:210)

Indeed, the notion of *topic* as something that discourse units are "about", or as information that persists in discourse, has no need of further division. But the equally prevalent notion of *sentence topic* treats form as well as function; and prosodically and syntactically oriented approaches to *topic* tend to treat TOPICS as in (5), whether this restriction is made explicit or not. Whether backgrounded information can be seen in such approaches as marginally *topical*, or part of *comment*, depends on the approach. Some approaches (e.g. Lee 2002) distinguish *contrastive* from *non-contrastive topics* – the latter encompassing both presentational topics and antitopics. Most subsume *focus/presupposition* structure itself to *comment*, tacitly grouping background under *presupposition*. (*Topicalization* as considered by both Gundel and Ward resembles a branching structure of this sort.) Mithun's, Chafe's and Erkü's approaches (see above) and Vallduví & Vilkuna's (1998) stand out in devoting analysis to *background/antitopic*. V&V's *theme/rheme* distinction, indeed, is isomorphic with my GROUND/FIGURE distinction. However, (1) V&V's top branching feature [\pm rheme] is not explicitly psychological, but rather a blend of stative GIVENNESS with active language on *assertion* and *presupposition*; and (2) their subdivision of *theme* into *topic* and *tail* (c.f. TOPIC and ANTITOPIC) is purely formal and misses the functional [\pm COA] and [\pm CONTRAST] features which I consider inextricable.

ANTITOPICS are palpably distinct from TOPICS, and function distinctly. While TOPICS needn't be GIVEN – only specific so that they can have an identifiable projection in an open proposition – ANTITOPICS are superlatively GIVEN. While TOPICS are *intended* to persist in discourse, ANTITOPICS are always already persistent. TOPICS and ANTITOPICS are universally associated with disparate prosodic and syntactic forms. Indeed, TOPICS are far more difficult to discern formally from FOCI than from ANTITOPICS, since TOPICS and FOCI are often expressed with the same syntactic (though usually not prosodic) devices (Ward 1985, Gundel 1974, 1975), and TOPICS' status as part of GROUND is not absolute like that of ANTITOPICS, but rather *relative* to the status of the respective FOCUS (c.f. Gundel & Fretheim 2004). A clearer approach than that in the extant literature is called for, wherein ANTITOPIC has theoretical status independent of both *topic* and *presupposition* as well as emergent formal properties such as *background*.

Armed with the necessary terminology and theoretical background, the reader may now confront the thesis of this paper:

- (7) In Ponca and Omaha, only ANTITOPICS may be referred to by postverbal referring expressions.

2.0. Methodology

I coded 142 referring expressions from a body of stories, historical texts and letters transcribed in the late nineteenth century by missionary James Owen Dorsey (1890). Of these data, 83% were

obtained by analysing randomly selected texts in their entirety, 49% by directly searching for expressions with particular determiners using regular expressions². The latter search was carried out to fill the many cells required by OP's array of determiners³, demonstratives, genitives and combinations thereof. Both methods retrieved many of the same data, hence the overlapping 32%. Additionally, studies with data from native speakers are consulted where possible, notably Eschenberg (2005a) and Rudin (1998).

For each discourse unit⁴, I first identified the FOCUS or FOCI according to (3) and its/their concomitant open proposition as in Ward (1985). The referring expressions⁵ not coded as FOCUS were then gauged according to Hirschberg's (1985) definition of *scalar relations*, the TOPIC criteria (5), and the ANTITOPIC criteria (4) – in terms of both their P- and their R-information. The criteria for the [±COA] feature referenced in (4) and (5) are taken, with adaptations, from the working coding protocol of the Minnesota Cognitive Status Working Group (Gundel, Bassène, Gordon, Humnick & Khalfaoui 2007), based on the category in [*psychological*] focus from Gundel, Hedberg & Zacharski 1993:

(8) Sufficient criteria for [+COA]:

- a. used as subject of immediately preceding clause/sentence
- b. used earlier in same clause/sentence
- c. a higher-level topic that is part of the interpretation of the preceding clause/sentence
- d. used in both of the two immediately preceding clauses/sentences
- e. the event denoted by the immediately preceding clause/sentence

Necessary criterion for [-COA]:

- f. meets none of the criteria (a)-(e)

Of the three terms in (3)-(5), the present work only maps the second (4) to a syntactic form; but all three are coded for, because TOPIC/FOCUS status is sufficient for non-ANTITOPIC status, and the identification of ANTITOPICS is made easier by the identification of FOCI and (where possible) TOPICS.

2.1. GIVENNESS and the postverbal position

In the following three subsections, I show how the data satisfy thesis (7). One prediction is that

2 Searches were performed using Kate (K Advanced Text Editor) upon a databased version of Dorsey (1890) prepared by John E. Koontz and converted by me to Unicode characters. Please contact me for a copy of my version, or to ask Koontz for his.

3 OP possess a typologically extreme eleven definite articles. These are specified for animacy, affectedness/case (Shea 2007), position/shape, motion, discourse-centrality (Eschenberg 2005a), and possibly number as well. They double as auxiliaries and evidentials (Eschenberg *ib.*, 2005b), and their semantics is the same when they serve as auxiliaries as when they serve as articles.

Other typological rarities, less pertinent but possibly familiar or interesting to readers, are (1) a four-mode distinction in the oral stops (e.g. /d t: tʰ tʰ/ – orthographically <d t tʰ tʰ>); (2) an interdental lateral glide (/ʎ/ – no standard IPA symbol – orthographically <th>); (3) a split-transitive pronominal system (both ergative and accusative) alongside a nominal system with a nominative/accusative system (Shea 2007), a discourse-marked system (Eschenberg 2005a) or both; (4) a deixis-based motion verb paradigm (Cumberland 2006).

4 Discourse units considered include full sentences as well as tight-knit groups of fragments and smaller units. This impressionistic division is a likely source of coder variability should someone replicate my methodology.

5 Propositional expressions, adjunct adverbials, locatives, quotatives and nonsubcategorised referring expressions were not coded for TOPIC/ANTITOPIC status. Including these expression types would have added complexity not directly pertinent to the present analysis. Watam, of the Lower Ramu valley in Papua New Guinea (Foley 1999), provides an example of an OV language which is liberal in allowing nonsubcategorised adjuncts to be expressed postverbally. In fact, in Watam, *only* nonsubcategorised adjuncts may be expressed postverbally.

the referents of postverbal referring expressions be in the centre of attention, and that less GIVEN referents cannot be referred to postverbally.

- (9) Táxti-gík^hidabi ak^ha ededí ak^ha-ama. (22.1)⁶
 Deer-they.kill.for.him the there AUX-REPORT⁷
 There was They-kill-deer-for-him.

In (9), *Táxti-gík^hidabi* (an evil giant) was coded as familiar (in long-term memory) to discourse participants [-COA]. This is therefore an example of an expression that cannot be postposed. The same referent in (10), on the other hand, was coded as [+COA], and this expression is postposed:

- (10) (later on, during an exchange between the protagonist Rabbit and Táxti-gík^hidabi)
 “Wami aⁿwaⁿbit^haⁿga ehe ha, Táxti-gík^hijabi-a!”
 blood push.me.down.in I.say DECL Deer-they.kill.for.him.DIMIN-VOC
 Ku[?]e athab’ egaⁿ wami ubit^haⁿbi-ama Táxti-gík^hidabi ak^ha. (23.14-16)
 Rushing went having blood pushed.down.in-REPORT Deer-they.kill.for.him the
 [Rabbit taunted,] “Push me down in the blood, I say, little They-kill-deer-for-him!” So they say They-kill-deer-for-him rushed over and pushed Rabbit down in the blood.

2.2. CONTRAST and the postverbal position

From (7) similarly follows that postverbal referring expressions refer exclusively to NON-CONTRASTIVE information. In (11), the referent of the second underlined expression (the *Urstone*, so large it blocked the sun) was coded as [+COA], being the subject of the immediately preceding sentence, and thus could be a candidate for postverbal position. But it also contrasts with the referent of the *third* underlined expression in a *part/whole* or *source/product* scalar relation, which is why it is evoked *in situ*.

- (11) íⁿ?e thaⁿ ugashneshn’ égaⁿ gatubextiáⁿbi-ama. K^hi edit^haⁿ íⁿ?e thaⁿ
 stone the.ROUND crack.REDUP having by.falling.ground.very-REPORT and from.there stone the.ROUND
 gatube ugaextiáⁿbi-ama, mazhaⁿ bthúga águdishtewaⁿ íⁿ?e ge. (331.4-5)
 by.falling.ground scattered.far-REPORT land all wherever stone the.SCATTERED
 They say the stone cracked all over and was ground very finely by the fall. They say that from there that one round stone was ground and scattered far, and became all the scattered stones of all the lands everywhere.

2.3. Exceptions?

One postverbal referring expression in the data⁸ was analysed as [-COA] (none as [+CONTRAST]):

6 The orthography used here is the phonemic official orthography of the Ponca Tribe of Oklahoma. Although the source texts are both Omaha and Ponca, I have used just one orthography, since the differences between the two orthographies are small.

Examples are given in the font *AboriginalSerif*, from indigenous-friendly www.languagegeek.com.

7 Small-caps glosses used: DECL – declarative particle; DIMIN – diminutive; HORIZ – horizontal-marked article; REDUP – reduplication; REPORT – reportative/hearsay; ROUND – round-marked article; SCATTERED – scattered-marked article; VOC – vocative.

8 Other exceptions, not analysed, were postverbal referring expressions occurring in locative phrases and quotative constructions.

- (12) P^hahaⁿga t^hédi athabi-ama Mashchiⁿge ama. (25.1)
past the.at went-REPORT Rabbit the
 Once upon a time Rabbit was going along.

Example (12) begins a story, so Rabbit is certainly NEW. However, this formulaic, performative (c.f. Austin 1962) construction is a frequent introduction to stories, and there is reason to analyse it as a fossilised idiom, which poses no threat to generalisations about productive grammar.

3. Quantitative evidence

This work combines qualitative microanalysis with quantitative methods (c.f. Mendoza-Denton 2007, Johnson 1997, Schegloff 1987), paying close attention to the formal and pragmatic features of each utterance in its context. Thus, it was unfeasible and undesirable to automate the coding routine over the entire corpus, and consequently quantitative evidence should be seen as supportive rather than primary.

Data were selected as described in §2.0. I found 98% of postverbal referring expressions (n=43) to refer to ANTITOPICS. (The one exception is (12) in §2.3.) Of *in-situ* expressions (n=99), 91% referred to FOCI OR TOPICS. (The exceptions are detailed in §4.1-§4.4.) Without further qualitative analysis, these figures provide evidence for phrasing (7) only as a necessary but insufficient condition.

4.0. ANTITOPICS *in situ*

Not all expressions in the data coded as ANTITOPICS were postverbal. Indeed, my claim thus far stops short of demanding this, imposing (7) as a necessary, not sufficient, condition. But in the following sections I present rationales for each *in-situ* ANTITOPIC. The existence of such clear rationales lends support to rewording (7) as a bidirectional implication.

4.1. Multiple ANTITOPICS *in situ*

- (13) Sithemak^haⁿ ak^ha p^háhaⁿ át^hiathabi-ama. “Kaⁿha, shék^he tashniⁿgthishka há
Sithemak^haⁿ the arose stood.suddenly-REPORT Grandmother.VOC that.the spotted.fawn skin
 úzhiha íⁿ?i thága,” ábi-ama. Gí?i théthabi-ama. Sithemak^haⁿ bthúga
bag to.me.give right.away said-REPORT to.him.gave right.away-REPORT Sithemak^haⁿ all
 uginazhiⁿbi-ama, táxti gaghabi-ama. (57.8-10)
in.his.stood-REPORT deer made-REPORT

They say Sithemak^haⁿ stood suddenly and said, “Grandmother, hand me that spotted-fawn-skin bag right away.” They say she gave it to him right away. They say Sithemak^haⁿ stood completely inside it, and made himself into a deer.

The referent of the underlined expression in (13) is certainly [+COA, -CONTRAST], but he is referred to *in situ*. Why? One possibility is that we should expand our notion of FIGURE to include resolution of referential ambiguity (c.f. Sneed German & Pierrehumbert's (2008) *attentional shift* and Mithun's (1999) *topic shift*). At a point in discourse at which there are three persistent ANTITOPICS, (Sithemak^haⁿ, his grandmother and his bag), "reduced" forms such as zero or the postverbal ANTITOPIC position create ambiguity, even if only temporarily. Certainly this is a fact that is compatible with a bidirectional variant of (7).

4.2. Repetition and ANTITOPICS *in situ*

The arguments presented in this section and in §4.3 and §4.4 refer to the following section of data, which it is important to present in its entirety for context's sake. The reader's attention is directed to the English translations, so that an idea of the context is obtained.

- (14)a. T^háⁿwaⁿgthaⁿ wiⁿ edí-thaⁿ-ama. K^hi wa^ʔu wiⁿ údaⁿx^ti athiⁿbi-ama t^háⁿwaⁿgthaⁿ ak^hadi.
K^hi shénuzhiⁿga gáⁿtha ahí-hnaⁿbi-ama. K^hi thi^ʔa-hnaⁿbi-ama.

It is said that there was a village. And they say they had a woman in this village who was very beautiful. And the young men would come desiring her, they say. And, they say, they would fail.

- b. K^hi níashiⁿga shénuzhiⁿga wiⁿ, “Hiⁿda! wa^ʔu gáⁿthai thi^ʔa-shnaⁿi éde wa^ʔu k^háⁿbtha⁹
bthe t^hé-na!” ethegaⁿbi-ama. Gaⁿ shénuzhiⁿga ak^ha athabi-ama.

And they say one person, a young man, thought to himself, “We’ll see! Sure, they may have wanted the woman and failed, but I want her and I will go anyway!” And they say the young man went.

- c. Shaⁿ pahe wiⁿ taⁿgax^ti edi-thaⁿ-ki níashiⁿga wiⁿ agthiⁿ ak^ha-ama. Shénuzhiⁿga
míⁿthigthaⁿ athe ama níashiⁿga pahadi gthiⁿ thiⁿk^he tape athabi-ama.

Now, there was a very large hill there, and they say a person was sitting upon it. The young man who was walking along with the woman on his mind came up close by the person sitting on the hill, they say.

A lot happens in (14b). The second and third underlined expressions are both IP's whose predicates are *want the woman* but whose subjects differ. Considering that postverbal – *éde káⁿbtha bthe t^he-na wa^ʔu (thiⁿk^he)* – or zero – *éde káⁿbtha bthe t^he-na* – expression of the object should be possible, and that these orders surface for other ANTITOPIC objects, why do they not surface here? One possibility is repetition. The resonant frame in this case is phonological, semantic, syntactic and referential all at once: almost all of the information in each of these components of language is preserved across the two IP's. Only morphologically is there contrast, and contrast has meaning in *failure to repeat*. The repeated IP is a part of the open proposition – *A ∈ {other suitors, protagonist} does X with respect to wanting the woman* – in both instantiations, and it is also plausible that reordering of a repeated open proposition is dispreferred in practice.

Fully 56% of *in-situ* ANTITOPICS in the data exhibit repetition of this sort. Repetition is a wild card that overrides much of what can otherwise be generalised about grammar. An emerging consensus identifies repetition and priming as a fundamental building block of all areas of language (c.f. Bock 1986, Bock & Loebell 1990, Weiner & Labov 1983, John du Bois's work).

4.3. Special discourse structures and ANTITOPICS *in situ*

The protagonist referred to in the first underlined expression of (14b) is still an ANTITOPIC at the time of his mention as the fourth underlined expression of (14b), having been the subject of the intervening *think* predicate. Yet his referring expression is *in situ*. The construction *Gaⁿ + sentence* marks a turn of events, a conclusory statement, a resolution of narrative tension, or otherwise directs the hearer to get ready for something new. Another discourse particle which accomplishes this effect quite often in the texts is *égithe*. After such particles, expressions coded as ANTITOPICS are found *in situ*.

Chafe (1994) argues that the centre of attention is constantly shifting from one

9 Although the correct form of this verb is *kaⁿbtha*, I have retained the phones represented in Dorsey's transcriptions. The "incorrect" form may be a speaker error, a transcription error, an instance of phonetic change in progress, a spelling adjustment, or any number of other things.

[psychological] focus to another. In this way, discourse particles such as *gaⁿ* and *égithe* direct the hearer to ignore her current centre of attention and pave the way for a new one. Consequently, the young man is actually [-COA] at the moment of his last mention: the centre of attention is instead on the resolution of the discourse tension created by use of the particle *gaⁿ*.

4.4. Elaboration and ANTITOPICS *in situ*

Even the most trivial elaborative or supplementary P-information, by virtue of being NEW, is FIGURE-like, even if signified by an expression whose R-information is entirely in GROUND. Ponca and Omaha bar all NEW information from postverbal expressions, regardless of whether it is P- or R-information. This may be seen in (14c). Despite the fact that the *identity* of the young man is clearly still an ANTITOPIC piece of information, and that the hearer has probably already inferred the content of the elaboration – *he is walking along with the woman on his mind* – the elaboration does not concern his *identity*, only his *type*, and is NEW to the linguistic context and a component of FIGURE, albeit a marginal one.

None of the rationales for ANTITOPICS in situ – elaboration, special discourse structures, repetition and ambiguity – pose a significant difficulty for a rewording of thesis (7) as a bidirectional implication, and therefore I do so here:

(15) *In Ponca and Omaha, referring expressions occur postverbally precisely whenever they refer to an ANTITOPIC.*

5. Afterthoughts

The reader may have objected by now: *OP are head-final and predominantly SOV, so why not analyse these postverbal referring expressions as extrasentential afterthoughts?* In Somali, for instance, there are no postverbal referring expressions within the sentence: all postverbal referring expressions are afterthoughts (Saeed 1984). In the Dorsey texts we can never be fully certain that a postverbal expression is *not* an afterthought, because OP have null argument pronominals which confound the English afterthought test (e.g. *He's foolish, your boyfriend*, which would be rendered in OP as *foolish, boyfriend yours the*). Also, extrasentential afterthoughts are often distinguished on the basis of prosodic information, and 115-year-old texts are an unreliable source of prosodic data in general.

Nonetheless, we can rely on comparative data from other OV languages. Kaiser (1999) finds that Japanese speakers position "old, non-prominent" (a good paraphrase of ANTITOPIC as in (4)) information postverbally. In Turkish, similarly, "non-activated" (NEW) referents at the right periphery are afterthoughts, while the "activated" (GIVEN) ones are intrasentential, according to Erkü (1983:154). Erkü also finds that these two patterns are prosodically discriminable, and therefore are formally and functionally distinguished.

As evidence that the same discrimination can be made in the Dorsey texts, note first that there are no postverbal FOCI or TOPICS in the Dorsey texts other than the sole exception (12), and second that there are commas in the texts which offer at least a somewhat reliable source of the particular type of prosodic information we need. Observe the following:

(16) "Thazhai ki daⁿbaia hé, thék^{he}," ábi-ama. (57.2)
you.doubt when look.at DECL this.the.HORIZ said-REPORT
 They say she said, "If you don't believe me, look at him, this one here."

The referent of the underlined expression, *this one here*, is referred to by overt deixis with additional semantic information about his horizontal posture. Clearly, he is not an ANTITOPIC, so he should not be referred to postverbally *within the sentence*. But the comma indicates that this expression is an afterthought, so we can rest at ease. Indeed, while ANTITOPICS, TOPICS and FOCI all occur in comma-marked afterthoughts in the Dorsey texts, postverbal referring expressions without commas in the data refer solely to ANTITOPICS.

6. Comparative evidence

We have already seen examples from phenotypically similar languages which either diverge strongly from Ponca and Omaha in the area of postverbal referring expressions (Somali), are very similar (Japanese, Turkish), or are somewhere in the middle (Watan). Turkish probably presents the most similar case among the non-related languages. As described by Erkü (1983) and Erguvanlı-Taylan (1984), the Turkish postverbal ANTITOPIC position is functionally and formally identical to the OP position in every way but that it allows ANTITOPIC referents with elaborative information attached to them, in contrast to the phenomenon described in §4.4.

The genotypically similar languages at first glance provide conflicting information. Quintero (2004) finds that the postverbal referring expressions in her data on Osage (a member of the Osage-Kansa subgroup closest related to OP) are all afterthoughts, and attributes the more frequent occurrence in Dorsey's Osage-language texts to storytelling style. Rudin (1998), however, finds that postverbal order for referring expressions is *more* common in conversational style than in formal registers in Omaha, so this analysis does not work for OP. Rankin (p.c.) points out that Quintero may not have distinguished between intrasentential ANTITOPICS and afterthoughts as is done here, so OP-like features may turn up if the data are scrutinised in that way. Indeed, Rankin (2005) notes that OVS is a common order for all Dhegiha languages.

A more distant relative, Lakota (Siouan: Central: Mississippi Valley: Dakotan), exhibits little postposition, with 0.9% occurrence of postverbal arguments in Mithun's (1999:196) survey, compared to 20.4% for Omaha! A still more distant Siouan language, Mandan (Siouan: Central), exhibited a 0.8% postposition rate (Mithun *ib*), although Sara Trechter (p.c.) notes a somewhat higher occurrence in her more recently obtained data. Assiniboine, a member of the Sioux-Assiniboine-Stoney (Dakotan) dialect continuum that includes Lakota, does postpose highly GIVEN information (Cumberland 2005:420), although it is not clear how frequently.

The comparative evidence, tentatively, is that the Dhegiha family developed a postverbal syntactic slot for ANTITOPICS, which may have developed further in significantly disparate ways within the family. This construction distinguishes Dhegiha from the rest of Siouan, although perhaps not all of Mississippi Valley Siouan: Proto-Siouan has been decisively reconstructed as SOV (Rankin 2003:201).

7. A grammaticisation explanation

A postverbal syntactic slot for ANTITOPICS which developed in one community but not in another should have a valid grammaticisation¹⁰ trajectory, and the proposal hereunder satisfies this demand. (A comma represents a prosodic pause; parentheses are sentence boundaries.):

¹⁰ I assume a basic framework for variation, change and grammaticisation that draws i.a. on Guy (forthcoming), Hopper (1991) and Mithun (1991).

- (17)a. Originally, only SOV existed. Speakers sometimes uttered ANTITOPICS as afterthoughts:
 (S O V) ~ (O V), S¹¹
- b. This discourse pattern was pragmatically unsatisfactory, because TOPICS and FOCI could also be expressed by afterthoughts. An alternand emerged with a shorter pause, and the pause became shorter and shorter over historical time until it was gone:
 (S O V) ~ (O V), S ~ (O V) S
- c. At this point, since no intrasentential subject expression remained in the third alternand (as noted in §5, the English afterthought test is inapplicable to OP), speakers reanalysed this structure as sentential:
 (S O V) ~ (O V), S ~ (O V S)
- d. At this point, Grice's Maxim of Quantity (1975) pressured speakers to abandon the use of the preverbal pattern for ANTITOPICS, leaving it the exclusive domain of TOPICS and FOCI. This forced the transition from the necessary condition in (7) to the bidirectional implication in (15). At this point (O V S) had grammaticised to the acceptable syntactic pattern for intrasentential ANTITOPIC subjects (and, later, objects as well), reflecting the state of Ponca and Omaha at the time of the Dorsey texts:
 (O V), S ~ (O V S)

Mithun (1999:199) claims that *ambiguity* may have inhibited Lakota from developing an OP-like information/syntax interaction. Lakota has only one definite article compared to OP's eleven, which carry much more semantic information than Lakota's *kiŋ*. Mithun (ib): "[The OP definite articles'] development apparently permitted speakers of Dhegiha languages to exploit word order for pragmatic purposes more often than speakers of Lakota, with less danger of ambiguity." (Where, however, does this explanation leave Assiniboine, which as previously noted *does* exhibit postposition, but has no articles at all?)

8. Discussion

Many directions for future research and implications for extant findings will already be apparent to the reader from the literature component of §1. Note also the success with which *pragmatic* criteria, with little attention to formal prominence or background, have triangulated a syntactic generalisation over the data. The form-function relationship between *background* as a syntactic/prosodic category, and ANTITOPIC as a relational and referential pragmatic category, demands a more sophisticated understanding; and referential scales such as Gundel, Hedberg & Zacharski's (1993), Ariel's (1990) and Chafe's (1976) which explicitly distinguish the centre of attention from mere working memory are light-years beyond other scales which do not.

The notion that prominent, marked TOPICS include at least some FIGURE information (which distinguishes them from ANTITOPICS) is a major departure from exclusively GROUND-oriented "aboutness" definitions such as (6), and to some extent parallels Ward's work specifying the types of functions that such *topics* (caveat: Ward disprefers this term) are required to embody. This notion is very *easily* empirically testable, and should be matched against other languages.

Like Gundel's (1974) *topic/comment* structure and her *topic* definition repeated in (6), my analysis concerns sentence structure and necessarily foregrounds sentences, but is more broadly concerned with the use of these sentences within intentional speech acts. As was seen in §4.3, even high-level discourse phenomena have relational and referential pragmatic information

¹¹ Although this trajectory is valid for both S and O, I trace here the special case of S only. There is syntactic reason to believe the development of the ANTITOPIC position may have begun with S only, and later generalised to O.

attached to them, with consequences for grammatical surface structure. Gordon (2008) and Jackson (2008) argue for syntactic-pragmatic models in which surface order is determined by informational, social and other high-level factors. Feist (2007), on the other end of the scale of syntactic size, applies similar distinctions *within* referring expressions. These methods, including analysis of both P- and R-information, should be expanded to granularities besides just sentences, and the formal consequences of TOPIC, FOCUS and ANTITOPIC structure on non-referring expressions (e.g. VP, IP, non-constituent units, etc.) certainly merit further study.

Effects of pragmatic structure on surface syntax are the object of a large body of research, but explicit attention needs to be paid to reconciling the inherently non-lexically-projected nature of pragmatic structure with models of narrow syntax. As argued elsewhere (c.f. Gordon 2008 and Jackson 2008), discourse-pragmatic information interacts directly with narrow syntax, at phase boundaries, to produce observable conventions best modelled by a combination of top-down (e.g. ranked constraints or constructions) and bottom-up (lexically projected) models of syntax.

Such a combined model accurately predicts a typology of languages with discretely distributed syntax-pragmatics generalisations, such as OP, and languages where this surfaces as probabilistic tendencies, such as English. For OP, it is feasible to pursue even the notion of a word order in total synchrony with information structure – TOPIC-FOCUS-ANTITOPIC is a plausible word order for OP (from which would follow the empirically testable prediction that ANTITOPIC verbs such as in (14b) prevent postposition of ANTITOPIC expressions). Good (2008) argues for a similar template, *topic-predicate-focus* (presumably *predicate* subsumes ANTITOPIC) in some Bantoid languages. On a different tact, Shibatani (2008) argues for a fully syntactic, predictable and interpretable *topic* position alongside and disparate from *subject* in the Sasak and Sumbawa (Austronesian: Western Malayo-Polynesian: Malayo-Sumbawan) subfamilies in Indonesia. Such generalisations demand serious attention from within mainstream syntax. Happily, some commentators within the Minimalist Program have begun to abandon the exclusively bottom-up approach (e.g. Frascarelli 2000, Newmeyer 2003a, 2003b, 2004, forthcoming).

Another promising direction is surveying the boundaries of TOPIC, FOCUS and ANTITOPIC in different languages vis-à-vis syntactic and prosodic facts. As already seen, Turkish allows elaborative information in its ANTITOPIC position, and not all languages categorise information in exactly the same way. My coding criteria, in fact, are in part an artifact of the distribution of data in Ponca and Omaha, and may need tweaking to apply to other languages. The level of GIVENNESS demanded of the ANTITOPIC category might differ from language to language, as might the types of relations that qualify as CONTRASTIVE. The *usage* of the categories may be inconsistent across languages, but their *existence* is universal. There are grey areas in between FIGURE and GROUND, as well as gradient degrees of prominence in the signal. Particular analyses of the same data may differ as well: what Gussenhoven (2007) calls *reactivating focus* is a type of *topic* for most authors. Most significantly, this work challenges binaries in the literature, such as *topic/comment*, *topic/focus*, *focus/presupposition* and *subject/predicate*, which do not adequately chart observable form-function mappings. Formal and functional distinctions between TOPIC and ANTITOPIC are real; and many generalisations have failed to be captured in syntax, pragmatics and child language acquisition due to glossing over them.

Implications of pragmatic statuses upon syntax and prosody are among the most difficult areas for second-language learners (and educators) to master. It is my ardent hope that these findings can be seized upon by Ponca and Omaha teachers and learners, and that the methods used can be extended to other sociolinguistic contexts involving language revitalisation or large-scale second-language learning.

9. Conclusion¹²

OP word order is discourse-configurational to a certain extent, but not to the extent attributed e.g. to Iroquoian, Algonquian and Pama-Nyungan: *OP word order is not free*, but rather *determined* by pragmatic considerations, and a rigid distinction between two types of GROUND usually conflated in the literature is crucial to describing this determination. All, and only, ANTITOPICS are referred to postverbally, and similar claims are possible with respect to TOPICS and FOCI too. There is both qualitative and quantitative support for this claim, and a plausible grammaticisation explanation (17) for the synchronic state represented in the Dorsey corpus. Repetition, special discourse structures and topic competition can override the constraint. There is an observable difference between intrasentential postverbal referring expressions and afterthoughts. Due to the discrete distribution of the data, the bidirectional implication in (15) and paraphrased in this paragraph merits inclusion in the grammar, specifically a component which interfaces between the lexically projected component (narrow syntax) and top-down constraints and/or constructions.

References

- Ariel, Mira (1990) *Accessing noun phrase antecedents*. London: Routledge/Croom Helm.
- Austin, John L. (1962) *How to do things with words*. Cambridge, Mass.: Harvard U. Press.
- Bock, Kathryn (1986) Syntactic persistence in language production. *Cognitive psychology* 18: 355-387.
- Bock, Kathryn, & Helga Loebell (1990) Framing sentences. *Cognition* 35: 1-39.
- Chafe, Wallace (1994) *Discourse, consciousness and time*. Chicago: U. of Chicago Press.
- Chafe, Wallace (1976) Givenness, contrastiveness, definiteness, subjects, topics and point of view. In C. N. Li (Ed.) *Subject and topic*, pp. 25-55. New York: Academic Press.
- Chomsky, Noam (1971) Deep structure, surface structure and semantic interpretation. In D. Steinberg & L. Jakobovits (Eds) *Semantics*, pp. 183-216. Cambridge: Cambridge U. Press.
- Cumberland, Linda A. (2006) Motion verbs. Read at the Comparative Siouan Workshop, Billings, Mont. (19 Jun.)
- Cumberland, Linda A. (2005) *A grammar of Assiniboine: A Siouan language of the Northern Plains*. (Indiana University dissertation)
- Dorsey, James Owen (1890) *The Cegiha language: Contributions to North American ethnology* 4. Washington: US Department of the Interior, Government Printing Office.
- Erguvanlı-Taylan, Eser (1984) *The function of word order in Turkish grammar*, pp. 43-71. Berkeley: University of California Press. (originally published as Erguvanlı 1984).
- Erkü, Feride (1983) *Discourse pragmatics & word order in Turkish*. (U. of Minn. dissertation)
- Eschenberg, Ardis (2005a) *The article system of Umo^hoⁿ*. (SUNY Buffalo dissertation)
- Eschenberg, Ardis (2005b) The development of evidentiality in Omaha. Read at ICHL, Madison, Wisc. (5 Aug.)
- Feist, Jim (2007) Information structure in English nominal phrases. Read at Mid-America Linguistics Conference, Lawrence, Kans. (28 Oct.)

¹² Thanks to Gregory Ward, Betty Birner and Jeanette Gundel; to MALC attendees in Lawrence and to Kevin Schluter, Hiroki Nomoto, Sara Schmelzer, Paula Chesley, Mark Wicklund, Ellen Lucast, Kaitlin Johnson, Kateryna Kent and Norma Mendoza-Denton for their suggestions; and to Jordan Niebur, Roland Noear, the Niebur, Peniska and Williams families, and the Páⁿka and Umóⁿhoⁿ people.

- Foley, William A. (1999) Grammatical relations, information structure and constituency in Watam. *Oceanic linguistics* 38: 115-138. (Jun.)
- Frascarelli, Mara (2000) *The syntax-phonology interface in focus and topic constructions in Italian*. Dordrecht: Kluwer Academic.
- Good, Jeff (2008) An information structure template in Bantoid. Read at LSA, Chicago (8 Jan.)
- Gordon, Bryan James (2008) Postposition and information structure in an SOV language: Theoretical implications. Read at LSA, Chicago (7 Jan.)
- Grice, H. Paul (1975) Logic in conversation. In P. Cole & J. L. Morgan (Eds) *Speech acts: Syntax and semantics* 3, pp. 41-58. New York: Academic Press.
- Gundel, Jeanette K. (2003) Information structure: How much belongs in the grammar? Read at International Conference on HPSG, East Lansing, Mich. (18-20 Jul.)
- Gundel, Jeanette K. (1999) On different kinds of focus. In Peter Bosch & Rob van der Sandt (Eds) *Focus*, pp. 293-305. New York: Cambridge U. Press.
- Gundel, Jeanette K. (1988) Universals of topic-comment structure. In Hammond, Moravcsik & Wirth (Eds) *Studies in syntactic typology*, pp. 209-239. Amsterdam: John Benjamins.
- Gundel, Jeanette K. (1975) Left dislocation and the role of topic-comment in linguistic theory. *Ohio State working papers in linguistics* 18: 72-131.
- Gundel, Jeanette K. (1974) *The role of topic & comment in linguistic theory*. (U. of Texas diss.)
- Gundel, Jeanette K., Mamadou Bassène, Bryan James Gordon, Linda Humnick & Amel Khalfaoui (2007a) Reference & cognitive status: Results from corpus studies of Eegimaa, Kumyk, Ojibwe & Tunisian Arabic. Read at IPrA, Göteborg. (12 Jul.)
- Gundel, Jeanette K., Mamadou Bassène, Bryan James Gordon, Linda Humnick & Amel Khalfaoui (2007b) Coding protocol for statuses on the givenness hierarchy. (Working ms. unpub., updated Jul. 2007)
- Gundel, Jeanette K., Nancy Hedberg & Ron Zacharski (1993) Referring expressions in discourse. *Language* 69: 274-307.
- Gundel, Jeanette K., & Thorstein Fretheim (2004) Topic and focus. In Laurence Horn & Gregory Ward (Eds) *Handbook of pragmatics*, pp. 175-196. London: Blackwell.
- Gussenhoven, Carlos (2007) Types of focus in English. In Chungmin Lee, Matthew Gordon & Daniel B. Ring (Eds) *Topic and focus: Cross-linguistic perspectives on meaning and intonation*, pp. 83-100. Amsterdam: Springer.
- Guy, Gregory (forthcoming) Language variation & change: The temporal locus of variation. In G. Guy & Anna Grimshaw (Eds) *Language variation & linguistic theory*. Oxford: Blackwell.
- Hirschberg, Julia (1985) *A theory of scalar implicature*. (U. of Penn. dissertation)
- Jackson, Scott (2008) Stress, phases & information structure. Read at LSA, Chicago (7 Jan.)
- Johnson, Keith (1997) Speech perception without speaker normalization: An exemplar model. In Keith Johnson & John W. Mullenix (Eds) *Talker variability in speech processing*, pp. 145-165. San Diego: Academic Press.
- Kaiser, Lizanne (1999) Representing the structure-discourse iconicity of the Japanese post-verbal construction. In M. Darnell (Ed.) *Case studies: Functionalism vs. formalism in linguistics* 2, pp. 107-129. Amsterdam: John Benjamins.
- É. Kiss, Katalin (1998) Identificational focus versus information focus. *Language* 74: 245-273.
- Lee, Chungmin (2002) Contrastive topic and proposition structure. In Anne-Marie di Sciullo (Ed.) *Asymmetry in grammar*, pp. 345-368. Amsterdam: John Benjamins.
- Mendoza-Denton, Norma (2007) Linguistic variation and entrainment in micro-time. Plenary talk, Conference on Culture, Language and Social Practice, Boulder, Col. (7 Oct.)

- Mithun, Marianne (1999) *The languages of Native North America*. Cambridge: Cambridge Language Surveys.
- Mithun, Marianne (1991) The role of motivation in the emergence of grammatical categories: The grammaticization of subjects. In E. C. Traugott & B. Heine (Eds) *Approaches to grammaticalization 2*, pp. 159-184. Amsterdam: John Benjamins.
- Newmeyer, Frederick J. (forthcoming) On split CP's and the "perfectness" of language. In Benjamin Shaer, Philippa Cook, Werner Frey & Claudia Maienborn (Eds) *Dislocated elements in discourse: Syntactic, semantic and pragmatic perspectives*. London: Routledge.
- Newmeyer, Frederick J. (2004) On split-CP's, uninterpretable features and the "perfectness" of language. In Benjamin Shaer, Werner Frey & Claudia Maienborn (Eds) *ZAS working papers 35: Proceedings of the Dislocated Elements Workshop*, pp. 399-421. Berlin: ZAS.
- Newmeyer, Frederick J. (2003a) Against the split-CP hypothesis. In Piotr Bański & Adam Przepiórkowski (Eds) *Generative linguistics in Poland: Morphosyntactic investigations*, pp. 93-108. Warszawa: Instytut Podstaw Informatyki, Polskiej Akademii Nauk.
- Newmeyer, Frederick J. (2003b) Grammar is grammar & usage is usage. *Language* 79: 682-707.
- Quintero, Carolyn (2004) *Osage grammar*. Lincoln, Neb.: University of Nebraska Press.
- Rankin, Robert L. (2005) Quapaw. In Heather K. Hardy (Ed.) *Native languages of the southeastern United States*, pp. 454-498. Lincoln, Neb.: U. of Neb. Press.
- Rizzi, Luigi (2000) The fine structure of the left periphery. In Luigi Rizzi. *Comparative syntax and language acquisition*, pp. 241-296. London: Routledge. (Originally published 1997).
- Rudin, Catherine (1998) Postverbal constituents in Omaha-Ponca. Read at Siouan and Caddoan Languages Conference, Bloomington, Ind. (Jun.)
- Saeed, John Ibrahim (1984) *The syntax of focus and topic in Somali*. Hamburg: Buske.
- Schegloff, Emanuel A. (1987) Analyzing single episodes of interaction: An exercise in Conversation Analysis. *Social psychology quarterly* 50: 101-114.
- Shea, Kathleen (2007) Another look at the role of the definite articles in (Omaha-)Ponca. Read at Mid-America Linguistics Conference, Lawrence, Kans. (28 Oct.)
- Shibatani, Masayoshi (2008) Focus constructions without focus morphology. Read at LSA, Chicago (8 Jan.)
- Sneed German, James, & Janet Pierrehumbert (2008) Conditions for accenting pronouns: Contrastive focus vs. attentional shift. Read at LSA, Chicago (8 Jan.)
- Vallduví, Enric, & Maria Vilkuna (1998) On rheme and kontrast. In Peter W. Culicover & Louise McNally (Eds) *The limits of syntax*, pp. 79-108. San Diego: Academic Press.
- Ward, Gregory (1985) *The semantics and pragmatics of preposing*. (U. of Penn. dissertation)
- Weiner, E. Judith, & William Labov (1983) Constraints on the agentless passive. *Journal of linguistics* 19: 29-58.
- Zimmermann, Ilse (1999) Die Integration topikalischer DPs in die syntaktische und semantische Struktur von Sätzen. In Monika Doherty (Ed.) *Studia grammatica 47: Sprachspezifische Aspekte der Informationsverteilung*, pp. 41-60. Berlin: Akademie Verlag.

Author contact information:

Bryan James Gordon: linguista@gmail.com