Information structure, variation and the Referential Hierarchy

Jane Simpson

Australian National University, Canberra

Silverstein (1976)’s hierarchy of features and ergativity (Referential Hierarchy) was proposed to capture apparent systematic variation with respect to word-class (pronouns versus nouns) in the expression of the grammatical functions Subject and Object and the semantic roles Agent and Undergoer linked to these functions. An assumption of the original hierarchy was obligatoriness of marking, rather than optionality (i.e. choice of marker or its absence). Optionality is often associated with a semantic/pragmatic force additional to straight expression of grammatical function. This additional meaning may determine reanalysis and subsequent change in the morphosyntactic expression of Subject/Object/Agent/Undergoer. Along the way, apparent counter-examples to the Referential Hierarchy may be created. To understand the counter-examples, and test the descriptive adequacy of the Referential Hierarchy, better language documentation is needed.

1. INTRODUCTION1. In the 1970s Michael Silverstein made a proposal about the linguistic means of expressing arguments and the accessibility of arguments for other processes. This was the hierarchy of features and ergativity (later called the Referential Hierarchy) (Silverstein 1976). It engendered considerable typological work as linguists sought to test the claims cross-linguistically, both within languages which they were researching and across languages. Grammars were the main source of publicly available data on many languages, although a few authors, especially those influenced by Franz Boas’s approach (cf. Voegelin & Hymes 1953), also published text collections and dictionaries. The language data was primarily available as transcribed texts without accompanying audio and had to be searched by hand rather than digitally. Thus testability of the predictions of the hierarchies was limited to what could be found by hand in the available data. Modern language documentation has expanded the number of languages that can be searched, as well as the range of language documentation (more texts and dictionaries and linked multimedia) and the ways in which language documentation can be searched (different search possibilities along with greater reliability).

1 I thank for their comments the participants in the workshop on Potentials of Language Documentation (November 3–4, 2011), Max Planck Institute for Evolutionary Anthropology, Leipzig, Frank Seifart and Geoffrey Haig, and an anonymous referee.
In this paper, I illustrate the need for electronically searchable text corpora linked to sound with respect to understanding how changes in the case-marking of pronouns in the Australian language Arrernte could create a counter-example to the Referential Hierarchy.

2. **BACKGROUND: THE REFERENTIAL HIERARCHY.** Silverstein (1976)’s hierarchy of features and ergativity (Referential Hierarchy) was proposed to capture apparent systematic variation in the morphosyntactic expression of the grammatical functions Subject and Object and the linked semantic roles Agent and Patient with respect to word-class (pronouns versus nouns).

\[
\begin{align*}
1/2 \text{pron} & > 3 \text{pron} > \text{proper N} > \text{human N} > \text{animate N} > \text{other} \\
\text{Agent Subject} & & \text{Undergoer Object}\end{align*}
\]

It was intended to capture two aspects of the linguistic expression of Subject and Object arguments relating to case and verb agreement respectively. On Case, Silverstein proposed that the higher an element is on the hierarchy, the less likely it is to have Ergative case if it is acting as an Agent Subject, and conversely that the lower an element is on the hierarchy, the less likely it is to have Accusative case if it is acting as an Undergoer Object. In terms of verb agreement, Silverstein proposed that the lower an element is on the hierarchy, the more likely there is to be special verb agreement such as inverse marking if it is acting as an Agent Subject and if the Undergoer Object is higher on the hierarchy. Thus it is less marked for first person to act on third person than vice versa, and the more marked situation may be expressed by use of special inversion markers or reversal of agreement marker order.

Silverstein observed that the Australian language, Arrernte was a counter-example to his proposal that first person would be less likely to be marked Ergative than second or third person. Table 1 gives the modern Arrernte pronouns, showing that nouns and first person singular follow an Ergative-Absolutive pattern, while plural pronouns and second and third person singular pronouns follow a Nominative-Accusative pattern. The odd behavior of first person singular pronouns is striking.

Since the Referential Hierarchy reflects strong statistical tendencies, a counter-example such as this needs discussion. One type of explanation lies in their historical origins – what were the pressures in the system that allowed them to emerge?

At some time there must have been a bridging context, that is, where two variants of a form arose in slightly different meaning contexts (semantic or pragmatic) to express the same idea (or part of the same idea) in different meaning-contexts. One form-variant+meaning pair followed the hierarchy in expressing grammatical relationships; the other was used in a context where the grammatical relations were downplayed (e.g. a case form was used to express a non-case grammatical concept, such as emphasis). At a later stage, the form-variant+meaning pair that expressed the grammatical relation stopped being

---

2 Abbreviations and conventions: 1, 2, 3 – 1st, 2nd, 3rd person; A – Subject of transitive verb; ABS – Absolute; ACC – Accusative; ALY – Alyawarr; ARR – Arrernte; C – underspecified consonant; CV – sequence of consonant and vowel; ERG – Ergative; IS – Subject of intransitive verb; N – noun; NOM – Nominative; NPST – non-past tense; O – Object; PST – Past Tense; PREA – pre-Arandic; PRS – present tense; PRON – pronoun; S – Subject; SG – singular; V – verb, underspecified vowel; ‘-’ – clear morpheme boundary; ‘=’ – clitic boundary; ‘*’ – frozen morpheme boundary; ‘=*’ – reconstructed form.
used, leaving the other form-variant+meaning pair either on its own, or to expand to cover both contexts. This then created a violation of the hierarchy.

In order to see how statistically unlikely systems arise, we need to be able to observe the variation at an earlier stage of the language, or, failing that, in a genetically and typologically similar language. The main sources for such observations are grammars and text corpora.

Grammar writers, especially those writing grammars of small languages of which they are not native speakers, naturally focus on the statistically most prevalent forms, and not on variants. As well, the descriptive capability of grammar-writers increases as we learn more about the expressive capabilities of languages. Topics such as information structure are treated in much more detail in modern grammars than in grammars even of the 1980s. So, the grammarian may not have made the needed observations of variation. Therefore, it is essential to have large balanced text corpora in order to observe the variation, so as to determine what the bridging contexts are.

3. **Background: Information Structure.** To understand the situation that gives rise to counter-examples to case-expression such as the Arrernte example, it is helpful to look at counter-examples to verb agreement, the other means of expression of argument relation that Silverstein linked to the hierarchy. These have been explored typologically (Bickel 2008, Filimonova 2005), and in general such counter-examples seem to be more numerous than the counter-examples to case-expression. Bickel observes that the Referential Hierarchy may work differently for case-expression and verb agreement because “these two ways of marking arguments have a fundamentally different relationship to referential distinctions”.

Bickel’s point is important. A hierarchy which has pronouns at one end and nouns at the other is a hierarchy of information structure, of how the Speaker organizes information about reference for the Hearer. **Pronouns** assume that the Hearer knows who is being referred to and are likely to be used for pragmatic functions such as continuing topic: *John came in and he sat down*. They may also be used for contrast: *It was him that did it, not me*, but very often languages have special means for marking these (as strong pronouns as opposed to the weak pronouns of continuing topic). **Proper names** are allegedly unique identifiers, but their use also assumes that the Hearer knows in general who is being referred to (*John* is a proper name, but in the sentence *John came in and he sat down*, it is assumed that the Hearer knows which *John* is being referred to). **Nouns** are more likely to be used for

<table>
<thead>
<tr>
<th><strong>Arrernte</strong></th>
<th><strong>Agent Subject</strong></th>
<th><strong>Intransitive V Subject</strong></th>
<th><strong>Undergoer Object</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 SG</td>
<td><em>(a)the</em></td>
<td><em>ayenge</em></td>
<td><em>ayenge</em></td>
</tr>
<tr>
<td>2 SG</td>
<td><em>unte, nge</em></td>
<td><em>unte, nge</em></td>
<td><em>nge-nhe</em></td>
</tr>
<tr>
<td>3 SG</td>
<td><em>re</em></td>
<td><em>re</em></td>
<td><em>re-nhe</em></td>
</tr>
<tr>
<td>Plural Pronouns</td>
<td><em>A/IS</em></td>
<td><em>A/IS</em></td>
<td><em>Onhe</em></td>
</tr>
<tr>
<td>Nouns</td>
<td>Ergative affix</td>
<td>unmarked</td>
<td>unmarked</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>(Absolutive)</em></td>
<td><em>(Absolutive)</em></td>
</tr>
</tbody>
</table>

**Table 1: Modern Arrernte singular pronouns**
new information, switch of referent for subject or object. Thus the order on the Referential Hierarchy reflects an ordering of information structure roles from those that the Speaker expects the Hearer to recognise, to those that the Speaker expects to be new to the Hearer.

‘Verb agreement’ relates to this hierarchy of information structure because in many languages ‘verb agreement’ is not just agreement with a co-present noun; it is the linguistic means of expression for referents. Consider the following two examples from the Central Australian language Warlpiri:

(1) Warlpiri [v agreement only]

\[
\begin{align*}
\text{Wangka-mi} & \quad \text{ka=RNA} \\
\text{talk-NPST} & \quad \text{PRS=1S} \\
& \quad \text{‘I am talking.’}
\end{align*}
\]

(2) Warlpiri [v agreement + pronoun]

\[
\begin{align*}
\text{Ngaju} & \quad \text{ka=RNA} \quad \text{wangka-mi} \\
\text{1S} & \quad \text{PRS=1S} \quad \text{talk-NPST} \\
& \quad \text{‘I am talking. / It’s me that’s talking.’}
\end{align*}
\]

The ‘agreement marker’ can express continuing topic on its own as well as agreeing with another element. In this respect, verb agreement lines up with ‘weak’ pronouns as opposed to ‘strong pronouns’. Therefore verb agreement may actually be an element on the hierarchy, rather than something the choice of which expresses the hierarchy – like case. That is, the elements of the Referential Hierarchy are actually part of the information structure resources of a language.

The information structure resources of a language involve many linguistic means of expression other than part of speech and morphological properties as represented in the Referential Hierarchy. A Hearer’s attention may be directed to a referent via other morphemes, e.g. clitics and demonstratives, or by word order, by prosodic means such as pauses and intonation changes, and by non-verbal means such as gesture and eye-gaze. It follows then that to understand putative counter-examples to the proposed typological universal of the Referential Hierarchy, we need to know how information structure as a whole works in the language.

Labelling information structure functions is hard to do precisely and reliably. To see what the information flow is, and to understand why something is considered old or new information or contrastive, we need continuous text, not just example sentences. To understand the linguistic means of expression of information structure, we need a corpus which covers conversational speech and dialogue talk as well as the monologues, narratives, descriptive texts, and procedural texts that make up most of the traditional text collections accompanying grammars. A corpus which contains transcriptions of text and intonation, and is linked to audio-visual files, is essential for providing empirical support for claims about information structure. Claims about the linguistic expression of hanging topics, contrastive topics, afterthoughts, and emphasis need to be backed up by reference to examples from such a corpus, and by quantifying the use of particular means of expression. Corpora produced
before the software became available for linking audio-visual files to text usually lack the fine-grained representations of intonation so necessary for understanding how information structure is expressed.

So, in sum, elements of the Referential Hierarchy (parts of speech/morphological properties) are part of the means speakers have for expressing information structure. How they operate within a language must be seen in the light of all means of structuring information in the language. For example, intonation often provides cues to information structure, and so audio-visual information is needed, not just transcripts.

4. ARRERNTE AS A COUNTER-EXAMPLE. With this as background, we return to the putative counter-example from Arrernte. When finding something that at first glance seems to run counter to a fairly well established typological universal such as the Referential Hierarchy, the first step is to establish how this counter-example fits into the patterns found in the neighboring and genetically close languages. Arrernte is a member of the Arandic language family. Examination of the neighboring, and closely genetically related, language, Alyawarr, reveals that it differs in having a three-way split for both first and second person singular (Wilkins 1989). This split is reconstructed for pre-Arandic as well (Koch 2004).

<table>
<thead>
<tr>
<th></th>
<th>Agent (Subject)</th>
<th>Intransitive V (Subject)</th>
<th>Undergoer (Object)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 SG ARR</td>
<td>(a)the</td>
<td>aye.nge</td>
<td>aye.nge</td>
</tr>
<tr>
<td>A Ly</td>
<td>ahe</td>
<td>aye-nge</td>
<td>aye-nhe</td>
</tr>
<tr>
<td>*PREA</td>
<td>*ngathu</td>
<td>*ngay(V)-nge</td>
<td>ngay(V)-hna</td>
</tr>
<tr>
<td>2 SG ARR</td>
<td>unte, nge</td>
<td>unte, nge</td>
<td>nge-nhe</td>
</tr>
<tr>
<td>A Ly</td>
<td>ntwe</td>
<td>nge</td>
<td>nge-nhe</td>
</tr>
<tr>
<td>*PREA</td>
<td>*nyuntu</td>
<td>*nyun.-nge</td>
<td>i.nge-nhe</td>
</tr>
<tr>
<td>3 SG ARR</td>
<td>re</td>
<td>re</td>
<td>re-nhe</td>
</tr>
<tr>
<td>A Ly</td>
<td>re</td>
<td>re</td>
<td>re-nhe</td>
</tr>
</tbody>
</table>
| *PREA | *CV.re | *CV.re | *CV.re-nhe

Table 2: Modern Arrernte and Alyawarr compared with Koch’s reconstructions for an earlier stage of Arandic

In Arrernte, Alyawarr, and pre-Arandic, -nhe (*-nha) represents an Accusative case-marker. In Alyawarr and pre-Arandic there is some evidence for a morpheme -nge on intransitive subjects, but it appears to be frozen in Arrernte.

From the point of view of the Referential Hierarchy, the pre-Arandic Ergative-Absolutive-Accusative of the Speaker-Hearer singular pronouns and the Ergative-Absolutive of the nominals would be not unexpected (and see Haig 2008 for a similar pattern in Vafsi). What is unexpected is the proposed break in the hierarchy through the third person Nominative-Accusative pattern, which retains the marked Accusative object. Since the evidence for the absence of Ergative-marking on the pre-Arandic reconstructed pronoun is basically absence of evidence for its presence, we leave this and focus on the changes in the modern systems.

CV stands for a reconstruction of an indeterminate C followed by an indeterminate V.
In sum, the three-way pre-Arandic system for Speaker/Hearer singular pronouns has been retained in Alyawarr but in Arrernte has moved to two different two-way systems: a two-term Ergative-Absolutive for first person singular (creating an unmarked Object), and, for second person singular, a three-term Nominative-Accusative system with two different terms for the Nominative.

The change appears to have taken place in two stages. The collapse of the second person appears to have been relatively recent (Wilkins 1989), who cites evidence from an early recorder, R. H. Mathews:

“There are two district [sic] forms of the first person in the singular number, namely, ta and yinga. Ta is always used when connected with a transitive verb…Yinga is employed when connected with an intransitive verb…In the second person singular there are also two forms of the pronoun [sic] – unta for use with transitive verbs, and nga with intransitive verbs.” (Mathews 1907: 325)

That is, in the late nineteenth century, when Mathews carried out work on Arrernte, both first and second person singular had Ergative forms. So how did a three-way pre-Arandic system change to the modern system with the discontinuity between first person singular (ERG-ABS), other pronouns (NOM-ACC), and nouns (ERG-ABS)? There must be bridging contexts which allow the change from one meaning to another.

5. BRIDGING CONTEXT: OPTIONAL ERGATIVE MARKING. I propose that the likely bridging context for the second person change is optional Ergative marking. When the hierarchy was first proposed, little attention was paid to the possible optionality of marking for case. The situations discussed were those of obligatory marking. But it has become clear that in many Australian languages, Ergative case marking is optional (McGregor & Verstraete 2010). Verstraete shows that in Umphithamu Ergative marking is obligatory for inanimate transitive subjects, but optional for animate transitive subjects (as predicted by the Referential Hierarchy), and that contrast and answers to questions are likely to favour the use of Ergative. Thus, Ergative case-marking has several interrelated associations: indicating the grammatical function of subject, indicating the semantic roles of agent, and identifying referents that the hearer needs to be alerted to, most likely as new information.

At any time a speech community may highlight one or other of these associations. If the grammatical function of Subject is highlighted, then Ergative may be obligatory for marking the subject of transitive clauses. If the semantic role is highlighted, then there may be splits in how the subjects of transitive clauses are marked (e.g. using Absolutive for the subjects of perception verbs). If discourse function is highlighted, then subjects of transitive verbs may have Ergative marking in certain non-prominent discourse contexts, or lack it in positions that are otherwise discourse-prominent (because Ergative marking would be redundant).

An association between optional Ergative marker and discourse prominence is found in the neighboring language Warlpiri (which is related to Arrernte, but not closely). In traditional Warlpiri spoken in Yuendumu, Ergative marking is optional on first and second person singular pronouns in initial position. Elsewhere (demonstratives, nouns), Ergative marking appears to be obligatory (Mary Laughren, p.c.).
Initial position in Warlpiri is a position where discourse prominent elements are placed (Laughren 2002, Legate 2002, Simpson 2007). This association of lack of Ergative marking with a discourse prominent position has several possible explanations. One is that the word order is changing so that Subject (whether Agent or Intransitive V Subject) is becoming fixed in initial position, and so only Subjects in non-initial position need Ergative marking to identify their function. But this needs testing to find out what the discourse contexts are that favor Ngaju V over Ngajuuru V and V ngajuuru. Unfortunately this is hard to do. The Warlpiri corpus (which was transcribed or composed between 1959 and the early 1990s, and has been the basis for the Warlpiri Dictionary) is large by Australian standards, containing at least 600,000 words. But it has certain gaps. There is limited meta-data on speakers. There is as yet almost no interlinear glossing and no linking to sound. The genres include elicitation, mythological and personal narratives, meta-linguistic discussion, and written texts, but no natural conversation.

Recent changes in Warlpiri further support the association of position and Ergative case. In Lajamanu, a Warlpiri community to the north of Yuendumu, a new mixed language, Light Warlpiri, has developed, based on the interaction of Warlpiri and Kriol (O’Shannessy 2006). This language has Ergative marking, but it has become generally more optional than in traditional Warlpiri. Using spontaneous texts, and comprehension and production tasks, O’Shannessy has shown that Ergative marking occurs 60% or so of the times when it could be expected on nouns in traditional Warlpiri, where it is categorical. Ergative marking is more likely when subjects of transitive clauses appear post-verbally (the language is moving to SVO order), and on unusual Agents (i.e. when the hierarchy is inverted, that is, when Subjects of transitive verbs are inanimate and Objects are animate).

The fact that Ergative marking appears post-verbally fits with its optional absence in traditional Warlpiri on initial first person pronouns. There is less need to mark elements in initial position as being discourse prominent because the initial position itself indicates discourse prominence. By contrast, post-verbal position is less prominent, and so a post-verbal expression which represents an Agent is more likely to receive special marking to indicate that it is expressing an Agent. Thus in Light Warlpiri, Ergative marking has shifted from being mostly a marker of grammatical function to being a marker of discourse prominence or of inverted expectations (cf. Umpithamu (Verstraete 2010), where inanimates must receive Ergative case, but animates are optionally marked with Ergative).

---

4 ngaju ‘Absolutive’; ngaju-rlu ‘Ergative’, where -rlu is a common allomorph of Ergative
The end state (=current state) of Light Warlpiri is well understood from O’Shannessy’s description (although she notes the difficulty of operationalizing pragmatic factors such as contrast and continuing topic for testing their relevance to the choice of Ergative case marking in comprehension and production tests). However, the transitional states and the reasons for change cannot be tested properly since we cannot easily find answers without time-consuming hand-coding of the existing corpus.

Having established that, in a neighboring language, Ergative case-marking may change to being associated with discourse prominence when this is not marked by other means, such as word order, we return to Arrernte. Wilkins writes:

“The two forms [unte and nge], are basically equivalent in meaning, although, of the two second singular pronouns, nge may be considered the unstressed and less emphatic form.” (Wilkins 1989: 125)

Thus, in his view, the second singular formunte (the old Ergative) is more associated with discourse prominence (like a strong pronoun, and like the use of Ergative in Light Warlpiri), while the second singular form nge (the old Absolutive) is more associated with continuing topic (like a weak pronoun).

We can posit the following sequence of changes:

- **Pre-Arandic**: Ergative only used to mark grammatical functions: three-way split in first and second person with -nhe as Accusative on all singular pronouns.
- **Pre-modern Arrernte** [Mparntwe Arrernte 1907]
  - 1/2 singular pronouns athe/unte ERG, aye.nge/nge ABS, (Mathews does not give the forms for Object)
- **Mparntwe Arrernte** 1989
  - 1 singular pronoun athe ERG, ayenge ABS
  - 2 singular pronoun: 2 pronouns representing NOMINATIVE: unte Discourse-prominent, nge Continuing topic
- **Proposed direction**:
  - Pre-Arandic → pre-Modern Arrernte
    Accusative form of 1st person is lost, possibly by conflation of the Accusative marker -nhe, with the augment -nge of the form used for Subject of Intransitive verb. This resulted in an Ergative-Absolutive system for 1st person singular (cf. Koch 2004).
  - pre-Modern Arrernte → Modern Arrernte
    Ergative form of 2nd person singular unte moves from marking grammatical function to marking discourse function of emphasis. The difference between unte (Subject of transitive verb) and nge (Subject of intransitive verb) is reanalyzed as a difference between unte as discourse prominent Subject and nge as continuing topic Subject, regardless of whether the verb is transitive or intransitive. This results in a
Nominative-Absolutive system for 2nd person singular parallel to that of 3rd person singular.

The explanation of the changes in second person requires finding out more about the discourse function of Ergative-marked elements. For its time, Wilkins' grammar is rich in including 12 complex interlinear texts (approximately 360 clauses) covering instruction, morality fables, mythological story, description of natural phenomena, and co-constructed written texts. But even despite the variety of genres, it is hard to find examples. The second person nge does not occur at all in his texts, and the second person unte only occurs twice in an instructional text (“If you have a cold then you inhale…”) and four times in polite commands presented as reported speech in a story. This is not enough to test patterns.

6. CONCLUSION. Putative counter-examples to proposed typological universals can often be understood in part by considering paths of grammatical change. But to understand and analyze how languages change from one typological state to another, we need more than just grammars, even rich ones such as Wilkins’ grammar, because they inevitably reflect the descriptive preoccupation of the time. Information structure is an area of the language whose importance in grammatical description only became widely accepted in the 1990s. Grammars written before that time frequently lack a discussion of information structure; for example, they may tell the reader how to ask a question (because that involves special pronouns or word order or clitics) – but not how to answer it (and so they lack examples of the clearest means of expressing new information).

We need, in addition, large corpora of texts linked to audio-visual recordings to observe the variation and determine the bridging contexts that lead to the emergence of counter-examples. Ideally, these would cover a wide range of genres and range of speakers, be annotated for pragmatic factors, and would be refreshed over time.

REFERENCES


Jane Simpson
jane.simpson@anu.edu.au