# Workshop on Information structure, referential status and referent type in discourse and grammar

27-28 June 2018

## 27 JUNE

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:00</td>
<td>COFFEE / RECEPTION (W1.13)</td>
</tr>
</tbody>
</table>

### Introduction / Annotation session

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:45</td>
<td>Stefan Schnell, Ruth Singer &amp; Eva Schultze-Berndt (U Melbourne/U Manchester) Introduction to the workshop, and to annotation schemes for referential expression (based on GRAID), and relevant theoretical decisions and applications</td>
</tr>
</tbody>
</table>

### Factors conditioning the choice of referential expressions

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:30</td>
<td>Ruth Singer, Peter Hurst, Sasha Wilmoth, Jason Johnston &amp; Rebecca Defina (U Melbourne) Exploring the use of nominal classification systems in discourse: animacy, grammatical function and referent expression in four languages</td>
</tr>
<tr>
<td>12:15</td>
<td>Stefan Schnell (U Melbourne) Conditions on object pronouns and object indexing</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>13:00</td>
<td>LUNCH</td>
</tr>
<tr>
<td>14:00</td>
<td>Adam Schembri, Jordan Fenlon &amp; Kearsy Cormier (U Birmingham/UCL) Variable argument expression and indicating verbs in British Sign Language</td>
</tr>
<tr>
<td>14:45</td>
<td>Gabrielle Hodge, Lindsay N. Ferrara &amp; Benjamin Anible (UCL) The semiotic diversity of doing reference in a deaf signed language</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>15:30</td>
<td>COFFEE</td>
</tr>
</tbody>
</table>

### The status of noun phrases: referential or predicative?

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>16:15</td>
<td>Delia Bentley (U Manchester) Information structure and the subject vs. predication partition: A case study in Gallo-Italian</td>
</tr>
<tr>
<td>17:00</td>
<td>Wout van Praet &amp; Tine Breban (U Namur/U Manchester) The referential status of the noun phrase in English copular constructions</td>
</tr>
</tbody>
</table>
28 JUNE

Identification of topics

9:00 Dina El Zarka (U Graz)  
Factors influencing the prosody of thematic constituents in Egyptian Arabic

9:45 Karolina Grzech (SOAS / U Stockholm)  
Distributional properties of the ‘topic’ marker =ga in Amazonian Kichwa

10:30  

11:15 Julio Villa-García (U Manchester)  
On information structure and different preverbal positions for the subject

Interaction of referential status and negation

12:00 Jessica Iubini-Hampton (U Manchester)  
Negative markers and their interaction with information structure: the case of Modense

12:45 LUNCH

Referential status, information structure, and differential agent marking

14:00 Gertraud Schneider-Blum (U Köln)  
Reference Tracking, Information structure and Split Ergative Marking in Tima

14:45 Charlotte Hemmings (U Oxford)  
Information Structure and Differential Actor Marking in Kelabit

15:30 COFFEE

16:00 Eva Schulze-Berndt (U Manchester)  
Competing motivations in differential agent marking: the (ergative) case of Jaminjung

16:45 FINAL DISCUSSION
Exploring the use of nominal classification systems in discourse: animacy, grammatical function and referent expression in four languages

Ruth Singer\textsuperscript{1}, Peter Hurst\textsuperscript{2}, Sasha Wilmoth\textsuperscript{3}, Jason Johnston\textsuperscript{4} & Rebecca Defina\textsuperscript{5}

Nominal classification systems are found in the majority of the world’s languages, yet there is no agreement as to whether these systems have a function (Singer 2018) or are just ‘junk’; some kind of historical accretion (Trudgill 2011). In this study we look at four languages with very different nominal classification systems, tracing how their classes are employed in referent expressions in four corpora annotated with the GRAID system (Haig and Schnell 2014, Haig and Schnell (eds.) 2016). We also look at measures of lexicality and referential density (Haig and Schnell 2016, Bickel 2003, Bickel and Stoll 2010), comparing the extent and nature of referent expression across the languages. While it is well known that animacy and grammatical function influences the expression of nominal classification for example in clause-level agreement, we find that the expression of nominal classification within NPs is also sensitive to animacy and grammatical function. Features that emerge as correlates of referent expression in discourse such as animacy, grammatical function and person also constrain the expression of nominal classification in many languages which suggests that more detailed studies of the use of nominal classification in discourse can show us more about its function.

References

\textsuperscript{1} University of Melbourne and Australian National University
\textsuperscript{2} University of Melbourne
\textsuperscript{3} University of Melbourne/Appen\textsuperscript{TM}
\textsuperscript{4} Appen\textsuperscript{TM}
\textsuperscript{5} University of Melbourne
Conditions on object pronouns and object indexing

Since Moravscik (1978) it has been repeatedly confirmed that argument indexing systems are typologically asymmetrical: indexing is more often attested with subjects than with objects (Siewierska 2013), and object indexing – but not subject indexing – is typically conditioned, and thus never fully grammaticalised (Siewierska 1999; Haig fc2018). Object indexing is in most cases indexing-related DOM (Iemmolo & Klumpp 2014; Iemmolo 2010; Aissen 2003). The factors involved commonly pertain to saliency features of objects, such as person, animacy, definiteness, individuatedness, specificity (Siewierska 2004, Croft 2003, Fedden et al 2013, a.m.o.) and factors pertaining to topicality (Dalrymple & Nikolaeva 2011, Iemmolo & Klumpp 2014).

While it is often believed that argument-indexing systems arise through a grammaticalisation process that turns free personal pronouns into person indexes / agreement markers (Siewierska 2004, Givón 2015, 1976), the use of pronouns as opposed to a full NP or zero anaphor is typically attributed to discourse factors like accessibility / activation (Ariel 2014[1990], Chafe 1994, 1976; Gundel et al 1993), or topic continuity (Givón 1983). However, a number of corpus-based studies into the use of object pronouns suggest that the factors governing the use of object pronouns are essentially similar to those behind indexing-related DOM: for instance, based on variationist corpus studies of Spanish and Portuguese varieties of South America, Schwenter (2014, 2006) attributes the use of object pronouns to animacy and specificity which also govern case-marking DOM in Spanish. Similarly, Genetti & Crain (2013) find that in narrative discourse from Nepali, the use of object pronouns is restricted to human referents. Given that objects are relatively rarely human, they induce an underlying constraint on object realisation, Avoid pronominal P.

In my talk I present data and findings from recent corpus studies of Vera’a (Oceanic; Schnell & Barth 2018) and across corpora from different languages that are more suggestive of pronoun use in objects to be mainly a matter of saliency rather than discourse factors. An interesting finding from the Vera’a study is that the overall strongest predictor is not animacy / humanness, but global discourse topicality (what a text / passage is mainly about), a finding that can only be established once extended and varied discourse data is taken into account. More generally, it seems that the conditions on object indexing are not so much the result of a truncated putative grammaticalisation process, as suggested by Dalrymple & Nikolaeva (2011) or Haig (fc2018), but may already be a matter of referential choice in discourse. Any grammaticalisation process would thus involve primarily a change in position and morphological boundedness of relevant person markers, but no significant changes at the semantic and/or pragmatic level. This seems to make the grammaticalisation of object indexing essentially different from that often claimed to underlie subject agreement, where pronouns lose semantic and pragmatic distinctiveness when turning into cross-indexes, or agreement markers.
REFERENCES


Haig, Geoffrey. fc 2018. The grammaticalisation of object pronouns. Why Why Differential Object Indexing is the natural endpoint. *Linguistics* 53.4


Variable argument expression and indicating verbs in British Sign Language

Indicating verbs in British Sign Language (BSL), and other sign languages, can be directed towards locations in space associated with their arguments. This directionality appears to work primarily as a reference-tracking device, and is widely analysed as a form of agreement marking (e.g., Lillo-Martin & Meier, 2011). Data elicited from native signers in a number of sign languages appeared to suggest that such directionality was obligatory for the marking of object arguments (e.g., Morgan et al., 2006, for BSL). Work drawing on spontaneous data over the last decade has revealed, however, that there is in fact significant variation in the use of directionality. In a recently published major study, Fenlon et al. (2018) investigated a range of linguistic and social factors in 1,436 indicating verb tokens collected from conversational data in the BSL Corpus. The analysis shows that directionality in BSL is not obligatory (as had also been reported for the related variety, Auslan, see de Beuzeville et al., 2009), and is conditioned by several linguistic factors, such as constructed action (a type of enactment used to represent referents), animacy and co-reference. In this paper, we investigate the relationship between variable argument presence in the clause and the use of modification in indicating verbs. In the existing literature, it has been claimed by linguists working within the generative paradigm that argument drop is ‘licenced’ by the use of spatial modification of indicating verbs, as is also claimed for spoken languages with person agreement marking (Lillo-Martin, 1986; Kimmelman, 2018). To explore this question, the clauses in our BSL Corpus indicating verb dataset were coded for the presence of a subject or object argument noun phrase. We also coded the data for factors including co-reference (whether the subject or object argument had the same referent as arguments in the preceding clause), person (first, second or third), animacy (human, non-human animate, inanimate), the use of constructed action, and a number of macro-sociological participant characteristics (age, gender, region and native/non-native signer background). While our results confirm that the use of directionality in indicating verbs is indeed a significant factor, with null arguments more likely when spatial modifications of verbs are present, we find that the most important factor is co-reference: arguments which represent the same referents as in the previous clause (i.e., in which reference is maintained) are more likely to be null. We also find person to be significant, with non-first person subject arguments more likely to be dropped, while the reverse is true for object arguments. Lastly, we also find evidence for syntactic priming, with clauses with argument presence more likely to follow clauses with argument presence, even when co-reference is controlled for. We discuss the implication of our work for an understanding of variable argument presence and indicating verbs, as well as for cognitive/functional and variationist approaches to sign language linguistics in general.
Research suggests deaf signer’s choice of referring expression is primarily motivated by information management, i.e., lexicalised nouns and noun phrases are typically used to introduce new referents, whereas maintained and reintroduced referents are often indexed and/or depicted via pointing signs, indicating verbs, depicting signs and/or enactment (e.g. Morgan, 2006; Frederiksen & Mayberry, 2016). Signers may also create ‘invisible surrogates’, whereby a confluence of indexing actions enables interactants to conceptualize an entity as located in the signing space and behave as if it were present (Engberg-Pedersen, 1993; Liddell, 2003). For example, in Figure 1 an Auslan signer introduces agent frog into her retelling by: (a) fingerspelling and mouthing the English word “frog”; (b) signing the Auslan sign FROG while again mouthing the English word “frog”; (c) pointing with her right hand to an imagined frog located inside an imagined jar in front of her body, which she ‘holds’ by depicting this action with her left hand. In this way, the signer simultaneously depicts the visible actions of the boy holding the jar, while indexing the invisible frog and jar referents to the space in front of her body.

However, the interaction of these diverse semiotic strategies for doing reference, along with additional factors potentially influencing signer choices, have not yet been investigated using a signed language corpus. Using twenty retellings of Frog, Where Are You? and twenty retellings of The Boy Who Cried Wolf archived in the Auslan Corpus, we analysed 4,699 tokens of referring expressions with respect to: (a) activation status (introduced, maintained or reintroduced); (b) semiotic form(s) (e.g. lexical manual sign, English mouthing); and (c) animacy (humans, animals or inanimate objects). Statistical analysis confirmed choice of strategy was most strongly motivated by activation status: new referents were expressed with more conventionalised forms (especially lexical manual signs and English mouthing), whereas maintained and reintroduced referents typically involved fewer and more emergent semiotics. However, animacy was also a motivating factor: humans and animals were often depicted via visible surrogates (not pointing signs), whereas inanimate objects favoured depicting signs and invisible surrogates. We describe these patterns and suggest they may be explained by discourse topicality and signer preferences for embodying specific referents. These findings highlight the role of animacy in signed language discourse, while demonstrating the ‘pretend world’ indexicality of signed language use and the pluralistic complexity of face-to-face communication.
Figure 1 Introducing ‘frog’ (Auslan Corpus, PCNB2c7a_CLU#004)

References


Information structure and the subject vs. predication partition: A case study in Gallo-Italian

In a number of Gallo-Italian dialects of Italy, out-of-the-blue utterances exhibit VS order (if declarative and intransitive). While being focal, S can have the referential status identifiable, accessible or new. Relevant examples are given here.

(1)  
a. Gh’è rivà i to surèi. (Milan (Àffori), Lombard)  
LOC be.3SG arrived the your sisters  
‘Your sisters have arrived (here).’

b. Gh’è mort tant suldà.  
LOC be.3SG died many soldiers  
‘Many soldiers have died.’

c. (*Gh’è / *l’ a) An ciàmà i to gent.  
LOC be.3SG SCL have.3SG have.3PL called the your parents  
‘Your parents called (here/us).’

d. A la festa (*gh’è / *l’ a) an balà tant invità.  
At the party LOC be.3SG SCL have.3SG have.3PL danced many guests  
‘Many guests danced at the party.’

In addition to occurring in post-verbal position, which is not the unmarked position of the subject in these dialects (assuming that this is detected on the basis of SVO order in out-of-the-blue utterances with declarative transitive clauses), the post-verbal S may fail to control number agreement on the conjugated verb, for example, the perfect auxiliary ‘be’ in (1a-b). Cross-dialectal investigation has uncovered several patterns of V-S agreement (Bentley 2018), suggesting that this is sensitive to verb class, and hence to the thematic role, or macrorole, of S. In the dialect of Milan (Àffori), V-S agreement is only controlled by the core argument of intransitive activities. This argument is an Actor (cf. 1c-d).

In this presentation we address the following question: what is the effect of discourse on the grammatical role of S and on the subject vs. predication partition? Even though S behaves as an object (Lambrecht 2000) in occurring post-verbally, and, in some cases (cf. 1a-b), failing to control agreement, it does not pattern with objects in being replaceable by a nominative pronoun, which does control agreement.

(2) Pö in rivà lur. (Milan (Àffori), Lombard)  
then be.3PL arrived they  
‘Then they arrived.’

Observe also that the perfect auxiliary ought to be ‘have’, if the post-verbal argument were a transitive object. This is shown in the transitive VS structure in (3).
(3) Chel film li, l’ an vist tütü me amis / tanti person.
that film DEM OCL have.3PL seen all the my friends many people
‘That film, all my friends / many people have seen it.’

However, the perfect auxiliaries ‘be’ and ‘have’ alternate in intransitive VS constructions (cf. 1a-d) in the same way as in their SV counterparts.

In the absence of subject agreement, the construction under observation exhibits an etymologically locative proclitic hosted by the auxiliary (see ghe in 1a-b). If the verb is a verb of inherently directed motion, the etymologically locative clitic can receive a speaker-oriented interpretation (‘here’) (cf. 1a). In addition, this clitic takes the position of a subject auxiliary clitic (Poletto 2000, Poletto & Tortora 2016) and can co-occur with a locative phrase in preverbal position. While lacking the locative clitic, the VS constructions with V-S agreement normally require a locative or temporal phrase in preverbal position (cf. 1d). In the presentation we shall return to the very few exceptions to this requirement (cf. 1c).

Relying, on the one hand, on the notion of argument integration (Fuchs 1980, Sasse 2006, Bentley 2018), and on the other hand, on evidence suggesting that the VS structures under discussion have a Subject of Predication, which can be manifested by the locative clitic, or by a locative or temporal phrase (Saccon 1992, Tortora 1997, a.o., compare locative inversion in Bantu languages, Bresnan & Mchombo 1987), we claim that the post-verbal argument of (1a-d) is part of the predication. Since the utterance is discourse-new, the whole event described by the verb and its core argument is predicated of a Subject of Predication, which can be a location or a non-referential event argument (Bentley & Cruschina 2018, drawing on Bianchi 1993).

A common distinguishing criterion for English copular constructions is their communicative function, i.e. what the speaker means to convey by uttering the sentence. Thus, a distinction is made between predicative and specificational clauses (e.g. Declerck 1988; Keizer 1992; Mikkelsen 2005; Patten 2012), illustrated in (1) and (2) respectively.

(1) Becks is a working-class man from East London.
(2) Envelope please… and the winner is… Darryl Wakelin.

The function of a predicative clause is to give a description, e.g. a working-class man from East London in (1), of the subject referent, e.g. Becks. The function of a specificational clause is to specify which concrete instance(s), e.g. Darryl Wakelin in (2), correspond(s) to the criteria set up by the variable, e.g. the winner.

An important difference between the two construction types is the referential status of their constituents, a question that has stirred up quite a bit of disagreement in the literature. Some argue that both clause types have one referential and one nonreferential term (e.g. Mikkelsen 2005; Patten 2012): in predicative clauses, the nonreferential term functions as complement; in specificational clauses, the nonreferential term is the subject, or variable, e.g. the winner in (2). Others, however, point out that the difference between the two clause types is essentially that predicative clauses have a nonreferential complement (i.e. the description), but that specificational clauses have two referential terms (e.g. Heycock & Kroch 1999; Rothstein 2001). An intermediary position is taken by Declerck (1988) – and tentatively adopted by Keizer (1992) – who argue that the specificational variable is neither nonreferential nor fully referential, but instead only has ‘weak reference’.

In this paper we aim to elucidate the referential status of the copular NPs from a cognitive-functional perspective, which we argue provides a better account of the actual conceptual mechanisms behind the issue at hand. Following Langacker (1999), we will show that the predicative NP is nonreferential only in that it does not denote an actual instance but rather a ‘virtual’ one, “conjured up solely for the purpose of ascribing a property” and “has no mental status outside the mental space created to do so” (ib.: 273-274). Specificational variables, on the other hand, denote ‘generalised instances’, i.e. higher-order entities characterized by their generalising over potential spatio-temporal instances (i.e. ‘competitor values’). By abstracting over potential values, the variable’s generalised instance specify the ‘role’ that the actually specified value plays in the speech context (cf. Fauconnier 1994; Willemse 2005).

In the final analysis we argue that the differing referential statuses of the predicative complement vs. specificational variable are assigned by the clauses’ different argument structures (cf. Halliday 1985, 1994), as a result of the different orders of assembly in which the various layers of conceptual dependencies within the predicative vs. specificational constructions are combined. Predicative clauses code intransitive relations, in which only the subject referent is assigned a participant role: the predicative complement is construed as relational, as it combines with the copula to form a second-order predicate of which the subject/describee is the sole participant (Langacker 1991). Specificational clauses, on the other hand, code transitive relations, in which the variable and value both combine directly with the copula be, because of which both have a participant status in the specificational relation.
In sum, we propose that the referential statuses of predicative and specificational NPs are better accounted for by the cognitive-functional analysis given here. Not only does the analysis of predicative complements as virtual instances explain their ‘nonreferentiality’ more insightfully; but the explanation of specificational variables as generalised instances also provides a conceptually and descriptively more sound characterisation of their ‘weak reference’ that reconciles the various intuitions raised in the literature.

References
FACTORS INFLUENCING THE PROSODY OF THEMATIC CONSTITUENTS IN EGYPTIAN ARABIC

It has been proposed that topics/themes or their subtypes such as contrastive topics, shifted topics or continuous topics are coded by specific prosodic contours in various languages (Jackendoff 1972; Steedman 2000; Büring 2003; Frascarelli & Hinterhölzl 2007; Brunetti et al. 2010).

The incongruent results of the different studies as well as the massive variation found in spontaneous speech data (Hedberg and Sosa 2008) rather suggest that there is no one-to-one mapping between information structural categories and prosodic categories, but that the actual prosodic shape of an utterance and consequently an information structural constituent such as topic/theme is influenced by a number of semantic, pragmatic and formal factors. These factors put constraints on the prosodic realization of a linguistic utterance, which interact in a language-specific and speaker-specific way (Calhoun 2010).

To provide evidence for this claim, I will discuss some results from a corpus study of themes in Egyptian Arabic (El Zarka and Schuppler submitted). For this study, we annotated naturalistic conversational data for a number of factors (information status, contrast, syntactic and phonological factors) and submitted the annotated data to statistical analysis. Crucially, we not only investigated the individual factors that are expected to influence prosodic shape as separate independent variables, but we also examined their impact on different prosodic features (accentuation, phrasing, register, tonal features) by running logistic regression models for each of these dependent variables separately. The results of the study suggest that the individual "shaping" factors have an impact on different prosodic features, and to different degrees.

In the present talk, I will give a short overview of the most important results. The focus of the talk will be on the discussion of examples from the corpus consisting of thematic (i.e. utterance-initial) constituents referring to entity topics to illustrate the interaction between information structure, syntax and prosodic shape.

References:


El Zarka, Dina and Barbara Schuppler (submitted). On the interplay of formal and pragmatic factors in the prosodic realization of themes in Egyptian Arabic.

Frascarelli, Mara and Roland Hinterhölzl (2007). Types of Topics in German and Italian. In K. Schwabe & S. Winkler (Eds.), On Information Structure, Meaning and Form: Generalizations across Languages (pp. 87–116). Amsterdam, Netherlands: Benjamins.


Distributional properties of the 'topic' marker =ga in Amazonian Kichwa

Karolina Grzech, SOAS, University of London

This presentation describes the distributional properties of the enclitic =ga, as attested in Amazonian Kichwa (Quechuan, Ecuador). In other Quechuan varieties, the cognates of the enclitic are analysed as markers of topicality (Parker 1969; Cerrón-Palomino 1976; Cusihuamán 1976/2001; Muysken 1995; Sánchez 2010; 2015; Muntendam 2015). On the other hand, most grammatical descriptions (except Weber 1989) provide little information on the discourse and grammatical contexts in which the marker occurs. Typical examples involve =ga attaching to a referential NP, which leads the readers to believe that the use of =ga is a mirror-image of ‘aboutness topic’ marking in better-described languages (cf. e.g. Lambrecht 1994).

However, upon a closer look, we discover that certain aspects of distribution of =ga, such as its relatively low frequency, or the possibility of its occurrence on non-nominalised verbal hosts (cf. Grzech 2016a; 2016b) call this straightforward topic-marking analysis of =ga in Amazonian Kichwa into question. This presentation aims to propose a more comprehensive analysis of the enclitic.

On the basis of its distributional properties, I aim to answer the following questions:

- How does the occurrence of =ga correlate with other criteria for topic-hood, such as clause-initial position and prosodic separation from the clause;
- How does the (non)obligatoriness of =ga on different types of topical expressions influence its analysis as a topic marker;
- Is it possible to accounting for the occurrences of =ga on non-nominalised verbal hosts within its analysis as a marker associated with topicality?

The answers to the above questions will be based on an audio-visual corpus containing 13 hours of Amazonian Kichwa discourse, collected from over 40 native speakers of the language. Examples illustrating the use of the marker will be given in a broader communicative context, in line with the usage-driven focus of the workshop.

References


On information structure and different preverbal positions for the subject

Julio Villa-García, University of Manchester

One major debate in the field of Romance syntax is whether overt subjects are canonical subjects in the sense of English, or whether they always perform some discourse function beyond merely marking the subject (e.g., topic or focus). This controversy stems from the fact that as is well known, Spanish or Italian do not require an overt preverbal subject. Put another way, in such languages, the subject can be null, preverbal or postverbal, depending on factors that go beyond purely grammatical considerations.

Barbosa (2009) points to an interesting difference between non-null subject languages like English and Brazilian Portuguese on the one hand and Spanish and Italian on the other. Whereas the former can display subject doubling, as in (1), the latter are reported to disallow such instances of more than one subject, as in (2):

(1) John, he really loves us

(2) *Juan, él nos quiere

Authors working within Minimalism such as Barbosa (2009) claim that John in (1) is in a high, discourse-related position in the clausal spine, while he occupies the canonical subject position in the inflectional domain. In languages like Spanish (cf. (2)), however, there is no canonical subject position and thus doubling is precluded (the real subject would be the grammatical subject expressed through the verbal morphology, which would dispense with the need to project an overt subject position). The overt subject (i.e., Juan), for its part, is in a high position, on the assumption that it is dislocated. Thus, the general claim has been that whereas in English the subject is in the inflectional domain, in null-subject languages like Spanish the (overt) preverbal subject occupies a high position outside the inflectional domain.

However, I show that in embedded contexts featuring multiple-complementisers (cf. recomplementation, Villa-García 2015), an unambiguously dislocated subject (i.e., one flanked by overt complementisers) can be reprised by a strong pronoun in a position below the lowest complementiser, as in (3), from spontaneous written WhatsApp data (Asturias, Spain):

(3) Te va a decir que José, que él no es así cl. goes to say that José, that he not is thus ‘S/he is going to tell you that as for José, he is not like that.’

What the data in (3) suggest is that as long as the leftmost subject is unambiguously left-dislocated (i.e., an instance of hanging topic), it can be doubled by a strong pronoun in a lower position. Note that only embedded hanging topics require a second instance of the complementiser, as is the case in (3) (which would be optional with Clitic Left Dislocation); see Villa-García (2015) and references therein. More generally, data like (3) point out that preverbal subjects in Spanish can occupy more than one position, and can even occur simultaneously in different positions, contra the predominant view summarised above that overt subjects in Spanish are always related to the left-edge (i.e., discourse) portion of the clause.

This case illustrates the relevance of taking into account information structure in order to determine subject properties, which aligns well with the goals of the workshop.


NEGATIVE MARKERS AND THEIR INTERACTION WITH INFORMATION STRUCTURE:
THE CASE OF MODENESE

One common feature of the morpho-syntax of Romance languages is syntactic negation and the diachronic process undergone by negative markers. Jespersen (1917) was the first to document how negators go through a 3-stage cycle (Jespersen’s Cycle). While it is common for languages to exhibit both synchronic and diachronic variation in the form of two variants (from two stages of Jespersen’s Cycle) of sentence negation, they very rarely tolerate a system of negation which adopts all three strategies (or stages). It has been posited that Brazilian Portuguese be the first Romance language to express negation in a threelfold fashion (Schwegler 1991: 190). However, in this study I offer evidence from a second Romance tongue, Modenese, whose speakers also seem to adopt three different strategies to express sentence negation; each strategy crucially corresponding to the three stages of Jespersen’s Cycle (cf. 1,3 each representing the relevant stage):

(1)  \textit{en ghé nsun.} \hspace{1cm} (NEG1)
\text{ESCL}=\text{neg}\hspace{0.5cm} \text{PF}=\text{be.3SG} \hspace{0.5cm} \text{nobody}
‘Nobody’s there.’

(2)  \textit{en ghé mia nsun.} \hspace{1cm} (NEG2)
\text{ESCL}=\text{neg}\hspace{0.5cm} \text{PF}=\text{be.3SG} \hspace{0.5cm} \text{nobody}
‘Nobody’s there.’

(3)  \textit{ag turné mia.} \hspace{1cm} (NEG3)
\text{SCL.1SG}=\text{there}\hspace{0.5cm} \text{return.PST}\hspace{0.5cm} \text{neg}
‘I didn’t go back.’

Building on the work of Schwenter (2002, 2005), who, in turn, combined and developed the frameworks adopted by Prince (1981) and Dryer (1996), my analysis provides evidence in support of the correlation between information status and sentence negation strategy. Indeed, my data – a one-hour recording, fully transcribed and annotated providing 52 tokens of negation from free narrative and 10 from elicitation tasks – offer suitable grounds to support the pragmatic differences between the three different negation strategies exhibited by Modenese and further elaborate on the notion of DISCOURSE ACTIVATION. Denied propositions are in fact sensitive to their status in the discourse, that is whether or not they have been triggered or activated. Accordingly, a proposition can be new to the discourse (not activated or \textit{discourse-new}), inferentially derived from both linguistic and extralinguistic content (activated because \textit{inferable}), and explicitly expressed in the ongoing discourse (\textit{directly activated}). It is therefore clear that the framework here adopted steers away from that of Lambrech (1994), for example, insofar that it goes beyond sentence structure and considers aspects within the discourse that are external to the individual sentence.

<table>
<thead>
<tr>
<th>Form</th>
<th>Discourse-new</th>
<th>Inferrable</th>
<th>Directly activated</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEG1</td>
<td>OK</td>
<td>OK</td>
<td>OK</td>
</tr>
<tr>
<td>NEG2</td>
<td>#</td>
<td>OK</td>
<td>OK</td>
</tr>
<tr>
<td>NEG3</td>
<td>#</td>
<td>#</td>
<td>OK</td>
</tr>
</tbody>
</table>
Schwenter demonstrated how each information status of activation correlates with the negation strategies of Brazilian Portuguese. Although the paradigm (cf. Table 1 in Schwenter 2005: 1452) is largely consistent with the data collected from Modenese, my preliminary analysis indicates that a fourth proposition-type should be added, namely, *internally activated* – that is whenever the proposition is directly activated internally to the speaker’s mind (cf. 4, where the proposition of ‘the presence of men’ had not been activated in the discourse). Not only is this study a contribution to our current understanding of the correlation between negation strategy and information structure, but it also considers more technical aspects of the practices of coding for referential status in the discourse, which, in turn, is instrumental to the underpinning moral obligation of documenting an endangered language, Modenese, currently on the brink of extinction.

(4) _no perché di omne agn'era mia._

NO because of men  ESCL=PF=INDE=be.IMPF.3SG neg

‘Because there were no men around [they had all escaped].’

References


If the constituent order of Tima, a Niger-Congo language of the Sudan, deviates from the basic AVO order to OVA order, the postverbal agent is formally marked, unlike preverbal agents (or objects regardless of their position relative to the verb). Thus, Tima shows signs of split ergativity/ differential argument marking. One of the triggering factors of these marked constructions is a certain participant constellation in discourse: When the speaker has decided to keep a non-agentive participant, more specifically an object, as the centre of attention and thereby in sentence-initial position in connected speech, a newly introduced agent in a transitive predication occurs postverbally and receives ergative marking. The notion of ‘topic’ for the sentence-initial position is avoided here, instead, “attentional centre”, that is the cognitively focused upon entity (see Himmelmann and Primus 2015: 42), seems more appropriate. Focus marking interferes with constituent order in that a focus marked constituent has to be positioned before the verb. As will be shown, the constituent order rules going along with focus marking override the condition that an ergative marked agent comes postverbal. Whereas animacy does not appear to play a role in alternative constituent order configurations, we argue that the feature [+human] licenses the occurrence of [+human] A or [+human] O in the most prominent, i.e. sentence-initial, position.
Information Structure and Differential Actor Marking in Kelabit

This paper explores differential marking in the Kelabit language of Northern Sarawak. Like other Western Austronesian languages, Kelabit has symmetrical voice alternations in which the mapping of arguments to functions alternate without changes in syntactic transitivity (Riesberg 2014). Hence, the following mappings can be assumed:

<table>
<thead>
<tr>
<th>Table 1. Argument-Function Mappings in Symmetrical Voice</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>actor</strong></td>
</tr>
<tr>
<td><strong>subject</strong></td>
</tr>
<tr>
<td><strong>Actor Voice (AV)</strong></td>
</tr>
<tr>
<td><strong>Undergoer Voice (UV)</strong></td>
</tr>
</tbody>
</table>

For NPs, the function of arguments is shown by word order. However, for a subset of the pronominal paradigm (1SG, 2SG, 3SG and 3PL) there is also a system of case-marking: UV actors are typically expressed using GEN (or ERG), whilst NOM is used for all other core arguments:

(1) a. **Actor Voice (AV)**

\[
\begin{align*}
1SG.NOM & \quad \text{neni’er} \\
\text{AV.PFV.see} & \quad \text{3SG.NOM} \\
\text{‘I saw him’} & \quad t=\text{ieh}
\end{align*}
\]

b. **Undergoer Voice (UV)**

\[
\begin{align*}
\text{Seni’er} & \quad \text{kuh} \\
\text{UV.PFV.see} & \quad 1SG.GEN \\
\text{‘I saw him’} & \quad PT=3SG.NOM
\end{align*}
\]

Although the use of GEN in (1b) is highly preferred out of context, and typically given when eliciting examples of UV, NOM is also possible as a means of expressing UV actors:

(2) a. **Undergoer Voice**

\[
\begin{align*}
\text{Seni’er} & \quad \text{uih} \\
\text{UV.PFV.see} & \quad 1SG.NOM \\
\text{‘I saw him’} & \quad PT=3SG.NOM
\end{align*}
\]

Hence, the expression of UV actors appears to be a case of differential argument marking and prompts the question of what determines the choice of NOM vs GEN?

Cross-linguistically, differential marking is known to be affected by information structure in a variety of ways: specifically, differential object marking (DOM) is often linked to the topicality of the object (Dalrymple and Nikolaeva 2011), whilst differential actor marking (DAM) is linked to the status of the actor as focus/contrastive/unexpected (Witzlack-Makarevich and Seržant 2018, McGregor 2010). What makes Kelabit an interesting case study is that – if we assume the symmetrical voice analysis in Table 1 – the UV actor is both an actor and an object. Hence, it is interesting to explore whether information structure plays a role in the differential marking, and whether it patterns like DOM or DAM?

Evidence from elicited judgements and initial corpus analysis suggests that information structure does play a role: GEN marked UV actors are used for continuing topics, whilst NOM is preferred when the UV actor is focus, contrastive or unexpected. This is illustrated below:

(3) **NOM pronoun preferred when actor is questioned (Lambrecht's 1994 focus test)**

\[
\begin{align*}
\text{Q. Seni’er} & \quad \text{uih} \quad t=\text{ieh?} & \text{A. Seni’er} & \quad \text{uih} \quad (*\text{kuh}) \quad t=\text{ieh} \\
\text{UV.PFV.see} & \quad \text{who} \quad PT=3SG.NOM? & \text{UV.PFV.see} & \quad 1SG.NOM \quad (*\text{GEN}) \quad PT=3SG.NOM \\
\text{‘Who saw her?’} & \quad \text{‘I saw her’}
\end{align*}
\]
(4) **NOM pronoun preferred when actor is contrasted**
   a. Pinupu’ *uih* (*kuh*) t=ieh pu’un, am dih iko
      UV PFV hit 1SG.NOM (*GEN) PT=3SG.NOM first NEG DEM 3SG.NOM
      ‘I hit him first, not you’ (i.e. you didn’t hit him first))

(5) **GEN pronoun preferred when actor is a given topic**
    Paul kedieh, kinan *neh* (*ieh*) bu’a ebpuk
    Paul 3SG.EMPH UV PFV eat 3SG.GEN (*NOM) fruit passion
    ‘As for Paul, he ate the passion fruit’

Hence, the choice of NOM vs GEN does appear to depend on information structure and patterns similarly to DAM in other languages (particularly languages with “optional ergative” marking). However, the marked case (GEN) is used for topics and the unmarked case (NOM) for focus/contrast, morphologically speaking.

Importantly, the use of NOM to indicate a focused actor in UV seems to be dependent on the status of the undergoer as a topic (presumably a motivating factor in the choice of UV in such instances, though perhaps not in all). Otherwise, pronominal actors can also be focused by appearing in the initial position of an AV construction:

(6) **Focused Actor in UV**
    Kayu’ inih, senuuk *uih* *neh*.
    Like DEM UV PFV string 1SG.NOM DEM
    ‘Like that one, I strung that [pointing to the bead cap on the table].’
    [from a conversation about beads – the necklace or undergoer is the topic]

(7) **Focused Actor in AV**
    *Uih* teh ne-ngimet inih keneh
    1SG.NOM PT PFV AV hold DEM he said
    ‘I am the one holding this [the ceiling] up’
    [from a folk story where animals compete to be crowned the strongest – the speaker or actor is the topic]

Though the actor is contrasted and the undergoer is given in both (6) and (7), the two differ in terms of which argument is the primary topic. In (6), the utterance is about the undergoer, whilst in (7) it is about the actor. Hence, the use of NOM actors in UV constructions may be restricted to cases where the actor is focused and the undergoer is the topic, which could explain why it is relatively infrequent in discourse.

Consequently, Kelabit provides an interesting case study to look at variation in differential marking patterns, since it appears to have differential actor marking in a context where the actor is (at least arguably) an object rather than a transitive subject. The patterns are nonetheless similar to other cases of DAM in that they reflect the status of the actor as focus/contrastive but the felicity of the construction also depends on the status of the undergoer argument. Hence, differential marking may not only depend on information structure characteristics of the argument encoded, but also on other relevant referents in the clause.
References:
Competing motivations in differential agent marking: the (ergative) case of Jaminjung/Ngaliwurru

This paper presents results of a corpus-based study of the factors underlying the presence vs. absence of case-marking on transitive subjects in a fluid differential agent marking (“optional ergative”) language, Jaminjung/Ngaliwurru (J/Ng), a Mirndi language of Northern Australia. It is the first study of this kind which includes a quantitative account of the role of information structure, coded on the basis of independent prosodic and contextual criteria, alongside other factors.

In terms of overall frequencies, case marking of agents (i.e. of subjects of transitive clauses) is the default in J/Ng: only 27% of agents are unmarked. The three factors that emerge as statistically significant for the absence of case marking include two that are related to the referential status of the agent NP itself: 1st/2nd (vs 3rd) person status of the agent, and the status of the agent NP as a sentence topic (defined on the basis of prosodic and contextual criteria). These two tendencies can be regarded as manifestations of an overarching principle of Economy (e.g. Silverstein 1976; McGregor 1992; Aissen 1999; Malchukov 2008, 2015): since overt topics are expected to be agents rather than patients, and 1st and 2nd pronouns are more likely to be agents than patients, the marking of their role can be omitted. The higher frequency of topical agents and speech act participant pronouns as agents as opposed to patients is indeed confirmed within the J/Ng corpus. While some recent work has collapsed topicality and high position in the animacy hierarchy into a single prominence scale (e.g. De Hoop 1999; De Hoop & Narasimhan 2005; Bickel et al. 2015), it is argued here that the two factors are independent of and potentially competing with one another.

As predicted by Malchukov (2008, 2015), overt ergativity can however also be subject to a functional constraint which is orthogonal to Economy, labelled “Indexing”. This manifests itself in the tendency for less prototypical agents to not be marked as agents, e.g. in events of a low degree of impact on the second participant (cf. also Tsunoda 1981). The third factor identified for J/Ng – a statistically significant tendency for the agents of verbs of transitive possession and speech verbs to be unmarked – directly reflects this constraint.

In sum, differential agent marking in J/Ng presents yet another instance of (competing) “soft constraints mirroring hard constraints” (Bresnan et al. 2001), in that the statistical tendencies identified for J/Ng mirror some of the factors identified as underlying strict ergative splits cross-linguistically (McGregor & Verstraete 2010). Crucially, in addition, these findings underline the importance of information structure for differential case marking.
References


