<table>
<thead>
<tr>
<th>Manuscript Number:</th>
<th>LAB-D-12-00041R1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Title:</td>
<td>Grammar in parsing and acquisition</td>
</tr>
<tr>
<td>Short Title:</td>
<td></td>
</tr>
<tr>
<td>Article Type:</td>
<td>Squib/Short article</td>
</tr>
<tr>
<td>First Author:</td>
<td>Vasiliki Chondrogianni</td>
</tr>
<tr>
<td>Other Authors:</td>
<td>Marco Tamburelli</td>
</tr>
</tbody>
</table>
| Corresponding Author: | Vasiliki Chondrogianni  
                     | Bangor University  
                     | Bangor, Gwynedd UNITED KINGDOM |
| Section/Category: | Basic Science Section |
| Keywords:         |                  |
| Abstract:         |                  |
| Author Comments:  |                  |
| Order of Authors Secondary Information: |  |
Grammar in parsing and acquisition

Vasiliki Chondrogiani and Marco Tamburelli

Bangor University

O’Grady’s keynote article comes at a timely moment when the acquisition and processing of semantic phenomena in a first and/or second language are gaining renewed interest (Gibson & Pearlmutter, 2011; Grüter, Lieberman, & Gualmini, 2010; Gualmini & Crain, 2005; Musolino & Gualmini, 2011; Unsworth, Gualmini, & Helder, 2008) and feed into the long-standing debate regarding the nature of language (Chomsky, 1993; Pinker, 1999). In this commentary, we will first discuss the theoretical issues raised in O’Grady’s article and then we will address acquisition and processing considerations.

A processor without grammar?

O’Grady suggests that children map sequences onto “semantic representation” without “making reference to conventional syntactic representations” (p. 4). This represents a “strong” version of Emergentism, as it does not simply claim that language acquisition occurs in the absence of pre-existent linguistic knowledge or UG; it claims that acquiring language involves acquiring no linguistic knowledge at all. It maintains, somewhat paradoxically, that a system can learn to process structural relationships without developing any knowledge about “structure” or “relationships”. O’Grady’s proposal appears not to satisfy its requirements, however, as the supposedly “grammar-free” process is associated with many a grammatical concept, not least that of hierarchical relations.
To begin with, O’Grady proposes that the processor develops a “routine” for interpreting basic SVO sequences (p. 6):

\[
\text{PRED} \begin{pmatrix} <a & b> \end{pmatrix}
\]

A major issue here is that \(<a>\) and \(<b>\) relate to the predicate in non-equivalent ways, as evident in the fact that they look suspiciously like an ordered pair, a well-known notational alternative to tree structure. In view of this, O’Grady appears to achieve one of two things: (i) relocating syntactic principles into what he calls “semantic representation”; (ii) providing a notational variant of syntactic relations. In the first case, O’Grady’s proposal would be plausible, though not new (Jackendoff, 2007; Lamb, 1966), while the second case would merely be a notational exercise. Syntactic relations are hardly abandoned, however. They are either embedded into semantics or recast through different notational conventions.

Hierarchical relations resurface when O’Grady introduces routines for interpreting the form NP-V-NP, a combination of grammatical entities, by definition. Abandoning the term “grammar” and simply calling these objects “hierarchical” would of course affect only terminology rather than the underlying issue, namely that language processing involves structural relations beyond individual words. It seems therefore clear that the processor cannot “subsume the duties of the grammar” (p. 8), or dispense with “grammatical rules and principles” (p. 3), at least to the extent that these are patterns of hierarchical relationships.

Notably, hierarchical relations are neither exclusive to UG nor equivalent to tree structures,\(^1\) an issue that is not always explicit in O’Grady’s exposition. Hierarchy can be successfully represented through dependency relations (J. Anderson, 1971; Hudson, 1984),

---

\(^1\) All tree structures represent hierarchical relationships but not all hierarchical relationships are represented by tree structures.
strata (Lamb, 1966), or any mechanism that handles non-linear relationships, which is ultimately what grammar is.

O’Grady does argue convincingly that the processor and the grammar need not be separate, though this idea has been previously formulated (Phillips, 1996). Nevertheless, showing that a processor can subsume some of the duties associated with UG is not equivalent to showing that we can discard grammar altogether, nor that language acquisition occurs without “construction […] of linguistic knowledge” (p. 2).

Indeed, O’Grady overtly contradicts this latter claim by borrowing Hawkins’ (2004) suggestion that universals arise from a working memory requirement to “minimize the distance between the verb and […] its dependent phrases” (p. 8). Crucially, for this explanation to work it is essential to credit the processor (or working memory) with knowledge of concepts such as “head”, “selection”, and “phrase”, all run-of-the-mill grammatical notions. This is clear from O’Grady’s examples, where the dispreferred option involves the intervention of a phrase (a syntactic constituent), and not of linear material, as a grammar-free working memory account would expect. Dispreferring “*read that are highly regarded books” does not entail dispreferring the equally long but structurally different “read fairly highly regarded academic books”. It is specifically structural interference, not linear length, that creates the working memory problem. Therefore, the explanation O’Grady embraces makes explicit reference to “conventional syntactic representations” (cf. p. 4). The phenomena that cannot be explained linearly are many, including V2, auxiliary inversion, island effects, etc.

It is therefore clear that even an initial description of a processing-based account must call upon relationships that fall entirely within the domain of grammar, whatever label we wish to use (e.g. structural/hierarchical/dependency), contrary to the claim that developmental events do not require “reference to conventional grammatical rules or
representations” (p. 9). While the Amelioration Hypothesis might reduce the number of grammatical principles we need to assume, it nonetheless presupposes (sometimes implicitly) a processor capable of identifying and decoding grammatical entities and relations. Whether this capability can develop in the absence of a Universal Grammar is an entirely separate question.

**Acquisition and processing of scope**

To test his predictions regarding linear parsing and routine efficiency, O’Grady examines a poverty-of-the-stimulus phenomenon, namely scope relations between negation markers and quantifiers. The quantifier-negation interactions are subject to cross-linguistic variation and, according to O’Grady, their acquisition should mirror dominant target language processing routines. However, cross-linguistic findings show that relative linear order does not predict scope preferences (Szabolcsi, 2002). Acquisition studies from Korean and Japanese further contradict a linear analysis. Both languages have an SOV word order whereby the quantifier precedes negation ($\forall > \neg$). Korean-speaking children have been shown to prefer this linear interpretation (Lee, Kwak, Lee, & O’Grady, 2011), whereas Japanese-speaking children opt for the narrow scope interpretation ($\neg > \forall$), against the linear surface order (Goro & Akiba, 2004).

Evidence that children parse quantifier-negation interactions on the basis of structural relations rather than linearly comes from studies with Kannada-speaking children (Lidz & Musolin0, 2002). Kannada, like Korean, is an SOV language where quantified objects are c-commanded by negation but are not linearly preceded by it. In a sentence such as *Cookie Monster didn’t eat two slices of pizza*, Kannada-speaking children followed the syntactic (c-command) structure rather than the surface order, thus behaving like English-speaking children. This finding is not in line with the results from Korean-speaking children.
reported by O’Grady, who unambiguously prefer the wide scope interpretation. As O’Grady
notes, the results from Kannada-speaking children do not necessarily contradict his
predictions, as the types of quantifiers in the two studies differed. A numeral is always
interpreted within the scope of negation, whereas a universal quantifier is always ambiguous
between the two interpretations. However, this distinction suggests that children are sensitive
to the interpretation of quantified NPs as a function of the lexical nature of the quantifiers and
of their syntactic position (Musolino & Gualmini, 2011). This sensitivity presupposes
grammatical knowledge that exceeds linearity. Furthermore, postulating that each scopal
pattern requires separate routines is against a notion of processing efficiency.

Routine efficiency also seems to be subject to experimental manipulations. Experimental studies have shown that adult speakers of English fail to access the felicitous
interpretation, when experimental conditions are manipulated (Conroy, 2008; Musolino &
Lidz, 2003). These findings challenge O’Grady’s argument regarding preferred routines. If
preferred routines were in place, then we would not expect adults to opt for the non-felicitous
interpretation against the dominant routine.

O’Grady makes the thought-provoking claim that parsing involves unidirectionality
and processing cost. “Unidirectionality” claims that interpreting an item as soon as it is
encountered enhances ease of processing. However, this assumption is problematic in relation
to well-known garden path (GP) sentences, such as “Without her contributions would be
impossible”. If the processor interpreted “her” immediately upon encountering it, it would
build the perfectly plausible unit [without her], then discard “her” from working memory and
proceed to the next unit without being garden-pathed. Contrary to the Unidirectionality claim,
however, such sentences show that the processor prefers a “wait and see” strategy (Frazier &
Rayner, 1982), favouring the sequence [without her contributions] instead of the equally
plausible but more incremental [without her].
By assigning scope-induced cost to the processor, O’Grady captures certain well-documented generalisations regarding surface scope. For example, non-linear scopal interpretations are seen as computationally more costly in theoretical models of scope (Fox, 2000; Musolino & Gualmini, 2011; Reinhart, 2006). Psycholinguistic experiments have also reported that inverse-scope configurations are dispreferred because they are computationally more costly (C. Anderson, 2004). Processing studies have shown that children have difficulties revising initial parsing interpretations (Snedeker, 2009; Trueswell, Sekerina, Hill, & Logrip, 1999). In his paper, O’Grady seems to conflate off-line accuracy on a truth-value judgment task, a metalinguistic and cognitively demanding task, with processing cost, which is more reliably measured using on-line methodologies that tap into sentence comprehension in real-time (Marinis, 2010; Sekerina, Fernández, & Clahsen, 2008). Nevertheless, his predictions regarding processing cost are potentially testable by applying more fine-grained on-line methodologies. The extent to which it is possible to tease apart O’Grady’s predictions from those of grammar-based accounts that evoke processing cost (C. Anderson, 2004), is in need of further investigation.

Turning to L2 acquisition, O’Grady’s *Transfer Calculus* argues that L2 learners will transfer dominant processing routines from the L1, unless the cost of implementing the L1 routine in the L2 is high. Recent bidirectional studies by Grüter et al. (2010) investigating scope ambiguities in the context of disjunction and negation in speakers of L2 English and Japanese provide more data to test O’Grady’s predictions. Grüter et al. (2010) examined the acquisition of scope and disjunction in Japanese L2 learners of English and English L2 learners of Japanese. According to O’Grady’s parsing model, the dominant routine in English would be the one where negation takes scope over disjunction, as dictated by the linear order (¬ > ∨). In Japanese the dominant routine is the exact opposite, since disjunction linearly precedes negation (∨ > ¬). Therefore, for English learners of L2 Japanese and Japanese...
learners of L2 English transferring the L1 interpretations on to the L2 would involve costly operations, as this would involve interpreting an operator (i.e. negation or disjunction) before it is encountered. O’Grady’s transfer calculus would thus predict that L1 transfer would be blocked in the case of Japanese learners of L2 English and English learners of L2 Japanese. Results from Grüter et al. (2010) seem to go against O’Grady’s predictions. In their study, Grüter et al. (2010) showed that both Japanese learners of L2 English and English learners of Japanese initially transferred the L1 settings on to the L2. These results seem to be compatible with the predictions made by the Full Transfer/Full Access model of L2 acquisition (Schwartz & Sprouse, 1994) combined with a Semantic Subset Principle (Crain, Ni, & Conway, 1994).

Overall, O’Grady’s article raises a number of challenging yet recurring issues in first and second language acquisition and processing by making testable predictions that future research will benefit from exploring further. Nevertheless, a full account is likely to need a stronger involvement of at least some grammatical relations that go beyond linear considerations.

Authors’ addresses

Vasiliki Chondrogianni  Marco Tamburelli
School of Linguistics and English Language  School of Linguistics and English Language
Bangor University  Bangor University
New Arts building  New Arts building
Bangor LL57 2DG  Bangor LL57 2DG
v.chondrogianni@bangor.ac.uk  m.tamburelli@bangor.ac.uk
References


