

How much is too much? The impact of redundancy on syntax

Giacomo Presotto – Goethe University Frankfurt

Redundancy in human grammars has been shown to ensure robustness against information loss [1, 2] and support learnability [3, 4]. At the same time, redundancy has been described as not economical [5, 6] and disadvantageous to learners [7], as it compromises transparent one-to-one relations between form and meaning [8, 9]. These perspectives produce conflicting predictions on the role of redundancy in language acquisition. The first goal of this study is to investigate which perspective is better supported by child language acquisition data. The second goal is to pin down the role of deficiency (i.e., the opposite of redundancy). To this end, we tested the effect of redundant and deficient argument-disambiguating cues on children's comprehension of Object Relative Clauses (ORCs) in Spanish, where the subject and the object can be told apart based on Differential Object Marking (DOM) and/or Word Order (WO). Interestingly, DOM can function as a disambiguating cue even when absent: if a [+animate, +specific] DP is not marked by DOM, it is necessarily the subject. I dub the latter "*in-absentia* cue", and argue that it conveys deficient information. The results show that children's performance improves when DOM and WO cues occur in isolation, decreases when they are simultaneously present, and is at floor when deficient information is provided. Furthermore, the interaction between WO and DOM gets significant once children have acquired the grammar of DOM. That is, if the preposition *a* that realises DOM is not yet associated with its function within the grammar, it is neglected by the parser, and, as such, does not count as redundant.

Spanish ORCs constitute an ideal domain to test whether redundancy facilitates learnability for two reasons: a) ORCs are challenging for children [10], who are therefore *a fortiori* expected to rely on redundancy if it eases the parser, and to disprefer it if it represents a burden; b) WO and DOM cues can be freely manipulated to gauge how their isolated/simultaneous occurrence impacts on children's comprehension. Specifically, I consider ORCs that vary with respect to the position of the subject, either postverbal (1, 2) or preverbal (3, 4), and the presence (2, 4) or absence (1, 3) of an overt *a* morpheme in front of the relativiser *que* (i.e. DOM).

- | | |
|---|--|
| (1) El niño que saluda el futbolista
The boy that greets the footballer
<i>The boy that the footballer greets</i> | (2) El niño al que saluda el futbolista
The boy DOM-the that greets the footballer
<i>The boy that the footballer greets</i> |
| (3) El niño que el futbolista saluda
The boy that the footballer greets
<i>The boy that the footballer greets</i> | (4) El niño al que el futbolista saluda
The boy DOM-the that the footballer greets
<i>The boy that the footballer greets</i> |

All ORCs in (1-4) are unambiguous, as the identification of the subject/object is always possible thanks to DOM and/or WO cues. The presence of DOM on the relativiser imposes that the relative head is the object (2, 4). The preverbal position within the relative clause (WO) is prevented to objects (under regular pragmatic and prosodic conditions), so that DPs in that position can be nothing but subjects (3, 4). Finally, (1) is disambiguated by the *absence* of DOM: because the postverbal [+animate, +specific] DP is *not* introduced by DOM, it can only be the subject. Thus, (2, 3) are disambiguated by one cue each (DOM and WO, respectively), (4) by two simultaneous cues (DOM+WO), and (1) solely by what I dub an *in-absentia* cue. If redundancy boosts learnability, the prediction is that children perform better with (4). If redundancy obstructs learnability, then (2, 3) should be preferred. Last, I predict (1) to be the most burdensome condition. Indeed, *in-absentia* cues, albeit equally disambiguating, are not as transparent and robust as cues that are manifestly present. In fact, children have been shown to struggle with covert morphological elements in general [11].

ORCs like (1-4) were arranged in 4 conditions in a 2x2 design and integrated with 16 fillers. All items featured two [+animate, +specific] DPs, serving as the agent and the patient of

transitive constructions. The sentences were recorded and presented as audio stimuli. Their comprehension was tested through a sentence-picture-matching-task, whereby each item is associated with two pictures, one representing the event described by the item, and the other representing the same event but with inverted agent-patient relation. Children were instructed to point to the picture described by the stimulus. The experiment was run with 57 Spanish monolingual children aged 4-6, who were additionally tested with a sentence-repetition-task (SRT). The SRT included 9 occurrences of DOM in different syntactic contexts. In the analysis, each child was assigned a score from 1 to 9 based on the number of successful repetitions of these target-structures. This score reflects the participants' degree of mastery of DOM-grammar at the time of the experiment. The results (**fig.1**) were analysed with a generalised linear mixed effect model with accuracy in the interpretation of ORCs as outcome variable and the interaction between presence of DOM, WO (preverbal vs. postverbal subjects) and DOM-score as predictor. We found an effect of WO – whereby the WO cue brings about the highest accuracy. The significant interaction between WO and presence of DOM shows that DOM enhances comprehension when the WO cue is absent (1, 2), but hinders comprehension when it co-occurs with WO (see 3 vs 4). Moreover, a higher degree of DOM-mastery predicts higher accuracy when DOM occurs in isolation, but lower accuracy when it co-occurs with WO (**fig.2**).

The observation that redundant morphosyntactic information, as realised by concurrent cues, hinders children's comprehension of complex syntactic structures shows that redundancy does not assist the acquisition of ORCs, but rather weighs on the child parser. At the same time, when disambiguating cues are not robust enough, as with *in-absentia* cues, children's comprehension is negatively affected. Altogether, evidence is provided that: a) Redundant morphosyntactic information does not support learnability but, rather, hinders children's comprehension of complex structures; b) the computation of redundancy in grammar is exclusively based on properties that are relevant to the system (i.e., they bear a specific function in the grammar), as shown by the fact that any non-acquired properties do not trigger redundancy effects; c) the child parser cannot rely on deficient disambiguating cues to resolve complex syntactic structures (as in the case of isolated *in-absentia* cues). In light of this, I argue that the child parser benefits from trade-offs between deficient and redundant morphosyntactic information.

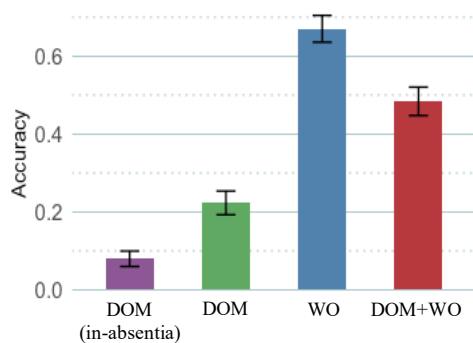


Fig.1
Children's accuracy across conditions

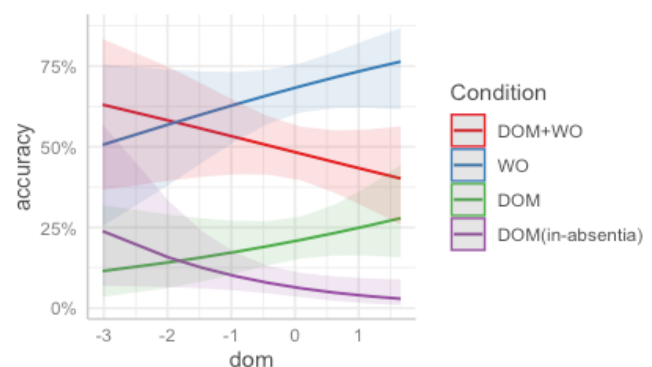


Fig. 2
Predicted probabilities of accuracies based on DOM-mastery level

Selected references

[1] Shannon, 1948; [2] Levshina, 2020; [3] Nicolaci-da-Costa and Harris, 1983; [4] Tal and Arnon, 2022; [5] Kusters, 2003; [6] Dahl, 2004; [7] Trudgill, 2011; [8] Sinnamäki 2009; [9] Leufkens, 2015; [10] Friedmann et al., 2009; [11] Biberauer and Roberts, 2012