Rethinking Overspecification in Terms of Incremental Processing

Raquel Fernández
Institute for Logic, Language and Computation
University of Amsterdam
raquel.fernandez@uva.nl

Abstract
Speakers often overspecify their referring descriptions by including more information than is required to uniquely distinguish a referent. Although overspecification has received a substantial amount of attention, the factors that play a role in determining this behaviour are not yet well understood. Given evidence of cross-linguistic difference between English and Spanish regarding overspecification with colour adjectives, we argue that a factor that contributes to the inclusion of arguably redundant properties in a description is their incremental informativity in comprehension. We sketch a generation model that can account for this phenomenon by allowing for incremental interaction between content selection and linguistic realisation and for interdependencies between generation and processing. Keywords: Referring expressions; overspecification; incrementality; production; comprehension.

Introduction
Contrary to what adherence to the Gricean Maxim of Quantity would predict (Grice, 1975), speakers often overspecify their referring descriptions by including more information than is required to uniquely distinguish a referent—a fact that has been attested by a large body of psycholinguistic and corpus-based research (amongst others, Pechmann, 1989; Engelhardt, Bailey, & Ferreira, 2006; Viethen, Dale, Krahmer, Theune, & Touset, 2008; Koolen, Gatt, Goudbeek, & Krahmer, 2011). Despite the growing amount of attention paid to overspecification, the reasons behind this phenomenon are not yet well understood. In this paper, we argue that the overspecification behaviour of a speaker who produces a description can partially be explained and modelled by considering the online processes involved in comprehending that description. From the point of view of comprehension, the description is incrementally processed from left to right as it is being produced. This means that properties that could count as redundant when considering the full description as a unit may in fact be informative during incremental processing. For instance, given the scenario in Figure 1, the description the red lamp may be considered overspecified, with the property red being redundant, i.e. not strictly needed to distinguish the target uniquely. From a comprehension perspective, however, red is incrementally informative: i.e. it allows the hearer to rule out possible referents at the point in time when the adjective is processed.

Taking incremental informativity into account allows us to make some interesting predictions regarding overspecification. Most importantly perhaps, it predicts that in languages where adjectives typically appear post-nominally—such as most Roman languages—rather than pre-nominally—as in English—, properties realised by adjectives will be used redundantly less often since whenever the head noun suffices to uniquely distinguish the referent they will not be incrementally informative. For example, given the visual scenario in Figure 1, both the English description the red lamp and the equivalent Spanish description la lámpara roja may be argued to be overspecified (the lamp/a lámpara would suffice in each respective case). In the English description, however, red is incrementally informative, while in the Spanish one roja is not. If speakers take into account the incremental informativity of the surface realisation of a property when planning their referring expressions, then we would expect Spanish speakers to produce fewer overspecified descriptions than English ones. And this is precisely what recent psycholinguistic data has shown. A set of experiments run by Fernández-Rubio and colleagues indicates that there are indeed cross-linguistic differences regarding overspecification, with English speakers producing a significantly larger proportion of redundant colour adjectives than Spanish speakers (Rubio-Fernández, 2011).

We argue that the sketched view of overspecification and the aforementioned psycholinguistic results call for a computational model of the generation of referring expressions where content determination and linguistic realisation take place incrementally and where there is interdependence between generation and processing. The need for these requirements has to some extent been acknowledged by the NLG community, but very seldom have they been incorporated into actual generation systems and algorithms (Krahmer & Deemter, 2012). We proceed by first giving an overview on how overspecification is most often modelled in current systems. After this we present the main features of our proposed model and go over an example in detail. We then discuss the implications of the approach and its connections to related research, before closing with some conclusions and suggestions for future work.

Overspecification in Current Systems
Systems for the generation of referring expressions that aim at emulating human behaviour typically incorporate mech-
anisms that enable the generation of overspecified descriptions. The Incremental Algorithm proposed by Dale and Reiter (1995)—which is considered a standard in the field—as well as most of its variants (e.g., Krahmer, Erk, & Verleg, 2003) employ property preference orders or cost functions, which are taken to reflect the relative salience of different types of properties in particular domains. This allows for salient properties to be included in the conceptual plan of a description as long as they have discriminatory power, regardless of whether they end up being redundant once all the properties are selected. For instance, a context model for the visual scene in Figure 1 may include the preference order <colour, type, position> indicating the relative salience of different attributes in this context: colour being the most salient property, followed by object type, followed by position (relative to e.g. the focal middle point of the scene). The Incremental Algorithm would go over each property in turn, incorporating it to the description plan if it has discriminatory power at the point in the order where the property is considered, and it would stop once the planned description uniquely identifies the target referent. In our context, the fact that colour (which rules out some possible referents but does not uniquely distinguish the target) is more salient than type (which does uniquely distinguish the target) would explain the "redundant" inclusion of red in the description the red lamp.

Algorithms such as Dale and Reiter's (1995) deal exclusively with the generation sub-problem of content selection, namely, deciding which properties of the target referent are to be included in the referring expression to be produced, independently of how such properties will end up being linguistically realised. They operate at the conceptual level, which presumably is invariant across languages, and thus are not able to account for cross-linguistic differences in overspecification behaviour. Besides salience, other factors have been identified as influencing the amount of information speakers include in referring description (Koolen et al., 2011), such as the complexity of the domain (the amount of properties available to describe a referent) and the cardinality of the target (with plural targets being more often overspecified). Again, however, these are factors that concern the conceptual level (the situation model) and cannot straightforwardly explain cross-linguistic differences.

It is worth pointing out that there have been a couple of cross-linguistic studies conducted by Krahmer and colleagues, who have looked into the possible differences in referential behaviour between English and Dutch speakers. However, the similarity between these two languages (especially regarding definite referring expressions) has not uncovered any substantial language-dependent factors (Theune, Koolen, & Krahmer, 2010; Koolen, Krahmer, & Theune, 2012). The cross-linguistic differences regarding overspecification between English and Spanish that we have mentioned in the Introduction are particularly interesting because they offer support for designing models of the generation of referring expressions and of production more generally that meet certain requirements. On the one hand, they indicate that the information included in a referring description does not only hinge on language independent factors related to content selection, but rather that language specific aspects may play a role as well and, hence, that content selection is interleaved with linguistic encoding. On the other hand, since the cross-linguistic differences observed involve what we have referred to as incremental informativity stemming from surface word order, we argue that they support close coupling between production and comprehension. The first of these requirements has often been emphasised in the psycholinguistics and the NLG literature (e.g., Levelt, 1989; Kilger & Finkler, 1995; Stone & Webber, 1998). The second one may be a bit less obvious, although it has recently been gaining importance in research on interactive dialogue settings. We shall go back to this issue in the discussion section, after sketching our model.

A Model Sketch

In a nutshell, in our model the generator operates in an incremental manner, interleaving semantics and syntactic realisation and imposing constraints on linguistic linearisation that take into account the left-to-right processing of the description from the comprehension side. Figure 2 sketches the architecture we proposed. It what follows we specify in more detail its components and the way they operate.

The model distinguishes between two types of resources: a language-independent situation model and a language-dependent grammar. The situation model includes a target referent representing the speaker’s communicative goal (the goal being to distinguish the target from other objects in the scene) and a representation of the context. We may want to encode several aspects of the context here, from a model...
of salience to a representation of the common ground between the interlocutors. For simplicity, we assume a model of salience in the form of a property preference order. As for the grammar, the key feature we require of it is full (i.e. word-by-word) incrementality as we want to be able to monitor at each derivation step whether the communicative goal is fulfilled (Levelt, 1989). Dynamic Syntax (Kempson, Meyer-Viol, & Gabbay, 2001), which seems to be fully incremental and has previously been applied to generation (Purver & Otsuka, 2003; Purver & Kempson, 2004), could be a candidate grammatical framework. Another option could be Lexicalised Tree Adjoining Grammar (LTAG), which has often been used to model incremental generation (e.g., Stone, Doran, Webber, Bleam, & Palmer, 2003). However, standard versions of LTAG are head-driven and thus not fully incremental (Ferreira, 2000). Here, to illustrate the main points in our example, we will assume a word-by-word incremental version of LTAG (although we note that specifying a wide-coverage incremental LTAG is provably not trivial). Importantly, we assume that incremental syntactic construction comes with an incremental semantics (Dynamic Syntax would again be an appropriate framework here).

These two resources (the situation model and the grammar) are leveraged by three processes—content selection, lexical choice, and syntactic-semantic construction—that are organised sequentially but operate incrementally: as soon as a process produces some output, this is acted upon by the next process. The content selection process is a version of Dale and Reiter’s (1995) Incremental Algorithm that outputs properties on the basis of their salience and their discriminatory power (and possibly other parameters in the situation model). The lexical choice process takes as input the properties generated by content selection as they become available and progressively outputs LTAG elementary trees. The syntactic-semantic construction process attempts to build derived trees given the available elementary trees at each incremental step. It does so following standard LTAG principles (adapted to full incrementality) but, crucially, it disfavors trees with lexical gaps on the fringe, i.e. prefixes of the in-progress description that contain non-terminal symbols. This favours surface, left-to-right incrementality and brings into the picture the comprehension side: even if some properties and corresponding lexical elementary trees are available from the production side, if they cannot be smoothly integrated into the surface form of the ongoing description they will be put on hold. Upon carrying out each incremental step, each process checks whether the communicative goal has been achieved. For instance, upon selecting a property that has discriminatory power, content selection will check whether the properties selected so far already single out the target and if they do it will stop; similarly syntactic-semantic construction will monitor whether the communicative goal has been achieved after each derivation step. If the description constructed so far is syntactically well formed and semantically distinguishing, the process will stop.

### An Example
Let us go over our example from the Introduction to illustrate how the model would generate descriptions in English and Spanish that differ in their degree of overspecification. For the case at hand, let us assume that the situation model includes a representation of the scene in Figure 1, the lamp in that scene as the target referent, and the property preference order <colour, type, ...> which takes colour to be the most salient property. We assume furthermore that the context initially triggers the plan to produce a definite description, which results in the initial elementary tree in Figure 3(a). This tree is now available to the syntactic-semantic construction process. Content selection then computes whether the most salient property (colour) of the target (red) has discriminatory power and since it does (it eliminates three distractors) it outputs red. Lexical choice can now act on this property to generate an appropriate elementary tree, 3(b) in the case of English and 3(c) in the case of Spanish. This tree now becomes available to the syntactic-semantic construction process, which may adjoin it to the initial tree.\(^1\) However the resulting derived tree would have different properties across languages: while the English tree respects surface incrementality (Figure 4), the Spanish one doesn’t since there is a lexical gap within the prefix derived so far (Figure 5). In our model derived trees that do not adhere to surface incrementality are not licensed.

\(^1\)Note that here we deviate from standard version of LTAG where auxiliary trees are typically adjoined to phrasal heads. For the sake of incrementality, we allow adjunction to a node that “is expecting” a head.

---

Figure 3: Elementary trees

<table>
<thead>
<tr>
<th></th>
<th>NP</th>
<th></th>
<th>N'</th>
<th></th>
<th>A</th>
<th></th>
<th>N'</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>Det</td>
<td>N'</td>
<td></td>
<td>the</td>
<td>red</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b)</td>
<td>Det</td>
<td>N'</td>
<td></td>
<td>the</td>
<td>red</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c)</td>
<td>Det</td>
<td>N'</td>
<td></td>
<td>the</td>
<td>roj-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d)</td>
<td>Det</td>
<td>N'</td>
<td></td>
<td>the</td>
<td>lamp</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4: Derived trees in English adhering to surface incrementality.

<table>
<thead>
<tr>
<th></th>
<th>NP</th>
<th></th>
<th>N'</th>
<th></th>
<th>A</th>
<th></th>
<th>N'</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>Det</td>
<td>N'</td>
<td></td>
<td>the</td>
<td>N'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b)</td>
<td>Det</td>
<td>N'</td>
<td></td>
<td>the</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c)</td>
<td>Det</td>
<td>N'</td>
<td></td>
<td>the</td>
<td>red</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d)</td>
<td>Det</td>
<td>N'</td>
<td></td>
<td>the</td>
<td>lamp</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---
who also make use of Tree Adjoining Grammar. Besides Doran, 1997; Stone & Webber, 1998; Stone et al., 2003),
of this is the work of Stone and colleagues (Stone &
content selection and linguistic realisation has often been ad-
consider one of the main features of Levelt’s seminal work
incrementality in production is of course one of the main features of Levelt’s seminal work
progress. These key features, which we have adopted to
The model we have proposed is similar to earlier models in
build on this body of work, our model is also inspired by research on incremental architectures for dialogue sys-
tended to overspecify their descriptions with colour adjectives
due to the syntactic properties of their language, which criti-
cally, we argue, cancel out the incremental informativity that
This is also semantically distinguishing and therefore the process
Thus, even though a property such as colour can be particu-
(these adjectives have (from the comprehension side) when
Thus we assume it is less salient in the situation model.
Besides being able to account for the cross-linguistic dif-
redundantly when they do not correspond to easy-to-name basic
can attribute this effect to a delay in the lexical choice process caused by difficulty with the lexical retrieval of
in pre-nominal position. Besides being able to account for the cross-linguistic dif-
factor of the partner’s comprehension.
For instance, (a) NP
(b) NP
Figure 5: Derived trees in Spanish: (a) violates surface incrementality, (b) doesn’t.
stituted into the first derived tree leading to the syntactically
complete tree in Figure 4(b). In Spanish the new tree can be
substituted into the initial elementary tree (Figure 5(b)), this
time without violating surface incrementality. At this stage,
we have constructed a syntactically complete description that
is also semantically distinguishing and therefore the process
may stop without need to further adjoin the adjectival tree.
Thus, even though a property such as colour can be particu-
larly salient to speakers regardless of the linguistic resources
they use, we predict that Spanish speakers will be less in-
clined to overspecify their descriptions with colour adjectives
due to the syntactic properties of their language, which criti-
cally, we argue, cancel out the incremental informativity that
these adjectives have (from the comprehension side) when
they appear in pre-nominal position.
Besides being able to account for the cross-linguistic dif-
fences described above, we believe that our model also has
the potential to account for other aspects that influence over-
specification. For instance, Viethen, Goudbeek, and Krahmer
(2012) found that in English colours are less often used re-
dundantly when they do not correspond to easy-to-name basic
colours. We can attribute this effect to a delay in the lexical
choice process caused by difficulty with the lexical retrieval of
a property generated by the content selection process. In
our incremental model, upon encountering difficulty with the
retrieval of an adequate word, the lexical choice process may
be able to operate on other properties that may have become
available in subsequent incremental steps by the content se-
lection process; and the syntactic-semantic construction pro-
cess, in turn, may be able to construct a well-formed disting-
ishing description before ever receiving as input a lexical
tree with a suitable colour word. This would explain why
colour ends up not being used in such cases, without need to
assume it is less salient in the situation model.

Discussion
The model we have proposed is similar to earlier models in
several respects. Strong incrementality in production is of

Conclusions
With this paper, we hope to have drawn attention to the need
for cross-linguistic studies of linguistic production in order to
get a broader view of requirements for NLG models. Ini-
tiatives such as the Dutch Tuna Corpus (Koolen & Krah-
mer, 2010) are a step in the right direction, although com-
parisons between languages that are less similar than English
and Dutch would potentially bring in more insights. We have
sketched a model that operates incrementally from content se-
lection to syntactic construction, interleaving generation with
processing. These key features, which we have adopted to
account for cross-linguistic differences regarding overspeci-
fication, seem to be required by any model of generation in
interactive dialogue settings.
References


