Every Boy Bought Two Sausages Each: Distributivity and Dependent Numerals

Lucas Champollion

1. Introduction

It has been reported that some English speakers accept sentences like Every boy bought two sausages each (Szabolcsi, 2010). Analogous sentences that involve distance-distributive elements in the scope of distributive universals are grammatical in languages like German, Japanese and Korean. This suggests that adnominal each (for the relevant speakers) and its crosslinguistic counterparts across languages require covariation with distributive universal quantifiers and other licensors, but that they are not themselves distributive operators, contrary to standard accounts. I analyze two sausages each as a dependent numeral, analogous to dependent indefinites (Farkas, 1997). My analysis is couched in terms of dynamic plural predicate logic with postsuppositions (Brasoveanu, 2013; Henderson, 2014), which I present in terms of a novel river metaphor.

2. Description of the Phenomenon

Distance-distributive elements often occur in the scope of universal quantifiers. The following sentences all can be translated as ‘Every boy ate two sausages’, on its surface scope reading:

(1) sonyen (-tul) -mata sosici twu- kay- ssik- ul mek- ess- ta. (Korean)
    boy   PL every sausage two CL each ACC eat PAST DECL
(2) Jeder Junge hat jeweils zwei Würstchen gegessen. (German)
    Every boy has each two sausage eaten.
(3) Subete-no danshi-ga sosseg-i o fu-tatsu-zutsu tabeta. (Japanese)
    Every-Gen boy-Nom sausage-Acc two-CL-each ate.

As these examples show, distance-distributive elements in many languages can also be licensed by distributive quantifiers (Choe, 1987; Szabolcsi, 2010:132ff.). English may be another case in point. Szabolcsi (2010) reports that “the native speakers [she] consulted overwhelmingly accepted examples like [(4)], and some even accepted [(5)]”. Moreover, examples like (6) are easy to find on the web.

(4) Every boy had one apple each.
(5) %Each boy had one apple each.
(6) Every square costs one Cent each at the beginning (...).

English sentences like these are often judged ungrammatical in the semantic literature (e.g. Choe, 1987; Zimmermann, 2002). But linguists and nonlinguists sometimes have different intuitions when it comes to acceptability judgments (Spencer, 1973). As Szabolcsi (2010:217) notes with respect to sentences like (4): “The only speakers who systematically rejected them were semanticists. I generally have no qualms about using experts as informants, but my suspicion is that rejection here is a normative judgment.” Whatever the status of the English sentences in question, the grammaticality of their counterparts in other languages as shown above is beyond doubt, and this will be my starting point.

* Lucas Champollion, New York University, champollion@nyu.edu. I would like to thank Adrian Brasoveanu, Robert Henderson, Anna Szabolcsi and the audience of WCCFL 32 for helpful comments and discussion. Thanks to Songhee Kim and Yohei Oseki for their help with the Korean and Japanese sentences.

† http://ricegraineffect.wordpress.com/category/ricegraineffect/

Distance-distributive elements are usually analyzed either as universal quantifiers or as distributive operators (e.g. Zimmermann, 2002; Dotlačil, 2011; Champollion, 2012, 2014). Since in the sentences above, the subject already distributes, these accounts predict that there is nothing left to do for the distance-distributive element, or in other words, that it is vacuous in these sentences. But as Dotlačil (2011) notes, there is no explanation why vacuous distributivity should be acceptable in sentences like the ones above given that it is unacceptable in sentences like this one:

(7) *Alex bought two sausages each.

3. Proposal

In this paper I take and implement the view that for those speakers and languages that conform to the pattern above, distance-distributive elements do not themselves introduce distributivity but instead are licensed by it. Formally, I will treat them as dependent numerals, analogously to dependent indefinites (Farkas, 1997; Henderson, 2012, 2014). Dependent indefinites and numerals are required to be in the scope of an operator with respect to which they can covary, their “licensor”. They can be licensed by every in Hungarian, Romanian, and Russian:

(8) Minden gyerek olvasott hét-hét könyvet. (Hungarian)
   ‘Every child read seven books.’ (Farkas, 1997)
   \( \forall > 7, *7 > \forall \)

(9) Fiecare băiat a recitat câte un poem. (Romanian)
   ‘Every boy has recited a poem.’ (Brasoveanu & Farkas, 2011)
   \( \forall > 1, *1 > \forall \)

(10) Každyj malčik vstretil kogo-nibud’ iz svoix odnoklassnic. (Russian)
    ‘Every boy met one of his girl classmates.’ (Yanovich, 2005)
    \( \forall > 1, *1 > \forall \)

Distributive numerals (Gil, 1982) can also be subsumed under this paradigm:

(11) Her çocuk ikiser sosis aldı. (Turkish)
    ‘Every child bought two sausages.’ (Tuğba Çolak-Champollion, p.c.)

I claim that for the relevant speakers of English, adnominal each turns its host into a dependent numeral, and similarly for the other languages we have seen. For those speakers that accept sentence (12) in the first place, it is false if the same two sausages keep changing hands. This suggests that it can only have a surface scope reading, like all the other sentences we have seen so far.

(12) Every boy bought [two sausages each].
    \( \forall > 2, *2 > \forall \)

To formalize this claim, I will use dynamic plural predicate logic (van den Berg, 1996; Nouwen, 2003) enriched with postsuppositions (Brasoveanu, 2013). This framework is well-suited to the modeling of dependent indefinites because it gives us a way to state that a given indefinite needs to be in the scope of something with respect to which it can covary. For this reason, it has been used by the most recent one of the authors who have worked on the topic of dependent indefinites (Henderson, 2014).

Since there are already several introductions to this framework, instead of simply repeating them I will introduce a metaphor that informally illustrates the point of dynamic plural predicate logic and of postsuppositions (Section 4). This metaphor lays the groundwork for the formalization (Section 5).

If the view developed here is correct, distance-distributive elements themselves do not contribute distributivity, but are licensed by it. Section 6 shows that this predicts a kind of “distributive concord” and that it is indeed found in German and Japanese.
4. The River Metaphor

Information flow in ordinary formal semantics as it is found for example in textbooks like Heim & Kratzer (1998) can be thought of as a river that flows downward, that is, from the c-commanding expressions to the c-commanded ones. (See Table 1 for an overview of the metaphor.) This corresponds to the ways in which variable assignments are used to transmit information about antecedents. An expression with anaphoric potential, like every boy, can load cargo on a boat and send it down this river (this corresponds to assigning a value to a variable), and a downstream pronoun can pick up its cargo (this corresponds to retrieving the value of the variable). So antecedents are departure harbors and pronouns are destination harbors.

In this metaphor, it becomes easy to give a high-level description of how different semantic frameworks relate to one another. The essence of dynamic semantics, as embodied for example in dynamic predicate logic (DPL, Groenendijk & Stokhof, 1991), is that it reroutes the river so that it not only flows downwards but also sideways, for example from restrictors of quantifiers to their nuclear scopes, from antecedents of conditionals to their consequents, and from sentences to subsequent sentences in the text. Ordinary predicates like run and kick (“tests” in DPL terms) are placed like sentinels along the river. They inspect the boatloads and either let them pass or reject them, depending on whether these boatloads satisfy the predicate in question.

In DPL, a universal quantifier like every boy cycles through all the boys in the domain, and launches each boy on a different boat. The boats do not interact with each other downstream. One can imagine that the river is so narrow that the boats travel one at a time. In dynamic plural logic (DPIL) (van den Berg, 1996; Nouwen, 2003), the river is larger, so that several boats can travel next to each other. DPIL represents a distributive universal quantifier like every boy as a combination of a maximization operator and a distributivity operator. The maximization operator still launches each boy on a different boat as before. The distributivity operator temporarily splits up the river into a different arm for each of these boats, like a river delta. This makes sure that collective predicates are incompatible with every boy. Each boat sails down one of these arms. The nuclear scope of the quantifier is then applied at each arm. For example, if the nuclear scope is talked, then at each of the arms of the river a sentinel is positioned that inspects the boat on this arm and checks if its passenger talks. (Sentinels can only see what is going on on their river arm, at the place where they are positioned.) After the nuclear scope has been passed, the river arms are reunited into a big river. This makes it possible for subsequent operations to affect all the boats at once. For example, plural pronouns can pick up all of the boats at once as antecedents, as in:

(13) Every boy talked. Then they left.

Let us now add plural entities to the domain, which may be modeled as sums as in mereology and algebraic semantics (see Champollion & Krifka (to appear) for an overview). Then it becomes possible for an antecedent like two boys to load a sum of two boys onto a boat, and a pronoun like they can pick it up. So there are two ways for plural pronouns to get their antecedents. Either they come on different boats launched by the same antecedents, as before, or they come on the same boat as a sum entity, as here. These two possibilities are known as “evaluation-level” and “domain-level” plurality (Brasoveanu, 2013). In what follows, I will speak more generally about evaluation-level and domain-level cardinality.

In the systems described so far, sentinels (tests) are stationary. DPIL with postsuppositions (Brasoveanu, 2013; Henderson, 2014) allows them to travel down the river as well. Postsuppositions can be thought of as traveling sentinels which are equipped with sealed instructions. They can move away from the word that launches them and follow the course of the river until some operator tells them to stop. Then they open their sealed instructions and execute them.

I will extend the theory by Henderson (2014), based on DPIL with postsuppositions, or DPILP as I will call it here. Henderson deploys that theory for dependent indefinites and I will follow suit. Here is an illustration of how a DPILP analysis applies to the surface scope reading of the sentence Every man, loves a woman. The dynamic semantic component of DPILP ensures that the constituents of these sentences are evaluated in the following way: subject, object, verb.

We begin by evaluating [Every man,]. The universal quantifier corresponds to the instruction to launch each man on a different boat. All these boats sail under the same flag “i”. The (silent) distributivity operator splits the river into as many arms as there are men, and sends each boat
downstream on a different arm. The nuclear scope of the sentence is now evaluated separately on each river arm. We evaluate \([\text{a woman}_j]\) by launching a woman on a boat under the “\(j\)” flag. (Since these launches happen independently of each other, it might happen by chance that the same woman is launched on different arms of the river.) We now get to the word \([\text{loves}_j]\). This is evaluated by placing a sentinel by the river arm with instructions to check the passengers on the two boats and let them pass only if the man loves the woman. Finally, the river arms join back to form a big river.

Suppose now that the next sentence is “They \(i\) married them \(j\).”. Again, these constituents are evaluated in the order: subject, object, verb. The word \([\text{They}_i]\) corresponds to an instruction to collect all the boats sailing under the \(i\) flag, and assemble the men on them to a sum. Likewise the word \([\text{them}_j]\) instructs us to collect all the boats sailing under the \(j\) flag, and to assemble the women on them to a sum. Finally, the word \([\text{married}_j]\): A sentinel checks if the men married the women.

The majority of theories of each model it as a distributivity operator. My main claim is that adnominal each does not distribute (it does not split up the river into arms). Instead, it sends out a traveling sentinel (a postsupposition) with sealed orders that ensure that its host noun phrase covaries.

Let me illustrate this claim by describing informally the analysis I will give to “Every boy \(i\) bought two sausages \(j\) each”. The noun phrase \([\text{every boy}_i]\) launches every boy on a different boat. All of them set sail under the “\(i\)” flag. A silent distributivity operator that is introduced by every temporarily splits the river into a different arm for each boy. On each river arm, the nominal \([\text{two sausages}_j]\) loads two sausages onto a boat and launches it under the “\(j\)” flag. The adnominal element \([\text{each}_j]\) launches a postsuppositional traveling sentinel with sealed instructions that say: “Check that the boats you see sailing under the \(j\) flag don’t all carry the same sausages.” The verb \([\text{bought}_i]\) checks if the boy on the \(i\) boat bought the sausages on the \(j\) boat. As we leave the scope of the distributivity operator, the river arms join back into a big river. A postsupposition plug at the end of the river stops the wandering sentinels. They unseal their instructions and carry them out by checking that the boats sailing under the \(j\) flag do not all carry the same two sausages. This will be true, for example, if different sausage pairs were loaded at the different river arms (if every boy bought a different pair of sausages).

Consider now the sentence *Alex bought two sausages each*. The noun phrase \([\text{Alex}_i]\) launches a boat under the “\(i\)” flag and places Alex on it. Everything is as before. When the sentinel launched by each is stopped at the end, it unseals its instructions and checks that the boats sailing under the \(j\) flag don’t all carry the same sausages. But this time there is only one such boat, so the sentinel reports failure. Since there is no way to avoid this fate, the sentence is predicted deviant.

<table>
<thead>
<tr>
<th>Metaphor</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Riverboat</td>
<td>Anaphoric dependency</td>
</tr>
<tr>
<td>Departure harbor</td>
<td>Antecedent</td>
</tr>
<tr>
<td>Destination harbor</td>
<td>Pronoun</td>
</tr>
<tr>
<td>Stationary sentinel</td>
<td>Test</td>
</tr>
<tr>
<td>River</td>
<td>Assignment function (DPL)</td>
</tr>
<tr>
<td>River or river arm</td>
<td>Set of assignment functions (DPL)</td>
</tr>
<tr>
<td>Branching river (e.g. river delta)</td>
<td>Distributivity operator (DPL)</td>
</tr>
<tr>
<td>Traveling sentinel</td>
<td>Postsupposition</td>
</tr>
<tr>
<td>Sentinel stopper</td>
<td>Postsupposition plug</td>
</tr>
</tbody>
</table>

Table 1: The river metaphor

5. Formalization

The basic innovation of DPIL compared with DPL is that instead of evaluating formulas with respect to pairs of assignments \(\langle i, o \rangle\), they are evaluated with respect to sets of assignments \(\langle I, O \rangle\). This is what I described above in terms of branching rivers. In DPILP, contexts are furthermore enriched by sets of propositions \(\zeta\), the postsuppositions. Formally, a context \(\langle C, \zeta \rangle\), or \(C[\zeta]\) as we will write it, consists of a set of assignments \(C\) and a set \(\zeta\) of propositions. By default, postsuppositions are passed on unchanged. DPIL with postsuppositions is defined as follows (for a full definition see Henderson (2014), which is based on Brasoveanu (2013)).
A singular indefinite’s global evaluation cardinality correlates with the number of atoms its domain. I will reserve the use of cardinality bars \(|·|\) for the latter. A variable’s domain cardinality is determined by its number of the mereological atoms of its referents. We can add a domain cardinality test into the object language as follows. Since this is a test, I constrain its input and output context to be identical.

\[
\| x > n \|_{I[\zeta],O[\zeta']} \text{ iff } I = O \text{ and } \zeta = \zeta' \text{ and } \forall i \in I \rightarrow |\text{atoms}(i(x))| > n
\]
As for the evaluation cardinality of a variable $x$, we can access it by checking how many distinct entities are referenced by $x$ across the various assignments in the input set $I$. This corresponds to checking all of the boats that sail under the flag $x$ and seeing how many distinct entities they carry as cargo (or passengers):

$$[EC(x) > n]^{I[O]}_{\zeta'} \text{ iff } I = O \text{ and } \zeta = \zeta' \text{ and } |\{i(x) : i \in I\}| > n$$

As I said above, I claim that a nominal $each$ does not distribute. Instead, it sends out a postsupposition that ensures that its host noun phrase covaries. Formally, $two \text{ sausages each}$ translates as follows. I represent postsuppositions a formulas with overlines.

$$[\text{two sausages each}] \sim [y] \wedge EC(y) = 1 \wedge *\text{SAUSAGE}(y) \wedge |y| = 2 \wedge EC(y) > 1$$

What this does is the following. $[y]$ is DPIL random variable assignment. Next we have a test that constrains the evaluation cardinality of $y$ to be 1. This means that $y$ has to have the same value in all the assignments of the context. The next two tests make sure that this value is a mereological sum of two sausages. The last test requires the evaluation cardinality of $y$ to be greater than 1. As indicated by the superscript, this last test is a postsupposition, so it does not have to be satisfied immediately (and that would be impossible given the first test).

More specifically, if we interpret $[(20)]$ in an input context $G[\zeta]$ and an output context $H[\zeta']$, we get the following. Here, $I[y]O$ is random assignment:

$$I[y]O \land \forall i \in I[^{*}\text{SAUSAGE}(i(y)) \wedge \text{atoms}(i(y))] = 2 \land \zeta \cup \{EC(y) > 1\} = \zeta'$$

The $EC(y) = 1$ test is not a postsupposition, so it is interpreted immediately and therefore it constrains the local evaluation cardinality of $y$. The $EC(y) > 1$ postsupposition is passed on downstream, and it will lead to a contradiction unless it is interpreted in a context that differs from the current one. In this sense, it constrains what I have called the global evaluation cardinality of $y$.

Above, I have described the behavior of a distributive universal quantifier $every \text{ boy}$ as splitting the river into as many arms as there are boys, and launching each boy on a different boat on one of these arms. This corresponds to the way in which universal quantifiers are modeled in DPlLP (and in some of its ancestors, such as DPlL). A distributive quantifier like $every \text{ boy}$ introduces the sum of all boys as a variable and then distributes over them (22). The maximality operator $\max$ and the distributivity operator $\delta$ will be defined immediately afterwards.

$$\text{every boy } \phi \sim \left[\max^*(\text{BOY}(x) \land \delta(\phi))\right]^{I[O]}_{\zeta'}$$

The $\max$ operator from Brasoveanu (2013), defined in (23), introduces a new variable and stores in it, spread out across the input set of variable assignments, the maximal set of individuals that satisfy the formula it takes scope over.

$$\left[\max^*(\phi)\right]^{I[O]}_{\zeta'} \text{ iff } [[x] \land \phi]^{I[O]}_{\zeta'} \text{ and there is no } O' \text{ such that } \{o(x) : o \in O\} \subset \{o(x) : o \in O'\} \text{ and } [[x] \land \phi]^{I[O]}_{\zeta'}$$

The $\delta$ operator from Henderson (2014), defined in (24), essentially splits up its input context into singleton sets of assignments. These will serve as “local” contexts, from the perspective of anything contained in the nuclear scope of the universal quantifier (which coincides with the scope of the $\delta$ operator). The operator then applies the expression in its scope to each of these singleton sets, collects the outputs back together into a global context, and applies any postsuppositions passed up from its scope:

$$\left[\delta(\phi)\right]^{I[O]}_{\zeta'} \text{ iff } \zeta = \zeta', \text{ and there exists a partial function } \mathcal{F} \text{ such that } I = \text{Dom}(\mathcal{F}) \text{ and } O = \bigcup \text{Ran}(\mathcal{F}), \text{ and there is a set of tests } \zeta'' \text{ such that for all } i \in I, \text{ we have } \left[[\phi]^{I[O]}_{\zeta''}\right]\text{ and } \left[\left[\zeta''\right]^{O[\zeta]}\right].$$

When we combine $every \text{ boy}$ with $ate \text{ two sausages each}$, the effect of the distributivity operator is that the test $EC(y) = 1$ is applied within its scope (since this test is not a postsupposition) and the test $EC(y) > 1$ outside of its scope (since this test is a postsupposition). This is shown in (25).
Every boy bought two sausages each.

\[ \max_x (\text{BOY}(x) \land \delta([y] \land ^*\text{SAUSAGE}(y) \land |y| = 2 \land \text{BUY}(x, y) \land EC(y) = 1) \land EC(y) > 1) \]

When there is no distributor to create a local context separate from the global one, \( EC(y) = 1 \) and \( EC(y) > 1 \) are applied in the same context, as in (26). The two requirements cannot be met in the same context, so (26) will always fail.

*Alex bought two sausages each.

\[ \text{BOY}(\text{ALEX}) \land [y] \land ^*\text{SAUSAGE}(y) \land |y| = 2 \land \text{BUY}(\text{ALEX}, y) \land EC(y) = 1 \land EC(y) > 1 \]

6. Final Remarks

In many languages, it is not distance-distributive elements themselves that contribute distributivity, but the universal quantifiers in whose scope they occur. The distance-distributive elements are merely licensed by these universal quantifiers or by other items that induce covariation. This makes them more similar to dependent indefinites than to universal or distributive quantifiers. I have suggested that for some speakers of English, namely those who accept the title of my paper as grammatical, adnominal each is an example of such an item that is licensed by distributivity. This is similar to an idea by Oh (2001, 2006), who sees distance-distributive elements as “distributive polarity items” which are licensed within the scope of a distributivity operator. The difference is that the licensing relationship in that work is a sui generis configurational constraint on LFs, while I have implemented it in the semantics and proposed a connection with the phenomenon of dependent indefinites (Farkas, 1997; Henderson, 2014).

On the view presented here, distance-distributive elements – at least for the relevant languages and speakers – do not introduce their own distributive force but are licensed in the scope of a distributive quantifier. This predicts that a kind of “distributive concord” should be possible. This prediction is borne out, as shown in these examples. Both of them mean ‘For every boy \( x \), there are three books and three girls such that \( x \) gave them to them’.

1. *Every day John bought two sausages each.

2. *John always bought two sausages each.


