Typicality Effects and Distributivity

**INTRODUCTION.** Non–Atomic Distribution (NAD) involves distribution over a sub–plurality of a plural expression: (1a) can be true in a situation where each pair of shoes costs $50. This is a NAD reading: distribution does not happen down to all the atomic individuals in the denotation of the plural shoes. Similarly, (1b) is only true under a NAD reading: Rodgers and Hammerstein wrote a musical together, and Rodgers and Hart wrote another musical together, but they never wrote a musical neither all three together nor separately.

(1) a. The shoes cost $50. [Lasesohn 1995]
   b. Rodgers, Hammerstein and Hart wrote operas. [Gillon 1987]

**GOAL.** The goal of this paper is to provide an account of the conditions under which a quantifier has access to the relevant sub-groupings of a plural expression and allows for a NAD reading. These are cases like (2), which have not received much attention in the literature even though they are problematic for virtually every theory of distributivity. These are the facts: a plural NP can be quantified over without losing the relevant NAD reading only in some cases. The sentences in (2a) can be true of pairs of shoes, but the ones in (2b) cannot be true of pairs of composers.

(2) a. {All the/Most/No/} shoes cost $50. ✓ pairs of shoes
   b. {All the/Most/No/} composers wrote operas. *R&H and R&H

**CLAIM.** The paper claims that NAD readings of plural quantified NPs can be explained by means of a context–independent notion of “typicality”. Given what we know about the world, the nature of the contrast in (2) is due to the fact that shoes, but not composers, can form Typical Groups, conceptually salient sub–pluralities of atomic individuals that are relevant for the grammar.

**BACKGROUND.** The domain $E$ of individuals is closed under a sum formation operation ‘⊕’ denoted by a pluralization operation *∗*, and the $i$(individual)-sums are partially ordered through a relation $≤$ on $E$: $a ≤ b$ iff $a ⊕ b = b$. For any 1-place predicate $P$, $[∗P]$ is a complete join-subsemilattice in $E$ generated by operating over atoms $At$, where $At(x) = −∃y[y < x]$ (atoms are entities with no proper parts). Distributive readings are derived by means of a distributive operator $D$ (Link 1983 a.o.; defined in (3)). The difference between collective and distributive readings is captured by the optional application of $D$ on the VP (4).

(3) a. $D ≡ λPλx∀y[y ∈ x → P(y)]$
   b. $D P ↔ ∀x(Px → At(x))$

(4) a. DISTRIBUTIVE: $D(∗cost−$50$)(s_1 ⊕ s_2) ≡ (cost−$50$)(s_1) ∧ (cost−$50$)(s_2)$
   b. COLLECTIVE: $D(∗cost−$50$)(s_1 ⊕ s_2) ≈ ⟨cost−$50$⟩(s_1) ∧ (cost−$50$)(s_2)$

Because there are only atoms and $i$-sums in $E$, $D$ cannot distribute over sub–pluralities. To account for NAD readings, Gillon (1987) introduced covers, sets of sets of individuals that restrict the domain of $D$ by functioning as context–dependent domain selection variables (as defined in (5)). Covers derive NAD readings by picking the relevant non–atomic individuals directly from the context (6):

(5) a. $Cov(X, Y) ≡ \emptyset \notin X \land \forall x \in X \exists y \in Y [x \lessdot y]$
   b. $C D ≡ λPλx∀y[y ∈ Cov^n ∧ y ≤ x → P(y)]$ [Schwarzschild 1996]

(6) a. $[(1)] = ∀y[y ∈ Cov^n ∧ y ∈ [shoes] → [cost−$50$](y)]$
   b. i. $Cov^1 = \{\{a, b\}, \{c, d\}, \{e, f\}\}$
   ii. $Cov^2 = \{\{a\}, \{b\}, \{e\}, \{d\}, \{c\}, \{f\}\}$
   iii. $Cov^3 = \{\{a, b, c, d, e, f\}\}$

**THE PROBLEM.** Current theories of distributivity cannot account for the following two facts: first, not all NAD interpretations are context-dependent. Out of the blue, (2a) is true of pairs of shoes, and so NAD is context–independent here, but even knowing that (1b) is true does not help in rescuing a NAD interpretation for (2b). Thus, NAD readings over a quantified expression cannot be triggered by context alone. Second,
only context–independent NAD can survive quantification, as in (2a), but not in (2b). These two facts point out that the NAD interpretations of composers and shoes are fundamentally different.

**Proposal.** I suggest that the distinction between the two kinds of sub–pluralities in (2) is that **pairs of composers** are contextually salient, whereas **pairs of shoes** are conceptually salient: to know what a shoe is is to know that they usually come in pairs. These salient sub–pluralities can be arranged in Typical Groups (TGs), pluralities whose cardinality is established by world–knowledge and not by the context. By virtue of being conceptually salient, TGs can denote impure atoms (Landman 2000) and, as such, they are visible for composers being conceptually salient, TGs can denote impure atoms (Landman 2000) and, as such, they are visible for

**Rationale.** A concept $\mathcal{C}$ has a probabilistic structure that measures the amount of properties that are true of an object $\text{obj}$ and evaluates whether they suffice for $\text{obj}$ to fall under some concept $\mathcal{C}$, This allows us to evaluate the degree to which $\text{obj}$ is representative of $\mathcal{C}$, and also to compare two objects in order to determine which one satisfies more properties relative to $\mathcal{C}$ (e.g., apple > coconut relative to FRUIT; Margolis & Laurence 2014).

**Assumptions.** I suggest that similar mechanisms may be applied to the domain of pluralities to account for NAD readings in (2a). TGs are defined in terms of a function $\Theta_{\mathcal{C}} : \mathbb{N} \rightarrow [0, 1]$, which maps cardinalities $n$ of objects instantiating a concept $\mathcal{C}$ onto real numbers in $[0, 1]$. $\Theta_{\mathcal{C}}$ allows us to compare different cardinalities of the same object with respect to some concept $\mathcal{C}$. It follows that some cardinalities can (but need not) be more typical than others; in the case of shoe, shoes in twos are more typical than shoes in threes, and so $\Theta_{\text{SHOE}(2)} > \Theta_{\text{SHOE}(3)}$. TGs are conceptually salient pluralities in the lattice structure of a plural NP (example in gray in Figure 1 for shoes), and can be modeled in a mereological system by means of notions like atomicity and parthood relations (e.g., Champollion 2014).

**Application.** The key idea is that TGs are relevant for the grammar (cf. Kerem et al. 2009 for reflexives), as stated in (7), which is formally implemented by the TYPG function in (8).

7. **Typical Cardinality Hypothesis:** A (sub–)plurality denoting NP with a typical cardinality $n$ of more than 1 can denote impure atoms by grouping $n$ atoms together.

8. **Typical Group Function:** given a plural predicate $\oplus$ and a individual–part $x$ of $\ominus P$,

\[ \mathrm{TYPG}(x) = \forall x[ x \leq \ominus P \land | x | = n \land \Theta_{P}(n) > \Theta_{P}(1) \rightarrow \exists y(y \neq x \land x = \uparrow (y)) ] \]

According to (8), TYPG forms groups $\uparrow (y)$ (after Landman 2000) of sub–pluralities $x$ of $P$s such that: the cardinality of $x$ associated to the concept expressed by $P$ is more typical than an atom of $P$ (i.e., than cardinality 1). The value attributed by $\Theta_{\#}$ to 1 is the “golden standard”, and so it is possible that there is more than one $x$ ranking higher than 1, as it seems to be the case:

9. All the Matryoshka dolls fit well together. **$\sim$ all the dolls in a single set of dolls fit well together**

If no such $x$ exists, the premise in (8) does not hold and no TG can be formed. If $x$ exists, a TG is available out of the blue, and so the grammar has access to three types of individuals: two atoms (pure and impure) and i-sums. A quantifier can pick either atom and derive the correct meanings for sentences like **all the shoes cost $50**; there is no need for $D$, and distributivity is still atomic (Winter 2001). For $\{\text{shoes}\} = \{a, b, c, d, e, f\}$,

\begin{align*}
\text{(10) a.} & \forall x [\star \text{shoe}(x) \rightarrow \star \text{cost-$50$}(x)] \\
\text{b.} & \forall x [\star \text{shoe}(\text{TYPG}(x)) \land \star \text{cost-$50$}(\text{TYPG}(x))], \\
& \Rightarrow \forall z [\star \text{shoe}(z) \land | z | = 2 \rightarrow \exists y(y \neq z \land z = \uparrow (y)) \land \star \text{cost-$50$}(\uparrow (y))] \\
& \text{because $\Theta_{\text{SHOE}(2)} > \Theta_{\text{SHOE}(1)}$, $\uparrow$ applies to pairs: $\uparrow (a \oplus b) = \uparrow (c \oplus d) = \uparrow (e \oplus f) = 1$.}
\end{align*}

**Conclusions:** Three important facts about NAD readings have been established: (i) sentences of the form $Q$-Det $A$ $B$ are predicted to have NAD readings only if $A$ can form a TG, (ii) NAD over TGs is context–independent, and (iii) Q-Dets must have access to TGs. The general theory of distributivity is relieved from the need to account for cases like (2a), since the reason why NAD can survive quantification is independent from any version of $D$. 