An impossibility theorem of aggregating semantic rankings under non-monotonic quantification

Claire RONG

March 2026

Contents

1 Introduction

Previous work on universal quantifiers with scalar items in their scope examined graded acceptability judgments for a sentence matched against different situations and modeled these judgments using mixed-effects regression. For instance, Rong (2025) did so with sentences like "each box contains books", considering bare plurals to be a scalar item with a literal *at least one* denotation that may be pragmatically strengthened to *at least two*. The statistical model that best predicted the data had the form:

$$\text{response} \sim c_{\text{vrf}} + c_{\text{lit}} + c_{\text{str}} + (1 \mid \text{participant})$$

where c_{lit} and c_{str} are binary variables encoding whether given semantic readings (respectively, literal and strong) are satisfied in the depicted situation, and c_{vrf} measures the number of *strong verifiers* in the situation, where a strong verifier is an individual of the quantificational domain verifying the pragmatically strengthened meaning of the predicate. Intuitively, in the case of the universal quantifier, the greater c_{vrf} is, the closer we are to the 'best' scenario. Crucially, the best-fitting model suggests that felicity judgments reflect both satisfaction

1

of discrete semantic readings (c_{lit} and c_{str}) and gradient proximity to a 'best' scenario (c_{vrf}, or rather its complement to the size of the domain).

We ask whether this picture extends to sentences with scalar items under the non-monotonic quantifier *exactly n*, e.g. "exactly two boxes contain books". Unlike in the monotonic case, situations can satisfy the locally strengthened interpretation (exactly two boxes contains several) without satisfying the literal one (exactly two...at least one), for instance, when exactly two boxes contain several books while other boxes contain exactly one. Moreover, greater proximity to an intuitive 'best' case may also move a situation closer to falsity by overshooting the exact cardinality requirement. Against this background, we consider two competing but independently motivated conjectures about how felicity judgments should be organized in the non-monotonic case.

The first conjecture, formalized below as the *trivalent distance-based ordering*, assumes a partition with no readings involved: situations make the sentence true, or false, or undefined. Undefined situations are ordered by their distance to the closest situation that makes the sentence true. The second conjecture, formalized below as the *reading-satisfaction-based ordering*, assumes that the sentence gives rise to several distinct bivalent readings. On this view, a situation is judged better with an increasing number of readings that it verifies, and assumes that the literal reading is more salient than the locally strengthened one (based on empirical data, see footnote 3).

A speaker's felicity judgment may well correspond to a cognitive unifying ranking over situations, integrating both types of rankings. The aggregation problem concerns how a single speaker's judgments might reconcile the rankings generated by multiple semantic theories, when the speaker's hypothesis space contains several plausible analyses and the linguistic input available during acquisition does not uniquely determine the correct one (for an example of such a phenomenon in syntax, see Amaral and Roeper, 2014). It asks whether a speaker who implicitly entertains several plausible semantic analyses at once can have a principled way to form a single felicity judgment that reconciles the induced rankings while satisfying a small set of natural consistency constraints. To address this question, we adopt a strategy inspired by Arrow-style impossibility results in social choice theory (Arrow, 1951), which have been applied beyond preference aggregation (e.g. Zwart and Franssen, 2007; Morreau, 2010; Stegenga, 2013; Süskind, 2026). In the analogy with social choice, the two semantic theories play the role of "voters", each inducing a ranking of situations (the "candidates"). Applying choice-theoretic aggregation results to questions in formal semantics is an approach that, to our knowledge, has not previously been explored. As noted in the literature on the non-monotonic quantifier *exactly n* (e.g. Gotzner and Benz, 2022), the main theoretical difficulties in semantic predictions arise precisely for $n \geq 2$, independently of the theoretical framework. We therefore restrict our attention to $n \geq 2$. We show that, under minimal and independently motivated constraints on a unifying ranking (Unanimity, Independence, Non-dictatorship), there is no way of combining the reading-satisfaction-based and trivalent distance-based orderings.

2 Situations, readings, and semantic orderings

2.1 Situations

We consider sentences of the form:

> *Exactly n individuals are P.*

where P contains a scalar item. For instance, in a toy scenario where $n = 2$ and the scalar item is a bare plural:

> *Exactly two boxes contain books.*

Intuitively, the situation that seems best described by this sentence is one where exactly two boxes each contain at least two books, and all other boxes are empty.

Formally, each individual in the domain can stand in one of three relevant relations to the predicate P:

- **F** (falsifier): the individual does not satisfy P. Here, a falsifier is an empty box, if we assume, for simplicity, that a box cannot contain any object other than a book.

- **W** (weak verifier): the individual satisfies the literal meaning of P but not the strengthened meaning. Here, the literal meaning of a bare plural is ≥ 1 and the strengthened meaning is > 1. Therefore, a weak verifier is a box that contains contains exactly one book.

- **S** (strong verifier): the individual satisfies the strengthened meaning of P. Here, a strong verifier is a box that contains several (i.e. at least two) books.

A *situation* is an assignment of one of these values to each individual in the domain. For instance, if we imagine that there are a total of four boxes in the toy scenario, situations are represented by strings of length four over the alphabet {F,W,S}. We reason about situations up to permutation. For example:

- FFSS denotes a situation in which exactly two boxes contain several books and the other two are empty;

- FWSS denotes a situation in which one box is empty, one contains exactly one book, and two contain several books;

- FWWW denotes a situation in which three boxes contain exactly one book and the last box is empty.

To make the structure of the situation space more concrete, Figure 1 represents the case of four individuals in the quantificational domain.

Adjacent vertices correspond to one minimal change in verifier status. This allows us to visualize proximity to the intuitively 'best' case (here FFSS).
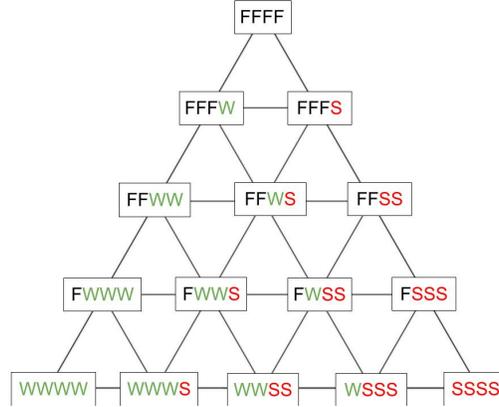
Figure 1: Situation space for four individuals.

There are several caveats. First, we use the distance corresponding to the minimal number of verifier-status changes, for expository simplicity. Other metrics could be used, and different verifier statuses might be weighted unequally. Second, nothing in the argument requires that all vertices be present, but in our proofs we will assume the existence of all vertices. Third, the triangular shape reflects the assumption of a fixed finite domain of four individuals. For larger (possibly infinite) domains, the structure extends naturally by expanding the base of the triangle.

2.2 Trivalent ordering with distance to the best case

We define a first ordering, called *trivalent distance-based ordering*, based on the trivalent framework of Strong Kleene (semantically implemented in e.g. Bassi et al. (2021); Križ (2016)) combined with a notion of distance.

Definition 1. *Each situation s is assigned a trivalent truth-value:*

- **T** *(true) if exactly n individuals are S and all others are F.*

- **F** *(false) if either less than n individuals in total are S or W, or than more than n individuals in total are S or W and among them not exactly n individuals are S;*

- **U** *(undefined) otherwise.*

Definition 2. *Define a weak relation $\succeq_t$ on situations by $s_1 \succeq_t s_2$ iff:*
(i) $V(s_1) \geq V(s_2)$, where the order on truth-values is $\mathbf{T} > \mathbf{U} > \mathbf{F}$,
(ii) when both s_1 and s_2 are $\mathbf{U}$, s_1 is closer than s_2 to the closest $\mathbf{T}$-situation.[1]

[1] The choice of metric is not crucial, what matters is that situations that make the sentence undefined can be ordered by proximity to a best case. Also note that both orderings rank all false situations at the same level. Although it would be natural to assume that 'all these

4

Definition 3. *The strict relation $\succ_t$ is defined by:*

$$s_1 \succ_t s_2 \quad \textit{iff} \quad s_1 \succeq_t s_2 \textit{ and not } s_2 \succeq_t s_1.$$

The induced equivalence relation $\approx_t$ is defined by:

$$s_1 \approx_t s_2 \quad \textit{iff} \quad s_1 \succeq_t s_2 \textit{ and } s_2 \succeq_t s_1.$$

The choice of metric is not crucial, what matters is that situations that make the sentence undefined can be partially ordered by proximity to a best case.

Also note that both orderings rank all false situations at the same level. Although it would be natural to assume that 'all these situations are false, but some are more false than others' (for instance, ranking by their distance to the closest situation that makes at least one reading true), this refinement is not necessary for any of the proofs here.

2.3   Reading-satisfaction-based ordering

We now define a second ordering, which we will call reading-satisfaction-based ordering. To define it, we assume three bivalent readings associated with a sentence:

- the *literal reading*, satisfied iff exactly n individuals are at least weak verifiers (W or S);

- an *intermediate reading*, satisfied iff exactly n individuals are at least weak verifiers (W or S) and it is not the case that all of the n individuals are weak verifiers;

- a *strong (i.e. local) reading*, satisfied iff exactly n individuals are strong verifiers (S), with no requirement that the remaining individuals be falsifiers.

Let $\#(s)$ denote the number of bivalent readings verified by a situation s.

Definition 4 (Reading-satisfaction ranking)**.** *Let s_1, s_2 be two situations. Define the weak reading-satisfaction relation $\succeq_r$ by:*

$$s_1 \succeq_r s_2 \quad \textit{iff} \quad \begin{cases} \textit{(i)} \quad \#(s_1) > \#(s_2), \\ \textit{or} \\ \textit{(ii)} \quad \#(s_1) = \#(s_2), \\ \qquad \textit{and if } \#(s_1) = 1, \textit{ then } s_1 \textit{ verifies the literal reading} \\ \qquad \textit{whenever } s_2 \textit{ verifies the local reading.} \end{cases}$$

situations are false, but some are more false than others', this refinement is not necessary for any of the proofs here.

Clause (i) implements the idea that situations satisfying more readings are preferred.[2] This idea has served as the basis for much experimental work on exhaustification in the scope of quantifiers (e.g. Chemla and Spector, 2011; Stateva et al., 2016). Clause (ii) resolves ties by giving priority to the literal reading over the strong local one. This reflects an independently motivated salience assumption, supported by introspection and experimental data (on universal as well as non-monotonic quantifiers), but is made explicit here as part of the definition. [3]

Definition 5. *The strict relation $\succ_r$ is defined by:*

$$s_1 \succ_r s_2 \quad \textit{iff} \quad s_1 \succeq_r s_2 \textit{ and not } s_2 \succeq_r s_1.$$

The induced equivalence relation $\approx_r$ is defined by:

$$s_1 \approx_r s_2 \quad \textit{iff} \quad s_1 \succeq_r s_2 \textit{ and } s_2 \succeq_r s_1.$$

2.4 Constraints on unifying rankings

Let n, m such that $2 \leq n < m$, let S be the corresponding situation space. Let $\mathcal{O}$ denote the set of all rankings over S[4], where a *ranking* is understood as a total preorder (i.e. a reflexive, transitive, and connected relation). A *unifying function* is a function $F : \mathcal{O} \times \mathcal{O} \to \mathcal{O}$ which takes as input two rankings (here the reading-satisfaction ranking $\succeq_r$ and the trivalent ranking $\succeq_t$) and returns a single ranking over situations.

A *unifying function* is a function

$$F : \mathcal{O} \times \mathcal{O} \longrightarrow \mathcal{O},$$

which takes as input two rankings (here the reading-satisfaction ranking $\succeq_r$ and the trivalent ranking $\succeq_t$) and returns a single ranking over situations.

Definition 6 (Unanimity, Independence, Non-dictatorship)**.** *A unifying function F satisfies the following conditions:*

---

[2]The reason why we chose to present ranking $\succ_t$ before ranking $\succ_r$ is in order for readers to not be confused about why we did not define the trivalent truth conditions using the bivalent readings. The primary motive is that the two orderings have different conceptual fondations. However, it would be logically equivalent to rewrite $\mathbf{T} > \mathbf{U} > \mathbf{F}$ using the bivalent readings. For each situation, we consider whether it satisfies all ($\mathbf{T}$), some but not all ($\mathbf{U}$), or none ($\mathbf{F}$) of the three bivalent readings. To order two situations, $\mathbf{T} > \mathbf{U} > \mathbf{F}$ could then be rewritten informally as *all* readings > *some but not all* readings > *no* readings.

[3]Readers interested in the data on universal quantifiers can refer to Rong (2025), and more precisely, to the relative weights of coefficients in the best model fitting the experimental data of either bare plurals or *some NPs*. Our data relative to non-monotonic quantifiers only comes, so far, from non-preregistered pilot experiments.

[4]Strictly speaking, the situation space and the set of rankings depend on n and m. We would therefore write $S_{n,m}$ and $\mathcal{O}_{n,m}$, but we suppress these subscripts for readability, as they play no role in the argument.

- **Unanimity.** *For all situations* φ, χ*, if*

$$\varphi \succ_r \chi \quad and \quad \varphi \succ_t \chi,$$

  then

$$\varphi \succ_F \chi,$$

  where $\succ_F$ *is the strict relation induced by the ranking* $F(\succeq_r, \succeq_t)$*.*

- **Independence.** *For all situations* φ, χ *and for all pairs of rankings* $(\succeq_r, \succeq_t)$ *and* $(\succeq'_r, \succeq'_t)$*, if* φ *and* χ *have the same relative ordering under* $\succeq_r$ *and* $\succeq_t$ *as under* $\succeq'_r$ *and* $\succeq'_t$*, then they have the same relative ordering under the corresponding unifying rankings:*

$$F(\succeq_r, \succeq_t)|_{\{\varphi,\chi\}} = F(\succeq'_r, \succeq'_t)|_{\{\varphi,\chi\}}.$$

  That is, the ranking of φ *and* χ *under* F *depends only on their pairwise rankings under* $\succeq_r$ *and* $\succeq_t$*.*

- **Non-dictatorship.** *It is not the case that* F *always follows one input ranking. In particular, it is not the case that:*

$$\forall \varphi, \chi \ (\varphi \succ_r \chi \Rightarrow \varphi \succ_F \chi),$$

  nor that:

$$\forall \varphi, \chi \ (\varphi \succ_t \chi \Rightarrow \varphi \succ_F \chi).$$

These conditions are minimal.[5] Unanimity ensures that the unifying ranking respects strict agreement between the two semantic perspectives. Independence restricts characterization to pairwise information, and Non-dictatorship rules out trivial resolutions that simply privilege one ordering across the board.

3 Impossibility result

While the Lemma and the Theorem generalize to any $n \geq 2$, note that they do not hold for $n = 1$.[6]

---

[5]A standard assumption in social choice theory is that of *Unrestricted Domain*, according to which an aggregation function must be defined for all logically possible profiles of individual rankings. In our setting, this would amount to requiring that every combinatorially possible way of ordering situations by the reading-satisfaction-based ordering and by the trivalent distance-based ordering be instantiated.

We do not make this assumption and our results do not rely on unrestrictedness. As in related work applying Arrow-style arguments outside of preference aggregation, it suffices that a small number of structurally simple profiles are instantiated for the theorem proof to hold.

[6]For $n = 1$, the intermediate and strong readings coincide, therefore it represents a degenerate case for the ranking $\succeq_r$. It has been observed in the literature on non-monotonic quantifiers (e.g. Gotzner and Benz, 2022) that the main theoretical difficulties arise precisely for $n \geq 2$, since the mechanisms available for $n = 1$ do not generalize in a straightforward way, independently of the semantic framework adopted. For this reason, we restrict attention to $n \geq 2$ in what follows.

Lemma. *One of these two profiles is always instantiated:*

$$\langle \phi \succ_r \chi \succ_r \psi \approx_r \omega, \phi \approx_t \psi \succ_t \chi \approx_t \omega \rangle \tag{P_1}$$

$$\langle \phi \approx_r \chi \succ_r \psi \succ_r \omega, \phi \approx_t \psi \succ_t \chi \approx_t \omega \rangle \tag{P_2}$$

Proof. Let m denote the (possibly infinite) cardinality of the quantificational domain. (P_1) is realized only when $n = 2$ (regardless of m), by taking:

$$\phi = \text{FFWS}, \quad \chi = \text{FFWW}, \quad \psi = \text{FWSS}, \quad \omega = \text{WWSS}$$

When (and only when) $n = 2$, under the reading-satisfaction view, ϕ satisfies the literal and the intermediate readings, χ satisfies only the literal reading, ψ and ω satisfy only the local reading. Under the trivalent view, all four situations are undefined.

For the cases $n > 2$, we first give here the example of $n = 3$, for ease of exposition. For $n = 3$, (P_2) is realized by taking:

$$\phi = \text{FWSS}, \quad \chi = \text{FWWS}, \quad \psi = \text{WSSS}, \quad \omega = \text{WWSS}$$

Under the reading-satisfaction view, ϕ and χ satisfy the literal and the intermediate readings, ψ satisfies only the local reading, ω does not satisfy any readings. Under the trivalent view, ϕ, χ and ψ are undefined, and ω is false.

Figure 2 illustrate the cases $n = 2$ and $n = 3$ when $m = 4$.

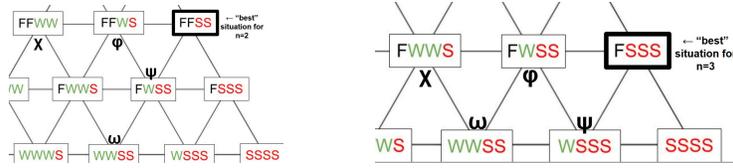

Figure 2: (P_1) for $n = 2$ and (P_2) for $n = 3$ in the situation space when $m = 4$.

Generally for $n > 2$, the corresponding situations are expressed by:

$$\phi = \underbrace{\text{F...F}}_{m-n}\text{W}\underbrace{\text{S...S}}_{n-1}, \quad \chi = \underbrace{\text{F...F}}_{m-n}\text{WW}\underbrace{\text{S...S}}_{n-2}, \quad \psi = \underbrace{\text{W...W}}_{m-n}\underbrace{\text{S...S}}_{n}, \quad \omega = \underbrace{\text{W...W}}_{m-n+1}\underbrace{\text{S...S}}_{n-1}$$

$\square$

Theorem. *There exists no unifying function F from $\succeq_r$ and $\succeq_t$ to a total preorder $\succeq$ over situations that satisfies Unanimity, Independence and Non-dictatorship.*

Proof. Suppose there exists a unifying function $F : (\succeq_r, \succeq_t) \mapsto \succeq_F$ which satisfies *Unanimity* and *Independence*. Our goal is to investigate which further constraints these two axioms impose on the form of F.

A *situation-profile* is an ordered pair $P = \langle \succeq_r, \succeq_t \rangle$ of rankings over the set of situations. For ease of notation, we write $F(P) = (\succeq, \succ, \approx)$ for the ranking

returned by F at profile P, and we henceforth omit the subscript F on the induced order relations whenever no confusion can arise.

The strategy of the proof is standard in social-choice theory. By *Independence*, the ranking that F assigns to any pair of situations depends only on how that pair is ranked by the two input orderings. Therefore, the behavior of F is fully determined by its action on *two-situation profiles*. We therefore fix two situations φ and χ, and consider all possible ways in which they can be comparatively ranked by the reading-satisfaction-based ordering $(\succ_r, \approx_r)$ and the trivalent ordering $(\succ_t, \approx_t)$. We begin by eliminating cases that are already settled.

First, we need not consider profiles in which the two orderings agree strictly on the direction of preference, i.e. profiles of the form

$$\langle \varphi \succ_r \chi, \ \varphi \succ_t \chi \rangle$$

By Unanimity, any unifying ranking must satisfy $\varphi \succ \chi$ in this case. This outcome is consistent with a dictatorship of either ordering and therefore places no constraint on the remaining cases.

Second, we need not consider profiles in which both orderings rank the two situations as equal, i.e. profiles of the form

$$\langle \varphi \approx_r \chi, \ \varphi \approx_t \chi \rangle$$

Any unifying ranking that sets $\varphi = \chi$ in this case is again compatible with a dictatorship of either ordering and has no bearing on the rest of the argument.

The only profiles that matter are therefore the remaining three:

A. $\langle \varphi \succ_r \chi, \ \chi \succ_t \varphi \rangle$, **B.** $\langle \varphi \succ_r \chi, \ \varphi \approx_t \chi \rangle$, **C.** $\langle \varphi \approx_r \chi, \ \varphi \succ_t \chi \rangle$

For each of these profiles, the unifying ranking must select one of the following three outcomes:

1. $\varphi \succ \chi$, **2.** $\chi \succ \varphi$, **3.** $\varphi = \chi$

Thus, there are $3^3 = 27$ formally possible ways of defining a unifying ranking on these profiles.

We write F_{A1} as a shorthand for any unifying function that maps profile A to outcome 1, and analogously for profiles B and C and outcomes 2 and 3. The remainder of the proof consists in showing that, once Unanimity and Independence are imposed, all but one of these 27 possibilities either lead to inconsistency or to a dictatorship of $\succ_r$ or $\succ_t$.

We proceed by a case distinction on A1, A2, and A3. Suppose that F maps A to 1. Note that any unifying ranking combining A1 with B1 would be a $\succ_r$-dictatorship. Therefore we need only consider the combinations of A1 with B2 and B3. We show that any combination involving A1 is incompatible with B2 or B3. For this, consider the two profiles described in the Lemma:

$\langle \phi \succ_r \chi \succ_r \psi \approx_r \omega, \phi \approx_t \psi \succ_t \chi \approx_t \omega \rangle$ and $\langle \phi \approx_r \chi \succ_r \psi \succ_r \omega, \phi \approx_t \psi \succ_t \chi \approx_t \omega \rangle$

We have:
$$F_{A1}(\langle \chi \succ_r \psi, \psi \succ_t \chi \rangle) = (\chi \succ \psi) \tag{1}$$

$$F_{B2 \cup B3}(\langle \phi \succ_r \psi, \phi \approx_t \psi \rangle) = (\psi \succeq \phi) \tag{2}$$

$$F_{B2 \cup B3}(\langle \chi \succ_r \omega, \chi \approx_t \omega \rangle) = (\omega \succeq \chi) \tag{3}$$

By transitivity:

- (1) and (2) give $F_{A1,B2 \cup B3}(P_1) = (\chi \succ \psi \succeq \phi)$, contradicting $\phi \succ \chi$ which follows from Unanimity in (P_1);

- (1) and (3) give $F_{A1,B2 \cup B3}(P_2) = (\omega \succeq \chi \succ \psi)$, contradicting $\psi \succ \omega$ which follows from Unanimity in (P_2).

Therefore, none of the nine combinations involving A1 is possible.

Suppose now that F maps A to 2. Note that any unifying ranking combining A2 with C1 would be a $\succ_t$-dictatorship. Therefore we need only consider the combinations of A2 with C2 and C3.

We first show that B2-C2, B2-C3 and B3-C2 are incompatible combinations, no matter what A is mapped to. This result will be useful again later, for eliminating cases under A3.

We have:
$$F_{C2}(\langle \psi \approx_r \omega, \psi \succ_t \omega \rangle) = (\omega \succ \psi) \tag{4}$$

$$F_{C2}(\langle \phi \approx_r \chi, \phi \succ_t \chi \rangle) = (\chi \succ \phi) \tag{5}$$

By transitivity:

- (2) and (4) give $F_{B2 \cup B3,C2}(P_1) = (\omega \succ \psi \succeq \phi)$, contradicting $\phi \succ \omega$ which follows from Unanimity in (P_1);

- (3) and (5) give $F_{B2 \cup B3,C2}(P_2) = (\omega \succeq \chi \succ \phi)$, contradicting $\phi \succ \omega$ which follows from Unanimity in (P_2).

Likewise, we have:
$$F_{B2}(\langle \phi \succ_r \psi, \phi \approx_t \psi \rangle) = (\psi \succ \phi) \tag{6}$$

$$F_{B2}(\langle \chi \succ_r \omega, \chi \approx_t \omega \rangle) = (\omega \succ \chi) \tag{7}$$

$$F_{C3}(\langle \psi \approx_r \omega, \psi \succ_t \omega \rangle) = (\omega = \psi) \tag{8}$$

$$F_{C3}(\langle \phi \approx_r \chi, \phi \succ_t \chi \rangle) = (\chi = \phi) \tag{9}$$

By transitivity:

- (6) and (8) give $F_{B2,C3}(P_1) = (\omega = \psi \succ \phi)$, also contradicting $\phi \succ \omega$;

- (7) and (9) give $F_{B2,C3}(P_2) = (\omega \succ \chi = \phi)$, also contradicting $\phi \succ \omega$.

Among the combinations involving A2, the possible ones remaining are therefore B1-C2, B1-C3 and B3-C3. We have:

$$F_{A2}(\langle \chi \succ_r \psi, \psi \succ_t \chi \rangle) = (\psi \succ \chi) \tag{10}$$

$$F_{B1}(\langle \chi \succ_r \omega, \chi \approx_t \omega \rangle) = (\chi \succ \omega) \tag{11}$$

$$F_{B1}(\langle \phi \succ_r \psi, \phi \approx_t \psi \rangle) = (\phi \succ \psi) \tag{12}$$

$$F_{B3}(\langle \chi \succ_r \omega, \chi \approx_t \omega \rangle) = (\chi = \omega) \tag{13}$$

$$F_{B3}(\langle \phi \succ_r \psi, \phi \approx_t \psi \rangle) = (\phi = \psi) \tag{14}$$

By transitivity:

- (4), (8) and (11) give $F_{B1,C2\cup C3}(P_1) = (\chi \succ \omega \succeq \psi)$, which contradicts (10);

- (5), (9) and (12) give $F_{B1,C2\cup C3}(P_2) = (\chi \succ \phi \succeq \psi)$, also contradicting (10);

- (8) and (13) give $F_{B3,C3}(P_1) = (\chi = \omega = \psi)$, also contradicting (10);

- (9) and (14) give $F_{B3,C3}(P_2) = (\chi = \phi = \psi)$, also contradicting (10).

Therefore, none of the nine combinations involving A2 is possible.

Suppose now that F maps A to 3. As shown before, we do not need to consider B2-C2, B2-C3 and B3-C2, as they are incompatible combinations. In fact, as we will now show, any unifying ranking involving A3 in incompatible with B3. We have:

$$F_{A3}(\langle \chi \succ_r \psi, \psi \succ_t \chi \rangle) = (\psi = \chi) \tag{15}$$

$$F_{B3}(\langle \phi \succ_r \psi, \phi \approx_t \psi \rangle) = (\psi = \phi) \tag{16}$$

$$F_{B3}(\langle \chi \succ_r \omega, \chi \approx_t \omega \rangle) = (\omega = \chi) \tag{17}$$

By transitivity:

- (15) and (16) give $F_{A3,B3}(P_1) = (\phi = \psi = \chi)$, contradicting $\phi \succ \chi$ which follows from Unanimity in (P_1);

- (15) and (17) give $F_{A3,B3}(P_2) = (\psi = \chi = \omega)$, contradicting $\psi \succ \omega$ which follows from Unanimity in (P_2).

Among the combinations involving A3, the possible ones remaining are therefore B1-C1, B1-C2, B1-C3 and B2-C1. We have:

$$F_{B2}(\langle \chi \succ_r \omega, \chi \approx_t \omega \rangle) = (\omega \succ \chi) \tag{18}$$

$$F_{B2}(\langle \phi \succ_r \psi, \phi \approx_t \psi \rangle) = (\psi \succ \phi) \tag{19}$$

$$F_{C1}(\langle \psi \approx_r \omega, \psi \succ_t \omega \rangle) = (\psi \succ \omega) \tag{20}$$

$$F_{C1}(\langle \phi \approx_r \chi, \phi \succ_t \chi \rangle) = (\phi \succ \chi) \tag{21}$$

By transitivity:

- (4), (8) and (11) give $F_{B1,C2\cup C3}(P_1) = (\chi \succ \omega \succeq \psi)$, which contradicts (15);

- (5), (9) and (12) give $F_{B1,C2\cup C3}(P_2) = (\chi \succeq \phi \succ \psi)$, also contradicting (15);

- (18) and (20) give $F_{B2,C1}(P_1) = (\psi \succ \omega \succ \chi)$, also contradicting (15);

- (19) and (21) give $F_{B2,C1}(P_2) = (\psi \succ \phi \succ \chi)$, also contradicting (15).

The only remaining possible mapping for F is A3-B1-C1. So far, we have shown uniqueness of a unifying function under condition of existence. Moreover, if such a function exists, then for all situations φ, χ, the resulting ranking $\succeq_F$ satisfies:

- if $\varphi \succ_r \chi$ and $\chi \succ_t \varphi$, then $\varphi \approx_F \chi$;

- if $\varphi \succ_r \chi$ and $\varphi \approx_t \chi$, then $\varphi \succ_F \chi$;

- if $\varphi \approx_r \chi$ and $\varphi \succ_t \chi$, then $\varphi \succ_F \chi$.

In other words, the resulting unifying ranking implements a form of mutual veto: whenever the two orderings disagree, neither is permitted to override the other; each ordering can only determine the outcome when the other is neutral.

To determine whether this unique candidate does define a total preorder, we test it on full situation spaces. Recall that for each pair of integers n, m with $2 \leq n < m$, the corresponding situation space is $S_{n,m}$, though we omitted the subscript. The candidate aggregation rule derived above is defined purely in terms of the pairwise comparisons delivered by $\succeq_r$ and $\succeq_t$, and therefore applies uniformly to every such $S_{n,m}$. For each pair (n, m), we computationally generate in Python all situations in $S_{n,m}$ up to permutation, compute the two semantic orderings they induce, and apply the candidate unifying rule pairwise across situations (see code in appendix).[7]. The resulting relation is then checked against the requirements of a ranking (being reflexive, complete, and transitive). We conduct this verification for all domain sizes $3 \leq m \leq 10$ and $2 \leq n < m$. It appears that, except for the cases with $n = 2$, the induced relation fails to be transitive. This confirms the impossibility result. [8] Tables below (1, 2 and 3) show the full verification results. $\square$

4 Conclusion

Sentences with scalar items in the scope of non-monotonic quantifiers yield diverging predictions across theories. Empirical evidence suggests that speakers' felicity judgments may integrate aspects of two theories, which we model as two rankings. Using a novel combination of social choice-theoretic aggregation

---

[7]With thanks to Benjamin for suggesting this method.

[8]We might add that for $n = 2$, the result could be a *local possibility* result, but we yet have to provide a general proof, or provide a reason why unification becomes impossible for $n \geq 3$.

methods and computational verification, we address the meta-theoretical question of whether these two rankings can be cognitively reconciled into a single ranking underlying speakers' judgments. We find that there exists no unifying function that aggregates uniformly, for all *exactly n* with $n \geq 2$, the reading-satisfaction-based ranking and the trivalent distance-based ranking into a total preorder while satisfying Unanimity, Independence and Non-dictatorship. Thus, if speakers entertain several semantic representations simultaneously, the judgments that would result from a compromise ranking face a structural incompatibility, as they cannot satisfy all three Arrow-style constraints at once.

Several directions remain open. First, our modeling idealizes the structure of the rankings. More refined models could allow weights on readings or alternative notions of proximity to a best case. Second, the impossibility result relies on specific aggregation constraints. Relaxing Independence (the most natural target for relaxation among the three constraints) or the requirement of a total preorder may yield different possibilities for aggregation. More generally, this result does not show that hybrid semantic theories are impossible. Rather, it shows that any such theory must either relax a constraint or abandon the idea that the two rankings are aggregated into a single ranking.

m	n	status	witness
3	2	PASS	–
4	2	PASS	–
4	3	FAIL	(F1-W3-S0, F0-W1-S3, F1-W2-S1)
5	2	PASS	–
5	3	FAIL	(F0-W2-S3, F2-W3-S0, F1-W1-S3)
5	4	FAIL	(F1-W3-S1, F0-W1-S4, F1-W2-S2)
6	2	PASS	–
6	3	FAIL	(F1-W2-S3, F3-W3-S0, F2-W1-S3)
6	4	FAIL	(F0-W2-S4, F2-W3-S1, F1-W1-S4)
6	5	FAIL	(F1-W3-S2, F0-W1-S5, F1-W2-S3)
7	2	PASS	–
7	3	FAIL	(F2-W2-S3, F4-W3-S0, F3-W1-S3)
7	4	FAIL	(F0-W3-S4, F3-W4-S0, F1-W2-S4)
7	5	FAIL	(F0-W2-S5, F2-W3-S2, F1-W1-S5)
7	6	FAIL	(F1-W3-S3, F0-W1-S6, F1-W2-S4)
8	2	PASS	–
8	3	FAIL	(F3-W2-S3, F5-W3-S0, F4-W1-S3)
8	4	FAIL	(F1-W3-S4, F4-W4-S0, F2-W2-S4)
8	5	FAIL	(F0-W3-S5, F3-W4-S1, F1-W2-S5)
8	6	FAIL	(F0-W2-S6, F2-W3-S3, F1-W1-S6)
8	7	FAIL	(F1-W3-S4, F0-W1-S7, F1-W2-S5)
9	2	PASS	–
9	3	FAIL	(F4-W2-S3, F6-W3-S0, F5-W1-S3)
9	4	FAIL	(F2-W3-S4, F5-W4-S0, F3-W2-S4)
9	5	FAIL	(F0-W4-S5, F4-W5-S0, F1-W3-S5)
9	6	FAIL	(F0-W3-S6, F3-W4-S2, F1-W2-S6)
9	7	FAIL	(F0-W2-S7, F2-W3-S4, F1-W1-S7)
9	8	FAIL	(F1-W3-S5, F0-W1-S8, F1-W2-S6)
10	2	PASS	–
10	3	FAIL	(F5-W2-S3, F7-W3-S0, F6-W1-S3)
10	4	FAIL	(F3-W3-S4, F6-W4-S0, F4-W2-S4)
10	5	FAIL	(F1-W4-S5, F5-W5-S0, F2-W3-S5)
10	6	FAIL	(F0-W4-S6, F4-W5-S1, F1-W3-S6)
10	7	FAIL	(F0-W3-S7, F3-W4-S3, F1-W2-S7)
10	8	FAIL	(F0-W2-S8, F2-W3-S5, F1-W1-S8)
10	9	FAIL	(F1-W3-S6, F0-W1-S9, F1-W2-S7)

Table 1: Exhaustive verification results for $3 \leq m \leq 10$ and $2 \leq n < m$.

m	3	4	5	6	7	8	9	10
n	2	2	2	2	2	2	2	2

Table 2: Passed cases.

m	n	witness
4	3	(F1-W3-S0, F0-W1-S3, F1-W2-S1)
5	3	(F0-W2-S3, F2-W3-S0, F1-W1-S3)
5	4	(F1-W3-S1, F0-W1-S4, F1-W2-S2)
6	3	(F1-W2-S3, F3-W3-S0, F2-W1-S3)
6	4	(F0-W2-S4, F2-W3-S1, F1-W1-S4)
6	5	(F1-W3-S2, F0-W1-S5, F1-W2-S3)
7	3	(F2-W2-S3, F4-W3-S0, F3-W1-S3)
7	4	(F0-W3-S4, F3-W4-S0, F1-W2-S4)
7	5	(F0-W2-S5, F2-W3-S2, F1-W1-S5)
7	6	(F1-W3-S3, F0-W1-S6, F1-W2-S4)
8	3	(F3-W2-S3, F5-W3-S0, F4-W1-S3)
8	4	(F1-W3-S4, F4-W4-S0, F2-W2-S4)
8	5	(F0-W3-S5, F3-W4-S1, F1-W2-S5)
8	6	(F0-W2-S6, F2-W3-S3, F1-W1-S6)
8	7	(F1-W3-S4, F0-W1-S7, F1-W2-S5)
9	3	(F4-W2-S3, F6-W3-S0, F5-W1-S3)
9	4	(F2-W3-S4, F5-W4-S0, F3-W2-S4)
9	5	(F0-W4-S5, F4-W5-S0, F1-W3-S5)
9	6	(F0-W3-S6, F3-W4-S2, F1-W2-S6)
9	7	(F0-W2-S7, F2-W3-S4, F1-W1-S7)
9	8	(F1-W3-S5, F0-W1-S8, F1-W2-S6)
10	3	(F5-W2-S3, F7-W3-S0, F6-W1-S3)
10	4	(F3-W3-S4, F6-W4-S0, F4-W2-S4)
10	5	(F1-W4-S5, F5-W5-S0, F2-W3-S5)
10	6	(F0-W4-S6, F4-W5-S1, F1-W3-S6)
10	7	(F0-W3-S7, F3-W4-S3, F1-W2-S7)
10	8	(F0-W2-S8, F2-W3-S5, F1-W1-S8)
10	9	(F1-W3-S6, F0-W1-S9, F1-W2-S7)

Table 3: Failed cases with a non-transitivity witness (x, y, z) such that $x \succeq y$ and $y \succeq z$ but not $x \succeq z$.

References

Luiz Amaral and Tom Roeper. Multiple grammars and second language representation. *Second Language Research*, 2014.

Kenneth Joseph Arrow. *Social choice and individual values*. Wiley, 1951.

Itai Bassi, Guillermo Del Pinal, and Uli Sauerland. Presuppositional exhaustification. *Semantics and Pragmatics*, 14(11):1–48, 2021.

Emmanuel Chemla and Benjamin Spector. Experimental evidence for embedded scalar implicatures. *Journal of semantics*, 28(3):359–400, 2011.

Nicole Gotzner and Anton Benz. Implicatures in (non-) monotonic environments. In *Proceedings of Sinn und Bedeutung*, volume 26, pages 340–358, 2022.

Manuel Križ. Homogeneity, non-maximality, and all. *Journal of semantics*, 2016.

Michael Morreau. It simply does not add up: Trouble with overall similarity. *The Journal of Philosophy*, 107(9):469–490, 2010.

Claire Rong. Plurals under quantification: a comparison of english and mandarin. *ESSLLI 2025 Student Session*, page 125, 2025.

Penka Stateva, Sara Andreetta, and Arthur Stepanov. On the nature of the plurality inference: Ladybugs for anne. *Papers dedicated to Anne Reboul. Lyon: CNRS*, 2016.

Jacob Stegenga. An impossibility theorem for amalgamating evidence. *Synthese*, 190(12):2391–2411, 2013.

Jakob Süskind. An impossibility theorem for biodiversity. *Biology & Philosophy*, 41(1):1, 2026.

Sjoerd D Zwart and Maarten Franssen. An impossibility theorem for verisimilitude. *Synthese*, 158(1):75–92, 2007.

```python
import itertools
import pandas as pd


def generate_states_up_to_permutation(m: int):
    states = []
    for f in range(m + 1):
        for w in range(m + 1 - f):
            s = m - f - w
            states.append({"F": f, "W": w, "S": s})
    return states

def state_id(st):
    return f"F{st['F']}-W{st['W']}-S{st['S']}"

def min_hamming_multiset(a, b, m: int):
    matches = min(a["F"], b["F"]) + min(a["W"], b["W"]) + min(a["S"], b
                                        ["S"])
    return m - matches


# ----------------------------
# Reading-satisfaction ranking
# ----------------------------
def readings(st, n: int):
    nonF = st["S"] + st["W"]
    L = (nonF == n)
    I = (nonF == n) and (st["S"] >= 1)
    S = (st["S"] == n)
    return {"L": L, "I": I, "S": S}

def spv_key(st, n: int):
    r = readings(st, n)
    count = sum(r.values())  # number of true readings
    tie = 0
    if count == 1:
        if r["L"]:
            tie = 2
        elif r["I"]:
            tie = 1
        # else (S) tie stays 0
    return (count, tie)


# ----------------------------
# Trivalent ranking
# ----------------------------
def tri_value(st, n: int, m: int):
    # designated "best" state for T
    if st["S"] == n and st["W"] == 0 and st["F"] == m - n:
        return "T"
    nonF = st["S"] + st["W"]
    if nonF < n:
        return "F"
    if nonF > n and st["S"] != n:
        return "F"
    return "U"

TV_ORDER = {"F": 0, "U": 1, "T": 2}
```

```python
def tri_key(st, n: int, m: int, best):
    return (TV_ORDER[tri_value(st, n, m)], -min_hamming_multiset(st,
                                            best, m))


# ----------------------------
# Pairwise comparison + mutual veto
# ----------------------------
def pair_rel(kx, ky):
    return ">" if kx > ky else "<" if kx < ky else "~"


def unify(rs, rt):
    if rs == rt:
        return rs
    if rs == "~":
        return rt
    if rt == "~":
        return rs
    return "~"


# ----------------------------
# Build final relation from two rankings
# ----------------------------
def build_final_relation(states, n: int, m: int):
    """
    Returns: (ids, rel) where rel[(i,j)] is '>=' if i >= j, else '<'
    using unify(pair_rel(spv), pair_rel(tri)).
    """
    best = {"F": m - n, "W": 0, "S": n}
    ids = [state_id(st) for st in states]

    s_key = {state_id(st): spv_key(st, n) for st in states}
    t_key = {state_id(st): tri_key(st, n, m, best) for st in states}

    rel = {}
    for i in ids:
        for j in ids:
            if i == j:
                rel[(i, j)] = ">="
                continue

            rs = pair_rel(s_key[i], s_key[j])   # '>', '<', '~'
            rt = pair_rel(t_key[i], t_key[j])   # '>', '<', '~'
            out = unify(rs, rt)                 # '>', '<', '~'

            rel[(i, j)] = ">=" if out in (">", "~") else "<"

    return ids, rel


# ----------------------------
# Check whether it's a total preorder
# ----------------------------
def check_total_preorder(ids, rel):
    # Reflexive
    for x in ids:
        if rel.get((x, x)) != ">=":
            return False, ("not reflexive", x)
```

18

```python
    # Total
    for x, y in itertools.permutations(ids, 2):
        if not (rel.get((x, y)) == ">=" or rel.get((y, x)) == ">="):
            return False, ("not total", x, y)

    # Transitive
    for x, y, z in itertools.product(ids, repeat=3):
        if rel.get((x, y)) == ">=" and rel.get((y, z)) == ">=" and rel.\
                                              get((x, z)) != ">=":
            return False, ("not transitive", x, y, z)

    return True, None

# ----------------------------
# Quick single (m,n) demo
# ----------------------------
n = 2
m = 4
states = generate_states_up_to_permutation(m)

ids, rel = build_final_relation(states, n, m)
ok, witness = check_total_preorder(ids, rel)
print("Ranking holds:", ok, "witness:", witness)

best = {"F": m - n, "W": 0, "S": n}
rows = []
for st in states:
    sid = state_id(st)
    rows.append({
        "state": sid,
        "counts": (st["F"], st["W"], st["S"]),
        "#readings": spv_key(st, n)[0],
        "spv_key": spv_key(st, n),
        "V": tri_value(st, n, m),
        "tri_key": tri_key(st, n, m, best),
        "dist": min_hamming_multiset(st, best, m),
    })

df = pd.DataFrame(rows).sort_values(
    by=["#readings", "spv_key", "V", "dist"],
    ascending=[False, False, False, True],
)
print(df)

# ----------------------------
# Test for more (n,m) pairs
# ----------------------------
MAX_M = 10
results = []

for m in range(3, MAX_M + 1):
    states = generate_states_up_to_permutation(m)
    for n in range(2, m):  # n = 2,...,m-1
        ids, rel = build_final_relation(states, n, m)
        ok, witness = check_total_preorder(ids, rel)
        results.append({
```

```python
            "m": m,
            "n": n,
            "status": "PASS" if ok else "FAIL",
            "witness": None if ok else witness
        })

df_results = pd.DataFrame(results).sort_values(by=["m", "n"])
print("\nFull results table:")
print(df_results.to_string(index=False))

passed_df = df_results[df_results["status"] == "PASS"]
failed_df = df_results[df_results["status"] == "FAIL"]

print("\nPassed cases:")
print(passed_df.to_string(index=False) if not passed_df.empty else "
                                        None")

print("\nFailed cases (with witnesses):")
print(failed_df.to_string(index=False) if not failed_df.empty else "
                                        None")
```