The text and figures in this PDF are approved by the author(s) for publication. Any mistakes in this PDF will not be corrected by the publisher. This PDF was created on July 8, 2020.

A QUD-based Theory of Quantifier Conjunction with *but*

Jing Crystal Zhong and James N. Collins

1. Introduction

Barwise & Cooper 1981 point out that not just any pair of DP quantifiers may be conjoined with *but*. They suggest that monotonicity is the relevant semantic property that determines the acceptability of quantifier conjunction with *but*: 'in order to use *but* in this way it seems necessary or at least preferable to mix increasing and decreasing quantifiers' (Barwise & Cooper 1981:p196). In (1), a decreasing quantifier, *no mathemticians*, may not be *but*-conjoined with *few linguists*, another decreasing quantifier.

(1) No mathematicians but { many | ??few } linguists have worked on natural language conjunction.

By 'increasing' and 'decreasing,' Barwise & Cooper 1981 refer to a quantifier's monotonicity behavior, that is, whether the quantifier licenses inferences about supersets or subsets of its scope. Monotonicity is defined as in (2). Let X and Y be arbitrary sets of individuals, and U be the domain of individuals.

(2) a. **Monotone increasing** ($\mathsf{mon}\uparrow$): A quantifier Q is monotone increasing if $X \in Q$ and $X \subseteq Y \subseteq U$ implies $Y \in Q$ (if $X \in Q$, then Q contains all supersets of X).
b. **Monotone decreasing** ($\mathsf{mon}\downarrow$): A quantifier Q is monotone decreasing if $X \in Q$ and $Y \subseteq X \subseteq U$ implies $Y \in Q$ (if $X \in Q$, then Q contains all subsets of X).

We can say a quantifier is non-monotone ($\mathsf{mon}\not\updownarrow$) if it is neither $\mathsf{mon}\uparrow$ nor $\mathsf{mon}\downarrow$. Below are some examples of DPs which are interpreted as quantifiers with various flavors of monotonicity.

(3) a. $\mathsf{mon}\uparrow$: *some alpacas*, *every beaver*, *most cassowaries*, *many dingoes*, *at least two emus*
b. $\mathsf{mon}\downarrow$: *no ferrets*, *few goannas*, *at most five hermit crabs*, *neither iguana*
c. $\mathsf{mon}\not\updownarrow$: *exactly seven jackasses*, *only three kangaroos*, *some but not all llamas*

Why is monotonicity important in natural language? The monotonicity of a quantifier determines not only the truth of the proposition formed by applying the quantifier to its scope, but can also determine the truth of applying the quantifier to its scope's subsets or supersets. This comes in handy when making inferences about quantificational statements in natural language. Assuming *hula* lexically entails *dance*, the monotonicity properties of quantifiers give rise to the following inference patterns.

(4) a. $\mathsf{mon}\uparrow$: Every student hulas. $\Rightarrow$ Every student dances.
b. $\mathsf{mon}\downarrow$: No student hulas. $\Leftarrow$ No student dances.
c. $\mathsf{mon}\not\updownarrow$: Exactly two students hula. $\not\Leftrightarrow$ Exactly two students dance.

Barwise & Cooper's hypothesis about *but*-conjunction, that *but* may only conjoin combinations of increasing and decreasing quantifiers, is an example of the productive application of the monotonicity

* Jing Crystal Zhong, the University of Hawai'i at Mānoa, jzhong@hawaii.edu. James N. Collins, the University of Sydney, jcol2036@uni.sydney.edu.au. We would like to thank Dylan Bumford, Lelia Glass, Thomas Kettig, Fred Zenker, audiences at WCCFL 38 (UBC) and the NINJAL-UHM Linguistics Workshop (UH Mānoa), and all experimental participants.

© 2020 Jing Crystal Zhong and James N. Collins
Cascadilla Proceedings Project
Completed July 8, 2020

behavior of quantifiers to analyzing grammatical phenomena.

But how empirically well-founded is Barwise & Cooper's hypothesis? What is the source of the grammatical sensitivity to the entailment patterns of quantifiers in *but*-conjunction? And how do non-monotone quantifiers fit into the picture?

In this paper, we investigate these questions and revisit Barwise & Cooper's hypothesis about *but*-conjunction. Based on experimental investigation, we demonstrate that quantifier conjunction with *but* is more complicated than first appears. Following from our empirical observations, we hypothesize that *but*-conjunction is only acceptable where the two determiners heading the two DP-conjuncts are *semantically disjoint*, i.e., they denote non-overlapping relations over properties.

Where conjoined DPs satisfy the *semantic disjointness* constraint on *but*-conjunction, they prototypically, but not exclusively, also satisfy Barwise & Cooper's differing monotonicity constraint. We discuss why *semantic disjointness* is relevant for *but*-conjunction, suggesting that the constraints on *but*-conjunction stem from a general theory of contrastive conjunction. Adapting the analysis of contrastive coordination in Toosarvandani 2014, we propose that *but* implies its two conjuncts resolve distinct sub-questions of the current *Question Under Discussion* with different polarities.

2. Revisiting mismatching monotonicity

Let DP_L and DP_R refer to the left conjunct and right conjunct respectively. In (5), DP_L is $\mathsf{mon}\uparrow$ and DP_R is $\mathsf{mon}\downarrow$, and therefore they mismatch in monotonicity. Barwise and Cooper's analysis therefore correctly predicts that (5) is acceptable. If we replace DP_R with *some dogs*, a $\mathsf{mon}\uparrow$ quantifier, we get infelicity, similarly correctly predicted by Barwise and Cooper.

(5) $\Big[[\text{ every cat }]_L \text{ but } [\text{ no dog }]_R\Big]_{\&}$ could behave themselves at Petco.

However, we find cases in which coordinations of two $\mathsf{mon}\uparrow$ quantifiers or two $\mathsf{mon}\downarrow$ quantifiers are acceptable. These constitute evidence that Barwise and Cooper's monotonicity-based constraint is *not necessary*. In our judgement, such coordinations are more acceptable if the semantically weaker determiner heads DP_L; in other words, it precedes the semantically stronger determiner. Given that $[\![\textbf{no}]\!] \sqsubseteq [\![\textbf{few}]\!]$ and $[\![\textbf{every}]\!] \sqsubseteq [\![\textbf{many}]\!]$, this generalization gives rise to the following ordering effects.

(6) a. *Many* pragmaticists ($\Uparrow$) but *every* phonetician ($\Uparrow$) attended the keynote. $\gg\gg$
b. ??*Every* pragmaticist ($\Uparrow$) but *many* phoneticians ($\Uparrow$) attended the keynote.

(7) a. *Few* phonologists ($\Downarrow$) but *no* syntacticians ($\Downarrow$) attended the keynote. $\gg\gg$
b. ??*No* syntacticians ($\Downarrow$) but *few* phoneticians ($\Downarrow$) attended the keynote.

We also find evidence that Barwise and Cooper's monotonicity-based constraint is *not sufficient*, either. We find cases in which the *but*-conjunction of two quantifiers with mismatching monotonicity is judged as unacceptable. For example, we observe the following contrast, even though in both cases *but* conjoins increasing and decreasing quantifiers.

(8) a. At least two thirds of the Democrats ($\Uparrow$) but fewer than half of the Republicans ($\Downarrow$) voted. $\gg\gg$
b. ??At least a third of the Democrats ($\Uparrow$) but fewer than half of the Republicans ($\Downarrow$) voted.

We hypothesize that the relevant constraint above, and on *but*-conjunction of quantifiers more generally, is based on *disjointness* of the determiners. We make the standard assumption that determiners refer to relations over properties, i.e., $[\![\textbf{det}]\!]^C \subseteq U \times U$, where C is a context and U is the universe of discourse referents. As such, two determiners are disjoint just in case $[\![\textbf{det1}]\!]^C \cap [\![\textbf{det2}]\!]^C = \emptyset$.

(9) *Disjointness condition on* but-*conjunction*:
Coordination of two DPs with *but* is unacceptable if the determiners overlap in reference.

In (8a) above, *at least two thirds* is semantically disjoint from *fewer than half*. In (8b), *at least a third* is not semantically disjoint from *fewer than half*. The overlapping and disjoint reference of the three relevant determiners is sketched in Figure 1, where the vertical axis represents proportions between 0

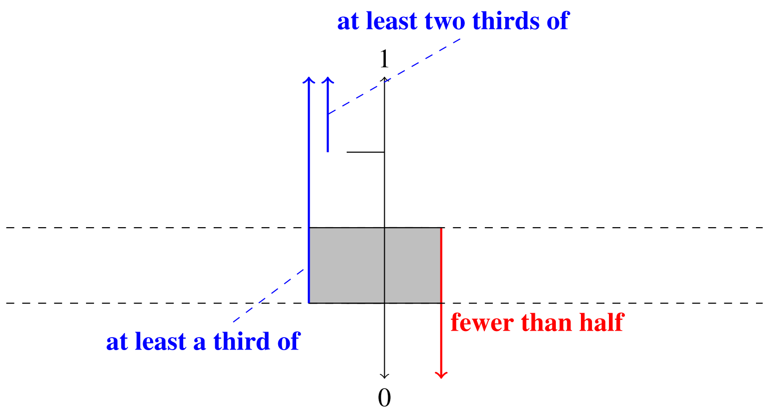


Figure 1: Overlapping and non-overlapping determiners

and 1. As *at least a third of* and *fewer than half* overlap in reference, they violate the condition in (9), and we predict the conjunction is unacceptable, as in (8b). This contrasts with Barwise and Cooper's monotonicity-based condition which predicts (8b) is acceptable.

To summarize our observations, we propose that *but*-conjunction is often, but not always, acceptable with quantifiers of differing monotonicity. A notable exception are cases such as (6a) and (7a), in which the DPs match in monotonicity, but the DP headed by the semantically weaker determiner precedes that headed by a semantically stronger determiner. Thus, mismatching monotonicity is *not necessary* for *but*-conjunction. We further find that DPs with mismatching monotonicity can be unacceptable in a *but*-conjunction. We hypothesize the relevant constraint is disjointness, i.e., where the two head determiners do not overlap in terms of reference. Thus, mismatching monotonicity is *not sufficient* for *but*-conjunction.

In the remainder of the paper, we provide experimental support for these observations and give an explanation as to why disjoint reference is the relevant factor in *but*-conjunction of quantifiers.

3. Experimental investigation

We ran two experimental studies. The first focused on the ordering effect discussed in the previous section, see (6) and (7). The second focused on the notion of disjointness, examining whether non-overlapping reference is a condition for the acceptability of *but*-conjoined quantifiers.

3.1. Experiment 1 Design

In Experiment 1, we examined whether there is an ordering effect for *but*-conjoined quantifiers. We hypothesized that order matters for quantifiers with matching monotonicity, but not for quantifiers with mismatched monotonicity. To test this hypothesis, we carried out an acceptability judgment task (AJT) with native English speakers. In the experiment, conjoined quantifiers either had matching monotonicity (*SameMono*) or mismatched monotonicity (*DiffMono*). We also varied the order of quantifiers: for matching monotonicity quantifiers, we varied the order of weaker quantifiers and stronger quantifiers (coded as *weak-strong* vs. *strong-weak*); for quantifiers with mismatched monotonicity, we varied the order of the $\mathsf{mon}\uparrow$ and $\mathsf{mon}\downarrow$ DP (coded as *up-down* vs. *down-up*).

The first two conditions varied order for *SameMono* quantifiers, as illustrated in (10). Each condition was comprised of six tokens, including three pairs of increasing quantifiers and three pairs of decreasing quantifiers. Our hypothesis predicts that Condition 1 will be rated higher than Condition 2.

(10) a. *Condition 1: SameMono, weak-strong order*
Many students (⇑) but every teacher (⇑) liked the book.

b. *Condition 2: SameMono, strong-weak order*
??Every teacher (⇑) but many students (⇑) liked the book.

Figure 2: Four-point ‘smiley face’ scale

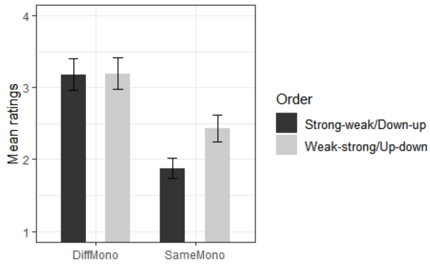


Figure 3: Mean ratings by condition in Experiment 1. Error bars represent 95% confidence interval.

The other two conditions included quantifiers with mismatched monotonicity, as illustrated in (11). We hypothesized that the variation in order in these two conditions should not have an effect. We included three tokens for each condition.

(11) a. *Condition 3: DiffMono, up-down order*
Every teacher ($\Uparrow$) but no students ($\Downarrow$) liked the book.
b. *Condition 4: DiffMono, down-up order*
No students ($\Downarrow$) but every teacher ($\Uparrow$) liked the book.

Altogether there were 18 critical items, which were then distributed across different lists in a Latin square design, so that participants saw each item in only one of the conditions. We also included an equal number of filler items for each list.

3.2. Experiment 1 Procedure

The experiment was administered online on Ibex Farm (Drummond 2020). The test sentences were presented on the screen, along with a 4-point ‘smiley face’ scale, see Figure 2. Participants were instructed to click on the faces to indicate their rating. If they could not rate the sentence for some reason, they could also click the ‘x’ key to skip the sentence.

Initially 28 native English speakers participated in this experiment. Four of the participants did not differentiate between the four conditions and rated them as equally good or bad. They were treated as outliers and removed from the analyses.

3.3. Experiment 1 Results

The ratings of the faces were converted into numbers 1 to 4, from left to right in Figure 2. The mean ratings of the remaining 24 participants are presented in Figure 3. Overall participants rated sentences with mismatched monotonicity higher than those with matching monotonicity. For sentences with mismatched monotonicity, the two orders were rated equally high ($M = 3.19, SD = 0.92$ for Condition 3 and $M = 3.18, SD = 0.92$ for Condition 4). For sentences with matching monotonicity, the weak-strong order ($M = 2.43, SD = 1.1$) received higher ratings than the strong-weak order ($M = 1.87, SD = 0.84$). These results align with our initial hypothesis.

To further explore the effects of monotonicity and order on the ratings, we constructed linear mixed-effects models with ‘monotonicity’ and ‘order’ as fixed effects, and with ‘participant’ and ‘item’ as random effects. The maximal model included a random intercept for by-participant and by-item variance, as well as by-participant and by-item random slopes for ‘monotonicity’ and ‘order.’ All models were fit using the lmerTest package (Kuznetsova et al. 2017) in R.

The model used in the analyses was entered as follows: $answer \sim monotonicity * order + (monotonicity * order|subj) + (monotonicity * order|item)$. The default dummy coding was used.

The fixed effects output from the model summary is shown in (12).

(12) Fixed effects from linear mixed-effects model output of Experiment 1.

	Estimate	SE	df	t-value	p-value
intercept	3.174	0.155	25.6	20.390	<**0.001*****
mono=SameMono	-1.295	0.147	23.5	-8.753	<**0.001*****
order=(weak-strong or up-down)	0.022	0.125	136.5	0.179	0.857
mono=SameMono & order=weak-strong	0.537	0.190	27.2	2.817	**0.009****

Note. The reference level is DiffMono for monotonicity and strong-weak/down-up for order.

Analyses of this model indicated that there was a main effect of monotonicity and an interaction effect of monotonicity and order. When all experimental items were considered, the effect of order was not statistically significant (estimate $= 0.022, SE = 0.125, t(136.5) = 0.179, p = 0.875$).

However, when focusing just on the experimental items with matching monotonicity, the effect of ordering was statistically significant. We constructed a second model with 'order' as a fixed effect, and with 'participant' and 'item' as random effects. This model was as follows: $answer \sim order + (order|subj) + (order|item)$. The analysis indicated that, when conjoined quantifiers with matching monotonicity were exclusively examined, weak-strong order was rated higher than strong-weak order, and the difference reached statistical significance (estimate $= 0.559, SE = 0.12, t(20.02) = 4.35, p < 0.001$).

3.4. Experiment 2 Design

Experiment 2 explored whether disjointness played a role in the *but*-conjunction of two quantifiers. Our hypothesis is that conjunction of disjoint determiners is better than determiners with overlapping reference, regardless of the monotonicity configuration.

To test this hypothesis, we carried out a second AJT with native English speakers. The variables 'monotonicity' and 'disjointness' were crossed in a 2 × 2 factorial design. Sixteen critical items ($k = 4$ per condition) were distributed across four lists in a Latin square design so that participants saw each item in only one of the four conditions. Each list also included 16 fillers. An example sentence in each of the four critical conditions is given in (13).

(13)
a. *Condition 1: SameMono, overlap*
??Exactly seven Democrats but an odd number of Republicans voted 'no'.
b. *Condition 2: SameMono, disjoint*
Exactly seven Democrats but an even number of Republicans voted 'no'.
c. *Condition 3: DiffMono, overlap*
??At least a third of Democrats but fewer than half of Republicans voted 'no'.
d. *Condition 4: DiffMono, disjoint*
At least two thirds of Democrats but fewer than half of Republicans voted 'no'.

The procedure of Experiment 2 was the same as that of Experiment 1. Initially 22 native English speakers participated in the experiment. One participant was an outlier and removed from the analyses.

3.5. Experiment 2 Results

The mean ratings of the critical items of the remaining 21 participants are shown in Figure 4. Overall participants rated sentences with different monotonicity higher than those with the same monotonicity. Within the DiffMono conditions, conjunctions of semantically disjoint determiners received higher ratings ($M = 3.34, SD = 0.92$) than conjunctions of semantically overlapping determiners ($M = 2.7, SD = 0.94$). Likewise, within the SameMono conditions, disjoint determiners received higher ratings ($M = 2.53, SD = 1.03$) than overlapping determiners ($M = 1.92, SD = 0.77$).

To provide more robust evidence for the effect of disjointness, we constructed linear mixed-effects models with 'monotonicity' and 'disjointness' as fixed effects, and with 'participant' and 'item' as random

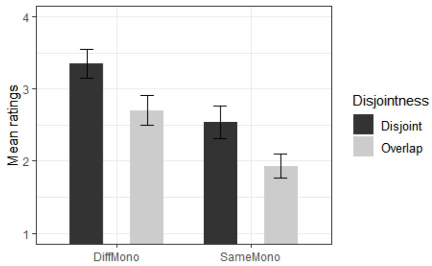


Figure 4: Mean ratings by condition in Experiment 2. Error bars represent 95% confidence interval.

effects. The maximal model included a random intercept for by-participant and by-item variance, as well as by-participant and by-item random slopes for 'monotonicity' and 'disjointness.' The model used in the analyses was entered as follows: $answer \sim monotonicity * disjointness + (monotonicity * disjointness|subj) + (monotonicity * disjointness|item)$. The default dummy coding was used.

The fixed effects output from the model summary is shown in the table in (14). Analyses of this model indicated that there was a main effect of monotonicity as well as disjointness, and there was no interaction of monotonicity and disjointness.

(14) Fixed effects from linear mixed-effects model output of Experiment 2.

	Estimate	SE	df	t-value	p-value
intercept	3.345	0.181	21.1	18.410	<**0.001*****
mono=SameMono	-0.809	0.207	21.0	-3.897	<**0.001*****
disjointness=overlap	-0.643	0.185	20.5	-3.477	**0.002****
mono=SameMono & disjointness=overlap	0.036	0.256	20.8	0.141	0.889

Note. The reference level is DiffMono for monotonicity and disjoint for disjointness.

The intercept estimate of 3.345 represented the mean rating for the *DiffMono* and *disjoint* condition. Changing the monotonicity from *DiffMono* to *SameMono* was associated with an estimated decrease of 0.809, which was a statistically significant difference ($p < 0.001$). Likewise, changing from *disjoint* to *overlap* was associated with an estimated decrease of 0.643, which was a statistically significant difference ($p < 0.01$). Subsequent pairwise comparison between the DiffMono conditions confirmed that the effect of disjointness reached significance (estimate $= -0.642, SE = 0.182, t(21) = -3.522, p < 0.01$). Pairwise comparison between the SameMono conditions also confirmed a significant effect of disjointness (estimate $= -0.607, SE = 0.198, t(21) = -3.064, p < 0.01$).

3.6. Summary of experimental results

Our experimental results suggest the following empirical generalizations. When the conjoined DPs are headed by determiners with **matching monotonicity**, participants judge the conjunction as more acceptable when the semantically weaker determiner precedes the semantically stronger determiner. For example, *many X but every Y* is rated better than *every X but many Y*.

Next, when two DPs headed by **non-monotone determiners** are conjoined, such as *an even/odd number of X* or *exactly seven Y*, participants prefer the conjunction of DPs headed by determiners with *non-overlapping* reference. For example, *exactly seven X but an* ***even*** *number of Y* is rated better than *exactly seven X but an* ***odd*** *number of Y*.

Finally, when the conjoined DPs are headed by determiners with **mismatched monotonicity**, participants prefer conjunctions of determiners which *do not overlap* in reference. For example, *at least two thirds of X but fewer than half of Y* is rated better than *at least one third of X but fewer than half of Y*.

4. Disjointness and determiner conjunction

Our experimental results shed light on the crucial factor determining the acceptability of quantifier conjunction with *but*. We find that the conjunction is more acceptable if the DPs are headed by determiners which *do not semantically overlap*. We suggest that non-overlapping reference is a better predictor of *but*-conjunction acceptability, as compared to Barwise and Cooper's proposal that such conjunctions require mismatched monotonicity. We therefore propose the generalization in (15) where non-overlapping reference is a necessary condition for *but*-conjunction. Note that only the determiners need to be non-overlapping, not the NP descriptions.

(15) **Revised generalization**:
Det1 X but Det2 Y is acceptable only if $[\![\mathbf{Det1}]\!]^C \cap [\![\mathbf{Det2}]\!]^C = \emptyset$

For example, participants rated a conjunction like *no X but every Y* as acceptable. We can demonstrate, in (16), that these determiners do not semantically overlap, assuming existential import.

(16)
a. $[\![\mathbf{no}]\!] = \{\langle P, Q\rangle : P \neq \emptyset,\ P \cap Q = \emptyset\}$
b. $[\![\mathbf{every}]\!] = \{\langle P, Q\rangle : P \neq \emptyset,\ P \subseteq Q\}$
c. *therefore*, $[\![\mathbf{every}]\!] \cap [\![\mathbf{no}]\!] = \emptyset$
d. *therefore*, 'no X but every Y' satisfies the constraint in (15).

Why is *but*-conjunction of DPs sensitive to the semantics of the determiners? Specifically, why is the acceptability of *but*-conjunction sensitive to whether or not the determiners semantically overlap?

4.1. 'but' and QUDs

Our perspective on the semantics of *but* is inspired primarily by Toosarvandani 2014. Toosarvandani argues that *but*-conjunctions are sensitive to the current structure of discourse. Discourse here is viewed as a structured hierarchy of *Questions Under Discussion* (QUDs), determined by the conversational goals of the interlocutors (see Büring 2003, Roberts 2012, Rojas-Esponda 2014, Cremers et al. 2020 etc. for various formulations). An expression of '*X but Y*' serves as the logical conjunction of X and Y. Additionally, for Toosarvandani, *but* signals partial resolution of an active QUD. Both X and Y must resolve sub-questions of the QUD. However, X and Y must resolve the questions with opposite polarities, i.e., one entails a 'no' answer, and the other a 'yes' answer.

(17) Toosarvandani (2014) on *but*:
Felicity condition on $[X\ but\ Y]$: there is a QUD Q, such that
a. For some sub-question of Q, $\sigma? \preceq Q$, $[\![X]\!] \models \sigma$.
b. For some sub-question of Q, $\tau? \preceq Q$, $[\![Y]\!] \models \neg\tau$.

For example, assume the current QUD, defined by the interlocutors' goals, is '*What kinds of cakes do you sell?*', with sub-questions such as those in (18). A relevant context might be a customer in a bakery.

(18) *What kinds of cakes do you sell?*

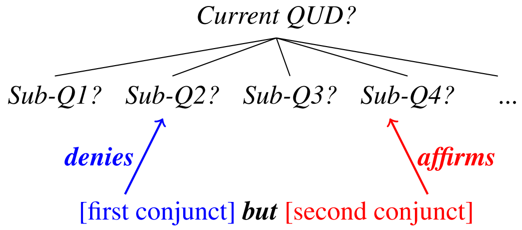
Do you sell carrot cake? *Do you sell chocolate cake?* ...

This conversational context licenses a sentential *but*-conjunction such as (19). Note the reduced acceptability of the second conjunct with positive polarity, by removing *don't*. This effect is predicted by the condition in (17). Under this analysis, the two conjuncts must resolve sub-questions in (18) with *differing polarity*, thus the version without the negation in the second conjunct is unacceptable.

(19) We sell carrot cake **but** we $^{??}$(*don't*) sell chocolate cake.

Thus our view is that *but*'s function is to conjoin two partial resolutions of the current QUD with opposing polarity. An example of a *but*-conjunction satisfying the requirement is depicted visually as in (20).

(20)



4.2. Conjoined generalization quantifiers and QUDs

How do conjoined DP quantifiers play into this view? We give the following semantics for *but* as encoding logical conjunction $\wedge$ as its 'at-issue' content, and the partial QUD-resolution component as a definedness condition. Below, Q refers to a contextually supplied current QUD, and the $\preceq$ relation follows the definition in Rojas-Esponda 2014:§4.2.

(21) $[\![\mathbf{but}(\phi)(\psi)]\!]^Q$ is defined iff there are sub-questions $\sigma?, \tau? \preceq Q$, s.t., $\phi \models \sigma$ and $\psi \models \neg\tau$, where defined $[\![\mathbf{but}(\phi)(\psi)]\!]^Q = [\![\phi \wedge \psi]\!]^Q$

As for *but*-conjoined DPs, we propose that the intonational structure of such conjunctions plays a role. We observe that the *but*-conjoined quantifiers under investigation in this paper bear contrastive focus, *both* on the quantificational determiners, as well as on the NP descriptions, as in (22).

(22) EVERY^F *cát* but NO^F *dòg* skateboarded.

The contrasting determiners and descriptions indicate the structure of the QUD: interlocutors are raising and resolving issues regarding the quantity of individuals of various types. (23) is a sketch of the overarching QUDs in which the utterance (22) is licensed. As 'every' and 'no' are assigned contrastive intonation, they serve as alternatives to each other (see Rooth 1992:§2.2, Vallduví 2016:§23.5). 'Cat' and 'dog' are also contrastive, and alternatives to each other, giving rise to the sub-question set in (23).

(23) $$\text{Did} \left\{ \begin{array}{c} \text{every} \\ \text{no} \\ \ldots \end{array} \right\} \left\{ \begin{array}{c} \text{cat} \\ \text{dog} \\ \ldots \end{array} \right\} \text{skateboard?}$$

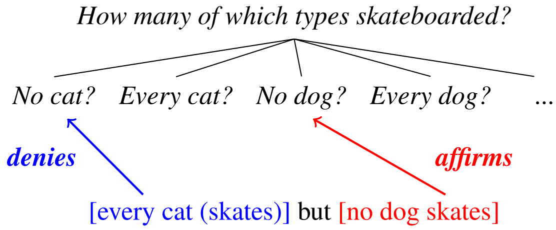
As a general rule, the overarching QUD might be expressed in English by the multiple-*wh* question 'How many of which types VP?', where the conjoined determiners serve as possible answers to the 'how many?' component of the question, while the conjoined NPs serve as possible answers to the 'which type?' component. With respect to the example (22), the intonational structure ensures that the licensing QUD can be decomposed into *at least* the following sub-questions.

(24) $$\left\{ \begin{array}{c} \textit{Did every cat skateboard?} \\ \textit{Did no cat skateboard?} \\ \textit{Did every dog skateboard?} \\ \textit{Did no dog skateboard?} \\ \ldots \end{array} \right\} \sqsubseteq \text{current QUD}$$

We are now in a position to explain the 'disjointness' effect required by the lexical entry for *but* in (21). The two conjuncts must resolve sub-questions of the current QUD with opposite polarity. The contrastive intonation expressed by the *but*-conjoined DP expression gives clues about the structure of the discourse in which the utterance is licensed. It is at least comprised of the questions in (24). (25) is an illustration of how the two conjuncts each resolve a sub-question with opposite polarity.

Due to the structure of the sub-questions in (24), one conjunct will *always* supply a positive polarity affirmation of a sub-question. Thus, the key requirement is that the *other* conjunct will supply a *negative* response to a sub-question, as in (25). Each conjunct is permitted to resolve multiple sub-questions, so long as there is at least one pair of resolutions with mismatched polarity.

(25)



Based on the proposed structure of the QUD-tree, we propose the following theorem regarding *but*-conjoined DP quantifiers. The property of semantic disjointness, $[\![\textbf{Det1}]\!]^C \cap [\![\textbf{Det2}]\!]^C = \emptyset$, outlined in (15), will *ensure* that the felicity condition on *but* is met, proof in footnote.[1]

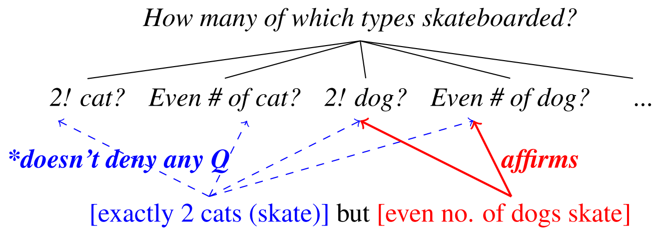
(26) **Theorem**: *any pair of Dets with disjoint reference will satisfy the felicity condition* of *but*

When *but* conjoins two DPs headed by semantically disjoint determiners, it will guarantee that the felicity condition in (21) is satisfied. But what goes wrong with non-disjoint determiners, such as in (27)?

(27) #**Exactly two** cats but **an even number of** dogs skateboarded.

Non-disjoint determiners like *exactly two* and *an even number of*, do not guarantee that the felicity condition of *but* is met. (28) illustrates how neither conjunct in (27) can provide a negative polarity response to any of the extant sub-questions of the QUD.

(28)



In (28), the felicity condition of *but* fails. It is false that one conjunct affirms a sub-question, while the other denies a sub-question. Therefore, the *but*-conjunction is infelicitous.

5. The ordering of *but*-conjoined quantifiers

The experimental results in §3.3 suggest that participants do not always reject co-ordinations of DPs with matching monotonicity. In fact, DPs headed by scale-mate determiners are judged as acceptable, when the semantically weaker determiner precedes the stronger one. For example, 'many X but every Y' is judged better than 'every X but many Y.' As of now, the lexical entry for *but* in (21) is symmetrical, not differentiating the left or right conjuncts.

Our working hypothesis is that this asymmetry can be explained via pragmatic strengthening. Specifically, the weaker determiner is exhaustified to exclude the stronger. For example, in the conjunction 'many X but every Y,' 'many' is strengthened to mean *many-&-not-every*. Thus, 'many X' is not upward monotone at all, but under a strengthened reading, it is non-monotone.

The conjunction 'many X but every Y Zed,' as per our QUD-based analysis, gives rise to the sub-question set *Many(X)(Z)?*, *Every(X)(Z)?*, *Many(Y)(Z)?*, *Every(Y)(Z)?*. Note that quantifiers in the QUD structure are blocked from exhaustifying due to the interrogativity (see Chierchia 2004). If the uttered determiner 'many' is pragmatically strengthened to mean *many-&-not-every*, it will negatively resolve a

[1] Let D_α and D_β be *disjoint determiners*, as per (15), and let A,B,C,X,Y be arbitrary properties.

a. For any X,Y, $D_\alpha(X)(Y) \models \neg D_\beta(X)(Y)$ and $D_\beta(X)(Y) \models \neg D_\alpha(X)(Y)$ by (15)

b. $\therefore$ for any A,B,C, $D_\alpha(A)(C)$ is an affimative answer to question $D_\alpha(A)(C)?$ and $D_\beta(B)(C)$ is a negative answer to question $D_\alpha(B)(C)?$. from (a)

c. $\therefore$ for any Q s.t. $D_\alpha(A)(C)?, D_\alpha(B)(C)? \preceq Q$, '$D_\alpha(A)(C)$ *but* $D_\beta(B)(C)$' is defined. from (b) and (21) sub-question of the current QUD, and the felicity condition of *but* in (21) is met.

(29) a. *many-&-not-every*$(X)(Z)$ negatively resolves Q: *every*$(X)(Z)$?
b. *every*$(Y)(Z)$ affirmatively resolves Q': *every*$(Y)(Z)$?

What gives rise to the ordering effects observed in §3.3? Why do participants prefer the weaker determiner to precede the stronger determiner? We take this effect to come from general pragmatic preferences about how to structure scale-mate items in *but*-conjunctions.

We stipulate a preference that in the conjunction of scale-mate alternatives in contrast, the *left* conjunct must negatively resolve a sub-question of the QUD, while the *right conjunct* must affirmatively resolve a sub-question. As *every* in (29) may only affirm the extant sub-question, it is therefore preferred as the right conjunct. This is a general effect for any QUD with sub-questions ϕ?, ψ?, where $\phi \models \psi$. An utterance of the stronger ϕ will affirmatively resolve both ϕ? and ψ?.

6. Conclusion

This paper has provided an account of the function of the connective *but*, specifically as it conjoins DP quantifiers. Using experimental data, we observe a general preference for *but*-conjoined DPs to be headed by determiners that are semantically disjoint. We offer this generalization as an alternative to Barwise and Cooper's (1981) generalization, which states that *but*-conjoined DPs must have mismatched monotonicity. We suggest our semantic disjointness generalization yields a closer match to the data.

We propose that the function of ***but*** is to signal the structure of the discourse as it relates contrastive expressions. *But* provides clues to interlocutors as to 'how the current QUD should be resolved.' We suggest, following a previous proposal from Toosarvandani 2014, that *but* signals that its conjuncts (at least partially) resolve sub-questions of the current QUD with opposite polarities.

We suggest that this study provides insight into how contrastive expressions are structured, and how they are sensitive to the discourse context, i.e., the hierarchy of Question(s) Under Discussion (QUDs). We hope to raise new questions and provide formal insight about how discourse-sensitive expressions, like *but*, formally interact with the discourse context.

References

Barwise, Jon & Robin Cooper (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy* 4:2, 159–219.

Büring, Daniel (2003). On D-trees, beans, and B-accents. *Linguistics and Philosophy* 26:5, 511–545.

Chierchia, Gennaro. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. Belletti, Adriana (ed.), *Structures and Beyond, Vol. 3*, Oxford University Press, Oxford.

Cremers, Alexandre, Elizabeth Coppock, Jakub Dotlacil & Floris Roelofsen (2020). Modified numerals: Two routes to ignorance. Ms., available at semanticsarchive.net/Archive/jk2M2Q3O/paper.pdf, last accessed June 2020.

Drummond, Alex (2020). Ibex 0.3.8 manual. Retrieved from github.com/addrummond/ibex/blob/master/docs/manual.md, last accessed June 2020.

Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82:13, 1–26.

Roberts, Craige (2012). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5:6, 1–69.

Rojas-Esponda, Tania (2014). A discourse model for *überhaupt*. *Semantics and Pragmatics* 7:1, 1–45.

Rooth, Mats (1992). A theory of focus interpretation. *Natural Language Semantics* 1:1, 75–116.

Toosarvandani, Maziar (2014). Contrast and the structure of discourse. *Semantics and Pragmatics* 7:4, 1–57.

Vallduví, Enric (2016). Information structure. Aloni, Maria & Paul Dekker (eds.), *Cambridge Handbook of Formal Semantics*, Cambridge University Press, Cambridge, 728–755.