Proceedings of SALT 28: 289–306, 2018

# The Thought Uniqueness Hypothesis[*]

Uli Sauerland
*Osaka University and ZAS, Berlin*

**Abstract** A novel principle, the Thought Uniqueness Hypothesis (TUH), unifies several restrictions on interpretation that work in theoretical semantics has observed, in particular the following: binding and scope economy of Fox (2000), and constraints on types (Heim 2017; Hirsch 2017). The principle not only derives these phenomena, but makes additional novel predictions such as a reduction of superiority and D-linking data and the interaction of i-within-i phenomena with coordination. Furthermore the principle exhibits close similarities to current work on exhaustification and efficiency (Meyer 2013) with a potential for further unification. The statement of the TUH is most natural on a realizational view of grammar, where conceptual representations are generated by a non-linguistic system, and then realized by the linguistic system. It therefore argues against the view that surface word order plays a role in interpretation (Chomsky 1970 and others).

**Keywords:** concepts, types, coordination, binding, scope, syntax-semantics

This paper attempts to unify several independently proposed principles of grammar and replace them with a single principle: the Thought Uniqueness Hypothesis (TUH). The TUH is similar to the Efficiency principle of Meyer (2013, 2015), but applies the level of concepts without reference to the linguistic form. The TUH requires a different view of grammar and its interaction with the conceptual system than the current mainstream view: Currently most researchers assume that the surface form possibly after some scope shifting operation serves as the input to interpretation.[1] I find myself forced to adopt a different view where the interpreted form is directly generated by the conceptual system, and then realized morphophonologically.[2]

---

[*] I am grateful to Irene Heim, Artemis Alexiadou, Aron Hirsch, Danny Fox, Serge Minor, Marie-Christine Meyer, Guillermo Del Pinal, and Kazuko Yatsushiro and to audiences at ZAS, at SALT 28 at MIT, the Tokyo Semantics Research Group at Keio University, and a the Universities of Hokkaido, Osaka, and Mie for their comments. Financial support was received from DFG grant SA 925/11-2 and an Osaka University visiting researcher fellowship.

[1] This view has been argued for by Chomsky (1972); Jackendoff (1972); Fox (2000) which we address in the following. Independently, work in formal semantics has adopted a view defined by strict adherence to surface word-order (sometimes called 'direct compositionality').

[2] The 'single-output model' of Bobaljik & Wurmbrand (2012) is related, but I don't view the generation

To introduce the novel perspective, consider first a simple model of a conceptual system – deliberately one that's overly simple for humans: a very bare categorial grammar. Assume that the conceptual system contains the following typed concepts.

(1) a. **John**, **Mary** of type $e$
    b. **love** of type $\langle e, \langle e, \langle s, t \rangle \rangle \rangle$
    c. **everyone** $= \lambda S\, \forall x\, .\, S(x)$ of type $\langle \langle e, t \rangle, t \rangle$
    d. **someone** $= \lambda S\, \exists x\, .\, S(x)$ of type $\langle \langle e, t \rangle, t \rangle$
    e. **not** $= \lambda p\, .\, \neg p$ of type $\langle t, t \rangle$
    f. **and** $= \lambda p\, \lambda q\, .p \wedge q$ of type $\langle t, \langle t, t \rangle \rangle$

The concept labels are mnemonic English words, but these are used only for convenience. For logical concepts, I state the meaning, while for non-logical concepts only the type matters and is all I state. Nonlogical concepts such as **love** have world argument. One feature of (1b) that will be relevant in the following is that the world argument position is innermost. This affects what concept combinations are possible – specifically, only the four propositions two of which are shown in (2): I assume here that the only possible combination of two concepts is function application, and use a flexible Polish notation to denote this that will be convenient later on. Namely, I take both $[xf]$ and $[fx]$ to correspond to $f(x)$ in the standard Euler notation.

(2) a. [**Mary** [**love John**]]
    b. [**Mary** [**love Mary**]]

Because the logical concepts are all purely extensional in (1), they cannot combine with the verb as long as their are no concepts denoting expressions of type $s$. I assume that in addition at least the expression **@** is contained in the intersection of types $s$ and $e$ (the possible sentient beings or centered possible worlds):[3]

(3) $\mathbf{@}, \mathbf{I} \in D_e \cap D_s$

The introduction of **@** allows the application of **not** and **and** as in (4), but not of the quantifiers. The expressions derived are of type $t$, but since they contain **@** can still be viewed as propositions by the **@**-convention in (5).

(4) [[**not** [**@** [**Mary** [**love John**]]]] [**and** [**@** [**John** [**love Mary**]]]]]

---

as a syntactic process in the same way. The realizational view is shared with Distributed Morphology of Halle & Marantz (1993), but in other respects my assumptions concerning the structure of grammar differ.

[3] I don't assume transworld individuals, primarily for convenience of exposition.

(5) ***@**-Convention:* A conceptual representation **X** of type $t$ represents the proposition $\phi$ if for any possible centered world $w$, $\phi(w)$ is true if and only if **X** interpreted with **@** $= w$ is true.

The total set of expressible propositions remains finite with **@** added (namely, the boolean closure of (2)). But since both **not** and **and** can apply recursively, the forms of expressions is infinite for every expressible proposition. This infinite overgeneration constitutes a difference to the system without **@** where each expressible proposition could be expressed uniquely. Independently of that, both systems are clearly insufficient as models of human cognition because they still lack expressive power.

Now consider one model of cognition that possesses sufficient expressive power: Lambda calculus (Church 1940) combined with a sufficient array of primitives. I will focus on lambda calculus since it is familiar to most linguists. There are well-known, equally expressive alternatives in combinatorial logic Schönfinkel (1924); Curry (1930), and the choice at this point has complex repercussions for the following.[4]But all alternatives also give rise to the same infinite overgeneration, which is my primary focus in the following.

Adding lambda-operators and variables makes it possible to integrate quantifiers in complex conceptual representations. One way of making this possible is to use individual variables in either the inner (object) or outer (subject) argument position bound from a higher position as shown in (6).

(6) a. [[$\lambda y$ [**@** [**Mary** [**love** $y$]]]] **everyone**]
b. [**someone** [$\lambda x$ [**@** [$x$ [**love John**]]]]]

As already noted for the propositional calculus system with **@**, the enriched system also allows infinitely many distinct representations of each expressible proposition. For example, I show some other representations equivalent to (6a) in (7): (7a) represents the world argument by a variable rather than the constant **@** and therefore results in a propositional interpretation without appeal to convention (5). In (7b), the outer argument of **love** in addition to the inner one is interpreted via variable binding which in this case is applied to Mary. In (7c), a complex expression composed entirely out of $\lambda$-operators and bound variables applies to the concept **love** to change its argument order.

(7) a. [$\lambda w$ [[$\lambda y$ [$w$ [**Mary** [**love** $y$]]]] **everyone**]]
b. [**Mary** [$\lambda x$ [[$\lambda y$ [**@** [$x$ [**love** $y$]]]] **everyone**]]]
c. [[**Mary** [@ [[$\lambda V$ [$\lambda w$ [$\lambda x$ [$\lambda y$ [[[$V$ $w$] $y$] $x$]]]]]] **love**]] **everyone**]

---

[4] I suspect that neither of the two is appropriate.

My broad claim in (8) states that our cognitive system is constrained so as to not allow multiple equivalent representations. The TUH may have a functional motivation since if it obtains it makes it easier to verify equivalence: equivalence can be examined simply by comparing representations.But I put such functional speculation aside for now, and instead focus on a more precise formulation of the TUH on the basis of the linguistic evidence available.

(8) *Thought Uniqueness Hypothesis (TUH):* Our cognitive system restricts us to have just one representation per thought.

The evidence for the TUH that I am aware of comes exclusively from linguistics. Much of it has been discussed in terms of constraints on LF-representations within an interpretive view of grammar as discussed above. In all of this cases, though, I show below that the constraint can be derived from the TUH and in some cases, I will furthermore show that the TUH derives additional, correct predictions that don't follow from the LF-based view. The TUH thus not only represents a unification of existing generalizations, but furthermore is corroborated by novel predictions. The specific linguistic constrains I address in the following are the binding economy condition of Fox (2000), the scope economy condition also of Fox (2000), and the inflexible types hypothesis discussed in ongoing work by Heim (2017) and Hirsch (2017). In addition, the TUH is closely related to the efficiency condition of Meyer (2013, 2015) and can furthermore be unified with current work on exhaustification by Chierchia (2013) and many others. The structure of the presentation is as follows: I first introduce a technical proposal making the broad intuition in (8) precise in section 1. In section 2, I illustrate the proposal by means of binding economy cases. In section 3, I address the core established cases of scope economy and then show that the TUH-account predicts in addition cases of overt scope economy. Section 3 address the core data in favor of the type inflexibility hypothesis. In Section 4, I examine the potential of unifying the TUH with work on exhaustivization by means of a super-exhaustivity operator. To conclude in section 5, I return briefly to the cognitive structure based perspective to show that the TUH perspective overcomes the arguments of Chomsky (1972) for the LF–perspective.

## 1 Thought Uniqueness – Concrete Hypothesis

In this section, I formulate a concrete proposal that spells out the intuition of the TUH (8)– namely the TUCH in (9).[5] This formulation is, I hope, consistent with the empirical data that I present in the remainder of the paper, though of course, the data underdetermines the details of the proposal. The TUCH specifies

---

[5] TUCH for Thought Uniqueness, Concrete Hypothesis.

that of all equivalent possible cognitive representations within the search space specified by (9b) only the best representation is actually available. To specify which representation is better, the TUCH relies on the notion of dependency complexity in (10). One notable decisions in the statement of the TUCH are that (9b) requires semantic equivalence of the two representations only for some variable assignment rather than all of them. This has some support in the account of i-within-i binding in section 4 below.

(9) *TUCH (version 1.0):* A cognitive representation **B** is banned if a cognitive representation **A** exists with:

a. $[\![\mathbf{A}]\!]^g = [\![\mathbf{B}]\!]^g$ for some assignment $g$,
b. **A** consists only of primitive concepts of **B**, $\lambda$-operators, variables, and combinations thereof, and
c. **A** has *lower dependency complexity (DC)* than **B** (i.e. $\mathrm{DC}(\mathbf{A}) < \mathrm{DC}(\mathbf{B})$).

The dependency complexity of a cognitive representation **A** depends on the distance between an occurrence of a bound variable and its binder.

(10) *Dependency Complexity (DC)* Let var(**A**) be the set of occurrences of bound[6] variables in A and $\mathrm{len}(x)$ be the number of complex concept units between a single occurrence $x \in \mathrm{var}(\mathbf{A})$ and its binder $\lambda_x$ within **A**. Then we define the dependency complexity of **A** as:

$$\mathrm{DC}(\mathbf{A}) = \sum_{x \in \mathrm{var}(\mathbf{A})} 2^{\mathrm{len}(x)}$$

In (10), a specific choice that I only justify later is the non-linear effect of distance.[7] The distance and the contribution to DC for an occurrence of a variable can be easily determined by counting the number of unpaired brackets between it and its binder as illustrated in (11).

(11) $\mathrm{DC}([\lambda_x\ x]) = 1$, $\mathrm{DC}([\lambda_x\ [\ x \cdots]]) \geq 2$, $\mathrm{DC}([\lambda_x\ [\ \cdots\ [\ x \cdots]\ \cdots]]) \geq 4, \ldots$

Applied to any of the three cognitive structures in (7), the TUCH blocks them because the dependency complexity of (6a) is lower. The TUCH is similar to the efficiency condition of Meyer (2013, 2015).[8]

---

[6] For the account of strict/sloppy ambiguities in the following section, I revise this aspect of dependency complexity slightly.

[7] More specifically, the effect is stated to be exponential but this is only to give a concrete number, and not determined by the data.

[8] Even earlier, (Magri 2009: p. 68) discusses (and ultimately rejects) a version of Grice's manner

## 2 Binding Economy Effects

Another principle very similar to the TUH is binding economy of Fox (2000). Recall that one central motivation for Fox's account is an observation of Dahl (1973) concerning the availability of strict and sloppy readings in (12). Specifically, Dahl observes that, if the two pronouns *he* and *his* in the first conjunct are interpreted as coreferring with *John*, the second conjunct can receive the three interpretations (12a), (12b), and (12c), but the fourth interpretation in (12d) isn't available.

(12) John said that he likes his mother. Bill does too.

a. *. . . and Bill said that John likes John's mother.*
b. *. . . and Bill said that Bill likes Bill's mother.*
c. *. . . and Bill said that Bill likes John's mother.*
d. **. . . and Bill said that John likes Bill's mother.*

We follow Fox's account in tracing the effect to an economy condition that blocks binding of *his* directly by *John* because the embedded subject *he* is closer, and both binding relations have the same interpretation. Before addressing (12), consider an account of the basic strict/sloppy ambiguity in (13) that is compatible with the TUH. I assume that representations (13a) and (13b) are both available and have the same dependency complexity.[9] Furthermore a representation where the proper name *John* is is interpreted in the argument position as any expression of type $e$ might. I assume that proper names like *John* correspond to conceptual representations of the generalized quantifier type $\langle e, \langle e,t \rangle \rangle$. This follows if the CR is **the 'John** where '$n$ represents the property of bearing name $n$ and **the**$_{\mathbf{GQ}}$ denotes the Russelian entry for the definite determiner, $\lambda R.\lambda Q.\exists x(\forall y(R(y) \leftrightarrow x = y) \wedge Q(x))$ of type $\langle e, \langle e,t \rangle \rangle$.[10]

(13) John likes his mother. Bill does too.

a. [[**the**$_{\mathbf{GQ}}$ **'John**] [$\lambda_x$ [[**the**$_{\mathbf{GQ}}$ [**his**$_x$ **mother**]] [$\lambda_y$ [**@** [$x$ [**like** $y$]]]]]]]
b. [[**the**$_{\mathbf{GQ}}$ **'John**] [$\lambda_x$ [[**the**$_{\mathbf{GQ}}$ [**his**$_z$ **mother**]] [$\lambda_y$ [**@** [$x$ [**like** $y$]]]]]]]

---

maxim for some cases of scalar implicature that bears a superficial similarity to the TUH/TUCH. I believe the reasons for this rejection don't apply to the TUH/TUCH, but cannot discuss them here in detail.

[9] An appropriate revision to dependency complexity (DC) relative to assignment function $g$ is: The DC of $A$ is equal to the DC of the binding-maximal for $A'$ variant of $A$ given $g$. $A'$ is a *binding-reduced* form of $A$ iff. $[\![A]\!]^g = [\![A']\!]^g$, $A'$ contains strictly more occurrences of bound variables than $A$ and $A'$ and $A$ are identical except for occurrences of variables. $A'$ is *binding-maximal* iff it has no binding-reduced variant.

[10] The advantage of adopting Russell's **the**$_{\mathbf{GQ}}$ is small and theory internal. The choice between the Russellian and Fregean concepts of **the** is a more general issue that I cannot exhaustively address here.

The dependency complexity of (13a) is 32 as determined by the length of the binding relations.[11] If the current assignment $g$ with $g(z) = \textbf{John}$, the dependency complexity of (13b) is also 32 as per footnote 9, and therefore both representations are ruled in by the TUH. (13a) gives rise to the sloppy reading, while (13b) leads to the strict interpretation.

For Dahl's example (12), two relevant representations are shown in (14) and (15) with the binding relations indicated by lines. Crucially, (15) contains a long dependency of length 7 for a total dependency complexity of 148, hence the TUCH blocks (15) as the equivalent (14) has dependency complexity 38.[12]

(14) $[[\textbf{the 'J}]\ [\lambda_x\ [\textbf{@}\ [x\ [\textbf{said}\ [\textbf{he}_x\ [\lambda_y\ [y\ [\textbf{like}\ [\textbf{the}\ [\textbf{his}_y\ \textbf{mother}]]]]]]]]]]]$

(15) $^*[[\textbf{the 'J}]\ [\lambda_x\ [\textbf{@}\ [x\ [\textbf{said}\ [\textbf{he}_x\ [\textbf{like}\ [\textbf{the}\ [\textbf{his}_x\ \textbf{mother}]]]]]]]]]$

In addition to (15), also representations such as (16) need to be considered, where an intermediate variable $x$ and a new binder $\lambda_x$ are inserted. The TUCH rules out (16) because, at 60, the dependency complexity of (16) is lower than that of (15), but still higher than (14).

(16) $^*[[\textbf{the 'J}]\ [\lambda_x\ [\textbf{@}\ [x\ [\textbf{said}\ [\textbf{his}_x\ [x\ [\lambda_x\ [\textbf{like}\ [\textbf{the}\ [\textbf{he}_x\ \textbf{mother}]]]]]]]]]]]$

Consider the effect of intermediate variables more generally: Assume $x$ occurs bound in a CR $C$ with $l = len(\mathrm{x})$, but is free within a subterm $P$ at a depth of $n$. Then replacing $P$ in $C$ with $\mathbf{[}x\ \mathbf{[}\lambda_x\ P\mathbf{]]}$ creates two dependencies of lengths $l-n+1$ and $n$ instead of a single one of length $l$. If there are further variable with $P$ that are bound from outside of $P$, the replacement also adds 2 to the length of each of these. But if there are no other dependencies this predicts that, in the case of length 3, an intermediate is thus neutral for TUCH whether an intermediate variable is added to create two dependencies of length 2: both (17a) and (17b) have dependency complexity 8.

(17) a. $[\lambda_x\ [\textbf{lex-a}\ \overbrace{[\textbf{lex-b}\ [\textbf{lex-c}\ x]]}^{P}\ ]]$

b. $[\lambda_x\ [\textbf{lex-a}\ [x\ [\lambda_x\ [\textbf{lex-b}\ [\textbf{lex-c}\ x]]]]]]$

---

[11] The first bound occurrences of $x$ and the one bound occurence of $y$ are three nodes away from their binders, the second bound occurrence of $x$ is four nodes away from its binder. The length of a binding relation can be easily determined by the counting the number of unmatched brackets between an occurrence of a bound variable and its binder.

[12] Fox argues the (15) is available when it's not equivalent to (14) as is predicted by his and present account.

For dependencies of length 4 and above insertion of an intermediate variables is predicted. Evidence for intermediate variables has been given from parasitic gaps (Nissenbaum 2000), reconstruction (Fox 1999), and cumulation (Sauerland 2001), corroborating the non-linear effect of dependency length.

## 3 Scope Economy Effects

In addition to binding economy, also the scope economy phenomena Fox (2000) discusses can be subsumed under the TUCH. The TUCH in this case, though, makes additional predictions – it predicts blocking to occur even when the overt word order differs, while Fox accounts only for cases with identical word order. In this section, I first present the TUCH accounts of two core examples of Fox's and then I argue that superiority phenomena in English question can be derived from the TUCH.

The contrast in (18) is one central case of scope economy (Fox 2000). The first sentence in (18a) allows only an interpretation with narrow scope of the object while (18b) is ambiguous.

(18) a. Some boy admires every teacher. Every girl does too. ($\exists \gg \forall$, $^*\forall \gg \exists$)
b. Some boy admires every teacher. Some girl does too. ($\exists \gg \forall$, $\forall \gg \exists$)

For the TUCH account of (18), consider in (19) two possible conceptual representations that could give rise to the second sentence in (18a).[13] The two representations are equivalent because both quantifiers are universal. Both representations contain two dependencies, but they are of lengths 3 and 4 in (19a) and lengths 2 and 5 in (19b). In both cases, the total lengths is 7, and therefore in this a linear version of the TUCH would not choose between the two. But the non-linear measure I adopted predicts that only (19a) should be available.[14]

(19) a. [[**every girl**] [$\lambda_x$ [[**every teacher**] [$\lambda_y$ [**@** [ $x$ [**admire** $y$]]]]]]]
b. $^*$[[**every teacher**] [$\lambda_y$ [[**every girl**] [$\lambda_x$ [**@** [ $x$ [**admire** $y$]]]]]]]

If (20a) and (20b) weren't equivalent, both representations would be available. This is the case with the first clause in (18a) and both clauses in (18b). But only a structure parallel to (19a) can license ellipsis, and therefore in (18a) the object wide scope representation of the first clause is unavailable. Therefore the TUCH predicts the contrast in (18). Note that the account is similar to Fox's (2000) in many respects, but his account appeals to the order in which words are pronounced to justify a

---

[13] I omit here operators involved in ellipsis licensing.

[14] Note that inserting an intermediate variables in (i) lowers both dependency complexities, but still a representation equivalent to (a) remains preferred.

preference for a representation equivalent to (19a) over one equivalent to (19b), while the TUCH account only makes reference to structural features.

Consider a second case of Fox's with different structural features, namely example (20) of scope economy with reconstruction. Fox argues that narrow scope of *an American runner* below *seem* is unavailable in (20), where in the second sentence the subject is a proper name.

(20) An American runner seems to Bill to have won a gold medal. Sergey does, too. (an American runner $\gg$ seem, *seem $\gg$ an American runner)

Assuming that **seem** is a primitive concept of type $\langle\langle s,t\rangle,\langle s,t\rangle\rangle$, the conceptual representations in (21) might underlie the second sentence of (20).[15] Because **the 'Sergey** has type $\langle\langle e,t\rangle,t\rangle$, assigning it low scope as in (21b) requires that the world variable argument position be represented and bound. Therefore (21a) has dependency complexity 8, while (21b) has 12. As the two representations are equivalent the subject **the 'Sergey**, the TUCH blocks (21b). The scope restriction in (20) then follows from the parallelism requirement of ellipsis.[16]

(21) a. [[**the 'Sergey**] [$\lambda_x$ [**@** [**seem** [$x$ [**to have won**]]]]]]
b. $^*$[**@** [**seem** [$\lambda_w$ [[**the 'Sergey**] [$\lambda_x$ [$w$ [$x$ [**to have won**]]]]]]]]

In this way, the TUCH explains (21) without reference to the overt word order unlike Fox's account. The TUCH account is based purely on features of the conceptual calculus.[17] This feature of the TUCH account predicts cases where blocking takes places between conceptual structures that are pronounced with different overt word orders. I argue that superiority in English is such a case.[18]

---

[15] I again omit operators for ellipsis licensing.

[16] Note that the TUCH also blocks the representation in (i), where a variable of the quantifier type $\langle\langle e,t\rangle,t\rangle$ is introduced predicting so called *semantic reconstruction*. This type of representation is blocked even when the subject is not scopally commutative with *seem* since (21b) has lower dependency complexity.

(i) $^*$[[**an American**] [$\lambda_Q$ [**@** [**seem** [$\lambda w$ [$Q$ [$\lambda x$ [$w$ [$x$ [**to have won**]]]]]]]]]]

[17] Note that in addition to the assumption that *Sergey* corresponds to a quantificational concept, also the **@**-convention is necessary for the account of (20). If pronouns or other expressions could be of type $e$, the TUH account predicts a reverse scope effect from (20).

[18] Independently, Fox (2012) discusses superiority data in an unpublished handout and proposes a version of the *attract closest* condition that is sensitive to interpretation. The TUH account seems preferable to me since it leads to a more uniform account.

### 3.1 Superiority

The TUCH account of scope economy predicts the following difference concerning word order in multiple questions in English observed by Pesetsky (1987). While in questions with the numberless interrogative phrases such as *who* and *what* only one word order is acceptable, two word orders are acceptable with *which*-phrases.

(22) a. Who invited who?
b. *Who did who invite?

(23) a. Which girl invited which boy?
b. Which boy did which girl invite?

The TUCH account relies on the observation that (24a) and (24b) exhibit a subtle difference in interpretation (Dayal 1996). Namely, (23a) presupposes that each girl invited exactly one boy, while (23b) presupposes that each boy was invited by exactly one girl. Of the three different scenarios indicated by the tables in (24), (23a) is only compatible with (24a), while (23b) is only compatible with (24b). The multiple *who* question (22a), however, it compatible with all three (24a), (24b), and (24c).

(24) a.

| | Abe | Ben | Cid |
|---|---|---|---|
| Ann | | * | |
| Bea | | * | |
| Cel | * | | |

b.

| | Abe | Ben | Cid |
|---|---|---|---|
| Ann | | | |
| Bea | | * | * |
| Cel | * | | |

c.

| | Abe | Ben | Cid |
|---|---|---|---|
| Ann | | * | |
| Bea | | * | * |
| Cel | | | |

The general nature of the TUCH account of superiority is probably compatible with different accounts of question semantics. In particular, the difference in interpretation between between (23a) and (23b) predicts that two distinct representations should be possible in this case. What needs to be formally verified though is that when in these two representations the *which*-phrases are replaced with *who*-phrases, the two representations are equivalent. To present these details I adopt a version of the higher-order questions approach (Fox 2012; Nicolae 2013; Xiang 2016). I assume two different operators of type $\langle\langle\langle s,t\rangle,t\rangle,\langle\langle s,t\rangle\rangle$ account for the presuppositions of multiple questions. The answerhood operator (25a) is adopted from Dayal (1996) and also active in single questions and selects the unique maximally informative element of a set of propositions. The big conjunction in (25b) reduces higher order sets of propositions to just sets of propositions by conjoining all true propositions. Finally, identity is understood as propositional identity.

(25) a. $\mathsf{Ans} = \lambda Q \lambda w \iota p \in Q[p(w) \wedge \forall q \in Q : \neg q(w) \vee p \rightarrow q]$

b. $\bigwedge = \lambda Q \lambda w\ .\ \iota p \in D_{\langle s,t\rangle} \forall w'[p(w') \leftrightarrow \forall q \in Q : \neg q(w) \vee q(w')]$
c. $= \in D_{\langle\langle s,t\rangle,\langle s,t\rangle\rangle}$

Furthermore, I assume that *who* quantifies existentially over quantifiers, while *which* phrases quantify of individuals (see Elliott, Nicolae & Sauerland 2018). This predicts that *which*-phrases in a single question (*Which person left?*) presuppose the existence of exactly one individual, while *who*-phrases do not presuppose existence since the quantifier *no one* can occur as part of the unique most informative answer.

(26) a. $\textbf{who} = \left[\lambda p\ \exists Q \in D_{\langle e,\langle e,t\rangle\rangle}\ p(Q)\right] \in D_{\langle\langle\langle e,t\rangle,t\rangle,t\rangle}$
b. $\textbf{which boy, which girl} = \left[\lambda p\ \exists x \in \{\text{singular boys/girls}\}\ p(x)\right] \in D_{\langle\langle e,t\rangle,t\rangle}$

The two non-equivalent representations with *which*-phrases that the multiple questions approach makes available are shown in (27) and (28).[19] (27) is interpreted with conjunction over the girl who each must have a invited a unique boy, which corresponds to (23a). (28) yields the other uniqueness condition of (23b). My assumption is hence that (25) is pronounced as (23a) and (26) as (23b).

(27) $[\bigwedge [\lambda_p [\textbf{wh girl} [\lambda_x [p [= [\mathsf{Ans} [\lambda_p [\textbf{wh boy} [\lambda_y [p [= [x [\textbf{invite}\ y]]]]]]]]]]]]]]$

(28) $[\bigwedge [\lambda_p [\textbf{wh boy} [\lambda_y [p [= [\mathsf{Ans} [\lambda_p [\textbf{wh girl} [\lambda_x [p [= [x [\textbf{invite}\ y]]]]]]]]]]]]]]$

To complete the account of superiority, we now need to verify that representations analogous to (27) and (28) are equivalent with multiple *who*-phrases. Because *who*-phrases bind quantifiers, the representations must be ones where they don't bind verbal argument positions directly. Two relevant representations are shown in (29) and (30). Since *who* introduces neither a uniqueness nor and existence presupposition, the two are indeed equivalent and essentially represent row-wise or column-wise complete descriptions of tables such as those in (24).

(29) $[\bigwedge [\lambda_p [\textbf{who} [\lambda_X [p [= [\mathsf{Ans} [\lambda_p [\textbf{who} [\lambda_Y [p [=$
$[\lambda_w [X [\lambda_x [Y [\lambda_y [w [x [\textbf{invite}\ y]]]]]]]]]]]]]]]]]]]$

(30) $^*[\bigwedge [\lambda_p [\textbf{who} [\lambda_Y [p [= [\mathsf{Ans} [\lambda_p [\textbf{who} [\lambda_X [p [=$
$[\lambda_w [X [\lambda_x [Y [\lambda_y [w [x [\textbf{invite}\ y]]]]]]]]]]]]]]]]]]]]$

---

[19] I leave it open why $\bigwedge$ and Ans occur in the order shown. If their order was switched, different predictions arise. Fox (2012) builds the ordering into a type system, but as far as I can see this also lack independent motivation.

Furthermore the dependency complexity of (30) is higher than that of (29) since the other binder $\lambda Y$ in (30) is associated with the innermost variable. The TUCH then predicts that only representation (29) is available. By analogy to (23a), I assume that (29) is pronounced as (22a), and cannot be pronounced as (22b). In this case, no conceptual representation exists that would be pronounced as (22b) and therefore (22b) is perceived to be ungrammatical. Below we see that there are also cases where only one conceptual representation is blocked for a surface form, but another remains possible.

In sum, the TUCH provides a principled account of the core cases of superiority in English. This argues that also structures that are pronounced with different surface word order can block one another. Specifically, it leads to ungrammaticality of (22b). I discuss one further case of the TUCH predicting a blocked association between a conceptual representation and a surface form.[20]

## 4 Type Inflexibility

In this section, I consider a further array of phenomena that provide evidence for the TUH. Heim (2017) and Hirsch (2017) have argued that in a number of cases the semantic types cannot be flexible in an unconstrained way. In particular, both address type shifts that can be expressed solely by $\lambda$-operators and variables.[21] In particular, I address the two instances of point-wise application of a function, both shown schematically in (31).

(31) a. unary pointwise application: $f'(g) = \lambda_x\, f(g(x))$, derived by $f' = \mathbf{G}(f)$ with $\mathbf{G}^{\gamma}_{\alpha,\beta}$ = $[\lambda_f^{\langle\alpha,\beta\rangle}\ [\lambda_x^{\langle\gamma,\alpha\rangle}\ [\lambda_y^{\gamma}\ [\ f\ [\ x\ y\ ]]]]]$
b. binary pointwise application: $J'(f)(g) = \lambda_x\, J(f(x))(g(x))$, derived by $J' = \mathbf{L}(J)$ with $\mathbf{L}^{\alpha} = [\lambda_J^{\langle t,\langle t,t\rangle\rangle}\ [\lambda_f^{\langle\alpha,t\rangle}\ [\lambda_g^{\langle\alpha,t\rangle}[\lambda_x^{\alpha}[[J[fx]][gx]]]]]]$

For (31a), Jacobson (1999) introduces the term *Geach rule*, while (31b) is sometimes referred to as lifting. Both operations can be defined in $\lambda$-calculus as shown by the definition of **G** and **L** (for result type $t$) above.

Jacobson (1990) effectively argues that (31a) cannot freely apply.[22] Specifically she addresses the proposal to allow raising verbs to belong to the same category as control verbs. Consider one concrete possibility how free application of **G** would

---

[20] There are at least two other cases that exhibit some, but as far as I can see, not all the predicted properties: *there* and raising in English and quantifier scope in scrambling languages (Bobaljik & Wurmbrand 2012). I cannot address either in detail at this point.

[21] I use the term *type shift* exclusively for injective functions from all elements of $D_\alpha$ to another domain $D_\beta$ with $\beta \neq \alpha$.

[22] I follow Heim (2017) in the exposition of Jacobson's work.

make allow raising verbs to be like control verbs: Specially, the control verb realizes the concept **try** of type $\langle\langle e,\langle s,t\rangle\rangle,\langle e,\langle s,t\rangle\rangle\rangle$, while **seem** has type $\langle\langle s,t\rangle,\langle s,t\rangle\rangle$. Then [$\mathbf{G}^{e}_{\langle st,st\rangle}$**seem**] has the same type as **try**. In the system we developed, the conceptual representations (32a) for *seem* with a finite complement, (32b) for *seem* with an infinite complement, and (32c) for *want* are interpretable.

(32) a. [**seem** [$\lambda_w$ [ [**the 'John**] [$\lambda_x$ [$w$ [$x$ **left**]]]]]]
It seems that John left.
b. (i) *[[**the 'John**] [**@** [[**G seem**] **leave**]]]
[**the 'John**[$\lambda_x$ [**@** [**seem** [$x$ **leave**]]]]
John seems to have left.
[**the 'John**[**@** [**tried leave**]]]
John tried to leave.

But Jacobson (1990) argues that the control representation (32b) should be ruled out. One of her empirical arguments (Jacobson 1990: p. 439) is that, while *try* like most control verbs allows bare complement anaphora as in (33a), *seem* doesn't.

(33) a. Mary tried to be friendly, and I think Bill tried too.
b. *Mary seems to be friendly, and I think Bill seems too.

The TUCH predictions for (32b) depend on the dependency complexity. Since **G** is an abbreviation for the complex of $\lambda$-operators and variables in (31a), it has a fixed dependency complexity, namely 20. This is the dependency complexity of (32bi), while (32bi)'s is 8. Therefore the TUCH blocks (32bi) and therefore predicts the ungrammaticality of (33b).

But there are two other possible ways (33b) could be predicted to be possible. An alternative primitive concept **seem₂** might exist, which is of type $\langle\langle e,\langle s,t\rangle\rangle,\langle e,\langle s,t\rangle\rangle\rangle$. This could occur in place of [**G seem**] in (32bi). This in general, I don't think should be ruled out: two different cases where such a point-wise lexical variant exists are i) **little** as a primitive equal to [**G not**] of type $\langle dt,dt\rangle$ (Heim 2006) and others analyze as, and ii) the passive morpheme **PASS** as a primitive equal to [$\lambda_V$ [$\lambda_w$ [**someone** [$w$ $V$]]]] of type $\langle\langle e,st\rangle,st\rangle$ (Alexiadou, Anagnostopoulou & Schäfer 2015). But in these two cases, the two related concepts are realized by different morphemes, while **seem** and **seem₂** are homophonous. I suggest that this is not a coincidence, and propose the following condition.[23]

(34) *Avoid Homophony:* The exponents of primitive concepts with different types cannot be homophonous.

---

[23] From this perspective, *threaten* and other verbs that seem ambiguous between a control and a raising concept must involve accidental homophony.

A second possible change rendering (33b) acceptable is to add $\mathbf{G}^{e}_{\langle st,st\rangle}$ to the inventory of primitive concepts. Since then the use of **G** wouldn't incur any dependency complexity, representation (32bi) above would be incorrectly predicted to be preferred by the TUCH. A generalization of the condition in footnote 9 maximizing occurrences of bound variables cover this case, but I leave the details for future work.

Now consider a second interesting argument to restrict type shifts due to Hirsch (2017) investigating coordination. I focus on the data point in (35). Hirsch points out that (35) is problematic for analyses that allow type shift to apply to **and** without constraint because, if the two quantifiers *a maid* and *a cook* could be combined by a type shifted **and** into a single quantifier, then the resulting quantifier should exhibit a scope ambiguity with the subject in (35). But in fact, (35) allows only wide scope of the subject.

(35) Some company hired a maid and a cook. (some ≫ and, *and ≫ some)

For the present account, Hirsch's argument raises the question how a lifted concept [$\mathbf{L}^{\langle et,t\rangle}$ **and**] is blocked with **L** as defined in (31b), specifically when **a maid L and a cook** receives wide scope. At an intuitive level, the analysis I propose is the following. I assume that the realization principles following Johnson (2009) have the following consequences: both a conceptual representation CR1 with some-≫-and scope and a different CR2 with and-≫-some scope could yield the realization (35). But, I argue that a third CR3 with and-≫-some scope can also be formed that has lower dependency complexity than CR2. Hence, CR2 is blocked by the TUCH. For this it plays no role that CR3 cannot be realized as (35), but must be realized as (36). This leaves CR1 as the only TUCH-compatible conceptual representation that is realized as (35), and therefore (35) is unambiguous.

(36) Some company hired a maid and some company a cook.

The specific representation are shown below: CR1 in (37), CR2 in (38), and CR3 in (39).

(37) [$\lambda_w$ [[**some company**] [$w$ [$\lambda_w$ [$\lambda_x$ [[[$\lambda_y$ [$w$ [$x$ [**hired** y]]]] [**a maid**]] [**and** [[$\lambda_y$ [$w$ [$x$ [**hired** $y$]]]] [**a cook**]]]]]]]]]

(38) *[$\lambda_w$ [[$w$ [$\lambda_w$ [$\lambda_y$ [[**some company**] [$\lambda_x$ [$w$ [$x$ [**hired** y]]]]] [**a maid** [ [$\lambda_J$ [$\lambda_{Q_1}$ [$\lambda_{Q_2}$ [ $\lambda_p$ [$Q_1$ $p$] [$J$ [$Q_2$ $p$]]]]]] **and** ] [**a cook**]]]]]

(39) [$\lambda_w$ [[**some company**] [$w$ [$\lambda_w$ [$\lambda_x$ [[$\lambda_y$ [$w$ [$x$ [**hired** y]]]] [**a maid**]]]]]] [**and** [[**some company**] [$w$ [$\lambda_w$ [$\lambda_x$ [[$\lambda_y$ [$w$ [$x$ [**hired** y]]]] [**a cook**]]]]]]]]

One case where the predictions of the TUCH account differ from Hirsch's proposal concerns noun phrase internal subjects. Hirsch (2017) proposes that *and* must always be of type $\langle t, tt \rangle$ and notes that this requires him to generally postulate noun phrase internal subjects to be able to analyze data such as (40).

(40) The linguist and philosopher is happy.

But to generally require noun phrase internal subjects opens the door to *i*-within-*i* binding in cases such as (41), which is odd if *her* and the wife are understood to be coreferent.

(41) #The wife of her childhood sweetheart is happy.

The TUCH-account makes slightly different predictions in this case. Consider first the two conceptual representations for (41) in (42):[24]

(42) a. $^*$[[**the** [$\lambda x$ [[**wife** [**the** [**her**-$x$ **childhood**-**sweetheart**]]] $x$]]] **is happy**]
b. #[[**the** [**wife** [**the** [**her**-$x$ **childhood**-**sweetheart**]]]] **is happy**]
c. [**she**-$x$ **happy**]

Representations (42a) and (42b) are Strawson equivalent assuming an assignment that maps $x$ to Lucy. Given their Strawson equivalence, the TUCH could be modified to predict that the less complex (42b) blocks (42a).[25] Then furthermore (42c) blocks (42b) as Marty (2017) argues.

Now consider the two conceptual representations in (43) for (40): with an NP-internal coordination. In this case, (43a) with the NP-internal subjects has lower dependency complexity.

(43) a. [[**the** [$\lambda x$ [**linguist** $x$] [**and** [**philosopher** $x$]]]]] is happy]
b. $^*$[[**the** [ **linguist** [$\lambda J$ [$\lambda n$ [$\lambda m$ [$\lambda x$ [[$J$ [$m$ $x$]] [$n$ $x$]]]]]] **and**] **philosopher**]]]] **is happy**]

In this way a suitably modified TUCH could rule out *i*-within-*i* cases while accounting for Hirsch's other data. One novel prediction such an account makes is that *i*-within-*i* reference should improve in noun phrases containing also coordination such as (44), a prediction my informal surveying of English speakers has found to be empirically supported.

(44) Every wife and colleague of her childhood sweetheart is happy.

---

[24] For ease of exposition, I assume here that **the** is of type $\langle et, e \rangle$.

[25] I leave the tension between the role of presuppositions for the TUH in this vs. the case of multiple questions for future work.

## 5 Conclusions

The TUH provides a new unified account of several independently established constraints (binding economy, scope economy, and type inflexibility). Briefly consider two further implications I consider important: a possible unification with current work relating to exhaustification, in particular Meyer's (2013; 2015) Efficiency, and implications for the data that drove Chomsky (1972) and others to adopt a architecture of grammar where surface word order plays a role in interpretation.

Exhaustification typically is stated as an operator that excludes logically stronger or at least non-weaker alternatives of a phrase $P$ from the interpretation of $P$. Two similarities to the TUH exist though: Katzir (2007) and others have argued that alternative must not be structurally more complex in a specific sense, and Chierchia (2013) and others have argued that exhaustification can block a representation if a logical contradiction arises. In cases blocked by the TUH, also a structurally simpler alternative is available and, since the alternative is equivalent to the blocked case, excluding the alternative would predict a contradiction. In other words, super-exhaustification as defined in (45) may potentially unify the TUH and other cases exhaustification.

(45) $[\![\textbf{s-exh}]\!](C)(p) = p \wedge \forall q \in C \wedge \mathrm{DC}(q) < \mathrm{DC}(p) : (q = \top) \vee \neg q$

Super-exhaustification negates not only non-weaker alternatives, but also equivalent ones and weaker ones as long as they are structurally simpler. Only the trivial case of $\top$ is excluded in (45). Super-exhaustivity thus predicts a stronger condition than the TUH: a sentence is also blocked if a weaker, non-trivial, and less complex alternative exists. In the case of scope economy, some support for precisely such a strengthening has been offered in work on generalized scope economy (Mayr & Spector 2010; Fleisher 2013).

Finally, consider briefly the scope preferences in (46) which played a major in role in the so called 'linguistic wars'. Chomsky (1972) and Jackendoff (1972) argued that quantifier scope here must depend on surface word order. But with the types $\langle\langle e, st\rangle, st\rangle$ for **PASS** and $tt$ for **not**, the representations in (47) have minimal dependency complexity (modulo intermediate variables).[26]

(46) a. Many targets weren't hit.
b. Someone didn't hit many targets.

(47) a. [[**many targets**] [$\lambda_y$ [**not** [**@** [**PASS** [**hit** $y$]]]]]]

---

[26] I assume that representations like (i) are realized with negation in a different position from (46a).

(i) [**not** [[**many targets**] [$\lambda y$ [**@** [**PASS** [**hit** $y$]]]]]]

b. [**someone** [$\lambda_x$ [**not** [[**many targets**] [$\lambda_y$ [**@** [$x$ [**hit** $y$]]]]]]]]

This observation in my view offers the promise of returning to an realizational architecture closer to the standard model of 1960s generative grammar.

## References

Alexiadou, Artemis, Elena Anagnostopoulou & Florian Schäfer. 2015. *External arguments in transitivity alternations: A layering approach*. Oxford University Press. doi:10.1093/acprof:oso/9780199571949.001.0001.

Bobaljik, Jonathan David & Susi Wurmbrand. 2012. Word order and scope: Transparent interfaces and the 3/4 signature. *Linguistic Inquiry* 43. 371–421. doi:10.1162/LING_a_00094.

Chierchia, Gennaro. 2013. *Logic in grammar: Polarity, free choice, and intervention*. Oxford, UK: Oxford University Press. doi:10.1093/acprof:oso/9780199697977.001.0001.

Chomsky, Noam. 1972. *Studies on semantics in generative grammar*. The Hague, Netherlands: Mouton. doi:10.1515/9783110814231.

Church, Alonzo. 1940. A formulation of the simple theory of types. *The Journal of Symbolic Logic* 5(2). 56–68. doi:10.2307/2266170.

Curry, Haskell B. 1930. Grundlagen der kombinatorischen Logik. *American Journal of Mathematics* 52. 509–536. doi:10.2307/2370716.

Dahl, Östen. 1973. On so-called 'sloppy identity'. *Synthese* 26(1). 81–112. doi:10.1007/BF00869757.

Dayal, Veneeta. 1996. *Locality in wh quantification*. Dordrecht, Netherlands: Kluwer. doi:10.1007/978-94-011-4808-5.

Elliott, Patrick D., Andreea C. Nicolae & Uli Sauerland. 2018. Who and what do who and what range over cross-linguistically? lingbuzz/003959.

Fleisher, Nicholas. 2013. Comparative quantifiers and negation: Implications for scope economy. *Journal of Semantics* 32(1). 139–171. doi:10.1093/jos/fft016.

Fox, Danny. 1999. Reconstruction, variable binding and the interpretation of chains. *Linguistic Inquiry* 30. 157–196. doi:10.1162/002438999554020.

Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, Mass.: MIT Press.

Fox, Danny. 2012. The semantics of questions. Class notes for MIT seminars.

Halle, Morris & Alec Marantz. 1993. Distributed morphology. In Ken Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, chap. 3, 111–176. MIT Press.

Heim, Irene. 2006. Little. *Proceedings of SALT* 16. doi:10.3765/salt.v16i0.2941.

Heim, Irene. 2017. Type economy. Unpublished handout.

Hirsch, Aron. 2017. *An inflexible semantics for cross-categorial operators*. Cambridge, Mass.: Massachusetts Institute of Technology, Ph.D. thesis.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press.

Jacobson, Pauline. 1990. Raising as function composition. *Linguistics and Philosophy* 13. 423–475. doi:10.1007/BF00630750.

Jacobson, Pauline. 1999. Towards a variable-free semantics. *Linguistics and Philosophy* 22. 117–184. doi:10.1023/A:1005464228727.

Johnson, Kyle. 2009. Gapping is not (VP-)ellipsis. *Linguistic Inquiry* 40. 289–308. doi:10.1162/ling.2009.40.2.289.

Katzir, Roni. 2007. Structurally-defined alternatives. *Linguistics and Philosophy* 30. 669–690. doi:10.1007/s10988-008-9029-y.

Magri, Giorgio. 2009. *A theory of individual-level predicates based on blind mandatory scalar implicatures. constraint promotion for optimality theory*. Cambridge, Mass.: Massachusetts Institute of Technology, Ph.D. thesis.

Marty, Paul P. 2017. *Implicatures in the DP domain*. Cambridge, Mass.: Massachusetts Institute of Technology, Ph.D. thesis.

Mayr, Clemens & Benjamin Spector. 2010. Not too strong! generalizing the scope economy condition. In *Proceedings of Sinn und Bedeutung 14*, 305–321.

Meyer, Marie-Christine. 2013. *Ignorance and grammar*. Cambridge, Mass.: Massachusetts Institute of Technology, Ph.D. thesis.

Meyer, Marie-Christine. 2015. Redundancy and embedded exhaustification. In S. D'Antonio, M. Moroney & C.R. Little (eds.), *Semantics and Linguistic Theory (SALT) XXV*, 491–511. doi:10.3765/salt.v25i0.3486.

Nicolae, Andreea C. 2013. *Any questions? polarity as a window into the structure of questions*. Cambridge, Mass.: Harvard University, Ph.D. thesis.

Nissenbaum, Jon. 2000. *Investigations of covert phrase movement*. Cambridge, Mass.: Massachusetts Institute of Technology, Ph.D. thesis.

Pesetsky, David. 1987. *Wh*-in-situ: Movement and unselective binding. In Eric Reuland & A. ter Meulen (eds.), *The representation of (in)definiteness*, 98–129. Cambridge, Mass.: MIT Press.

Sauerland, Uli. 2001. Intermediate cumulation. *Snippets* 4. 10–11.

Schönfinkel, Moses. 1924. Über die Bausteine der mathematischen Logik. *Mathematische Annalen* 92. 305–316. doi:10.1007/BF01448013.

Xiang, Yimei. 2016. *Interpreting questions with non-exhaustive answers*. Cambridge, Mass.: Harvard University, Ph.D. thesis.

Uli Sauerland, Schützenstr. 18, 10117 Berlin, Germany
uli@alum.mit.edu