A diachronic explanation for cross-linguistic variation in the use of inverse-scope constructions[*]

Omri Amiraz
The Hebrew University of Jerusalem

Abstract This study investigates the cross-linguistic variation in the use of inverse-scope constructions such as *All that glitters is not gold* for expressing "Not all X are Y" propositions. In particular, it seeks to explain why these constructions are in use and even common in some languages but lacking in others. It is argued that the cross-linguistic variation is explained by competition with alternative scope-transparent constructions, but only when the history of individual languages is taken into consideration. When a language develops a novel construction such as *Not all that glitters is gold*, which expresses scope relations transparently, it may take several centuries before this novel construction finally pushes a pre-existing inverse-scope construction out of use. In the time it takes for the blocking effect to manifest itself, inverse scope is used alongside its scope-transparent competitor. This diachronic path provides evidence for a bias for scope transparency and explains why inverse scope is blocked by competing constructions in some languages but not others.

Keywords: negation, quantifiers, scope, competition, diachronic change, typology, parallel corpus.

1 Introduction

There is a common claim in the literature that languages have a preference for expressing scope relations transparently by the surface word order—i.e., languages prefer to avoid inverse scope (see e.g., Pesetsky 1989; Beck 1996; Szabolcsi 1997; Bobaljik & Wurmbrand 2012). Inverse-scope constructions compete with alternative constructions that express the same meaning transparently by the surface word order. Thus, the availability of inverse scope is constrained in the following way: if scope relations can be expressed transparently, inverse scope is blocked or restricted to certain contexts; if the language's grammar does not allow the transparent word order, inverse scope is allowed. For instance, consider the inverse-scope reading

of (1), which is equivalent to 'Every student is such that at least one teacher talked to them (but not necessarily the same teacher)'. The claim is that inverse scope is allowed in (1) because the transparent word order in (2) is not possible in English. In contrast, languages like German, Hungarian, and Chinese, which have scrambling, tend not to allow inverse scope in this case since they can and prefer to express the scope relations transparently by the surface word order. Hence, scope rigidity is correlated with word order flexibility.

(1)   **At least** one teacher talked to **every** student.   [∃ > ∀; ∀ > ∃]

(2)   * To **every** student **at least one** teacher talked.

Bobaljik & Wurmbrand (2012) argue that scope freezing (meaning that only a surface scope reading is available) is construction specific rather than a general property of languages—i.e., there is no scope rigidity parameter. A language may exhibit scope freezing with respect to one construction and scope flexibility with respect to another since the availability of inverse scope depends on competition between alternative constructions, and in each case there is a different candidate set and a different set of relevant constraints. Bobaljik & Wurmbrand propose a violable constraint on form-meaning pairings called ScoT, articulated in (3), which favors scope transparency: it penalizes LF:PF pairings in which *A* takes scope over *B* at LF, but *B* precedes *A* at PF. ScoT can only be violated when it interacts with another constraint, which is respected by the candidate that violates ScoT (i.e., the inverse-scope candidate). In other words, inverse scope is only permitted as a last resort in cases where the scope-transparent alternative violates some constraint too (as in (2), which violates a certain constraint on movement).

(3)   *Scope Transparency* (ScoT) (Bobaljik & Wurmbrand 2012)
        If the order of two elements at LF is A≫B, the order at PF is A≫B.

This study investigates the cross-linguistic variation in the use of inverse-scope constructions such as (4), in which clausal negation linearly follows a universal quantifier in subject position but takes wide scope over it. The language sample that is used for this study contains 110 languages from diverse language families and linguistic areas. For 31 of these languages, diachronic data are also available. A definition of the type of inverse-scope construction under investigation is given in (5). These are sentences in which a quantifier phrase headed by a universal quantifier such as *all* or *every* occurs in subject position either as a determiner or as a pronoun and precedes clausal negation. Representative examples from the Corpus of Historical American English (COHA) are given in (6).

(4)   **All** that glitters is **not** gold.
        ≈ It is **not** the case that **all** that glitters is gold.   [¬ > ∀]

(5)   *An inverse-scope construction*
        A main declarative clause in which a universal quantifier in subject position linearly precedes clausal negation, but negation takes wide scope over the quantifier.[1]

(6)   a. **All** politicians are **not** opposed to reform. [COHA:MAG, 1876]

        b. **Every** thinker is **not** a philosopher. [COHA:NF, 1983]

        c. **Everybody** is **not** willing to take a joke, you know. [COHA:FIC, 1908]

It is often claimed that constructions like (4) and (6) are marked and require a special intonation contour (see e.g., Jackendoff 1969; Büring 1997). However, this study shows that this is only true of some languages. Inverse-scope constructions as defined in (5) are very common cross-linguistically—they are found in about half of the languages in the world-wide sample used for this study. In about a third of the languages in the sample, inverse scope is even the most common strategy for expressing "Not all X are Y" propositions (e.g., the most common and neutral way of saying *Not all of the students passed the exam* is a construction equivalent to *All of the students didn't pass the exam*). Furthermore, in many languages, the

---

1 The scope of this study is restricted to main declarative clauses because there are languages that normally do not allow an inverse-scope interpretation in sentences where a subject universal quantifier precedes clausal negation, but they do allow this interpretation when the clause is embedded in a downward-entailing context, e.g., Dutch (Zeijlstra 2017) and Hebrew (Amiraz 2019). Given that this type of data is not available for all the languages in the sample, only main declarative clauses are considered.

   Another type of sentence which is excluded from this study is ECHO DENIAL (sometimes referred to as emphatic denial or metalinguistic negation). Echo denial is a sentence that rejects a previous utterance by repeating it with the addition of a negative marker, e.g., in (i). In this case, too, an inverse-scope interpretation is possible even in languages that normally do not allow it (Zeijlstra 2017; Amiraz 2019). Another piece of evidence that echo denial differs from ordinary clausal negation is that it allows positive polarity items (PPIs) to occur in the scope of negation (Baker 1970: 169), e.g., in (ii), where the PPI *already* scopes below clausemate negation. Given that echo denial tends to preserve the word order and lexical choices of the previous utterance, I do not consider such sentences as evidence that a language uses inverse-scope constructions as defined in (5).

(i)   Polly Williams: They left the public schools, go into the private school, and the children are just like 100 percent changed. Children are happy. They love going to school. Their grades have improved. The parents are involved. **Everybody's happy**.
        Art Hackett: **But everybody is not happy**, perhaps least of all Wisconsin Superintendent of Public Instruction Herbert Grover. [COCA:SPOK, 1991; boldface added]

(ii)   No, Trump has **not already** won the election, and it is deeply irresponsible for him to say he has. [Shapiro, Ben. *Twitter*, 3 November 2020, twitter.com/benshapiro/status/1323890527767416834; boldface added]

surface-scope reading is unavailable or strongly dispreferred. That is, sentences like (4) cannot be interpreted as 'Nothing that glitters is gold'. This contrasts with other cases of scope ambiguity, where the surface-scope reading is typically preferred.

Assuming that there is a universal bias for scope transparency, one may ask why languages use inverse scope at all. I suggest that inverse scope is so common in constructions like (5) because in most languages, these constructions look just like plain negative sentences. For example, in English, standard negation is expressed by a construction in which the negative particle *not* follows the auxiliary (7). If the same negative construction is used to negate a sentence with a universal quantifier in subject position, the result is an inverse-scope construction (8). In order to avoid inverse scope, a different negative construction needs to be used such as one where *not* precedes the quantifier phrase (9). Yet, about half of the languages in the sample lack constructions like (9), and additional strategies (discussed in Section 3) are even rarer cross-linguistically. In other words, not all languages have a good alternative to inverse scope.

(7)   a. Mary saw John.

        b. Mary did**n't** see John.

(8)   a. **Everyone** saw John.

        b. **Everyone** did**n't** see John. $[\neg > \forall]$

(9)   **Not everyone** saw John. $[\neg > \forall]$

However, it is not the case that languages only use inverse scope as a last resort. According to the results of this study, about half of the languages in the sample have scope-transparent constructions such as (10), and about a third of these languages also use inverse-scope constructions like (4). In contrast, other languages that have constructions like (10) do not permit inverse scope. For example, Hebrew only allows a surface-scope reading in (11). This raises the following question: **if the availability of inverse scope depends on competition, why do we see a blocking effect in some languages but not others**?

(10)   **Not all** that glitters is gold. $[\neg > \forall]$

(11)   **kulam   lo**   heʔeminu li.
       everyone NEG believed   me

       'Nobody believed me.' $[\forall > \neg \,; *\neg > \forall]$

I propose that the availability of inverse scope is indeed influenced by competition with alternative scope-transparent constructions. However, the cross-linguistic variation is only explained when the history of individual languages is taken into

Strategy	Definition	Example
A	Inverse scope	***All*** *of the cookies did**n't** burn*
B	Negation precedes QP	***Not all*** *of the cookies burned*
C	VS word order	***Not*** *burned **all** of the cookies*
D	Cleft	*It's **not all** of the cookies that burned*
	Superordinate negation	*It's **not** (the case) that **all** the cookies burned*
E	Floating quantifier	*The cookies did**n't all** burn*

Table 1: Five strategies for expressing "Not all X are Y" propositions

consideration. The bias for scope transparency is a WEAK UNIVERSAL FORCE (a term coined by Seržant (2019))—its synchronic effect is weak, but it drives systematic historical developments in unrelated language families and linguistic areas. Languages tend to develop novel constructions that express scope relations transparently, and these gradually replace pre-existing inverse-scope constructions.[2] In the time it takes for the blocking effect to manifest itself, both the older inverse-scope construction and the novel scope-transparent construction are available. Hence, the effect of the bias for scope transparency is sometimes too weak to be observed on the synchronic level.

This paper is structured as follows: Section 2 presents the parallel corpus that was used for this study: contemporary and historical translations of the New Testament into 110 languages. Section 3 defines five major strategies that languages use for expressing "Not all X are Y" propositions, as presented schematically in Table 1, and provides representative examples for each strategy.[3] Section 4 presents the synchronic and diachronic results. It addresses both empirical and theoretical questions: (i) which languages use inverse-scope constructions and which do not? (ii) how to explain the cross-linguistic variation? Section 5 concludes.

---

2 I remain agnostic about the locus and mechanism of change. One possibility is that there is a bias for scope transparency in language acquisition, and its effect is magnified over time due to iterated learning (see e.g., Kirby, Cornish & Smith 2008; Reali & Griffiths 2009; Kirby, Griffiths & Smith 2014).

3 For the purpose of this study, I assume that the candidate set consists of constructions that express the same truth-conditional meaning but do not necessarily share the same LF. For instance, one may argue that (4) and (10) have different LFs, and so they would not belong to the same candidate set in certain models such as Bobaljik & Wurmbrand 2012. This working definition of a candidate set is hopefully suitable for diachronic competition, which is the primary focus of this study. However, in a model of synchronic competition, it would probably overgenerate inverse-scope readings because the likelihood of constraint ties would be significantly higher due to the size of the candidate set. I leave it for future research to propose a more restrictive definition of the candidate set.

Strategy	Text	Translation	Year
A	**All** men can**not** *receive this saying*	KJV	1611
B	***Not everyone*** *can accept this teaching*	NRSV	1989
C	***ne*** *under-foð* ***ealle*** *menn þis word*	Wessex Gospels	990
D	*It is* **not everybody** *who can live up to this*	Phi	1958
Paraphrase	*This teaching does* **not** *apply to* **everyone**	GNT	1976

Table 2: Comparison of English translations of Matthew 19:11

2 Methodology: The New Testament corpus

Inverse-scope constructions are rarely described in grammars. Therefore, this study requires collecting primary cross-linguistic data. For this purpose, I use translations of the New Testament (NT) as a parallel corpus.[4] The language sample contains 110 languages from diverse language families and linguistic areas. For 31 of these languages, historical translations are also available, which is a major advantage of this corpus. Whenever possible, at least two different translations were used for each language in order to have more data points. Each data point was categorized into one of the five strategies listed in Table 1, if possible. Otherwise, it is considered a paraphrase or a borderline case.

The Greek source text uses a *B* construction 12 out of 14 times in the relevant verses, e.g., in (12). If a translation uses a different construction instead, it is taken as evidence that this construction is in use in the language. For instance, Table 2 presents five English translations of (12) from different time periods.[5] Given that the translator(s) chose to depart from the source text by using a different construction, it is likely that the translation reflects the grammar of the target language.

(12)　**ou　pántes** xōroûsi　　　　tòn　　lógon　　toûton
　　　NEG all　　have.room.for.PRS.3PL DEF.ACC word.ACC this.ACC

　　　'Not everyone can accept this teaching' [Mt. 19:11, SBLGNT, 2010]

One may wonder whether the NT corpus is a good indicator of the relative frequency of the various strategies in a given language compared to the results that would be obtained from a larger and more balanced corpus. As an approximation, I compared the frequency of strategy *A* in the NT corpus to its frequency in the larger Europarl corpus for 16 languages. The Europarl corpus consists of the proceedings

---

4 The NT has been previously used in typological studies, and it has some advantages and disadvantages. See De Vries 2007 and Wälchli 2007 for an overview.

5 No translation was found where strategy *E* is used in this verse, but here is an example from another verse: *But they did* **not all** *hearken to the glad tidings.* [Rom. 10:16, ASV, 1901].

of the European Parliament from 1996 to 2012. While this corpus is mostly made up of translated texts too, it contains interactions between speakers, and thus it is more similar to natural discourse than the NT corpus. The results from the two corpora proved to be highly correlated—in languages where strategy *A* was common in the NT corpus (e.g., Finnish and Swedish), it was also common in the Europarl corpus; in languages where strategy *A* was unattested in the NT corpus (e.g., Spanish and Dutch), it was likewise unattested in the Europarl corpus. In two languages, namely Danish and Hungarian, strategy *A* was unattested in the NT corpus but rarely attested in the Europarl corpus (under 1% of the results). Such cases are expected since the NT corpus is relatively small, and some rare constructions are likely to be absent from it. Therefore, some instances of false negatives are expected due to languages where a certain strategy is rare but nonetheless existent. Overall, I conclude that the NT corpus provides a good estimate of the relative frequency of the various strategies used by a given language.

Another potential concern is that inverse-scope constructions might be more common in speech than in writing, which would mean that the results of the corpus study are not representative of the language at large. Unfortunately, I was not able to test this. However, Tottie & Neukom-Hermann (2010) studied the use of *all...not* sentences in British English using the British National Corpus (BNC), and they observe that there is no significant difference between spoken and written language in the frequency of the inverse-scope interpretation.[6]

3   Strategies for expressing "Not all X are Y" propositions

This section defines and illustrates five strategies for expressing "Not all X are Y" propositions distinguished in this study. For space reasons, this section abstracts away from dimensions of cross-linguistic variation that are tangential to the concerns of this study. However, it aims to provide enough information so that the reader can infer how a given construction would be categorized. For a detailed discussion of borderline cases, I refer the reader to Amiraz forthcoming.

3.1   Strategy *A*

Strategy *A* is an inverse-scope construction, as defined in (13).

---

6 Neukom-Hermann (2016) tentatively suggests that in German, inverse-scope constructions seem to be more common in speech than in writing (she did not have access to a corpus of spoken German, so the results are based on a private collection of examples). However, most of the examples she gives from spoken language fall under the category of echo denial (see Footnote 1), and these are not relevant to the present study.

(13) *Strategy A*
      A mono-clausal construction in which a universal quantifier in subject position linearly precedes clausal negation.

Strategy *A* is defined in terms of linear precedence rather than hierarchical relations—i.e., no distinction is made between the following two cases: (i) the subject is higher than negation at PF and undergoes scope reconstruction at LF; (ii) negation is higher than the subject at PF, and scope reconstruction is not required.

For instance, in English, the subject moves from its base position (e.g., the specifier of *v*P in transitive clauses) to its surface position at the specifier of IP. Negation is higher than the base subject position but lower than the surface subject position. Thus, in sentences like (14), an interpretation where negation takes wide scope over the quantifier requires scope reconstruction of the subject quantifier.

(14) for $[_{IP}$ **all** men_1 have_2 **not** $[_{vP}$ t_1 t_2 faith]]. [2 Thess. 3:2, KJV, 1611]

In contrast, there are languages where negation is argued to be higher than the surface position of non-topicalized/scrambled subjects. For example, Öztürk (2005: 131) argues that in Turkish, a non-scrambled subject remains low in its base position, where it is c-commanded by negation (15). Homer & Bhatt (2020) argue that in Hindi-Urdu, the subject does raise to a higher position (the specifier of Asp), but negation is even higher. In both cases, an interpretation where negation takes wide scope over the quantifier does not require scope reconstruction of the quantifier since the subject is lower than negation in its surface position.

(15) $[[$**bütün** çocuklar o   test-e t_1 $_{vP}]$ gir-**me**-di$_1$ $_{IP}]$.
      all   children that test-DAT   take-NEG-PST

      'Not all of the children took that test.' (Öztürk 2005: 131)

Treating sentences like (15) as involving inverse scope is controversial since the scope relations match the hierarchical relations. For the purpose of this study, the question is what matters for ScoT in this case—c-command or linear precedence?[7] If ScoT is only sensitive to c-command relations, the fact that negation can take wide scope over a subject quantifier phrase in sentences like (15) follows directly from the syntactic structure and does not require any further explanation—this construction respects ScoT and does not involve covert movement. However, my working assumption is that ScoT is violated in these sentences. Consider the contrast

---

7 Bobaljik & Wurmbrand (2012: 372-373) define ScoT in terms of linear precedence, but they take linear precedence to be a proxy for c-command, assuming that these two notions coincide in the cases they discuss. Most previous accounts of scope taking and scope rigidity are also based on c-command, e.g., the Scope Principle (May 1985; Aoun & Li 1989) and the Isomorphic Principle (Huang 1982).

in (16) in Hindi-Urdu. The universal quantifier *har* prefers to scope below negation from subject position (16a), whereas *bahut* 'many' prefers to scope above negation in the same configuration (16b).[8] Suppose that ScoT is sensitive to c-command relations but not linear precedence. In that case, it is puzzling that the reading where *bahut* 'many' scopes below negation is so unsalient given that it presumably respects ScoT. In contrast, if ScoT is sensitive to linear precedence, we expect to find cases such as (16b) where a wide-scope reading of negation is dispreferred even in languages in which negation is higher than the surface subject position since these sentences presumably violate ScoT.

(16)   a. **har**   laṛke=ne vo kitaab **nahĩ:** paṛhi:.
           every boy-ERG that book NEG read.PFV.F.SG

           'Not every boy read that book.' (Homer & Bhatt 2019)

       b. **bahut** bacchon=ne kitaab **nahĩ:** paṛhi:.
           many child-ERG   book NEG read.PFV.F.SG

           'Many children didn't read the book.' (Ashwini Deo, p.c.)

3.2   Strategy *B*

Strategy *B* is defined in (17).

(17)   *Strategy B*
      A mono-clausal construction in which a negative marker precedes a quantifier phrase in pre-verbal subject position.

This strategy consists of two types of constructions: (i) negated quantifier phrase constructions; and (ii) clause-initial negative markers. A negated quantifier phrase construction is a negative construction distinct from standard negation in which a negative marker precedes a quantifier phrase. For example, in English, standard negation is expressed by a construction in which the negative particle *not* follows the auxiliary (18). When the subject is a universal quantifier, *not* can precede the quantifier phrase (19). Following Collins 2020, I assume that negation combines

---

8 The reading where *bahut* scopes below negation becomes more salient if the object is scrambled (i) or occurs post-verbally (Ashwini Deo, p.c.). I leave it for future research to explain why changing the word order makes this reading more salient and why different types of quantifiers have different scope preferences. I thank Ashwini Deo and Rajesh Bhatt for a fruitful discussion of these data.

(i)   kitaab **bahut** bacchon=ne **nahĩ:** paṛhi:.
       book many child-ERG   NEG read.PFV.F.SG

       'Not many children read the book.' (Ashwini Deo, p.c.)

with the entire quantifier phrase rather than just with the determiner (i.e., *not every* is not a constituent).

(18) Mary did**n't** see John.

(19) [*IP* [*NegP* **Not** [*QP* **every** great restaurant]] has a Michelin star]. [COCA:BLOG]

Other languages have negative constructions in which negation occurs in clause-initial position without forming a constituent with the quantifier phrase. For example, in Ojibwe (20), the negative particle *gaawiin* occurs either before or after the subject (Tilleson 2019: 243, 253). Expectedly, the negative particle can also precede a subject quantifier phrase (21). Given that (21) is not distinct from standard negation, it is not considered a negated quantifier phrase construction. Also note that negation in Ojibwe is expressed by a discontinuous negative marker consisting of the pre-verbal particle *gaawin* and the verbal suffix *-siin*. Such constructions are considered to be strategy *B* as long as the first negative marker precedes the quantifier phrase.

(20) {Inini **gaawiin** / **Gaawiin** inini} o-gii-waabam-aa-**siin**   ikwe-wan.
      man   NEG   NEG   man   3SG-PST-see-DIR-NEG woman-OBV

      'The man didn't see the woman.' (Tilleson 2019: 253)

(21) [*CP* **Kaawiin** tahsh [*IP* **kahkina** o-kii-'otaahpin-an-**sii**-naawaa
          NEG   but   all   3-PST-accept-DIR-NEG-3PL
      minwaacimowin]].
      good.news

      'But not all have obeyed the good news' [Rom. 10:16, Ojibwe Kihci-masina'ikan, 1988]

3.3 Strategy *C*

Strategy *C* is defined in (22).

(22) *Strategy C*
      A mono-clausal construction in which a quantifier phrase occurs in post-verbal subject position (i.e., the word order is VS), and clausal negation occurs earlier in the clause.

In some languages, the basic word order is VS, and clausal negation precedes the surface position of non-topicalized/scrambled subjects. In Irish, standard negation is expressed by the pre-verbal negative particle *ní* (Stenson 2019: 42), and strategy *C* looks the same as standard negation or negative copular clauses (23).

(23) ach [$_{IP}$ **Níl**$_1$ [$_{PredP}$ **gach** ní t_1 fóinteach]].
      but   NEG.COP.PRS   every thing   helpful

      'but not all things build up.' [1 Cor. 10:23, An Bíobla Naofa, 1981]

In other languages, the basic word order is SV, but VS is a possible alternative in certain contexts. In Slovenian, the basic word order is SVO (Herrity 2015: 333), but VS tends to be used in "Not all X are Y" propositions (24). Note that in SV languages, strategy *C* often involves fronting of another constituent to clause-initial position (e.g., the object in (24)), so the verb is not the first constituent in the clause.

(24) Vendar [$_{CP}$ tega   spoznanja$_2$ [$_{IP}$ **nimajo**$_1$ [$_{vP}$ **vsi** t_1 t_2]]].
      however   this.GEN knowledge.GEN   NEG.have.3PL   all

      'But not everyone possesses this knowledge.' [1 Cor. 8:7, Slovenski Standardni Prevod, 1996]

Another type of *C* construction is negative fronting. In Icelandic, standard negation is expressed by the negative particle *ekki*, which follows the auxiliary or finite verb. In certain contexts, the negative particle *ekki* can be fronted to clause-initial position, where it precedes the auxiliary or finite verb (Thrainsson 2007: 82). The same negative fronting construction is used as strategy *C* in (25).

(25) [$_{CP}$ **Ekki**$_2$ hafa$_1$ [$_{IP}$ **allir** t_1 t_2 þessa þekkingu].
           NEG have   all   this   knowledge

      'But not everyone possesses this knowledge.' [1 Cor. 8:7, Icelandic Bible Society, 2007]

3.4 Strategy *D*

Strategy *D* is defined in (26).

(26) *Strategy D*
      A mono-clausal or bi-clausal construction in which the quantifier phrase is a cleft or focus constituent and/or negation is superordinate.

This strategy consists of several types of constructions containing an additional element other than the negative marker and the quantifier phrase. This element can be either a relativizer, a complementizer, a focus marker, or a copula. The first type of *D* construction is a cleft in which the quantifier phrase is the cleft constituent, and the nuclear scope of the quantifier is expressed by the cleft clause. Both negation and the quantifier phrase are in the main clause, as demonstrated in (27).

(27) It is **not everybody** [$_{CP}$ who can live up to this] [Mt. 19:11, Phi, 1958]

Some languages have mono-clausal focus constructions in which the focused constituent is marked by a focus marker. In Jamaican Creole, the focus marker *a* precedes the focused constituent (Durrleman 2008: 74), e.g., in strategy *D* in (28).

(28)   Bot [$_{CP}$ a   **no evribadi**$_1$ [$_{IP}$ t_1 nuo   dis]].
       but   FOC NEG everybody   know this

       'But not everyone possesses this knowledge.' [1 Cor. 8:7, Di Jamiekan Nyuu Testiment, 2012]

Another type of *D* construction is superordinate negation. In (29) in Japanese, the quantifier phrase is embedded inside a nominalized clause headed by the noun *wake* 'conclusion, story, excuse', and negation negates the main clause copula. This construction is not considered inverse scope since negation is in a higher clause.[9]

(29)   [seito-tati-no   **zenin**-ga siken-ni   ukat-ta $_{CP}$] wake-de-wa
       student-ASS-GEN all-NOM exam-DAT pass-PST   conclusion-COP-FOC
       nai.
       NEG

       'Not all of the students passed to exam.' (lit. 'It is not the conclusion/story that all of the students passed to exam.') (Akitaka Yamada, p.c.)

3.5   Strategy *E*

Strategy *E* is defined in (30).

(30)   *Strategy E*
       A mono-clausal construction in which the subject position is filled by a lexical noun phrase or a pronoun, and clausal negation precedes a floating, adverbial, or right-dislocated quantifier.

English has a floating quantifier construction in which *all* follows the negative particle and precedes the predicate, and the subject position is filled by the noun phrase or pronoun that functions as the restriction of the quantifier (31).

(31)   The Tea Party people$_1$ are$_2$ **not** [$_{PredP}$ **all** t_1 t_2 Republicans]. [COCA; SPOK]

---

9 Analyzing (29) as involving surface scope is at odds with the assumption that ScoT is sensitive to linear precedence (see the discussion in Subsection 3.1). However, I assume that such *D* constructions do respect ScoT. The difference between strategy *D* and strategy *A* is that in the former, the hierarchical relations can be read off the surface structure, while in the latter, the relative position of negation and the quantifier phrase cannot be inferred from the surface structure, and additional diagnostics are required (see e.g., Homer & Bhatt 2019, 2020). If this difference is indeed meaningful, it is predicted that the type of quantifier will not have an effect on the scope of negation in *D* constructions like (29), unlike in *A* constructions such as (16).

In other languages, the quantifier follows the predicate and occurs at the right edge of the clause, e.g., (32) in Zulu.

(32) Izingane$_2$ **a**-zi-thand-**i**$_3$ [$_{vP}$ t_1 t_2 t_3 amaswidi] **zonke**$_1$.
      10.children NEG-10.SM-love-NEG   6.sweets   10.all
      'Not all children like sweets.' (Buell 2008: 46)

Some languages have adverbial quantifiers which are not derived from or related to determiners. In Mandarin Chinese, the adverbial quantifier *dōu* always precedes the verb and cannot function as a determiner (Cheng 1995: 198-199). In (33), the negative particle *méiyǒu* precedes *dōu*.

(33) neixie ren$_1$ **meiyou dou** [$_{vP}$ t_1 kan-guo neiben shu].
      those person NEG   all   read-ASP that   book
      'Not all of these people read that book.' (Cheng 1995: 199)

Note that in some (and perhaps most) languages, strategy *E* is restricted to cases where the subject is a lexical NP or a referential pronoun. For example, (34) does not have a good strategy *E* paraphrase in English.

(34) **Not everybody** is going to be a musician. [COCA; NEWS]
      ≉ People are **not all** going to be musicians.

4 Results

The hypothesis that this paper sets out to test is that languages prefer to express scope relations transparently, i.e., there is a slight bias towards candidates that respect ScoT. This hypothesis makes diachronic predictions: (i) languages tend to develop novel constructions that express scope relations transparently; and (ii) these novels constructions gradually replace pre-existing inverse-scope constructions. This means that if a language uses both a strategy *A* construction, e.g., (35a), and a strategy *B* construction, e.g., (35b), the former is older because otherwise the language would have no functional motivation to start using *A*.

(35) a. **All** that glitters is **not** gold.

      b. **Not all** that glitters is gold.

These predictions are restricted to the competition between strategies *A* and *B*. My working assumption is that strategy *B* does not violate any major constraint. Given that ScoT does not interact with another constraint, the bias for strategy *B* is expected to lead to a blocking of *A* over time. In contrast, strategy *C* is not expected to block strategy *A* by itself since there can be an interaction between ScoT

and various word order constraints that favor SV word order in certain contexts—especially in languages where the basic word order is SV (e.g., when the quantifier phrase is a contrastive topic, see Büring 1997). Hence, it is expected that in contexts where ScoT conflicts with another constraint, both *C* and *A* would be possible (see, e.g., example (38) below). As for strategy *E*, in many (and perhaps most) languages, strategy *E* can only be used when the restriction of the quantifier is specified by a common noun or a personal pronoun, and this strategy is not an option when the restriction is not overtly specified, as in *everyone* or *everything* (see Subsection 3.5 for discussion). Therefore, strategy *E* is typically not an all-purpose strategy, and it is not expected to block strategy *A* by itself. Finally, I remain agnostic as to whether strategy *D* violates any major constraint, so no predictions about the competition between *A* and *D* were made beforehand.

4.1 Synchronic results

The language sample used in this study is imbalanced because European languages are vastly over-represented. The sample was constructed in this way because the diachronic data come exclusively from European languages, so as many European languages as possible were included in the sample. The Eurocentric bias in the sample was alleviated by subsampling without replacement:[10] random subsamples were created in which a single language from each genus was randomly chosen (e.g., one Germanic language, one Romance language, etc.). This process was repeated 1,000 times, and then the results were averaged.

Table 3 presents the proportion of languages in the balanced sample where each strategy is (i) attested; (ii) the most common strategy in the language. Strategies *A* and *B* are the most common strategies cross-linguistically, and they are both found in about half of the languages in the balanced sample. Despite being attested in approximately the same frequency, when a language has *B*, it tends to be its most common strategy. In contrast, strategy *A* is quite often a secondary strategy. Strategies *C*, *D*, and *E* are significantly less common cross-linguistically, but there are certainly languages that use them as their primary strategy.[11] Note that there are likely to be some instances of false negatives due to languages where a certain strategy is rare but nonetheless existent (see Section 2). Therefore, the numbers in the row of attested strategies are probably somewhat underestimated.

---

10 See Politis, Romano & Wolf 1999.

11 The prevalence differences between the strategies are partially explained by their correlation with basic word order since not all word orders are equally common cross-linguistically. For example, strategy *A* is very common in verb-final languages, whereas strategy *B* is uncommon in these languages except in the Americas. Strategy *C* is common in verb-initial languages but is also found in some verb-medial languages (see Subsection 4.3.1). For further discussion, see Amiraz forthcoming.

	A	B	C	D	E
Attested	54%	51%	18%	27%	14%
Most common	31%	41%	7%	20%	1%

Table 3: Proportion of languages in the balanced sample where each strategy is (i) attested; (ii) the most common strategy

4.2 Diachronic results

The diachronic results are based on a sample of 31 languages from Europe, belonging to three language families: Indo-European, Uralic, and Basque. Most of these languages have NT translations from at least as early as the 16-17th centuries, but for a few languages the corpus covers longer or shorter time spans. The strongest evidence for the diachronic effect of ScoT comes from Romance and Germanic languages. Most languages in these families went through the following path:[12]

- **Stage 1**: *A* is a common strategy—in most cases, it is in variation with *C* or *E*.

- **Stage 2**: The language develops a novel *B* construction. This happens around the 14th-17th centuries in most cases. Over several centuries, *B*'s frequency increases, and *A*'s frequency decreases.

- **Stage 3**: *A* either falls out of use or becomes extremely rare and used mostly in formulaic expressions.

The diachronic path is demonstrated for English. In Old English, the corpus only contains strategy *C*, where the negative clitic *ne* occurs clause-initially immediately before the finite verb (36a). This is the most common word order in standard negation in Old English (van Kemenade 1997; Cichosz 2020). In Middle English, this pattern becomes increasingly rare until it finally falls out of use (Wallage 2012: 7). At the same time, an *A* construction begins to be attested (36b).[13] In the late 14th century, a *B* construction appears, which is exemplified by a later example in (36c).

---

12 Three languages in these families do not conform to this path. Icelandic has not developed a *B* construction—the most common strategy is *C* (see e.g., (25)), and *A* is rare. Standard European French also lacks *B*—the most common strategy in the corpus is *A*, and *D* and *E* are common too. However, Quebec French developed a *B* construction (possibly due to contact with English), and *A* is no longer used in the spoken language (Justin Royer, p.c.). In Swedish, *B* was more common than *A* at a certain point in time, but the picture changed when the negative particle *inte* replaced the now obsolete *icke*, and *B*'s frequency dropped. Today, *A* is about as common as *B*.

13 There are earlier examples in other corpora dating back to the early 13th century, but (36b) is presented here for the sake of clarity since it is a translation of the same verse as (36a).

(36)   a. **Ne** gæð **ælc** þæra   on heofena rice þe cwyþ to me drihten drihten
              NEG goes all of.those on heaven's realm that says to me Lord   Lord

              'Not everyone who says to me, 'Lord, Lord,' will enter the kingdom of heaven' [Mt. 7:21, Wessex Gospels, circa 990]

       b. **al** þilke þat seyȝen to me lord lord **ne** schul **not** entren in to þe kyngdome of heuene [Mt. 7:21, Life of Soul, circa 1400]

       c. **Not all** they that saye vnto me Master Master shall enter in to the kyngdome of heven [Mt. 7:21, Tyndale Bible, 1526-1535]

However, the novel *B* construction does not immediately replace the older *A* construction. In fact, *A* appears to have remained more common than *B* for several centuries. The King James Version (1611) contains seven instances of *A* and but one instance of *B*, and a similar distribution is observed in other contemporary translations. Examples of *A* are easy to find at least as late as the 19th century, e.g., (37a). Today, in most dialects of English, *A* is not fully productive, and these constructions are mostly found in idiomatic and semi idiomatic expressions such as (37b) and (37c) (Tottie & Neukom-Hermann 2010). For instance, a sentence like *Everyone isn't hungry* is considered odd or even unacceptable by speakers of most dialects of English (except as an echo denial, see Footnote 1). In other words, strategy *B* all but pushed strategy *A* out of use in most dialects of English, but it took it half a millennium to do so. This shows that a frequently used *A* construction can be maintained for centuries even though there is a bias towards its competitor.

(37)   a. 'Hunters!' repeated Edward — 'But why must you have hunters? **Everybody** does **not** hunt.' [Sense and Sensibility, Jane Austen, 1811]

       b. **All** is **not** lost.

       c. Sweetie, **everything** is**n't** about money. [Big Little Lies, S02E05]

Given the unidirectionality of the diachronic path, one may ask how it came about that so many languages still use *A*. If a language can only lose *A*, one would expect that strategy *A* would become extinct. I argue that strategy *A* is still common because the diachronic path can be cyclic. A stage-3 language may lose its *B* construction for independent reasons, and so the cycle can be repeated. This may happen when the language develops a novel negative marker or universal quantifier, and the new item is banned or dispreferred in the pre-existing *B* construction. For instance, Latin had a *B* construction: *non omnis* 'not all'. Romance languages developed a universal quantifier from the Latin adjective *tōtus* 'whole' (Haspelmath 1995: 364), and the *B* construction was not extended to the new quantifier. It took several centuries for a *B* construction to rearise (14th-15th centuries in most cases). In the meantime, most of these languages used strategies *C* and *A*. In Greek, not only the quantifier was

replaced (Koine Greek *pántes* vs. Modern Greek *óli*, derived from *hólos* 'whole' (Haspelmath 1995: 365)), but also the negative marker was replaced (Koine Greek *ou* vs. Modern Greek *dhen*, derived from *oudhén* 'not at all'). To this day, Modern Greek lacks a *B* construction equivalent to Koine Greek *ou pántes* 'not all', and strategy *C* is by far the most common strategy.

4.3   The other strategies

The previous subsections focused on the competition between strategies *A* and *B*. This subsection discusses strategies *C* and *D*, which also provide evidence in some cases for the effect of ScoT (strategy *E* will not be further discussed for reasons of space—I refer the reader to Amiraz forthcoming). While languages usually develop strategy *B* constructions specifically for the purpose of respecting ScoT (but not always, see e.g., (21)), the other scope-transparent strategies typically make use of independently existing constructions.

4.3.1   Strategy *C*

Strategy *C* is common in verb-initial languages such as Celtic languages and Malagasy (in both cases, *D* is very common too), but it is also the preferred strategy in a few verb-medial languages. For instance, this strategy is common in certain Balkan languages, which are verb-medial. In this linguistic area, we observe a striking pattern. In languages where *B* is the most common strategy (Albanian, Bulgarian, and Romanian), *A* is not attested in the corpus. In languages where *C* is the most common strategy (Greek, Macedonian, Serbian-Croatian-Bosnian, and Slovenian), *A* is used as a secondary strategy. The fact that *C* tends not to block *A* in verb-medial languages is predicted since ScoT interacts with information structure in some contexts. Consider (38) in Slovenian, where *vse* 'all' is a contrastive topic. Assuming that information structure favors a clause-initial position for contrastive topics, information structure and ScoT impose conflicting requirements in (38): the former favors SV word order, while the latter favors VS word order. As a result, both *A* and *C* are possible.

(38)   »Vse je dovoljeno,« vendar **ni**   **vse** koristno! »Vse je dovoljeno,« toda
       all   is lawful   but   NEG.is all beneficial all   is lawful   but
       vse ne   izgrajuje.
       all NEG builds

   '"All things are lawful," but not all things are beneficial. "All things are lawful," but not all things build up.' [1 Cor. 10:23, Slovenski Standardni Prevod, 1996]

4.3.2 Strategy *D*

Strategy D is another case where an independently existing construction is recruited. The primary function of these constructions tends to be related to information structure, but in the context of "Not all X are Y" propositions, their use is sometimes due to ScoT rather than information structure.

In Japanese, the most common strategy for expressing "Not all X are Y" propositions is a strategy *D* construction in which the quantifier is inside a nominalized clause, which is negated by *wake de wa nai* 'it doesn't mean that ...'. The primary function of *wake de wa nai* is to deny an accessible, discourse-old proposition (Yoshimura 2013). Yet, in the context of "Not all X are Y" propositions, this construction is pragmatically neutral and does not have its usual discourse requirements—i.e., it can be used out-of-the-blue as a neutral description of a fact (Akitaka Yamada, p.c.), e.g., in (29), repeated here as (39). In other words, a construction whose function is typically related to information structure is recruited for the purpose of respecting ScoT.

(39) seito-tati-no   **zenin**-ga siken-ni   ukat-ta   wake-de-wa   **nai**.
       student-ASS-GEN all-NOM exam-DAT pass-PST conclusion-COP-FOC NEG

      'Not all of the students passed to exam.' (Akitaka Yamada, p.c.)

5 Conclusion

In this paper, I argued that ScoT is a weak universal force—its synchronic effect is weak, but it drives systematic historical developments. In order to respect ScoT, a language may: (i) develop a novel scope-transparent construction (typically, a strategy *B* construction); or (ii) recruit an independently existing construction (e.g., a strategy *D* construction). However, it may take several centuries until an older inverse-scope construction is finally pushed out of use. This means that the relative frequency of the competing constructions plays a major role, and it can even outweigh a universal force like ScoT, preventing a blocking effect on the synchronic level.

The claim that a weak universal force like ScoT can shape languages over time is similar in spirit to works on universal biases in other grammatical domains such as word order (e.g., Culbertson, Smolensky & Legendre 2012) and argument marking (e.g., Bickel, Witzlack-Makarevich, Choudhary, Schlesewsky & Bornkessel-Schlesewsky 2015; Seržant 2019). Thus, this study supports a long-standing claim in Greenbergian typology (see e.g., Greenberg 1966, 1969; Bybee 1988) that synchronic cross-linguistic distributions, e.g., the prevalence of inverse-scope constructions, are due to diachronic tendencies of change and retention. These diachronic tendencies are in turn explained by synchronic factors of grammar and usage, e.g., the bias for scope transparency.

References

Amiraz, Omri. 2019. Semi positive polarity items. In Maggie Baird & Jonathan Pesetsky (eds.), *North East Linguistic Society (NELS) 49*, vol. 1, 49–62. Amherst, MA: GLSA.

Amiraz, Omri. forthcoming. *The scope of negation*. Jerusalem: The Hebrew University of Jerusalem PhD dissertation.

Aoun, Joseph & Yen-hui Audrey Li. 1989. Scope and constituency. *Linguistic Inquiry* 20(2). 141–172.

Baker, C. L. (Lee). 1970. Double negatives. *Linguistic Inquiry* 1(2). 169–186. doi:10.1080/08351816909389104.

Beck, Sigrid. 1996. Quantified structures as barriers for LF movement. *Natural Language Semantics* 4(1). 1–56. doi:10.3403/30338106u.

Bickel, Balthasar, Alena Witzlack-Makarevich, Kamal K. Choudhary, Matthias Schlesewsky & Ina Bornkessel-Schlesewsky. 2015. The neurophysiology of language processing shapes the evolution of grammar: Evidence from case marking. *PLoS One* 10(8). e0132819. doi:10.1371/journal.pone.0132819.

Bobaljik, Jonathan David & Susi Wurmbrand. 2012. Word order and scope: Transparent interfaces and the 3/4 signature. *Linguistic Inquiry* 43(3). 371–421. doi:10.1162/ling_a_00094.

Buell, Leston. 2008. VP-internal DPs and right-dislocation in Zulu. *Linguistics in the Netherlands* 25(1). 37–49. doi:10.1075/avt.25.07bue.

Büring, Daniel. 1997. The great scope inversion conspiracy. *Linguistics and Philosophy* 20(2). 175–194. doi:10.4324/9780203448014-10.

Bybee, Joan L. 1988. The diachronic dimension in explanation. In John A. Hawkins (ed.), *Explaining Language Universals*, 350–379. Oxford: Basil Blackwell.

Cheng, Lisa Lai-Shen. 1995. On *dou*-quantification. *Journal of East Asian Linguistics* 4(3). 197–234. doi:10.1007/bf01731509.

Cichosz, Anna. 2020. Negation and verb-initial order in Old English main clauses. *Journal of English Linguistics* 48(4). 355–381. doi:10.1177/0075424220941911.

Collins, Chris. 2020. Outer negation of universal quantifier phrases. *Linguistics and Philosophy* 43(3). 233–246. doi:10.1007/s10988-019-09269-4.

Culbertson, Jennifer, Paul Smolensky & Géraldine Legendre. 2012. Learning biases predict a word order universal. *Cognition* 122(3). 306–329. doi:10.1016/j.cognition.2011.10.017.

De Vries, Lourens. 2007. Some remarks on the use of Bible translations as parallel texts in linguistic research. *Language Typology and Universals* 60(2). 148–157. doi:10.1524/stuf.2007.60.2.148.

Durrleman, Stephanie. 2008. *The Syntax of Jamaican Creole: A Cartographic Perspective*. John Benjamins Publishing.

Greenberg, Joseph H. 1966. Synchronic and diachronic universals in phonology. *Language* 42(2). 508–517. doi:10.2307/411706.

Greenberg, Joseph H. 1969. Some methods of dynamic comparison in linguistics. In Jaan Puhvel (ed.), *Substance and Structure of Language*, 147–204. Berkeley, CA: University of California Press.

Haspelmath, Martin. 1995. Diachronic sources of 'all' and 'every'. In Emmon W. Bach (ed.), *Quantification in Natural Languages*, 363–382. Springer Netherlands.

Herrity, Peter. 2015. *Slovene: A Comprehensive Grammar*. London: Routledge.

Homer, Vincent & Rajesh Bhatt. 2019. Licensing of PPI indefinites: Movement or pseudoscope? *Natural Language Semantics* 27(4). 279–321. doi:10.1007/s11050-019-09155-6.

Homer, Vincent & Rajesh Bhatt. 2020. Restructuring and the scope of negation in Hindi-Urdu. *Glossa: A Journal of General Linguistics* 5(1). doi:10.5334/gjgl.927.

Huang, C.-T. James. 1982. *Logical Relations in Chinese and the Theory of Grammar*. Cambridge, MA: Massachusetts Institute of Technology PhD dissertation.

Jackendoff, Ray S. 1969. An interpretive theory of negation. *Foundations of Language* 5(2). 218–241.

van Kemenade, Ans. 1997. Negative-initial sentences in Old and Middle English. *Studia Anglica Posnaniensia* 31. 91–104.

Kirby, Simon, Hannah Cornish & Kenny Smith. 2008. Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences* 105(31). 10681–10686. doi:10.1073/pnas.0707835105.

Kirby, Simon, Tom Griffiths & Kenny Smith. 2014. Iterated learning and the evolution of language. *Current Opinion in Neurobiology* 28. 108–114. doi:10.1016/j.conb.2014.07.014.

May, Robert. 1985. *Logical Form: Its Structure and Derivation*. Cambridge, MA: MIT Press.

Neukom-Hermann, Anja. 2016. *Negation, Quantification and Scope. A Corpus Study of English and German All. . . Not Constructions*. Zürich: University of Zurich PhD dissertation.

Öztürk, Balkız. 2005. *Case, Referentiality, and Phrase Structure*. Amsterdam: John Benjamins Publishing Company.

Pesetsky, David. 1989. Language-particular processes and the Earliness Principle. Unpublished manuscript, Massachusetts Institute of Technology.

Politis, Dimitris N., Joseph P. Romano & Michael Wolf. 1999. *Subsampling*. Springer Science & Business Media.

Reali, Florencia & Thomas L. Griffiths. 2009. The evolution of frequency distri-

butions: Relating regularization to inductive biases through iterated learning. *Cognition* 111(3). 317–328. doi:10.1016/j.cognition.2009.02.012.

Seržant, Ilja A. 2019. Weak universal forces: The discriminatory function of case in differential object marking systems. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis & Ilja Seržant (eds.), *Explanation in Typology: Diachronic Sources, Functional Motivations and the Nature of the Evidence*, 149–178. Language Science Press.

Stenson, Nancy. 2019. *Modern Irish: A Comprehensive Grammar*. New York: Routledge.

Szabolcsi, Anna. 1997. Strategies for scope taking. In Anna Szabolcsi (ed.), *Ways of Scope Taking*, 109–154. Dordrecht: Kluwer Academic Publishers.

Thrainsson, Hoskuldur. 2007. *The Syntax of Icelandic*. Cambridge, UK: Cambridge University Press.

Tilleson, Paul. 2019. *On Bipartite Negation*. Minneapolis, Minnesota: University of Minnesota PhD dissertation.

Tottie, Gunnel & Anja Neukom-Hermann. 2010. Quantifier-negation interaction in English: A corpus linguistic study of *all. . . not* constructions. In Laurence R. Horn (ed.), *The Expression of Negation*, 149–185. Berlin: Mouton de Gruyter.

Wälchli, Bernhard. 2007. Advantages and disadvantages of using parallel texts in typological investigations. *Language Typology and Universals* 60(2). 118–134. doi:10.1524/stuf.2007.60.2.118.

Wallage, Phillip. 2012. Negative inversion, negative concord and sentential negation in the history of English. *English Language & Linguistics* 16(1). 3–33. doi:10.1017/s1360674311000268.

Yoshimura, Akiko. 2013. Descriptive/metalinguistic dichotomy?: Toward a new taxonomy of negation. *Journal of Pragmatics* 57. 39–56. doi:10.1016/j.pragma.2013.07.002.

Zeijlstra, Hedde. 2017. Universal quantifier PPIs. *Glossa: a Journal of General Linguistics* 2(1). doi:10.5334/gjgl.220.

Omri Amiraz
Mandel Building, room 441
The Hebrew University of Jerusalem, Mt. Scopus
Jerusalem 9190401, Israel
omri.amiraz@mail.huji.ac.il