A typology of negation in a constraint-based framework of syntax and semantics

Henriëtte de Swart
Department of French/Utrecht Institute of Linguistics
Utrecht University

Proceedings of the HPSG04 Conference

Center for Computational Linguistics, Katholieke Universiteit Leuven

Stefan Müller (Editor)

2004

CSLI Publications

http://csli-publications.stanford.edu/

The full version of this paper will appear under the title 'Marking and interpretation of negation: a bi-directional OT approach' in: Zanuttini, Raffaella, Héctor Campos, Elena Herburger and Paul Portner (eds.). *Negation, Tense and Clausal Architecture: Cross-linguistic Investigations.* Georgetown University Press.

See also: http://www.let.uu.nl/~Henriette.deSwart/personal/pubs.htm.

0. Abstract

Negation and negative indefinites raise problems for the principle of compositionality of meaning, because we find both double and single negation readings in natural languages. De Swart and Sag (2002) solve the compositionality problem in a polyadic quantifier framework. All negative quantifiers are collected into an N-store, and are interpreted by means of iteration (double negation) or resumption (negative concord) upon retrieval. This paper extends the earlier analysis with a typology of negation and negative indefinites using bi-directional optimality theory (OT). The constraints defined are universal, but their ranking varies from one language to the next. In negative concord languages, the functional motivation for the marking of 'negative variables' wins out. Double negation languages value first-order iteration. The bi-directional set-up is essential, for syntactic and semantic variation go hand in hand.[1]

1. Introduction

Languages generally have ways to express negation, i.e. something that corresponds to the first-order logic connective ¬. In English this would be *not*. Many languages also have pronominal expressions negating the existence of individuals having a certain property, i.e. something that corresponds to ¬∃x. In English, this would be *nobody, nothing*. If we assume that knowledge of first-order logic is part of human cognition, we would seem to predict that negation and negative quantifiers behave alike across languages. From empirical research by typologists and theoretical linguists, we know that this is not the case. In particular, differences arise in the way languages express ¬∃x ∃y∃z. The variables *y* and *z* here indicate 'negative variables' in the sense of Corblin and Tovena (2003: 326). They correspond to arguments that must be interpreted within the scope of negation. The simplest way to realize such arguments would be to use (plain) indefinite pronouns. We find this case in Dutch:

(1)	a.	*Niemand* heeft *iets* gezien. Nobody has something seen. 'Nobody saw anything'	[Dutch]

But many languages treat (plain) indefinite pronouns like positive polarity items, and use a special class of negative polarity items within the scope of negation. English is a case at hand. (2a) is not ungrammatical, but it does not express the meaning ¬∃x∃y:

(2)	a.	#*Nobody* saw *something*.	[English]
	b.	Nobody saw *anything*.	

Languages like Romance, Slavic, Greek, etc. use so-called 'n-words', rather than negative polarity items (cf. Haspelmath 1997 for an overview):

(3)	a.	A:Qué viste? A: What did you see?	B: Nada B: Nothing	[Spanish]
	b.	Nessuno mangia. Nobody ate.		[Italian]
	c.	No vino nadie. Not came nobody.	 = Nobody came	[Spanish]
	d.	Nadie maraba a nadie Nobody looked at nobody.	 = Nobody looked at anybody	

[1] Many thanks for helpful comments and feedback from audiences at Utrecht University, Radboud University, the University of Amsterdam, Georgetown University, the University of California at Santa Cruz, Hopkins University, and the University of Leuven. All remaining errors are of course my own. The financial support of the Netherlands Organization for Scientific Research (grant 051-02-070 for the cognition project 'Conflicts in Interpretation') is hereby gratefully acknowledged.

Negative polarity items do not contribute a semantic negation, and require a licensor, whereas n-words can provide a semantic negation in elliptical answers (3a), and in sentences in which no other expression contributes a semantic negation (3b) (cf. Ladusaw 1992, Vallduví 1994). This paper is concerned with n-words, not with negative polarity items.

2. An HPSG analysis of double negation and negative concord.

The main semantic claims made by de Swart and Sag (2002) are that n-words are inherently negative, and that both double negation and negative concord involve polyadic quantification. Double negation involves iteration (function application), and is first-order definable. Negative concord is interpreted in terms of resumption:

- Resumption of a k-ary quantifier (Keenan and Westerståhl 1997).
  $Q'^{A1, A2, \dots Ak}_{E}(R) = Q_{Ek}^{A1xA2x\dots Ak}(R)$.

A sequence of *k* quantifiers Q' binding just one variable each and taking a k-ary relation R as its scope is interpreted as one quantifier Q binding *k* variables predicating over R. E.g. a sequence of quantifiers *No* x, *No* y, *No* z predicating over a three-place relation R(x,y,z) is interpreted as $No_{x,y,z}$ R(x,y,z), claiming that no triple <x,y,z> satisfies the relation R. At the first-order level, the resumptive quantifier is equivalent to $\neg\exists x\exists y\exists z$ R(x,y,z), so we obtain the NC reading, as desired.

The syntax-semantics interface defines how we obtain the DN and NC readings from the syntax. HPSG uses a notion of Cooper storage in which all quantifiers are collected into a store, and interpreted upon retrieval from the store (cf. Manning, Iida and Sag 1999). This mechanism is generally used to account for scope ambiguities, but de Swart and Sag (2002) extend it to account for polyadic quantification. All negative (anti-additive) quantifiers are collected into an N-store. Interpretation upon retrieval from the store is by means of iteration (leading to DN) or by resumption (leading to NC). The formal definition of retrieve is as follows:

- Retrieve: Given a set of generalized quantifiers Σ and a partition of Σ into two sets Σ_1 and Σ_2, where Σ_2 is either empty or else $\Sigma_2 = \{NO_{\sigma 1}^{R1}, \dots, NO_{\sigma 2}^{R2}\}$,
  $Retrieve(\Sigma) =_{def} iteration(\Sigma_1 \cup res(\Sigma_2))$.

So the grammar does not decide between DN and NC. This is what we need for a language like French, in which both readings are available. Consider the ambiguity of the following sentence in the HPSG analysis of de Swart and Sag (2002):

(4) Personne n'aime personne. [French]

(a) Arg-St<[Store $\{NO_{\{x\}}^{\{Person(x)\}}\}$], [Store $\{NO_{\{y\}}^{\{Person(y)\}}\}$]>
Content Quants <$NO_{\{x\}}^{\{Person(x)\}}$, $NO_{\{y\}}^{\{Person(y)\}}$>
Nucleus *Love(x,y)*
Semantic interpretation: NO(HUM, {x|NO(HUM, {y|x loves y})})
In first-order logic: $\neg\exists x\neg\exists x$ Love(x,y) [DN]

(b) Arg-St<[Store $\{NO_{\{x\}}^{\{Person(x)\}}\}$], [Store $\{NO_{\{y\}}^{\{Person(y)\}}\}$]>
Content Quants <$NO_{\{x,y\}}^{\{Person(x), Person(y)\}}$>
Nucleus *Love(x,y)*
Semantic interpretation: $NO_{E2}^{HUMxHUM}$(LOVE)
In first-order logic: $\neg\exists x\exists y$ Love(x,y) [NC]

(4a) and (4b) are identical as far as the argument structure, the storing mechanism, and the interpretation of the *love* relation is concerned. The difference resides in the interpretation of the polyadic quantifier: iteration in (4a), resumption in (4b). The main insights of this analysis are the following. The HPSG grammar assumes no lexical difference between negative quantifiers and n-words, so in the rest of this paper we use the term 'neg expression' to generalize over both. The analysis works for n-words in argument and adjunct position alike (so *nobody, nothing,* as well as *never, nowhere*). Finally, it does not involve empty elements or 'hidden' negations in the syntactic structure. These are major advantages of this proposal.

The OT analysis comes in when we try to relate the HPSG analysis to languages that do not allow double negation and negative concord as freely as French does. In general, the combination of two negative quantifiers in English leads to a double negation reading, and resumption is only marginally available as an interpretive strategy. On the other hand, Spanish, Greek, Polish, and many other languages are typically negative concord languages, which hardly ever realize the iteration version of the polyadic quantifier analysis. The analysis proposed by de Swart and Sag (2002) does not predict cross-linguistic variation where it arises (Spanish vs. English, for example). The aim of this paper is extend the earlier analysis with a bi-directional OT component in order to define a typology of negation.

3. A typology of negation within Optimality Theory

We need to study negation from two perspectives: the generation perspective (how does a speaker express a negative meaning in a particular language?) and the interpretation perspective (how does the hearer interpret a sentence with a sequence of negative expressions in a particular language?). In order to allow for variation in the answers to these questions, we use the framework of Optimality Theory (OT). OT uses universal, but violable constraints, and allows variation in the ranking of the constraints from one language to the next. In an OT syntax, the input is a meaning (a first-order formula), the set of candidates generated by GEN is a set of possible forms, and a ranked set of violable constraints selects the optimal form for the given meaning. In OT semantics, the input is a form (a well-formed sentence), the set of candidates is a set of possible meanings (first-order formulae), and a ranked set of violable constraints selects the optimal interpretation for the given form. Bi-directional OT looks at balanced ('harmonic') pairs of form and meaning.

The starting point of the analysis is the observation that negative sentences are formally and interpretationally marked with respect to affirmative sentences. This means that we expect to see the negative meaning reflected in the syntax, and the negative syntax to be reflected in the meaning. The constraint FaithNeg (Faith negation) accounts for this intuition:

♦ **FaithNeg**
Reflect the non-affirmative nature of the input in the output.

FaithNeg is a faithfulness constraint, i.e. a constraint that aims at a faithful reflection of input features in the output. Since negation is marked in all languages, we take FaithNeg to be universally ranked at the top. In OT, faithfulness constraints are balanced by markedness constraints, which are output oriented. The markedness constraint that plays a role in negative statements is *Neg:

♦ ***Neg**
Avoid negation in the output

*Neg is obviously in conflict with FaithNeg. Such conflicting constraints are characteristic of OT style analyses. FaithNeg and *Neg play a role in OT syntax as well as in OT semantics. In addition, we need two maximizing constraints, one aimed at the syntax (MaxNeg), the other one aimed at the semantics (IntNeg):

♦ MaxNeg
Mark 'negative variables' (i.e. the arguments that are interpreted within the scope of negation)

♦ IntNeg
Force Iteration (i.e. every neg expression in the form contributes a semantic negation at the first-order level in the output)

The functional motivation for the marking of negative variables (Haspelmath 1997, Corblin and Tovena 2003) explains why the use of n-words is widespread among natural languages. However, the use of n-words is not universal: languages like Dutch, English, Turkish, etc. do not use n-words. This suggests that MaxNeg is not a hard constraint, and its position in the constraint ranking is not the same for every language. We can account for the difference between languages with and without n-words by changing the position of MaxNeg relative to *Neg. If *Neg is ranked higher than MaxNeg, the optimal way to express the meaning $\neg\exists x_1\exists x_2\ldots\exists x_n$ is by means of indefinite pronouns. If MaxNeg is ranked higher than *Neg, n-words are used to express indefinites under negation. The following OT syntactic tableaux reflect this for the binding of two variables:

Tableau 1 (generation of indefinite, for Dutch, Turkish, etc.)

Meaning	Form	FaithNeg	*Neg	MaxNeg
$\neg\exists x_1\exists x_2$	Indef+indef	*		
☞	neg+indef		*	*
	neg + neg		**	

Tableau 2: (generation of n-word for Greek, Romance, Slavic, etc.)

Meaning	Form	FaithNeg	MaxNeg	*Neg
$\neg\exists x_1\exists x_2$	indef+indef	*		
	neg+indef		*	*
☞	neg + neg			**

The top ranking of FaithNeg makes it impossible to express indefinites under negation by indefinites exclusively (in the absence of a marker of sentential negation). In tableaux 1 and 2, the candidates that we need to compare are those that mark negation somehow in the output. This invariably leads to a violation of *Neg. Two neg expressions are 'worse' than one, so the combination of two neg expressions incurs two violations of *Neg.

As far as generation is concerned, we conclude that languages that allow indefinites under negation (Dutch, Turkish, etc.), and languages that use n-words (Romance, Slavic, Greek, etc.) differ in their ranking of the two constraints MaxNeg and *Neg. This approach immedediately raises the question of the interpretation of the expressions involved. In isolation, we cannot determine whether a particular expression is a negative quantifier or an n-word, because they both contribute the meaning $\neg\exists$ (cf. 3a, b). Following de Swart and Sag (2002), I assume that this question is decided in the grammar, not in the lexicon. The use of neg expressions in a generative OT system means that we run into the recoverability problem: from the expressions generated, we can derive multiple interpretations, not only the intended one. Recoverability is assured by the way the generation of negative sentences hangs together with their interpretation. So we need an OT semantic component.

In the interpretive system, FaithNeg outranks all the other constraints as usual. MaxNeg is a purely syntactic constraint that does not play a role in interpretation. So the constraints that need to be ordered are *Neg and IntNeg. If *Neg is ranked higher than IntNeg in the OT semantics, a sequence of multiple Neg expressions leads to a single negation meaning by resumption. If IntNeg is ranked higher than *Neg, a series of Neg expressions is interpreted as multiple negation by forcing iteration. The following tableaux illustrate the two possible rankings and their optimal output:

Tableau 3: double negation (interpretation of Dutch, English, etc.)

Form	Meaning	FaithNeg	IntNeg	*Neg
neg + neg	$\exists x_1 \exists x_2$	*	**	
	$\neg\exists x_1 \exists x_2$		*	*
☞	$\neg\exists x_1 \neg\exists x_2$			**

Tableau 4: negative concord (interpretation of Romance, Slavic, Greek, etc.)

Form	Meaning	FaithNeg	*Neg	IntNeg
neg + neg	$\exists x_1 \exists x_2$	*		
☞	$\neg\exists x_1 \exists x_2$		*	*
	$\neg\exists x_1 \neg\exists x_2$		**	

The top ranking of FaithNeg implies that we cannot interpret a statement involving two neg expressions without a reflection of the non-affirmative meaning. As a result, the relevant candidates we compare have at least one negation in the output, and always incur a violation of *Neg. The combination of two neg expressions leads to a double negation reading in languages like Dutch and English, for the constraint IntNeg is ranked higher than *Neg in tableau 3. Because *Neg outranks IntNeg in tableau 4, single negation readings win over double negation readings in NC languages such as Spanish, Italian, Greek, Polish, etc.

Collapsing the generation and interpretation perspective, we derive the following two rankings for negative concord and double negation languages:

Bidirectional grammar

- Negative concord languages: FaithNeg >> MaxNeg >> *Neg >> IntNeg
- Double negation languages: FaithNeg >> IntNeg >> *Neg >> MaxNeg

In the full paper, I argue that only rankings where MaxNeg and IntNeg are distributed on either side of *Neg reflect viable options for a linguistic system that balances generation and interpretation of negative statements. In sum:

- **Negative Concord**: if you mark 'negative variables' (MaxNeg >> *Neg in syntax), then make sure you do not force Iteration (*Neg >> IntNeg in semantics).
- **Double Negation**: if you force Iteration, (IntNeg >> *Neg in semantics), then make sure you do not mark 'negative variables' (*Neg >> MaxNeg in syntax).

4. Concluding remarks.

A bi-directional version of Optimality Theory offers new perspectives on the range of variation we find in natural language for the expression and interpretation of negation. Patterns that are frequently found in natural language, but do not display absolute tendencies can be fruitfully described in a framework that formulates universal constraints, but allows these constraints to be violable. Bi-directionality is a central feature of our analysis, because it relates the semantic compositionality problems raised by negative concord to the functional tendencies to formally mark the scope and focus of negation, in accordance with the view on compositionality advanced by Blutner, Hendriks and de Hoop (2003).

Many further questions arise in the domain of negative concord languages. As we know from French, double negation readings do arise in negative concord languages, and this requires the possibility of overlap between interpretive constraints. Furthermore, NC languages vary in their interaction between n-words and the marker of sentential negation. Slavic languages, Greek, Afrikaans, etc. always require the presence of a marker of sentential negation in negative sentences. Languages such as Spanish, Italian, Portuguese display an asymmetry between pre-verbal and postverbal n-words. The different subclasses can be accounted for with the help of two extra constraints (NegFirst and MaxSN), which are discussed in the full paper. These constraints only play a role in the OT syntax, they do not affect the interpretive system. This paper thus supports the conclusions from de Swart and Sag (2002), who argue that the grammar is responsible for the differences between negative concord and double negation languages by means of the interpretation mechanisms for polyadic quantification. The position and distribution of the marker of sentential negation in negative concord is relevant for syntax, but does not affect the semantics.

Selected References

Blutner, R. (2000). Some aspects of optimality in natural language interpretation, *Journal of Semantics* 17: 189-216.

Blutner, R., P. Hendriks and H. de Hoop (2003). A new hypothesis on compositionality, *Proceedings of ICCS 2003*, Sidney.

Boersma P. (1998). *Functional phonology*, PhD diss. University of Amsterdam.

Corblin, F. and L. Tovena (2003). L'expression de la négation dans les languages romanes, in: D. Godard (ed). *Les langues romanes: problème de la phrase simple*, Paris: CNRS Editions, 279-342.

Corblin, F., V. Déprez, H. de Swart and L. Tovena (2004). Negative concord, in: F. Corblin and H. de Swart (eds.). *Handbook of French Semantics*, Stanford: CSLI Publications.

Haspelmath, M. (1997). *Indefinite pronouns*, Oxford: Clarendon Press.

Herburger, E. (2001). The negative concord puzzle revisited, *Natural Language Semantics* 9, 289-333.

Horn, L. (1989). *A natural history of negation*, Chicago: University of Chicago Press.

Jespersen, O. (1917). *Negation in English and other languages*, Copenhagen: A.F. Høst. Reprinted in: *Selected writings of Otto Jespersen* (1962), London: George Allen and Unwin, 3-151.

Keenan, E. and D. Westerståhl (1997). Generalized quantifiers in linguistics and logic, in: J.van Benthem and A. ter Meulen (eds.). *Handbook of Logic & Language*, Amsterdam: Elzevier, 837-893.

Ladusaw, W. (1992). Expressing negation, *Proceedings of SALT* 2, Columbus: Ohio State University, 237-259.

Manning, C., I. Sag and M. Iida (1999). The lexical integrity of Japanese causatives, *Nordic Journal of Linguistics* 21, 107-144.

Swart, H. de and I. Sag (2002). Negation and negative concord in Romance, *Linguistics and Philosophy* 25: 373-417.

Vallduví, E. (1994). Polarity items, n-words and minimizers in Catalan and Spanish, *Probus* 6: 263-294.

Zanuttini, R. (1991). *Syntactic properties of sentential negation*, PhD. Diss. University of Pennsylvania.

Zeevat, H. (2000). The asymmetry of optimality theoretic syntax and semantics, *Journal of Semantics* 17: 243-262.