Published in final edited form as:

Linguist Inq. 2010 November 1; 41(4): 681–689. doi:10.1162/LING_a_00017.

When are Downward Entailing contexts identified? The case of the domain-widener *ever*

Charles Clifton Jr. and Lyn Frazier

University of Massachusetts Amherst

1 When is the monotonicity of a context identified?

It has been known for some time that Negative Polarity Items (NPIs) are often licensed in Downward Entailing (DE) contexts – contexts that support inferences from sets to subsets (Ladusaw, 1979). For example, the NPI *ever* is licensed in the negative (DE) context (1a) but not in the nonDE context (1b).

(1)

- a. Josh didn't ever play chess.
- b. *Josh ever played chess.

We examine sentences in which the NPI is eventually seen to be licensed, but which contain a tempting initial misanalysis under which the NPI does not appear to be licensed. We ask two questions: (a) are computations about the licensing of an NPI done locally and on-line, or are they only done in the global context of a full sentence, and (b) if such local computations are made, does a temporary apparent failure of licensing result in lowered acceptability for a globally-acceptable sentence?

There is evidence that the DE contexts that support an NPI are used word by word as a sentence is read to guide some aspects of processing. Scalar implicatures (e.g., *some* implicates *some but not all*) are more likely to be reported if the scalar term occurs in a nonDE context than in a DE context (Chierchia, 2004; Chierchia, Frazier, & Clifton, 2009; Schwarz, Clifton & Frazier, to appear). Panizza, Chierchia & Clifton (2009) provide eyetracking evidence that computation of scalar implicatures (for numerals) is done incrementally, immediately when the scalar term is read (which they take to support claims that implicatures are computed locally rather than globally as in Gricean approaches to implicatures; Grice, 1989, Geurts & Poussolous 2009, Russell, 2006).

ERPs have also been used to study the processing of NPIs. Absence of a licenser such as a negative morpheme has been reported to result in the appearance of an N400 (a marker associated with semantic predictability/anomaly) as well as a late positivity (perhaps a P600, often taken as a marker of syntactic processing difficulty) to a NPI, indicating that the licensing is done on-line. Interestingly, the presence of a licenser in a structurally inappropriate position also reduces the size of the N400, though not as much as a licenser in a structurally appropriate position (Saddy et al., 2004; Drenhaus et al., 2004, 2005, 2006, Vasisht et al., 2008, a.o.). Self-paced reading data do not show an immediate effect of spurious licensors, however, suggesting that the ERP effects might be due to semantic associations between words – a factor the N400 is known to be sensitive to (Xiang, Dillon & Phillips, 2009, though further assumptions are needed to address the effect of a spurious licenser on the P600).

2 Incremental vs delayed computations: the case of the domain widening NPI *ever*

To investigate when the monotonicity (downward entailingness or upward entailingness) of a context is identified, we examined the NPI *ever*, which has been analyzed as a domain-widener (Chierchia, 2006), as illustrated in (2).

(2)

- a. Ella didn't go to Paris.
- b. Ella didn't ever go to Paris.

Sentences without *ever* (2a) may be used to make a claim about a particular time period under discussion (e.g., last week) and thus (2a) might be true even if uttered in a situation in which Ella went to Paris at some time in the past. By contrast *ever* in (2b) serves to widen the domain, and thus (2b) could not be used to make a true assertion if Ella went to Paris at some point in the past. Because *ever* widens the domain, a sentence containing *ever* does not readily tolerate exceptions.

If NPIs such as *ever* are licensed semantically by DE contexts and monotonicity computations come into play only in terms of the global context of the entire utterance, then the nature of the context, DE or nonDE, should be computed with respect to the final syntactic analysis of the sentence. A temporary misanalysis, as occurs in a “garden-path” sentence (Bever, 1970; Frazier, 1987), shouldn't influence the processing of the NPI. In contrast, on a local approach to implicature computation and context determination, one might expect difficulties if an NPI (a domain widener like *ever*) appeared to be unlicensed on the first, erroneous, syntactic analysis of the sentence.

Consider the classic garden path sentence in (3) (Bever, 1970)

(3) The horse raced past the barn fell.

The sentence is initially parsed as a main clause up to the verb *fell*. At *fell*, the initial analysis becomes untenable because *fell* requires a subject and the preceding string has been analyzed as a clause, not a DP. Consequently, on the final correct analysis, the initial sequence *The horse raced past the barn* must be analyzed as a reduced relative clause.

Strings like *A teacher caught...* (4) are known to be parsed initially as main clause structures on a par with (3) above (Frazier, 1987). Affirmative main clauses are nonDE contexts (unless they contain a DE operator). Thus the local approach would lead us to expect (4b) to be difficult or odd compared to (4a), because the NPI *ever* will at first appear not to be licensed, i.e., before the syntactic garden-path is reanalyzed.

(4)

- a. A teacher caught with communist literature won't be allowed to teach Social Studies.
- b. A teacher ever caught with communist literature won't be allowed to teach Social Studies.
- c. A teacher who was ever caught with communist literature won't be allowed to teach Social Studies.

If this initial oddity persists in ratings of the final sentence, then reduced relatives containing NPIs will be rated worse than unambiguous relatives like (4c), where it is clear from the outset that the NPI is licensed. An NPI is licensed in a relative clause with a universal head

or with a generic interpretation because the structures are DE contexts, as shown by the fact that the truth of (5a) implies the truth of (5b).

(5)

- a. Every man/A man who eats vegetables will be healthy.
- b. Every man/A man who eats carrots will be healthy.

Apart from the question of incremental computation of the monotonicity of contexts, it is interesting to know under what circumstances acceptability judgments of sentences presented without time pressure reflect temporary difficulties in processing sentences. Previous findings suggest that ungrammatical sentences with temporary grammatical analyses receive higher acceptability ratings than counterparts without the temporary misanalysis (Fanselow & Frisch, 2006), and they show that sentence processing difficulties can be reflected in the final acceptability ratings of sentences (Sprouse, 2008). How temporary syntactic analyses interact with logical properties of context may also be reflected in acceptability judgments.

3 Experimental findings

Below we report two written acceptability judgment studies to test whether the initial incorrect (nonDE context) analysis of reduced relatives results in decreased acceptability when they contain an NPI. Experiment 1 tested sentences with reduced relative clauses (RRC) like those in (6). Sentences with full relative clause (FRC) structures served as controls.

(6)

- a. [RRC –ever] A man arrested in this country won't break the law here again.
- b. [RRC +ever] A man ever arrested in this country...
- c. [FRC, –ever] A man who was arrested in this country...
- d. [FRC, +ever] A man who was ever arrested in this country...

The prediction that (6b) should be rated as slightly odd due to the temporary problem of *ever* occurring in a nonDE context was tested using 24 sentences like those in (6), plus their counterparts with a plural subject (*Men ever arrested....*). There were eight versions of each sentence, generated by the factorial combination of +/-ever, full vs. reduced relative clause, and singular vs. plural subject: Half the versions contained *ever* (indicated as +ever); half did not (–ever). Half had a head noun followed by a reduced relative (RRC) (6a,b) and half had a head noun followed by a full relative clause (FRC) (6c,d). Half had a singular head noun, as illustrated; the other half had a plural head noun. The number of the head noun was manipulated to make sure that the oddness of (6b), if confirmed, could not be due to an independent bias against having an indefinite singular in a non-episodic context. All experimental sentences are available from the authors (cec@psych.umass.edu, lyn@linguist.umass.edu).

The resulting sentences were arbitrarily divided into two subsets of 12 sentences, each of which was divided into eight counterbalanced lists. Four of the lists contained just singular head nouns, and the other four, plural head nouns. Each list contained three sentences in each of the four versions illustrated in (6), and each sentence was tested in each of these versions in one singular and one plural list. Each sentence was followed by a 5-point rating scale, labeled “How acceptable was that sentence,” with “1” being “unacceptable” and “5” being “fully acceptable.” The sentences were combined with a total of 84 items testing

unrelated questions. Each of these sentences was followed by one of a variety of types of questions, including two-choice questions and plausibility rating scales.

Ninety-six University of Massachusetts undergraduates were tested in individual half-hour sessions. They were given instructions, including the instruction that they were to rate some sentences for acceptability, which was explained to them as how much “the sentence sounds like a sentence that a native speaker of English could say or understand without noticing anything peculiar or ungrammatical or odd or confusing about it.” Six subjects were assigned to each of the 16 (2 sets of 8 counterbalanced lists) lists described above. A computer presented six practice items, then the 12 experimental items plus the 84 other items, in individually randomized order. A subject saw a sentence on the computer screen, and pressed a key on the computer keyboard after reading it. The question and rating scale then appeared, and the subject pressed a number key, corresponding to the chosen acceptability rating.

The mean acceptability ratings appear in Table 1 (5 = fully acceptable). RRC was rated worse than FRC, +ever examples were rated worse than –ever examples. As predicted there was an interaction of ±ever and full/reduced RC, with the lowest ratings being the reduced relative +ever forms.

The significance of the effects of type of relative clause, presence vs. absence of *ever*, subject number, and their interactions, was tested using a mixed model analysis with subjects and items as partially crossed random factors, using the default contrast coding of the factors. We chose this analysis over the more familiar ANOVA framework, separately testing for generalization to subjects and items, for a variety of reasons detailed by Baayen, Davidson, and Bates, 2008. These reasons include the fact that the linear mixed models analysis permits simultaneous generalization to subjects and items and encourages comparisons of various models to identify the most adequate one. The R statistical package (<http://www.R-project.org>) was used for all analyses. Because there is no principled way of calculating the df to be used in significance tests of the coefficients of the linear mixed model, precluding the use of standard t-tests, the effects were tested using Markov Chain Monte Carlo sampling (again, see Baayen et al. for justification). The interaction of the presence of *ever* and the type of relative clause was significant ($t = 2.41$, $pMCMC < .02$). The penalty for having *ever* was 0.68 for reduced relative clauses, and a significantly smaller (and nonsignificant; $t < 1.0$) 0.25 for sentences with full relative clauses. Sentences with a reduced relative clause were rated as less acceptable than ones with a full relative clause ($t = 5.17$, $pMCMC < .001$, tested for sentences with *ever*, the intercept in the linear mixed model, but also approaching significance for sentences without *ever*, $t = 1.91$, $pMCMC < .06$). No effect involving singular vs. plural subjects approached significance.

The results of Experiment 1 confirmed the prediction that monotonicity is computed locally, not just with respect to the final correct analysis of a sentence. Even though *ever* was globally licensed by the final analysis of the sentence, the fact that it was apparently unlicensed in the initially-favored analysis of the reduced relative clause sentences resulted in its presence lowering acceptability ratings in those sentences.

To further test the hypothesis that the difficulty with an NPI in a reduced relative is due to the NPI appearing to be unlicensed, Experiment 2 tested the sentences from Experiment 1 with negation in the head of the relative, as illustrated in (7).

(7)

- a. [RRC –ever] No man arrested in this country will break the law here again.

- b. [RRC +ever] No man ever arrested in this country will break the law here again.
- c. [FRC, -ever] No man who was arrested in this country will break the law here again.
- d. [FRC, +ever] No man who was ever arrested in this country will break the law here again.

If the difficulty with NPIs in reduced relatives was indeed due to the NPI appearing not to be in a DE context, then the effect should go away in Experiment 2 since the negation will create a DE context. Consequently, there should be no difficulty with an example like (7a) relative to its (7b) counterpart without *ever*.

The singular sentences from Experiment 1 were altered by substituting *no* for *a* and changing the polarity of the predicate, as appropriate. The plural sentences used in Experiment 1 were not tested in Experiment 2. The four versions of each sentence defined by the factorial combination of +ever vs -ever and RRC vs FRC were divided into four counterbalanced lists of 24 sentences and combined with 80 sentences from unrelated experiments.

Forty-eight University of Massachusetts undergraduates were tested using a procedure similar to that for Experiment 1, except that 12 subjects were assigned to each of the counterbalanced lists, each subject saw all 24 experimental sentences, and the acceptability rating scale ranged from “1” = “Terrible” to “5” = “Perfect.”

The mean ratings appear in Table 2. A multi-level analysis with subjects and items as crossed random terms, and type of relative clause and presence of *ever*, similar to the analysis of Experiment 1, was conducted. No effect approached significance ($t < 1.0$ except for the apparent cost of having *ever* present, where $t = 1.25$, ns.). The pattern of results is clearly different from that observed in Experiment 1: the interaction observed in that experiment disappeared in Experiment 2. Apparently, the presence of the initial *No* immediately licensed the NPI *ever* in Experiment 2, while *ever* initially appeared to be unlicensed in the reduced relative clause versions of Experiment 1.

4 Conclusions

It has been claimed that humans come biologically equipped to acquire and use human languages and, further, that this remarkable ability includes a natural language deductive system (Chierchia, 2006). From this perspective it is not surprising that on an essentially word by word basis, the language processor may determine the monotonicity of the current sentence context. Consequently even temporary syntactic analyses which place an NPI in a nonDE context may give rise to an oddity which persists in final judgments of the acceptability of a sentence. (For other examples of acceptability judgments being influenced by temporary analyses, see Fanselow & Frisch, 2006, Sprouse, 2008).

The idea that the processor keeps track of the monotonicity of its current context fits well with the idea that implicature computation may be local, and not just determined by global properties of the utterance. Of course local monotonicity determination is only a pre-condition for local computation of implicatures, it does not entail local computation of implicatures.

On the other hand, it is certainly easy to imagine that the world might have been otherwise: logical properties of contexts, whether a context would support an inference from sets to subsets, vice versa, or neither, might have been properties that came into play only when

language was used in a logical mode for verification of inferences or the like. We thus take it to be an interesting property of natural language processing that close tabs are kept on the monotonicity of sentence contexts as the analysis of the sentence proceeds.

Acknowledgments

This research was supported by HD-18708 from the National Institute for Child Health and Human Development. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Science Foundation, the National Institute for Child Health and Human Development, or the National Institutes of Health. Order of authors is alphabetical. We are grateful to Gillian Ramchand, Curt Rice, Peter Svenonius and two anonymous reviewers for comments on an earlier draft of this paper.

References

- Baayen, Harold; Davidson, DJ.; Bates, Douglas. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*. 2008
- Bever, Thomas. The cognitive basis for linguistic structures. In: Hayes, JR., editor. *Cognition and the development of language*. Wiley; New York: 1970. p. 279-352.
- Chierchia, Gennaro. Scalar implicatures, polarity phenomena and the syntax/pragmatics interface. In: Belletti, A., editor. *Structure and Beyond*. Oxford University Press; Oxford: 2004.
- Chierchia, Gennaro. Broaden your views: Implicatures of domain widening and the “logicality” of language. *Linguistic Inquiry*. 2006; 37:535–590.
- Chierchia, Gennaro; Frazier, Lyn; Clifton, Charles, Jr.. When basic meanings are (not) enough: Processing scalar implicatures in adult comprehension. In: Biezma, Maria; Harris, Jesse A., editors. *UMOP 39: Papers in Pragmatics*. GLSA Publications; Amherst, MA: 2009. p. 21-39. Amherst
- Drenhaus, Heiner; beim Graben, P.; Frisch, Stefan; Saddy, Douglas. Diagnosis and repair of negative polarity constructions in the light of symbolic resonance analysis. *Brain and Language*. 2006; 96:255–268. [PubMed: 15975647]
- Drenhaus, Heiner; Saddy, Douglas; Frisch, Stefan. Processing negative polarity items: When negation comes through the backdoor. In: Kepser, S.; Reis, M., editors. *Linguistic Evidence -- Empirical, Theoretical, and Computational Perspectives*. Mouton de Gruyter; 2005. p. 145-165. *Studies in Generative Grammar* 85
- Drenhaus, Heiner; Saddy, Douglas; Frisch, Stefan. Intrusion effects in the processing of negative polarity items. In: Kepser, S.; Reis, M., editors. *Pre-Proceedings of the International Conference on Linguistic Evidence*. Tübingen: 2004. p. 41-46.
- Fanselow, Gisbert; Frisch, Stefan. Effects of processing difficulty on judgments of acceptability. In: Fanselow, Gisbert; Fery, Caroline; Vogel, Ralf; Schlesewsky, Matthias, editors. *Gradience in Grammar*. Oxford University Press; Oxford: 2006. p. 291-316.
- Frazier, Lyn. Sentence processing: A tutorial review. In: Coltheart, M., editor. *Attention and performance*. Lawrence Erlbaum Associates; Hillsdale, NJ: 1987. p. 559-586.
- Grice, Paul. *Studies in the Ways of Words*. Harvard University Press; Harvard, MA: 1989.
- Geurts, Bart; Poussoulous, Nausicaa. Embedded implicatures. *Journal of Semantics and Pragmatics*. 2009; 2:1–34.
- Ladusaw, William. *Polarity Sensitivity as Inherent Scope Relation*. University of Texas; Austin: 1979. Ph. D. Dissertation Distributed by the Indiana University Linguistic Club
- Panizza, Danielle; Chierchia, Gennaro; Clifton, Charles, Jr.. On the role of entailment and scalar implicature in the processing of numbers. *Journal of Memory and Language*. 2009; 61:503–519. [PubMed: 20161494]
- Russell, Benjamin. Against grammatical computation of scalar implicatures. *Journal of Semantics*. 2006; 23:361–382.
- Saddy, Douglas; Drenhaus, Heiner; Frisch, Stefan. Processing Polarity Items: Contrastive Licensing Costs. *Brain and Language*. 2004; 90:495–502. [PubMed: 15172566]

- Schwarz, Florian; Clifton, Charles, Jr.; Frazier, Lyn. Focus, downward entailing contexts, and scalar implicatures. In: Anderssen, J.; Moulton, K.; Schwarz, F.; Ussery, C., editors. *UMOP 37: Semantics and Processing*. GLSA Publications; Amherst, MA: To appear
- Sprouse, Jon. The differential sensitivity of acceptability judgments to processing effects. *Linguistic Inquiry*. 2008; 39:686–693.
- Vasishth, Shravan; Brüssow, Sven; Lewis, Richard L.; Drenhaus, Heiner. Processing polarity: how the ungrammatical intrudes on the grammatical. *Cognitive Science*. 2008; 32:685–712.
- Xiang, Ming; Dillon, Brian; Phillips, Colin. Illusory licensing effects across dependency types: ERP evidence. *Brain and Language*. 2009; 108:40–55. [PubMed: 19007980]

Table 1

Mean (and SE) acceptability ratings, Expt 1 (1 = 'unacceptable' 5 = 'fully acceptable')

Condition	Singular Head Noun	Plural Head Noun
RRC, -ever	3.87 (0.11)	3.69 (0.11)
RRC, +ever	3.17 (0.12)	3.06 (0.12)
FRC, -ever	4.01 (0.11)	3.97 (0.11)
FRC, +ever	3.61 (0.11)	3.85 (0.10)

Table 2

Mean (and SE) acceptability ratings, Experiment 2 (1 = 'terrible' 5 = 'perfect')

Condition	
RRC, -ever	3.59 (0.07)
RRC, +ever	3.43 (0.07)
FRC, -ever	3.58 (0.06)
FRC, +ever	3.49 (0.07)