The Influence of Prosody and Ambiguity on English Relativization Strategies

Ted Briscoe and Paula Buttery
Computer Laboratory and Research Centre for English and Applied Linguistics
University of Cambridge

Interdisciplinary Approaches to Relative Clauses Conf. RCEAL,
Cambridge, Sept07

1 Introduction

We present evidence that, for English, ambiguity is an active factor in the choice of relativization strategy and that, in speech, prosody plays a role in resolution of ambiguity over the internal role of the relativized constituent. The evidence is based on (semi-)automatic analysis and comparison of automatically-parsed written and spoken portions of the British National Corpus (BNC, Leech, 1992) and of the prosodically-transcribed Spoken English Corpus (SEC, Taylor and Knowles, 1988). The results are evaluated with respect to a model of parsing complexity and syntactic disambiguation (Briscoe 1987, 2000a) building on Combinatory Categorial Grammar (Steedman, 2000) and this model is in turn motivated by an evolutionary account of linguistic coevolutionary adaptation of the syntactic and phonological prosodic systems to a solution which minimizes processing cost. To our knowledge this is the first work which investigates linguistic adaptations aimed at reducing ambiguity while making testable predictions about linguistic organization.

We work within the framework of evolutionary linguistics where we expect, other things being equal, that language traits and linguistic usage will be selected via successive generations of language acquisition for learnability, expressiveness, and/or interpretability (see Briscoe, 2000a,b)

2 Psycholinguistic Data

It is well known that subject relative clauses (SRCs), where the relativized constituent is internally subject (see (1a)) are less complex than non-subject ones (NSRCs), such as (1b)).

(1) a The guy who/that likes me just smiled

b The guy who/that/0 I like *e* just smiled

This is explained by sentence complexity metrics which incorporate some notion of locality between 'filler' and 'gap' (Gibson, 1998; Hawkins, 1994; 2004). We use filler to refer to the interrogative pronoun, if present, or the nominal head modified by the NSRC, and gap to refer to the canonical position of the filler in the NSRC – e.g. *who/guy* and *e* respectively in (2b)). However, NSRCs exhibit unbounded dependencies, which are also known to be both potentially highly ambiguous (Church, 1980) and psycholinguistically complex (Gibson, 1998). (2a) and (2b) illustrate that NSRCs can contain multiple ambiguous gaps (*e?*) with unbounded material between filler and gap and between ambiguous gaps.

(2) a The guy who I think you want *e?* to succeed *e?* just smiled

b The guy who I want *e?* to think that the boss will succeed *e?* just smiled

The psycholinguistic consensus is that there is a parsing preference for early potential gaps because reading times after potential gap positions are slowed if the gap is filled locally or if the filler is semantically implausible (Stowe, 1986). Gibson (1998) argues that a locality-based complexity metric predicts this result if the human parser chooses the least complex analysis when lexical frequency or semantic plausibility considerations do not dictate otherwise. (3a) is a mild garden path, probably because *want* occurs five times more often with VPinf than NP+VPinf complementation.[1] Certainly, if we substitute *ask*, as in (3b), which exhibits a far stronger preference for NP+VPinf complementation, then the effect disappears.

(3) a The guy who I wanted to give the present to Sue refused

b The guy who I asked to give the present to Sue refused

In (4a) and (4b), there are clear garden path effects for most readers when the actual gap at the end of the RCs is incorrectly filled by *three books*.

(4) a I gave the guy who I wanted to give the books to three books

b I wouldn't give the guy who was reading three books

Once again, the frequency-based lexical preference for no direct object with *want*, and the fact that *read* is used transitively almost twice as often as in-

[1] This and the following estimates of the relative frequency of subcategorization frames are based on the VALEX lexicon (Korhonen *et al.*, 2006).

transitively might explain these preferences, overriding any (default) structural preference for the first possible gap. However, as *succeed* occurs about 4.5 times more often intransitively than transitively, frequency effects in (2a) between *succeed* and *want* are in conflict. Early resolution of the ambiguity at the point of the first potential gap and before the second verb has been processed therefore predicts at least an initial preference for the late gap attachment, but the preferred interpretation is for the early gap with *succeed* interpreted intransitively as 'win'. The lack of an apparent garden path effect here is unexplained under the Gibson/Stowell account.

3 Typological Data

Moving from psycholinguistic preferences of on-line interpretation to typology, Hawkins (1994:323f) explains the non-occurrence of initial subordinators in prenominal relatives crosslinguistically by arguing that the advantage of marking the onset of the embedded clause is offset by the remaining ambiguity over whether the embedded clause is a sentential complement or RC. Kuno (1974) considers the unattested strategy of marking both boundaries of RCs with subordinators and suggests this is dispreferred because it leads to patterns of unbounded nested dependencies similar to those in centre-embedded constructions. In the CCG model, placement of a single subordinator at the opposite end of the RC to the modified head creates equivalent complexity via creation of an additional unbounded dependency, if the subordinator must be syntactically linked to the head (i.e. has a CCG category like (N/N)/(S|XP)). Thus under our account of complexity (or that of Gibson or Hawkins), this is a non-optimal strategy for resolving such potential ambiguity. In English, this strategy applied to (4b) might look like (5a) where an additional subordinator *tath* occurs at the right boundary of the RC.

(5) a I wouldn't give the guy who was reading tath three books

b I wouldn't give the guy who was reading three books tath another one

If *tath* is the mirror image of *that* and has CCG category (S|XP)\(N/N) then this blocks any local ambiguity concerning the correct role of *three books* as illustrated in (5b), but it also increases the syntactic complexity of RCs potentially unboundedly by introducing an additional syntactic dependency between it and the head of the relative, *guy* here. Thus, there is a trade-off between resolving ambiguities syntactically and the overall syntactic complexity of RC constructions.

4 The Role of Prosody

In both Japanese prenominal RCs and English postnominal RCs there is evidence that in speech the RC boundary at the opposite end to the head is often marked by a prosodic boundary (PB, often a major tone group / intonational phrase boundary, but possibly a minor/intermediate one; Venditti, Jun and Beckman, 1996). Assuming the human speech processor generates a metrical analysis of the input independently of the parser, but the latter can take account of extrasyntactic information, the alignment of PBs with syntactically unmarked RC boundaries provides an efficient means for languages to mark the other RC boundary. Warren (1999) reviews psycholinguistic evidence that PBs are exploited by the human parser to resolve syntactic indeterminacies, and Nagel *et al.* (1994) argue that *actual* gaps are always marked by PBs. Thus, (5a) and (5b) would both be resolved in speech by the occurrence of a PB as indicated by (||) in (6a) and (6b).

(6) a I gave the guy who I wanted to give the books to
|| three books

b I wouldn't give the guy who was reading || the book

However, Straub *et al.* (2001) show that intonational/major PBs occur at the end of NSRCs and not medially, as would be required in one interpretation of (3) and in (4). On the other hand, Cooper and Paccia-Cooper (1980) and Warren (1985) provide some evidence from sentence production experiments that minor/intermediate boundaries, marked principally by syllable-lengthening, occur on the predicate preceding medial gaps in NSRCs as in (7a) versus (7b).

(7) a The guy who I want | to succeed || just smiled

b The guy who I want to succeed || just smiled

c The guy who I wanna succeed || just smiled

The lack of the medial PB when the actual gap is later licenses optional cliticization of *to* or reduction to *wanna* as in (7c) but blocks it in (7a) in the metrical framework assumed here, subsuming this well-known phenomenon into a more general account of ambiguity resolution.

5 The Model

We can account for the data discussed above in a model which integrates CCG with a (1,1)bounded-context parser which embodies default structural preferences for late closure and late gaps via a preference for shift over reduce whenever both parsing actions are possible in the current context, but which uses lexical frequency, semantic plausibility or prosodic information to override this preference *at the point when the parsing indeterminacy arises*

(see Briscoe 1987, 2000a) for further details).

A specific CCG (see e.g. Steedman, 2000 for further details) consists of a lexicon pairing CCG lexical categories with lexemes, a fragment of which for English is given below:

who(m)	(N\N)/(S/NP)		
I	S/(S\NP)		
want	((S\NP)/NP)/VP	(S\NP)/VP	
succeed	(S\NP)/NP	S\NP	
...			

More than one category can be associated with a given lexeme if there is lexical ambiguity. For instance, the two categories above for *succeed* indicate that it can function as a transitive and intransitive verb. A specific CCG may also include certain restrictions on the application of rules of application and composition such as those shown below:

Forward Application (FA):

X/Y Y ⇒ X λ y [X(y)] (y) ⇒ X(y)

Backward Application (BA):

Y X\Y ⇒ X λ y [X(y)] (y) ⇒ X(y)

Forward Composition (FC):

X/Y Y/Z ⇒ X/Z λ y [X(y)] λ z [Y(z)] ⇒ λ z [X(Y(z))]

These define a cancellation syntax over lexical and derived categories with an associated semantics. A sample derivation for a relative clause like: ... *who I want* **e** *to succeed* is given below:

who	*I*	*want*	*to*	*succeed*
(N\N)/(S/NP)	S/(S\NP)	((S\NP)/NP)/VP	VP/(S\NP)	S\NP
	---------- **FC**			
	(S/NP)/VP			
---------- **FC**				
((N\N)/S)/VP				
			---------- **FA**	
			VP	
---------- **FA**				
(N\N)/S				

Figure 1 illustrates the state of the parser at the onset of the shift-reduce conflict for (7). The relative pronoun in cell 2 can be combined with the constituent in cell 1 (forward composition), but the lookahead item can be combined (forward composition) with the constituent in cell 1, so shift is preferred. However, either a lexical preference for the (S/NP)/VP category for *you want* and/or a PB marked by lengthening of *want* could override the default parse action and force the early gap interpretation. The complexity

Stack Cells		Lookahead	Input Buffer
2	1	L	
(who)	(I want)	to	succeed
S/(S/NP)	(S/NP)/VP S/VP	VP/VP	

Figure 1: Shift-reduce Conflict for (7))

After each parse action (Shift, Reduce, Halt):

1. Assign any new Stack entry in the top cell (introduced by Shift or Reduce) a cost of 1 multiplied by the number of CCG categories for the constituent it represents
2. Increment every Stack cell's cost by 1 multiplied by the number of CCG categories for the constituent it represents
3. Push the sum of the current costs of each Stack cell onto the Cost-record

When the parser halts, return the sum of the Cost-record which gives the total cost for a derivation.

Figure 2: The Cost Algorithm

and ambiguity metric is given in Figure 2. For the configuration in Figure 1, ignoring earlier material, the cost associated with cell 1 is 4 (3 shifts and one reduce to reach this state), and that with cell 2 is 2 (reset after the previous reduce action to 1 multiplied by the 2 CCG categories).

Similarly to the metrics of Hawkins (1994, 2004) and Gibson (1998), the cost metric represents the load on working memory during language processing and predicts that costs increase with the length of grammatical dependencies and with the degree of ambiguity (i.e. the numbers of putative dependencies within a sentence) up to the point where extrasyntactic information can be deployed to resolve them (see Briscoe, 1987, 2000a for more details). However, the parser's default preferences (contra Gibson) select analyses which increase stack-depth and hence complexity. That is, in the absence of extrasyntactic information that a potential gap is the actual gap, the parser delays attachment. This strategy actually reduces processing cost *provided that language is organized to override parsing defaults when they lead to the wrong analysis*. So the model places adaptive pressure on grammatical systems to evolve in such a way that PBs (and/or lexical and semantic information) are available at the onset of ambiguities which require non-default interpretations.

The method of integration of PBs into the analysis makes different pre-

dictions from that of Steedman (2000), as it relates PBs to parse actions not to CCG categories. For instance, the 'adverbial' category, Steedman associates with PBs would not block combination of *you want* and *to..* in Figure 1, as required for the analysis of *you want simply to....* Our model predicts that the placement of PBs is mediated more by ambiguity resolution than by structural and informational mapping constraints *per se* and thus departs from the dominant tradition of Selkirk (1984), which Steedman largely follows, not by arguing that there are no such constraints on the syntactic-phonology interface, but by predicting that where these underdetermine the placement of boundaries an account based on ambiguity resolution makes more fine-grained and correct predictions (see also the experiments reported in Snedeker and Trueswell, 2003).

6 Corpus/Usage-based Predictions

Our model predicts a complexity hierarchy of (SRCs < NSRCs) *leq* (unambiguous NSRCs < ambiguous NSRCs) $\wedge$ (short NSRCs < long NSRCs) and thus that in speech NSRCs will mark an actual gap with a PB, particularly if it is ambiguous and not resolvable given effects of local semantic plausibility, lexical frequency or parsing preferences, and that in writing the lack of PBs may lead to avoidance of ambiguous NSRCs. We tested these predictions by automatically extracting RCs from parsed versions of the BNC and SEC corpora, by automatically categorizing wh-RCs into SRCs/NSRCs and manually analysing samples of that(-less) RCs, as well as the correlation of PBs with gaps in NSRCs in the SEC. We found that 1) there is an equal preference for SRCs over NSRCs in speech and writing ((ratios of SRCs:NSRCs are approximately 6.9:1 in speech and 6.4:1 in writing – the difference in ratios is not significant, $\chi^2_1 = 3.2p = 0.07$); 2) there is an equal preference for unambiguous (single verb group) NSRCs in writing and speech ((ratios of unambiguous:ambiguous NSRCs are 4.4:1 in speech and 6.3:1 in writing – again the difference in ratios is not significant, $\chi^2_1 = 1.61p = 0.20$); 3) longer NSRCs containing longer intervening NPs, parentheticals and so forth occur in writing (e.g. (8))

(8) The business that JR, director...of restructuring at M, sees *e* as promising

(measure of average length of NSRCs in spoken data is 2.82 and in writing is 4.07 – these averages are very significantly different, t-test shows $p = 0.0005$); 4) ambiguous medial gaps in NSRCs in the SEC are not marked with PBs where this would lead to the wrong interpretation (35 were found, 32 have no following PB, 3 are marked by minor/intermediate PBs but these occur in wh-adverbial RCs like (9) in which the CCG analysis predicts early 'non-configurational attachment to the verb – e.g. Pickering and Barry, 1991).

(9) ... where there are limited domestic reserves | of some non-renewable resource | as with ...

Actual but ambiguous medial gaps are marked with minor/intermediate PBs and RC-final ambiguous gaps are marked with major/intonational PBs (40 were found, 39 were followed by PBs in the annotation leaving one putative counter example which may be an annotation error). These corpus-based results suggest that ambiguity reduction and prosodic disambiguation play a role in the form of NSRCs observed in speech and writing. The fact that syntactically ambiguous NSRCs occur with equal frequency in writing and speech suggests that in writing there must be a greater reliance on contextual or semantic resolution of ambiguity in the absence of PBs and this needs investigating further. The results are also compatible with the predictions of the model presented, but nevertheless further spoken sentence production and comprehension experiments are needed to test the model more directly.

7 Discussion and Conclusions

Language interpretation involves a decoding step, based on properties of meaning conveyed grammatically, and an inferential step which further constrains and refines meaning by integrating contextual information, background knowledge, and so forth. In general, there is a trade-off between these two steps where more coding usually leads to increased articulatory or production costs, while less coding increases ambiguity and requires a greater degree of inference. For instance, Hoefler (2006) argues that ambiguity emerges as a result of lossy compression to ease the articulation bottleneck in language production. We have shown that enriched syntactic encoding in RC constructions to remove some ambiguities would lead to increased processing complexity. However, a strategy of parallel encoding of the same information in the prosodic phonological system (which is required independently as a component of speech processing) achieves the same effect with very little additional processing cost during the decoding step. It only requires that the parser have access to the location of PBs when faced with the onset of a syntactic ambiguity. This allows syntactic processing to proceed nearly-deterministically reducing the costs of ambiguity without increasing the need for inference.

The evidence reviewed here from psycholinguistic work, typological work and the novel corpus-based investigations we report suggest that human language processing does incorporate default syntactic ambiguity resolution strategies, that these can be overridden by extrasyntactic information, including PBs at the onset of ambiguities, and that language usage does support the model in that PBs do occur in speech in the predicted locations, and

written and spoken usage does reflect the predicted cost hierarchy. Briscoe (2000a) demonstrates that if a cost algorithm very similar to that of Figure 2 is incorporated into a simulation of language evolution, then languages adapt to reduce syntactic complexity in a manner which predicts many well-known typological implicational and/or statistical universals. We predict that an extended version of this simulation will show that languages will adapt to align prosodic and syntactic information to reduce ambiguity.

References

Briscoe, E. 1987. *Modelling Human Speech Comprehension: A Computational Approach*, Wiley.

Briscoe, E. 2000a. Grammatical acquisition: inductive bias and coevolution of language and the language acquisition device. *Language* 76.2, 245–96.

Briscoe, E. 2000b. Evolutionary Perspectives on Diachronic Syntax. In Pintzuk, S., Tsoulas, G. and Warner, A., editors, *Diachronic Syntax: Models and Mechanisms* OUP.

Church, K. 1980. *On Memory Limitations in Natural Language Processing*, Indiana Univ. Ling. Club.

Cooper, W. & Paccia-Cooper, J. 1980. *Syntax and Speech.* Harvard UP.

Gibson, E. 1998. Linguistic complexity: locality of syntactic dependencies. *Cognition* 68, 1–76.

Hawkins, J. 1994. *A Performance Theory of Order and Constituency*, CUP.

Hawkins, J. 2004. *Efficiency and Complexity in Grammars*, OUP.

Hoefler, S. 2006. Modelling Pragmatic Aspects of Language Evolution: The Case of the Articulation Bottleneck, LAGB meeting (paper on-line).

Korhonen, A., Krymolowski, Y. and Briscoe, E. 2006. A Large Subcategorization Lexicon for Natural Language Processing Applications. *Proc. of the 5th Int. Conf. on Language Resources and Evaluation*, Genova, Italy.

Kuno, S. 1974. The position of relative clauses and conjunctions. *Linguistic Inquiry*, 5, 117–136.

Leech, G. 1992. 100 million words of English: the British National Corpus. *Language Research*

Nagel, H., Shapiro, L. & Nawy, R. 1994. Prosody and the processing of filler-gap sentences. *J. of Psycholinguistic Research*, 23, 473–485.

Pickering, M. and Barry, G. 1991. Sentence processing without empty categories. *Language and Cognitive Processes*, 6, 229–259.

Selkirk, E. 1984. *Phonology and syntax: the relation between sound and structure*, MIT Press.

Snedeker, J. and Trueswell, J. 2003. Using prosody to avoid ambiguity: effects of speaker awareness and referential context. *J. of Memory and Language* 48, 103–130.

Steedman, M. 2000. *The Syntactic Process*, MIT Press.

Stowe, L. 1986. Evidence for on-line gap location. *Language and Cognitive Processes*, 1, 227–245.
Straub, K., Wilson, C., McCollum, C. & Badecker, W. 2001. Prosodic structure and wh-questions. *J. of Psycholinguistic Research*, 30, 379–394.
Taylor, L.J. & Knowles, G. 1988. Manual of information to accompany the SEC Corpus. Mimeo, UCREL, Lancaster University.
Venditti, J., Jun, S. & Beckman, M. 1996. Prosodic cues to syntactic and other linguistic structures in Japanese, Korean and English. In *Signal to Syntax*, eds. Morgan, J,. & Demuth, K., Lawrence Erlbaum.
Warren, P. 1985. The temporal organisation and perception of speech. PhD. Dissertation, Dept. of Linguistics, University of Cambridge.
Warren, P. 1999. Prosody and language processing. In *Language Processing* eds. Garrod, S., and Pickering, M. Psychology Press.