Data at the Grammar-Pragmatics Interface: the case of resumptive pronouns in English

Ronnie Cann, Tami Kaplan, Ruth Kempson

Ronnie Cann
Theoretical and Applied Linguistics
University of Edinburgh
Phone: +44 (0)131 651 1839
Fax: +44 (0)131 650 3962
E-mail: R.Cann@ed.ac.uk

Tami Kaplan
Dept. of Resource Development
MIT
E-mail: tkaplan@mit.edu

Ruth Kempson
Dept. of Philosophy
King's College London
Phone: +44 (0)20 7848 2655
Fax: +44 (0)20 7848 2270
E-mail: Ruth.Kempson@kcl.ac.uk

Abstract

This paper explores the relation of grammaticality to acceptability through a discussion of the use of resumptive pronouns in spoken English. It is argued that undergeneration by some grammar of observed linguistic phenomena such as these is as serious a problem for theoretical frameworks as overgeneration, and that it has consequences for the way in which grammaticality and acceptability are to be construed. Using the framework of Dynamic Syntax, a theoretical account of relative clauses and anaphora construal is provided from which the use of resumptive pronouns in English emerges as a natural consequence. The fact that examples are considered by native speakers to be unacceptable in neutral contexts is argued to follow from pragmatic effects, explicable from a Relevance Theoretic perspective.

Keywords: Resumptive pronouns; Dynamic Syntax; grammaticality; acceptability.

1 Acceptability and Grammaticality

It is a commonplace assumption that unfiltered linguistic data cannot constitute the basis for articulating grammar formalisms, not only because of simple errors, hesitations, etc, but because processing and general cognitive constraints may play a role in determining the unacceptability of strings that the grammar formalism may license as wellformed. Centre-embedding constructions are the most well-known type of case:

(1) The mouse the cat the dog chased chased ran away.

The concomitant separation of grammaticality and acceptability judgements has led to a subclass of strings being deemed, conversely, to be ungrammatical but acceptable. This category is used of strings evaluated by speakers as "perfectly straightforward to parse, hence acceptable, but I wouldn't say it that way". The following examples illustrate the point.

(2) What does himself$_i$ want for his$_i$ supper?

(3) There's every student of mine coming to my inaugural.

(4) This is the house that I don't know its name.

The acceptability of these examples may be given variously. For example, (2) might be explained by hypothesizing that the form *himself* is used — incorrectly? — not as a reflexive but as an emphatic homonym not subject to the binding principles.[1] (3) might be argued to be acceptable only in a context which supports the open proposition *There are n students coming to y's inaugural.* The determiner provides the value for n (and the pronoun provides that of y), giving rise to a focus effect, which, one might argue, is precluded by *every.* In example (4), the pronoun is commonly construed as

acceptable as a 'last resort' option to prevent the surfacing of an illicit trace in possessive position (Ross, 1967, Chomsky 1977, Shlonsky 1992).

A relatively common response to such examples is to 'let the grammar decide' the limits of grammaticality, resorting to ancillary explanations of acceptability such as those sketched above. However, such a response is not straightforward, because it brings into question the delimitation of the empirical domain of grammatical theory. If such examples as the above are excluded by the grammar, how is it possible to identify the limits of acceptability? What factors define the examples in (2-4) as acceptable, if they are, but those in (5,6) as not acceptable (if, indeed, they are not)?

(5) What does he$_i$ want for himself$_i$'s supper?

(6) There's every student in the garden.

The status of acceptable but ungrammatical data is particularly problematic: it means assuming that that interacting rules of the grammar formalism have to allow some joint effect to be suppressible by users of the language in a given context in order to allow some information to be extracted from some utterance. But if ungrammaticality can be ignored by hearers in context, how can one ever be certain that the data one is using as the basis of grammar construction is of this sort and not, in fact, genuinely grammatical? Indeed, without some principled basis for distinguishing acceptability and grammaticality, there is the risk that the claim of theoretical linguistics to have an empirical basis will be undermined, as the data that constitute that basis will of necessity be indeterminate. Similarly, the uncertainty over whether a string is grammatical, or 'merely' acceptable, inevitably undermines the task of determining what constitutes knowledge of language, since the limits of such knowledge cannot be strictly determined. Thus the fact

that theories of grammar often generate strings that no speaker would utter in normal circumstances, while simultaneously excluding many that would be so uttered, and indeed may even be frequent, poses real problems for two central questions of theoretical linguistics: what constitutes knowledge of language and what is the empirical basis of linguistics?

This issue is discussed in more general terms in Pullum and Scholz (2001) where grammaticality is discussed with reference to two approaches to syntax. One such approach, which the authors dub Generative Enumerative Syntax (GES), characterises grammars as generating some set of expressions (or structures) which is taken to constitute a language, anything not generated by the grammar is thus not an expression in that language. This characterisation is commonly found within theoretical syntax, applying to such frameworks as most variants of transformational grammar (including the Minimalist Program), Generalized Phrase Structure Grammar, Categorial Grammar and so on. As pointed out by a reviewer, the problem of undergeneration poses a much more serious problem for GES than overgeneration, the more common focus of attention. If certain strings are not generated by a particular grammar, they should not occur and certainly should not be understood by a hearer who should have no means of parsing them. All grammaticality judgements should thus be of the same sort and as categorical as responses to 'word salad' strings such as **the a of man.* Invoking general cognitive strategies to permit interpretation of ungrammatical, but acceptable, strings merely undermines the whole programme of identifying what constitutes knowledge of language, and, as noted above, brings into question the empirical basis of the enterprise.

The second approach to syntax discussed in Pullum and Scholz (2001) is Model Theoretic Syntax (exemplified by recent forms of Head-Driven Phrase

Structure Grammar) which does not set limits to a language by generatively defining some set of well-formed expressions, but 'merely states necessary conditions on the syntactic structure of individual expressions' (Pullum and Scholz 2001:20). In this paper, we follow such an approach and present the case that in situations where the line between grammaticality and acceptability is vague, the grammar formalism should be liberal, matching the reported judgement of acceptability as a hearer. Pragmatic principles may then be invoked to explain the apparent asymmetry between observable language data and acceptability judgments for speakers. As we shall see, this makes possible an explanation of a range of data for which acceptability judgements are unclear, while sustaining a clearcut concept of wellformedness. Despite the liberality of the system itself, the empirical base of the overall explanation is strengthened. In the event, the grammar to be proposed is novel, reflecting the dynamics of a parser; and we argue in favour of this much closer match between the concept of linguistic knowledge (competence) and language use (performance).[2]

2 Resumptive Pronouns in English

The particular case study we take is so-called resumptive pronouns in English. It is widely acknowledged that English pronouns can be used as a marker of the position relative to which a *wh* expression has to be construed (i.e. as marking the "gap", hence "resumptive"), but that such constructions are "substandard" or "marginal" (cf. Chomsky 1977, 1982; Sells 1984; Engdahl 1985; Safir 1986; Shlonsky 1992; Erteschik-Shir 1992):[3]

(7) I had some other point which I can't remember what it is.

(8) I have three people that I don't know how they are in other classes.

(9) He did a lot of things in high school in the 50s that if kids did them now....

There is, however, a lack of consensus as to how to treat such examples. In (7) and (8), the pronoun construed as resumptive occurs in the subject position of a *wh* question embedded within the relative clause, and in (9) it occurs in the object position of the antecedent of a conditional. As these are *wh* islands, extraction is marginal, but rescuable by the use of a pronoun, as noted above, as a "last resort" operation to save what would otherwise be irredeemable utterances, giving rise to the assumption that pronouns may be used resumptively only when *wh*-movement is not licit.[4]

Examples such as the following, however, are not widely discussed, yet are produced by native speakers:

(10) She got a couch at Sears that it was on sale.

(11) He's a professor that nobody liked him.

(12) ... who I was going to have lunch with him....

(13) ...Newton, Mass., where it's been pretty cold there

(14) This is the person that I told her about quitting my dissertation.

These cases all have corresponding constructions with gaps which are fully grammatical, and preferred by native speakers if asked directly:

(10') She got a couch at Sears that e was on sale.

(11') He's a professor that nobody liked e.

(12') ...who I was going to have lunch with e...

(13') ...Newton, Mass., where it's been pretty cold e this afternoon.

(14') This is the person that I told about quitting my dissertation.

The pronouns used resumptively in these examples occur in a variety of positions within the relative clause: subject (10), object (11), indirect object (14'), oblique (12), and locative (13), so one cannot argue for their existence based on idiosyncrasies in grammatical function. Moreover, though some of them can be interpreted either as restrictive or nonrestrictive (10), – (12), this is by no means true of all, as witness (13') which is only nonrestrictive, and (14'), which is only construed as restrictive.[5]

Erteschik-Shir's account (1992: 93-94) for cases like these as "coordinate resumptives" is problematic. She proposes that these resumptively construed pronouns are derived from coordinate structures when the relative clause involved could be considered to be a focus, i.e. providing "new" information. However, note that while this could work in (10) and (13) – indeed in all cases where the relative is naturally construed as nonrestrictive:

(10") She got a couch at Sears and it was on sale.

(13") ...Newton, Mass., and it's been pretty cold there this afternoon.

we still find cases like (11) where the relative clause is not a potential focus, so (11") is definitely odd:

(11") He is a professor and nobody liked him.

And in (14'), whose natural interpretation is restrictive, the conjunction analogue is clearly not equivalent:

(14") This is the person and I told him about quitting my dissertation.

In sum, then, resumptive pronouns appear to be available in English for a wide variety of speakers with no obvious differentiating characteristics.

The resumptive pronoun serves in all these cases to identify the role of the nominal as copied by the *wh* relativizer into the structure projected by the relative clause. This function can be provided by all pronouns, unlike the function of a pronoun as expletive, a property which has to be specifically defined for individual pronouns (eg *it* in English):

(15) It is likely that I am wrong.

The question posed by these resumptive pronoun data is their status. Are speakers of English licensed to provide sentences of this sort, or are they irregular in some way, having to be processed by general reasoning devices that enable a hearer to set aside constraints which their regular system imposes? Alternatively, are they licensed but dispreferred? And if so, what does this mean? In particular, why is such a strategy dispreferred in one language but a standard way of forming relative clauses in another language?

What is notable about previously proposed analyses is that the phenomenon of resumptive pronouns is treated as entirely separate from regular anaphoric processes, thus multiplying the different forms of analysis for what is morphologically a regular pronoun. Of analyses that address the problem in any detail, Chao and Sells 1983 (see also: Sells 1984) analyse such pronouns, not as regular gap-filling constructions that are A' bound by some operator, but as a discrete E-type pronouns, following the E-type form of analysis first introduced by Evans 1980 for cross-sentential anaphora as in (16):

(16) Three men entered. They were laughing.

But this analysis fails to account for those cases for which a restrictive interpretation is available, for which, as our analysis will in due course bring out, an E-type form of interpretation is not definable.

In this paper we outline an account of these phenomena in English as a peripheral but nevertheless systematic part of English syntax, and moreover one which emerges as an expected result of the interaction between structural constraints and anaphora construal. The shift in perspective is to analyse all structural properties of language in terms of the process of building up interpretation. We then show how English can be seen to differ from other languages in which resumptively used pronouns are more freely available, and finally we return to why these options remain in English a marked device, and what this tells us about the interaction between language-internal processes and general pragmatic constraints.

3 The Flow of Language Understanding

Our first step is to set out the background against which we set the explanation. We assume that what has to be built up as interpretation is a formula representing content established in an individual context;[6] and we describe a gradual process of structural growth whereby information is built up on a left-to-right word-by-word basis relative to some context against which choices may be made as the construction process proceeds.

The framework of *Dynamic Syntax* (Kempson et al 2001) models the process of natural language understanding as a monotonic tree growth process defined over the left-right sequence of words, with the goal of establishing some propositional formula as interpretation. Taking information from words, pragmatic processes and general rules, the theory derives partial tree structures that represent the content of a string as interpreted in context up to the current point in the parse. Intrinsic to this process are concepts of underspecification whose resolution is driven by requirements which determine the process of tree growth, having to be satisfied for a parse to be

successful.[7]

To get the model of the PROCESS of establishing such a structure as interpretation, all nodes in the semantic trees constructed during a parse are introduced with requirements to be fulfilled, reflecting the idea that the tree is underspecified with respect to some property that needs to be specified as the parse proceeds. Requirements may be to specify values for any of the labels that decorate a node, but the principal drivers of the parsing process are requirements to establish nodes of certain types, starting from the initial (universal) requirement to build a representation of the propositional content expressed by a string in context. This is expressed as $?Ty(t)$, an instruction to build a tree rooted in $Ty(t)$, the type of a proposition.

To satisfy such requirements, a parse relies on information from various sources. In the first place, there are general processes of construction which give templates for building trees that may be universally available or specific to a language. For example, the initial unfolding of a requirement $?Ty(t)$ may be to establish subgoals $?Ty(e)$ and $?Ty(e \rightarrow t)$, requirements to build the subject and predicate nodes, respectively, as shown in Figure 1, where the pointer, $\Diamond$, indicates the subgoal that is to be attempted next.[8]

[Figure 1 about here.]

Information about tree building may also come from packages of actions encoded in lexical entries which are accessed as words are parsed. An entry for a word contains conditional information initiated by a trigger (the condition that provides the context under which subsequent development takes place), a sequence of actions (possibly involving the building of nodes and/or the annotation of a node with type and formula information) and a failure statement (commonly an instruction to abort the parsing sequence) if the conditional action fails. For example, parsing the word *John* involves

annotating the current node with formula and type values, given that the pointer is at a node decorated with $?Ty(e)$:

		IF	$Ty(e)$
(17)	*John*	THEN	$put(Ty(e), Fo(John), [\downarrow]\bot)$
		ELSE	ABORT

Note the bottom restriction, $[\downarrow]\bot$, which prevents further elaboration of the node it decorates, restricting the predicate itself to decorating only terminal nodes in the resulting structure. This property reflects the compositionality of content as expressed in natural language: words project concepts which must be taken to decorate terminal nodes in the tree from which complex concepts are built up. More complex lexical actions are induced by parsing (non-intransitive) verbs. Parsing the word *upset*, for example, is relative to a predicate requirement ($?Ty(e \rightarrow t)$), and induces a complex sequence of actions that build, and annotate nodes, the final step imposing the requirement to construct a representation of the content of an object DP (see Figure 2).[9]

[Figure 2 about here.]

The parse will continue just in case the next word has a trigger of the appropriate type, i.e. $?Ty(e)$. A string like *John upset Mary* thus gives rise ultimately to the tree in Figure 3 where all terminal nodes are type and formula complete and the remaining type requirements on the predicate and propositional nodes have been satisfied through the compilation of the tree obtained through functional application over types.

[Figure 3 about here.]

Annotations in construction rules and lexical entries are expressed using the Logic of Finite Trees (LOFT, Blackburn and Meyer-Viol 1994) which

provides a means of referring to arbitrary nodes in a tree using the following modal operators (amongst others): $\langle\downarrow\rangle$ the general daughter relation; $\langle\downarrow_0\rangle$ and $\langle\downarrow_1\rangle$ the argument and functor daughter relations, respectively; $\langle\downarrow_*\rangle$ the dominance relation (the reflexive, transitive closure of the daughter relation); and the inverses of these using the mother relation, $\uparrow$.

The specific and novel advantage of LOFT emerges from the use of the LOFT operators in combination with a generalization of the concept of requirement $?X$ to any decoration X. This combination makes it possible to describe partial trees which have requirements on a treenode which are modal in form, which will be fulfilled by some other node having a given annotation. For example, while $\langle\downarrow_*\rangle Fo(\alpha)$ holding at a node n implies that n dominates a node m where $Fo(\alpha)$ holds, $?\langle\downarrow_*\rangle Fo(\alpha)$ holding at n implies that $Fo(\alpha)$ is REQUIRED to hold at a node m dominated by n. This provides an additional mechanism for pairing noncontiguous expressions according as one expression imposes some REQUIREMENT on a node which is secured by a decoration on some discrete node by the other expression.

These requirements encode the need to develop all aspects of underspecification, driving the parsing process in various ways. For example, every node in a tree is associated with an ADDRESS which is encoded as a value to the TREENODE predicate, Tn.[10] A requirement to establish some such address indicates positional underspecification; and this is used to account for long distance dependencies. Dislocated expressions are analysed in terms of initially unfixed nodes whose position in the emergent tree structure is fixed at some later stage in the parsing process. Such nodes are decorated with a modality $\langle\uparrow_*\rangle Tn(n)$, which merely indicates that the node currently under construction is dominated by treenode with address $Tn(n)$, and a requirement to establish the nature of this dominance by establishing a treenode

address, $?\exists \mathbf{x}.Tn(\mathbf{x})$. A construction rule of *Adjunction introduces such unfixed nodes, defining a transition from an incomplete tree of $Ty(t)$ with only a single node to a tree, adding an unfixed node with requirements both to identify its treenode address of the unfixed node and to construct a type e decoration.

Analysing the string *Mary, John upset* in these terms is illustrated in Figure 4 with an initially projected unfixed node and the pointer at the object position.

[Figure 4 about here.]

At the point in the parse at which all words in the string have been processed, there remains outstanding an unfixed node and a requirement to construct a node of type e. In this environment, a process of *Merge* may take place which unifies the unfixed treenode with the current node which satisfies both requirements.[11] Ultimately, completion of the tree yields a $Ty(t)$ Formula value, *Upset(Mary)(John)* decorating the topnode, with all requirements fulfilled.

Interacting with tree growth of this sort is the context-dependent processing of anaphoric expressions. This phenomenon of content underspecification, which we here take in a representationalist spirit (cf. Kempson et al 2001:ch.1 for arguments), involves lexical projection of a metavariable to be replaced by some selected term during the construction process. Such replacement is associated with a substitution process that is pragmatic, and system-external, restricted only in so far as locality considerations distinguishing individual anaphoric expressions preclude certain formulae as putative values of the projected metavariable (i.e. analogues of the Binding Principles, Chomsky 1981, etc.):[12]

(18) Q: Who upset Mary?

Ans: John upset her.

In processing the pronoun in (18), the object node is first decorated with a metavariable **U**, with an associated requirement, $?\exists \mathbf{x}.Fo(\mathbf{x})$, to find a contentful value for the formula label. Construed in the context provided by having processed the question in (18), substitution will determine that the formula **U** is replaced by $Mary$:[13]

		IF	$?Ty(e)$
(19)	*her*	THEN	$put(Fo(\mathbf{U}), Ty(e), ?\exists \mathbf{x}.Fo(\mathbf{x}), [\downarrow]\bot)$
		ELSE	ABORT

With the incorporation into the grammar formalism of the process of substitution to enable the construal of anaphoric expressions, the interpretation of a string is determined by the interaction of computational, lexical and pragmatic processes. A wellformed string is one for which at least one logical form can be constructed from the words in sequence within the context of a given class of computational and pragmatic actions with no requirements outstanding. This concept of well-formedness entails that, within Dynamic Syntax, grammaticality is necessarily sensitive to context. Thus, the formula requirement associated with pronouns has to be satisfied by the selection of some contentive term from the current context. This term may be local to the structure currently being built, or inferentially derived from the wider discourse context. We thus assume a notion of context that may be characterised, neutrally, as a database of terms, propositional and other tree structures, being minimally the set of representations of content established from recent utterances.[14] While articulation of context is needed for a fully formalised theory of interpretation, what is important at this juncture is that identification of interpretation relative to context

is essential for defining well-formedness. Different contexts affect the acceptability/grammaticality of a string and thus judgements made about the same string will vary from utterance to utterance. This in turn means that the grammar formalism must be liberal with respect to the *strings* that it defines as grammatical, since contexts are not (in anyone's theory) part of linguistic competence and so not within the domain of the grammar.

4 Linked Structures and Relative Clauses

We have so far seen how individual trees can be built up following information provided by both general rules and lexical instructions. However, the more general perspective is to model how multiple structures are built up in context, with context construed here as an arbitrary partial tree (or tree sequence). The simplest case is to build up a second structure once a first structure has been completed, as in coordination, but we may also build structures in tandem, constructing first one partial structure, and then another which uses the first as its context, and so interweaving the processes of building paired structures. This process is displayed in particular by relative clauses. The characteristic property of what we shall call "linked" structures is that they share a common term: the process of inducing the second structure involves, not only a transition from one tree to the other, but also a requirement imposed on the second tree that it contain an occurrence of a term that appears in the first.

Taking nonrestrictive relatives as the most transparent case, consider the steps involved in projecting the construal of (20), which involves a dual assertion about John that I like him and that he chain-smokes:

(20) John, who I like, chain-smokes.

The intuition is that the word *who*, correctly described by Jespersen (1927) as a relative "pronoun", provides the pivot from one structure to the other, copying information from one to the other. Having, that is, processed the word *John* to yield a partial tree in which the formula $Fo(John)$ annotates a subject node in some tree (the 'head' node), a transition is licensed by a rule of *Link Adjunction* which builds a relation from that node, introducing a new tree with topnode decorated with a requirement for an occurrence of the formula $Fo(John)$ at some node, without further specification as to where in the newly introduced tree that might be (Figure 5). **Adjunction* can then apply to introduce an unfixed node; and the relative pronoun *who* duly provides the necessary copy at such an unfixed node, in (20) construed as $Fo(John)$.

[Figure 5 about here.]

The process of tree construction then proceeds as in the simpler case of left dislocation, such as *Mary, John upset*, with the initially unfixed node having its position in that tree established in due course through the process of *Merge*. The two rules of *Link Adjunction* and **Adjunction* jointly provide, in conjunction with the lexical actions defining the relative pronoun, a formal reflex of how paired structures can be built subject to a requirement of overlap of content, how information can be introduced into a structure without initially being fixed within it, and how, within such structures, copy devices can be introduced which meet the requirements that dictate the overlapping development of such paired structures. In short, this is a formal matching of the informal observation that relative clauses are structures that are developed in tandem to yield the result of sharing a common term.

The *Link Adjunction* rule, perhaps surprisingly, applies to restrictive and nonrestrictive relative clause construals alike, given the Dynamic Syntax ac-

count of quantified NPs (Kempson et al 2001, ch.7). Within this system, both quantified and nonquantified noun phrases are projected as type e, with quantified expressions analysed as arbitrary names (such as epsilon terms) which denote witnesses of the containing formula (see Kempson et al 2001, chapter 7, Meyer-Viol 1995). Such terms once fully compiled over some constructed logical form have a restrictor which reflects their full quantificational force, following the pattern of the equivalence between the predicate-logic formula $\exists x.F(x)$, and its epsilon-calculus equivalent $F(\epsilon, x, F(x))$.[15] To project such terms, quantified NPs are projected as incomplete terms with accompanying lexically defined scope constraints on their final construal (allowing for lexical idiosyncracy), and these scope constraints jointly determine the interpretation of the logical form once constructed. For example, a string such as *A student smokes* is interpreted by a left-to-right construction process in which the determiner and noun together project an incomplete epsilon term, which then combines with some predicate to yield a logical form and accompanying scope statement, as shown in Figure 6.[16]

[Figure 6 about here.]

This logical form, with its scope statement indicating that the term binding the variable x takes narrow scope with respect to the index of evaluation, is subject to a final rule of interpretation which yields (21) as the output logical form (ignoring details of the temporal specification) which is truth conditionally equivalent to $\exists x.Student(x) \wedge Smoke(x)$.

(21) $Fo(S_i : \ Student(a) \wedge Smoke(a))$
$a \ = \ (\epsilon, x, Student(x) \wedge Smoke(x))$

On this analysis, all quantified NPs project two nodes of type e, one decorated with a variable (introduced by the noun), and one projecting the

composite information necessary for the resulting quantified terms. This enables us to define restrictive relatives as projecting a copy of the variable associated with the nominal, which decorates the internal node of type e in the structure. Nonrestrictive relatives, on the other hand, are defined as involving a copy of the formula decorating the containing node of type e. The result of completing the LINK structure for the restrictive relative clause in (22) is shown in Figure 7.

(22) A man who Sue likes smokes.

[Figure 7 about here.]

The requirement for a copy of the variable 'head' of the relative clause imposed by the rule of *Link adjunction* is satisfied by the relative pronoun, just as in the case of the non-restrictive relative in Figure 3, and the unfixed node bearing the copy is merged with the object node in the LINKed tree, the whole structure being compiled to give a Formula decoration $Fo(\epsilon, x, Man(x) \wedge Like(x)(Sue)))$, with both instances of the variable within the scope of the determiner as required.[17]

5 The interaction of long-distance dependency and anaphora

There is a more general significance to this analysis. Long-distance dependency, anaphoric expressions, and relative clause construal all share the property of projecting some weak specification of structure which is subsequently enriched. In the case of long-distance dependency, such as *Mary, I like*, the first constituent is analysed as providing some information identifying who is being talked about by the term *Mary* but not the role of that

term in the structure. So its structural role is initially underspecified and resolved later when the appropriate structure has been introduced. Anaphoric processes project an underspecified formula which is interpreted in context as taking the value of some antecedently selected term:

(23) Ruth smiled at Tom when she saw him.

Relative clauses then project a combination of these forms of underspecification: the relative clause is construed against the head nominal as context, the relative pronoun is identified as having the same value but at a position in the secondary structure which, as a long-distance dependency effect, is initially unfixed; and the underspecification of position of this copied term in the structure has to be subsequently fixed. All these analyses turn on underspecification at early stages in the construction process with their updating prior to the final structure.

5.1 Strong crossover

By seeing the construal of anaphora and long-distance dependency as two sides of the same coin, we might expect there to be interaction between the processes of interpretation, and indeed there is. Consider the pair of sentences:

(24) John, who I'm certain he said would be at home, is in the surgery.

(25) John, who I'm certain said he would be at home, is in the surgery.

In (24) the word *he* cannot be construed as picking out the same individual as the expression *John*, while, in (25), it can. Why? A natural answer is given by following the steps of processing. Suppose we process the string as a succession of steps as before. The question is, what happens when we get

to the pronoun *he*? Can we interpret this as 'John' and what happens if we do? The answer is that there is nothing to prevent the identification of 'he' as John, but such an update would arise through unifying the unfixed node associated with the relative pronoun and the node decorated by the pronoun (*Merge*), and not by *Substitution*. This is because the unfixed node decorated with $Fo(John)$ will count as local to the node decorated by the pronoun, precluding *Substitution*. So IF this choice of updated is selected for the pronoun, the effect will be to have provided the identification of the position in the resulting structure for the occurrence of the term 'John'. But if this is so, then there will be nothing left to provide the necessary construal of the subordinate structure where there is no subject provided. Since no subject is provided the parsing will break down, and on this interpretation no successful outcome will be achieved. If, however, the choice of interpretation for the pronoun *he* had been as picking out Tom or Dick or Harry, etc., indeed anything other than the name picking out John, then there would be no problem at any juncture in the parsing process: the unfixed node with a term picking out John would still be described as unfixed at the point at which the embedded predicate is reached, and so at this point, this unfixed node can be updated appropriately. In (25), no such problem arises. This is because the process which merges the unfixed node and the open node requiring a subject are unified before the pronoun is ever processed. So when the interpretation process reaches that point, there is no problem interpreting the pronoun – it can be interpreted as either John or Tom, Dick or Harry – any choice will lead to a wellformed interpretation.

5.2 Resumptive Pronouns

The immediate significance of this process-oriented account of what is elsewhere called strong crossover (Postal 1972, Chomsky 1981, etc.) is that we are led to expect that English pronouns used with resumptive construal in relative clauses will be wellformed in all positions:[18]

(26) ...those little potato things that you put 'em in the oven...

(27) There are people who I've had lots of ups and downs in my friendships with them.

(28) One of my cats had a litter that they were extremely wild.

(29) He's over Marshall county where flooding is a concern there.

(30) I have three people that I don't know how they are in other classes.

(31) ...he builds this house, which it's called Pandemonium....

In all cases, once an unfixed node has, through the relative pronoun, been provided with a copy of the head formula, it can be merged with a node which the pronoun decorates. In so doing it will provide a value for the metavariable projected by the pronoun, but crucially without violating the 'bottom restriction', $[\downarrow]\bot$, that ensures that a node is terminal in a tree. In this way, the model anticipates as wellformed the resumptive use of pronouns in English relatives, even though pronouns in English retain the property of contentive nouns in that they only decorate terminal nodes. This is because it is only the *formula* of the head that is copied by the relative pronoun and not any associated structure. Thus, *who* may project a node decorated as $\{\langle\uparrow_*\rangle Tn(n), Ty(e), Fo(John), ?\exists \mathbf{x}.Tn(\mathbf{x})\}$ which can Merge with a node decorated by a pronoun, e.g. $\{Tn(a), Ty(e), Fo(\mathbf{U}), ?\exists \mathbf{x}.Fo(\mathbf{x}), [\downarrow]\bot\}$ because the Merge process causes no growth of the tree below node $Tn(a)$.

English pronouns, therefore, lack the possibility of merging with structures of arbitrary complexity.[19] In particular, in *wh* questions, and in quantificational structures,where no linked structure analysis is available (see Kempson and Meyer-Viol 1999, Kempson et al 2001 for an account of topic structures in terms of a pair of linked trees) resumptive use of pronouns is debarred:[20]

(32) *Which book did John read it?

(33) *Every book, John read it.

As Figure 8 displays, these facts are as we would expect on the assumption that, in English, the pronoun remains a regular lexical item decorating a terminal node in the tree. For in order for the node projected by the pronoun to be unified with the node projected from *which book*, there would have to be no restriction on the decorations projected by the pronoun that the node it decorates be a terminal node in some tree. Should *Merge* apply in the process of interpreting (32), the formula would be replaced by the node dominating the pair of Determiner node and nominal node, and it is this that the terminal node restriction precludes.[21]

[Figure 8 about here.]

By contrast, in relative clause construal, all that is copied over in the introduction of a LINKed structure is the formula projected by the head, and not its internal structure. So, in relative clause construal, the terminal node restriction constraining possible construal of the pronoun is not violated. Hence the acceptability of resumptive pronouns in English relative clauses, but not in all long-distance dependency structures, specifically not in *wh*-questions (unlike their Arabic analogue which allows this (see Demirdache 1991)).[22]

5.2.1 Resumptives as E-type pronouns?

The analysis of resumptive pronouns offered here thus treats such pronouns *in English* as a regular pronoun: it projects a metavariable which is updated by some term as made available in context. There is no need to posit any ambiguity in the lexical specification of the pronoun itself: indeed, given the DS account of quantification, the contrary Chao and Sells 1983 analysis of resumptive pronouns, which invokes an E-type form of construal for an assumed discrete form of pronoun, cannot be sustained.

In order to project an E-type form of interpretation, there must be a specified quantificational term to provide the type of antecedent for the pronoun to give rise to this form of interpretation. The pronoun antecedent relation is invariably cross-clausal, as the term to be selected as antecedent is only made available the final step in some interpretation for a propositional formula, providing as it does for each term in that structure a restrictor which reflects the relative scope of all quantifying terms in that structure. For example, in the canonical case of an E-type form of anaphoric construal, the pronoun can indeed be identified with some term witnessing the truth of the preceding sentence as interpreted, for it takes as antecedent the term $(\epsilon, x, Man(x) \wedge Enter(x))$ constructed from the first sentence,[23] which then has the new predicate *Smoke* applied to it:

(34) A man enters. He is smoking.

A pronoun in a relative clause which is nonrestrictively construed can, in like manner, identify as its antecedent the head of the relative, since, in virtue of the LINK transition being defined from some top node of a noun-phrase projection, a term is made available as input to the construal of the relative clause sequence. However, in restrictive forms of relative-clause

construal, no such term can be available with which to identify a pronoun within the relative clause sequence, as the restrictor for which the relative clause acts as input is under construction at the point at which the pronoun is processed. Any attempt to retain an E-type form of analysis by defining some means of delaying the construal of the pronoun would not only violate compositionality, but would be circular, for the value of the metavariable projected from the lexical specification of the pronoun has by an E-type form of analysis to reflect the full content of the containing relative clause but the interpretation to be assigned to this clause turns in part on the value assigned to this metavariable.[24]

In the present analysis of such resumptive use of pronouns within a restrictively construed relative, no such problem arises. A fixed node decorated with a metavariable (as lexically projected from a pronoun) within the construction of some LINKed structure is, in principle, compatible with there being an unfixed node already introduced into that structure that is decorated with a fixed formula value of the same type. Furthermore, nothing prevents the merging of that node (decorated by actions of the relative pronoun with a copy of the variable projected by the noun which is head of the relative) and the fixed node decorated by the pronoun. Indeed, such an update resolves in one step the different forms of underspecification up to that point associated with the two nodes in question. Furthermore, since in relative clauses the formula value decorating that initially unfixed node in the emergent LINKed structure does so in virtue of a formula-copying device (and not in virtue of any internal structure it may have getting copied over from one tree to the next), there is no conflict with the restriction projected by the pronoun that the formula value assigned as its interpretation be a terminal node in the tree. Hence, the wellformedness of resumptive rela-

tive pronouns in English, either under a nonrestrictive or restrictive form of construal.

5.3 Acceptability

We have thus proposed an account of resumptive pronouns in English that is permissive with respect to the data, but this poses the question as to why native speakers often judge sentence with resumptive pronouns to be less acceptable than their gapped counterparts. Our answer turns on the presumption that all anaphora construal is a unitary pragmatic process, but pragmatic considerations will influence acceptability. If we adopt Relevance Theory assumptions, we would expect that all aspects of utterance interpretation not determined by rule will be constrained by a balancing of cognitive effort with some degree of cognitive effect (see Sperber and Wilson 1995). On this basis, any unnecessary morphological processing (parsing or production) will be avoided unless its use in some way benefits the interpretation process - either to secure additional pragmatic effects, or to ensure processing success that might otherwise be judged to be at risk. And this is exactly what we find.

On the one hand, resumptive use of pronouns in English becomes fully acceptable if the resumptive pronoun is stressed, and interpreted as inducing a contrastive interpretation:

(35) That friend of Mary's, who HE had to be the one to admit that we had low teaching loads, was roundly condemned.

(36) That friend of Mary's, who even HE admits he needs a holiday, was nevertheless at the conference.

In these examples, it is only the use of a stressed resumptive pronoun that

allows focus on the subject of the relative clause (stressing *who* has no such effect). Hence, there is sufficient pragmatic effect for the examples to be acceptable in an otherwise null (or unspecified) context.[25]

On the other hand, resumptive use of pronouns is invariably judged to be more acceptable when they occur in an embedded structure:

(37) I had some other point which I can't remember what it is.

(38) That friend of yours who Sue noticed that he was looking unwell has been taken off to hospital.

10' She got a couch at Sears that her brother told her he was sure it was on sale even though it wasn't.

Examples of this latter type show that acceptability improves just in case the resumptive either saves an otherwise potentially unacceptable string or that simply identifies a 'gap site' that is not local to a dislocated expression. These cases are thus explicable in terms of ease of processing rather than pragmatic effect.

However, the resumptive use of pronouns is judged not to be acceptable where there is no apparent justification of either type:

(10) ?She got a couch at Sears that it was on sale.

(11) ?He's a professor that nobody liked him.

Yet even these occurred in conversation, as did other comparable data:

(31) He builds this house, which it's called Pandemonium,....

(26) Those little potato things that you put 'em in the oven.

Given such examples, we would, given a relevance theoretic perspective, expect acceptability to improve in context. For example, the string in (39)

occurred as part of a discussion of child actors and what they get up to on stage. The pragmatic effect of the pronoun here appears to have been to emphasise the agentivity of the children and the example was perfectly natural within that context.

(39) I've had children that they've come in and ... (BBC Radio 3 21/07/2002)

In the other examples recorded above, we expect acceptability also to be affected by further specifying the context. (31) and (26) both (for example) improve when you extend the context:

(31") He builds this house, which it's called Pandemonium, and rightly so!

(26") Those little potato things that you put 'em in the oven and when you take them out again, they've turned into mush.

That this analysis is on the right track is suggested by other phenomena such as the parallel use of definite noun phrases as resumptive, a phenomenon which is widespread cross-linguistically:[26]

(40) That friend of yours, who the idiot had to be the one to admit that our teaching loads were low, was roundly condemned.

(41) John, who Sue tells me the poor dear is suffering dreadfully from overwork, is still at the office.

From a relevance-theoretic perspective, the use of definite noun phrases as resumptive can always be justified on the grounds that the very addition of the predicative content of the nominal accompanying the definite article ensures additional inferential effects can be derived from this addition, thus justifying the additional processing cost of parsing the determiner-noun configuration. It is notable that this resumptive use of definite noun phrases is

expected in this framework, given independent justification of an analysis in which definite noun phrases are analyzed as a type of anaphoric expression (see Kempson et al 2001, ch.1).[27]

This suggests that indeed pragmatic relevance constraints play a role in determining acceptability. Prosodic cues may be used to provide additional contrastive effects improving acceptability where there is no processing difficulty in establishing the requisite anaphoric construal. In the absence of any such special intonational effects, acceptability may be determined by the increased ease of retrieval of the requisite form of anaphoric construal when the structure under construction reaches a certain level of complexity. Otherwise, acceptability may be improved by embedding the examples within richer contexts which induce other pragmatic effects.

5.3.1 Quantification and resumptive pronouns

One reported restriction on resumptive pronouns not so far explained is the supposed prohibition of any resumptive form of construal should the head be a quantified expression:

(42) *Every passenger who I had to take him on one side and give him a thorough search complained to the authorities about my treatment of him.

(43) *We are told that any passenger who we have to report him to the authorities must have his rights read out to him.

The problem with this characterisation of the unacceptability of (42) is that all apparent counterexamples, where acceptability is not in doubt, are deemed not to be "properly quantificational" (as one referee put it):

(44) I suspect that there are other professors who I could be their technician.

(45) I had some other point which I can't remember what it is.

(46) There are people who I've had lots of ups and downs in my friendships with them.

(47) It's an antibiotic-like thing that you give it to them.

(48) It's the kind of place that it would need....

Given the analysis of quantification adopted here, this is not a possible stance to take. All such cases, indeed all indefinitely quantified expressions, constitute quantified expressions. In any event, the suggestion that the phenomenon of specificity involves an ambiguity of indefinites, with one interpretation corresponding to a referential expression (following Fodor and Sag 1982, and many others since) is not sustainable, as we and others have argued at length elsewhere (following Farkas 1981, and many others). On the analysis set out in Kempson et al 2001, Kempson and Meyer-Viol 2002, indefinites are existentially quantified expressions (represented as epsilon terms in the resulting logical structure). They are lexically idiosyncratic in that choice of scope assignment for these expressions is very free, licensing a pragmatically driven choice of dependency on any other term constructed within the same interpretation process, a choice that may give rise to very wide scope effects. Nevertheless, once a scope relation is fixed, the construction of logical form and its evaluation relative to a scope statement follows the pattern of all other quantifying terms. Whatever the apparent preclusion of (42), it cannot be due to a restriction that in principle no quantified expression can be followed by a restrictively modifying relative clause.

The restriction, we suggest, is a matter of agreement. For a singular determiner such as *any* and *every*, agreement between some pronoun and its antecedent, which by analysis on a restrictive construal is the variable which that determiner binds, would seem to demand a singular pronoun. The singular specification associated with the pronoun however is a filter on output, and this necessitates that the entity denoted be a singleton set, a condition not met by such universal quantifiers, hence the tension.

This leaves us with examples such as (44)-(48). The form of analysis we propose is pragmatic: the system itself licenses all such cases as wellformed, and pragmatic constraints act as an additional filter.[28] The issue is why a speaker might choose to use a pronoun identified with the head of some relative already realised as a variable when expressing the restrictor of an indefinite term, but would not do so if the head of the relative is a variable to be bound by a non-indefinite form of quantification? Here, the Fodor and Sag argument has a role to play.[29] An extremely salient form of justification for an indefinite assertion of the form $\exists x.F(x)$ is to have information that some fixed entity has the property in question, viz. $F(a)$, for some fixed a, with a step of inference to the existentially quantified statement, with all sorts of reasons possibly intervening to determine that the speaker wishes to express the weaker existentially quantified assertion, rather than the more specific assertion about that particular individual. Given the salience to the speaker in entering the production task of the representation of any such individual, the choice of pronoun may seem entirely natural, even though the hearer is not intended to recover this representation given the decision to express a quantified statement. Fortunately, this asymmetry between the context available to the speaker and that available to the hearer is harmless, since the hearer can recover the existential form of proposition expressed through

the regular construction of a LINKed structure with a variable as the shared term, together with simple application of Merge as a means of identifying the appropriate interpretation of the LINKed structure. Hence the reasonably natural, albeit peripheral use of pronouns resumptively construed in restrictive interpretation of relative clause sequences. To the contrary, however, should the speaker be deciding to utter a non-indefinite quantified statement, no such easy inference from attributing some predicate to a fixed individual to the corresponding existentially quantified statement is possible. The only entity in mind independent of the linearisation task itself is the set of individuals about whom the quantified assertion is made. Hence, in such cases, there cannot in principle be some entity in mind for the speaker in virtue of which she uses a pronoun as a reflection of its salience; so the restrictor for some nonindefinite term will normally not be linearised with a singular pronoun within the relative clause sequence.[30] However, as we would expect, as long as no agreement clash arises between pronoun and antecedent, resumptive pronouns remain as a borderline possibility, even without such pragmatic buttressing:

(49) ?All trainee parachutists, who the pilot had carefully drilled them in what they had to do, had to jump out of the plane at 200 metres.

5.4 Cross-linguistic variation

Taking the approach to resumptive pronouns that we have, where the grammar is liberal and acceptability is determined by pragmatics, naturally brings up the problem of cross-linguistic differences in the use of such devices in other languages. If the framework is so flexible as to license such data as wellformed in English, what basis does it provide for such cross-linguistic distinctions? Our analysis of resumptive pronouns in English has turned

on three properties: (i) the process of node unification (*Merge*) applying in two environments, both providing acceptable forms of update as input – when the fixed node has a requirement for a formula of a certain type, and when the fixed node is decorated with a metavariable of a certain type needing completion; (ii) the relativizing element being analysed as a form of pronominal, albeit one with an algorithmically fixed value – this relative pronoun provides a copy of the head from which the linked structure is induced; (iii) the resumptively construed pronoun being an entirely regular pronoun, in particular retaining the constraint indicative of a lexical element that it annotate a terminal node in the tree under construction.

If any one of these properties varies, distribution of resumptively used pronouns can be expected to vary. For example, if the process of *Merge* is restricted so that it only applies when the fixed node has no formula decoration but only a requirement (a simple restriction on the general rule of Merge available in languages like English, see Kempson et al. 2001:86,87), then no resumptively used pronoun will be licensed. This characteristic generally holds in Germanic languages, English apart. In both German and Dutch, resumptive pronouns in relative clauses are invariably judged to be illformed with greater severity than the corresponding English data:[31]

(50) **Das Buch dass ich es gelesen hat,*
the book that I it read have
the book that I have read it

Though in Germanic languages, the relativising complement is undeniably anaphoric, in other languages the relativizing element in the language appears not to constitute an anaphoric relative "pronoun" that creates a copy of the head, but rather, more weakly, to merely signal the presence of a LINK transition and the REQUIREMENT for a copy of the head. In this

event, the language has to make greater use of the regular copy device in the language to ensure the copy of the head in the LINKed structure, for there is no structure-specific term to provide it. And this leads to the canonical resumptive use of pronouns in relative clauses in languages such as Arabic: without the occurrence of a relative pronoun suitably construed as identical to the head and decorating an unfixed node within the newly introduced LINKed structure, the requirement for a copy of the formula at the head node from which the LINK transition is constructed will not be met as a result of processing the relativising expression. The condition on completion of the construal of the relative clause, that is, will not be satisfiable UNLESS an anaphoric device in the language is used to provide it. Hence, the obligatory use of resumptive pronouns in Arabic:[32]

(51) *il* *mudarris* *illi* *Magdi* *darab-u*
the teacher who Magdi hit him
'the teacher who Magdi hit' [Egyptian Arabic]

On this analysis, the pronoun itself does not need a specific lexical definition to ensure this distribution: the actions of the complementiser *illi* is what guarantee it – the pronoun is merely the vehicle for meeting the requirements the actions of the complementiser impose.

Finally, there is variation induced by a pronominal element losing the defining criterion of what one might dub "full-content" lexical items – that the item decorate what is a terminal node in the resultant tree. In such circumstances, one would anticipate freer optional use of resumptive pronouns, as unification of an unfixed node and a node decorated with a pronoun would then be applicable in all unfixed node structures (see Anagnastopoulou 1993, Tsimpli 1998, Alexopoulou 1999 for discussion of the Greek data). This is in effect the formal reflex of what it means for a pronoun to start shifting

into an agreement device, and, reflecting this, such account might be given to Greek object clitics (commonly said to be an agreement phenomenon) in clitic doubling in Greek:[33]

(52) *Ton Petro ton nostalgo poli* [Greek]
The Peter$_{ACC}$, Cl$_{ACC}$ miss-1sg much
'I miss Peter a lot'

Notice here what this stance leads us to expect with respect to acceptability. In Arabic, where resumptive pronouns are forced to occur because of the lack of a distinct copy-device within the relative clause, no pragmatic effects are expected with the appearance of resumptive pronouns (in non-subject positions) whereas their omission is perceived as categorically ungrammatical. However, the analysis does lead us to suppose that inferential effects may arise where a pronoun appears where it is, strictly speaking, not necessary, such as in subject position. Accordingly, the strong form of pronoun will be used in subject position whenever the resumptive element is, for whatever reason, phonologically stressed (see Tsiplakou 1998 for comparable Greek data):

(53) *?* *irra:gil illi* HUWWA *mabsu:t* [Egyptian Arabic]
the man who HE happy
'the man who HE is happy'

(54) *irra:gil illi (hatta)* HUWWA *mabsu:t*
the man who (even) HE happy
'the man who (even) HE is happy'

The use of a pronoun in subject position immediately following the relative complementiser can never be justified on the grounds of alleviating cost in processing. However, just as in English, their use can be justified on grounds of the additional inferential effects that can be achieved by using a strong pronoun.[34]

With respect to Greek, however, the significance of the resumptive construal of clitic or other phonologically weak forms of pronoun is that in virtue of having a reduced phonological form, parsed as a single unit with the verb, clitic pronouns involve minimal cognitive effort in establishing the regular structure, and accordingly their use relies less on the derivation of additional effects for justification. Indeed the combination of the minimal cognitive effort in parsing a clitic pronoun and its associated inability to bear stress go together as expected on a relevance-theoretic perspective. It is then no longer a surprise that it is only in a language such as English with pronouns that have retained their 'independence' that the effect of using a resumptive pronoun without any rhetorical effect to compensate the cognitive effort of processing will lead to reduced acceptability judgements as a reflection of less than optimal relevance of the form. In languages with a clitic variant, such considerations don't arise as long as the clitic form of pronoun is used.

6 Conclusion

Overall, by comparing English use of resumptive pronouns to their different uses in other languages, we have found additional grounds for concluding that the computational system for parsing English sentences freely allows resumptive construal of pronouns in processing relative clauses: their actual use in relative clause constructions is determined by whether they are perceived by the speaker as contributing to the task of optimally conveying some required interpretation.

The reason why informants very generally reject sentences containing resumptively used pronouns when asked to make acceptability judgements in isolation about them is, arguably, that they attempt to reconstruct prag-

matic considerations as part of their judgement of acceptability. This can be explained on the view in which parsing is the basic tool, and sentence production involves a hypothesis as to the best way to linearly package words for optimal (relevance-constrained) recovery of interpretation. If the speaker can envisage an apparently simpler way of expressing some apparently intended content than the string for which a context-independent acceptability judgement is called for, then this will influence the acceptability judgement to be assigned to the string in question. Hence the often-accompanying comment: "I can understand it easily enough, but I wouldn't say it that way." Presented with a null context, this response is to be expected, but in richer contexts we predict that responses will differ and the examples judged to be more acceptable.

In closing, let's take stock of the overall perspective. According to the orthodox methodology, evidence for linguistic competence can only be culled from judgements of grammaticality independent of language use, for there is no direct relation between the competence model to be defined and the phenomenon of language use, the former being logically prior. On such a view, gradient judgements of wellformedness have to be reflected in corresponding gradient concepts of grammaticality, with attendant difficulties over what it means for a string to be fairly ungrammatical.

On the new view, to the contrary, the characterisation of linguistic ability is not defined to reflect judgements of wellformedness: it is defined to reflect an abstraction from the parse process. The grammar articulates constraints on possible parse sequences, and these constraints together with general pragmatic considerations determine the retrievability of interpretation of sentences in context, hence their acceptability in context. The chasm between competence and performance in more orthodox frameworks is replaced

by a perspective in which the competence of a speaker is seen as a direct abstraction from performance data. Grammaticality judgements, required to be context-free, are replaced by judgements of acceptability relative to context. The methodological advantage of this move is that it provides a basis for explaining how it is that while some judgements of acceptability are clearcut, others are much more uncertain and blurred, without any carry over of unclarity or gradience into the concept of wellformedness itself.

On this new view, we leave on one side the concept of linguistic ability as a disembodied body of knowledge, witnessed only by evaluations of sentences as grammatical/ungrammatical which no more than partly correspond to reported intuitions of acceptability. In its place, we advocate a concept of linguistic ability which, errors and false starts apart, is directly reflected in observable facts of language use, in particular in language perception.[35] Linguistic ability is having a capacity to induce logical structure from sequences of words relative to contextually provided structure. The linguistic competence that reflects our judgements on our own language and the way we use it is then the combination of having accessible encoded specifications which guide language processing, and there being pragmatic constraints on general processing, these jointly determining language understanding. With this move, linguistic theorizing and the articulation of formal grammar specifications can return to the real world of modelling language as used in context.

Notes

[1]This would not, however, explain either the apparent accusative form or the fact that in many languages emphatic pronouns are homonymous with reflexive ones.

[2]In preparing this paper, we are grateful, as ever, to colleagues and students for their input. In particular, we thank Wilfried Meyer-Viol, Eleni Gregoromichelaki, and Theodora Alexopoulou for many useful conversations on the topics raised by resumptive pronouns, to Lutz Marten, Masayuki Otsuka and Caroline Heycock for conversations on general issues pertaining to this paper, and to Annika Thiem for the German resumptive-pronoun data, which await further study. We also thank two anonymous referees for helpful comments on an earlier draft.

[3]The data were collected by the second author between 1995 and 2001.

[4] Other analyses bear similar characteristics. Erteschik-Shir (1992), for instance, argues for these as "distance resumptives" which are not syntactically derived but spelled out at PF as a result of processing constraints due to the distance between the trace and its antecedent.

[5]Fabb 1990 (inter alia) reports that the relativiser *that* cannot occur with nonrestrictive construals, but (10) and (11) arguably can have interpretations in which they are interpreted as two separate assertions, hence with the relative construed nonrestrictively.

[6]This assumption reflects the tradition in the philosophy of psychology in which cognitive systems are taken to interpret incoming signals as providing information about the world around them by constructing from those signals some internalized representation which they take to denote those entities — the so-called representationalist view of the mind (see Fodor 1981, 1983 and many references thereafter). Though all details as to how to flesh out this programmatic statement about cognitive systems are fiercely contested, there is little disagreement that some such assumption has to be made (see eg Dennett 1993). The representations are thus logical forms in some selected calculus for which inference can be defined.

[7]This framework shares many of the conceptual ideas as another theory of the same name espoused in Tugwell 1999. There are, however, significant differences, not least of which is the fact that the representations that are incrementally built in the current theory are of content, not the dependency relations between words or the structural properties of strings.

[8]In this paper, we do not give formal rule definitions, simply showing their effect in

tree displays. See Kempson et al 2001.

[9]Here and below, all tense information is ignored as not germane to the current discussion. Note also that the order of functor and argument is irrelevant and does not reflect string order, because the trees represent only the *content* expressed by the string, not any phrasal structure. In this paper an arbitrary decision has been made to order trees so that arguments appear to the left of their functors, unless this seriously undermines readability. *Word order* in a string is defined by properties of the movement of the pointer ($\Diamond$) within the content trees, as determined by lexical and computational actions. Thus, in English, SVO order is determined by the (general) default assumption that arguments are analysed before predicates; by the trigger for parsing verbs being a predicate requirement; and having lexical entries for ditransitives leave the pointer on the internal object node.

[10]The topnode of a tree has an address $Tn(0)$ from which other addresses are constructed regularly: the functor daughter of a node with address $Tn(n)$ has an address $Tn(n1)$ while the argument daughter has an address $Tn(n0)$. In Figure 3, for example, the node labelled by $Fo(John)$ has an address of $Tn(00)$, the predicate node has address $Tn(01)$ and the node decorated with $Fo(Upset)$ has address $Tn(011)$ and so on.

[11]Formally, updating a tree with an unfixed node involves attempting to unify the unfixed node (by *Merge*) with each node along the path between the unfixed node and its point of resolution. Note that this process of *Merge* is not to be confused with Merge in variants of the minimalist Program (Chomsky 1995, etc.). Here, *Merge* merely unifies the descriptions of two nodes and is successful if (and only if) no contradictory annotations result.

[12]For pronouns, the provided value must not decorate an argument node of the same predicate as the metavariable.

[13]A more detailed specification of *her* would include a first sub-entry that caused the update sequence of actions to abort in an environment in which the node to be decorated was a subject node, but we ignore this complexity here.

[14] There is in this framework no analogue to the differentiation between discourse referents and conditions in Discourse Representation Theory (Kamp and Reyle 1993). The more general concept of context is made possible by the analysis of noun phrases as projecting expressions of type e (see below) and because of the embedding of the formalism within an inferential theory of pragmatics such as Relevance Theory (Sperber and Wilson 1986/1995), within which all interpretation is structural, and potential substituends are

constructed from such structured information as is available to the participants.

[15]The epsilon calculus was a formal system defined to provide explicit analysis of arbitrary names as used in predicate logic natural deduction proofs: Hilbert and Bernays 1939. See Meyer-Viol 1995 for discussion.

[16]S_i representing an index of evaluation.

[17] In the corresponding nonrestrictive construal, it is the constructed term $(\epsilon, x, Man(x))$ that is copied over as a requirement on the newly introduced root node for some LINKed tree. With an epsilon term specification induced as a requirement on the way the LINKed tree should unfold, the relative pronoun is duly identified as projecting the value $Fo(\epsilon, x, Man(x))$ and the relative clause projects the second of two conjuncts both containing this term (see: Kempson, Meyer-Viol and Otsuka 2003):

(i) $\quad S < x \quad Fo(Smoke(\epsilon, x, Man(x)) \wedge Like(\epsilon, x, Man(x))(I))$

[18]See Kempson et al 2001 for a detailed analysis of crossover phenomena, integrating what are otherwise said to be various different forms of crossover (Lasnik and Stowell 1992).

[19]Arguably the expletive *it* should be defined in these terms (see: Cann 2003).

[20]Sells 1984 reports the existence of (i), for which we would have to allow that NPs of the form *which of X* sequences in English can decorate an independent linked structure:
(i) Which of the linguists do you think that if Mary marries him, then everyone wll be happy.
But on the basis that (32) seems to be not wellformed, we resist this move as a general strategy.

[21]This analysis requires that *wh* question words such as *who* project a complex structure of the form:$[_e WH[_{cn} x, person(x)]]$. This structure reflects both their semantics and the status of *wh* words in questions as a type of indefinite, not needing a substituend within an individual interpretation process.

[22]We leave questions of Pied-Piping on one side here, partly because of space and partly because our dataset contains no examples of resumptive pronouns used with Pied-Piped relative phrases.

[23]In the DS account, the term results from the scope evaluation algorithm applying to logical-form scope-statement pairs.

[24]It is notable that no formal analysis is provided in Chao and Sells 1983.

[25]Stressing the pronoun in strong crossover examples such as (24) notably does not

allow the pronoun to be construed as *John* (24') while it is possible if the following gap is replaced by a resumptive pronoun as in (25'):

(24') John, who I'm certain HE said would be at home, is in the surgery.

(25') John, who I'm certain HE said he would be at home, is in the surgery.

Thus pragmatic considerations notably do not over-ride system-internal construction processes.

[26]Since the Kaplan collection was restricted to resumptive pronoun use, the collection did not contain epithet use of definite noun phrases.

[27] In all other frameworks, this use of definite noun phrases is problematic, requiring an analysis in terms other than the regular binding mechanism.

[28]This stance places the Dynamic Syntax formalism squarely within the set of model-theoretic syntactic frameworks, as characterised by Pullum and Scholz 2001.

[29]This does not however involve Fodor and Sag's assumption that indefinites are lexically ambiguous or lead to ambiguity at the level of logical form.

[30]See Cormack and Kempson 1990, where arguments against any ambiguity account of specificity using plural indefinites are set out, together with the above proposed pragmatic explanation for the supposed specificity effects, albeit not in the Dynamic Syntax terms here presupposed.

[31]It remains to be seen whether this would be the correct approach as our provisional culling of German examples of resumptive pronouns in weak island environments with data such as (i) suggests that the situation may not be more clearcut than in English:

(i) Das ist der Laden, von dem ich nicht weiss, ob er morgen auf hat.

'That's the store which I don't know whether it will be open tomorrow.'

[32]Resumptively construed pronouns are not required in subject position because the lexical actions given by the verb project a full template of structure for the predicate and its associated arguments, annotating the subject with a metavariable exactly as though a pronoun had been expressed.

[33]With such weakened forms of pronouns displaying the shift towards being an agreement device, with the loss of bottom restriction, we expect a contrast in these languages between these and strong forms of pronouns, which, by contrast, retain this bottom restriction. Indeed, in many languages, the distribution of the strong form of pronoun may become restricted so that it must co-occur with a clitic form of pronoun. See Kempson et al for analysis of this phenomenon in terms of a restriction on the strong pronoun that it

only decorates LINKed structures.

[34]Though these data go against the observations by Aoun and Choueiri (2000), who report that use of strong pronouns in subject positions is only licensed where the anaphoric construal of these is across a subjacency boundary, the observation by our Egyptian Arabic informants was that with appropriate contrastive stress these are fully well formed. There is reason to think that these judgements would transfer to all forms of Arabic, as it is notable that if suitably subordinate, such forms become fully acceptable irrespective of suitable contrastive stress. See also Aoun, Hornstein and Choueiri 2001.

[35]This stance requires an account of production, which makes use of the same parsing process. See Kempson and Otsuka 2002, Otsuka and Purver 2003.

References

Alexopoulou, T. 1999)The Syntax of Discourse Functions in Greek: a non-configurational approach. Unpublished PhD dissertation, University of Edinburgh.

Anagnastopoulou, E. 1993. On representations of clitic doubling in modern Greek. Paper presented at the ESF workshop on Clitics and Their Hosts. University of Trondheim, June 1993.

Aoun, J. and Choueiri, L. 2000. Epithets. Natural Language and Linguistic Theory 18. 1–39.

Aoun, J. Hornstein, N. and Choueiri, L. 2001. Resumption, movement and derivational economy. Linguistic Inquiry 32, 371–404.

Blackburn, P and Meyer-Viol 1994. Linguistics, logic, and finite trees. Bulletin of Interest Group of Pure and Applied Logics 2, 2-39.

Cann, R. 2003. Towards a dynamic account of 'be' in English. MS, University of Edinburgh.

Chao,W. and Sells, P. 1983. On the interpretation of resumptive pronouns. NELS 13.

Chomsky, N. 1977. On Wh-Movement. In P. Culicover, T. Wasow and A. Akmajian, eds. Formal Syntax. Academic Press, New York. 71-132.

Chomsky, N. 1981. Lectures on Government and Binding. Dordrecht: Foris.

Chomsky, N. 1982. Some Concepts and Consequences of the Theory of Government and Binding. MIT Press, Cambridge, Massachusetts.

Chomsky, N (1995) The Minimalist Program. Cambeidge, Mass., MIT Press.

Cormack,A. and Kempson, R. 1990. On specificity. In J.Garfield and M.Kiteley (eds), Meaning and Truth, . New York: Paragon House. 546-82.

Demirdache, H. 1991. Resumptive Chains in Restrictive Relatives, Appositives, and Dislocation Structures. Unpublished Ph.D dissertation MIT.

Dennett, D. 1993. Learning and labeling. Mind and Language 8, 540-593.

Engdahl, E. 1985. Parasitic Gaps, Resumptive Pronouns and Subject Extractions. Linguistics 23, 3-44.

Erteschik-Shir, N. 1992 Resumptive Pronouns in Islands. In H. Goodluck and M. Rochemont, eds. Island Constraints. Dordrecht, Kluwer Academic Publishers, 89-108.

Evans, G (1980) On pronouns. Linguistic Inquiry 11, 337-362.

Fabb, N (1990) The difference between English restrictive and non-restrictive clauses. Journal of Linguistics 26, 58-77.

Farkas,D. 1981. Quantifier scope and syntactic islands. In R.Hendrick et al (eds.), Chicago Linguistic Society Proceedings 17, 59–66.

Fodor, J.A. 1981. Representations. MIT Press.

Fodor, J.A. 1983. Modularity of Mind. MIT Press.

Fodor, J. D. and Sag, I. A. (1982) Referential and quantificational indefinites. Linguistics and Philosophy 5: 355-398.

Hilbert, D and Bernays, P. 1939. Grundlagen der Mathematik II, Berlin, Julius Springer.

Jespersen, O. 1927. A Modern English Grammar on Historical Principles. George Allen and Unwin.

Kamp, H. and Reyle, U. (1993) From Discourse to Logic. Dordrecht, Kluwer Publications.

Kempson, R. and Meyer-Viol, W. 1999. Topic and Focus: The Dynamics of Tree Growth. In Bouma, G. et al (eds.) Constraints and Resources in Natural Language Syntax and Semantics. Stanford: CSLI, 175-195.

Kempson, R., Meyer-Viol, W, and Gabbay, D. 2001. Dynamic Syntax: The

Flow of Language Understanding. Oxford, Blackwell.

Kempson, R., and Meyer-Viol, W. 2002. Indefinites and scope choice. in A. Bezuidenhout and M. Reimer Descriptions and Beyond: An Interdisciplinary Collection of Essays on Definite and Indefinite Descriptions and Other Related Phenomena. Oxford: Oxford University Press.

Kempson, R. and Otsuka, M. 2002. Dialogue as collaborative tree growth. Edilog Proceedings. Edinburgh University.

Kempson, R., Meyer-Viol, W., and Otsuka, M. (2003) Growth of logical form: the dynamics of syntax. In Peregrin, J. (ed.) Meaning in the Dynamic Turn. Oxford, Elsevier.

Lasnik, H. and Stowell, T. 1991. Weakest crossover. Linguistic Inquiry 22, 687-720.

Meyer-Viol, W. 1995. Instantial Logic. Unpublished Ph.D dissertation, Utrecht.

Otsuka, M. and Purver, M. 2003. Incremental generation by incremental parsing. EACL Proceedings.

Postal, P. 1972. Crossover Phenomena. Holt, Rinehart, Winston.

Pullum, G. K. and Scholz, B. (2001) On the distinction between model-theoretic and generative-enumerative syntactic frameworks. In G. Morrill, P. de Grote and C. Retoré (eds.) Logical Aspects of Computational Linguistics: 4th International Conference LACL. Berlin, Springer, 17-43.

Safir, K. 1984. Multiple Variable Binding. Linguistic Inquiry 15, 603-638.

Ross, J R (1967) Constraints on Variables in Syntax. Unpublished doctoral disssertation, MIT.

Safir, K. 1986. Relative Clauses in a Theory of Binding and Levels. Linguistic Inquiry 17, 663-690.

Safir, K. 1996. 'Derivation, representation and resumption. Linguistic In-

quiry 27, 313-340.

Sells, P. 1984. Syntax and Semantics of Resumptive Pronouns. Unpublished PhD dissertation, University of Massachusetts at Amherst.

Shlonsky, U. 1992. Resumptive Pronouns as a Last Resort. Linguistic Inquiry 23, 443-468.

Sperber, D. and Wilson, D. 1995. Relevance: Communication and Cognition. Oxford, Blackwell.

Tsimpli, I-M. (1998) Null operators, clitics and identification: a comparison between Greek and English. In A. alexiadou, G. Horrocks, and M. Stavrou (eds.) Studies in Greek Syntax. Dordrecht, Kluwer, 241-262.

Tsiplakou, S. 1998. Focus in Greek: Its Structure and Interpretation. Unpublished Ph.D dissertation, School of Oriental and African Studies, London.

Tugwell, D (1999) Dynamic Syntax. Unpublished PhD dissertation, University of Edinburgh

List of Figures

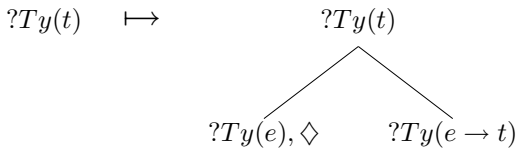

Figure 1: An initial expansion of $?Ty(t)$

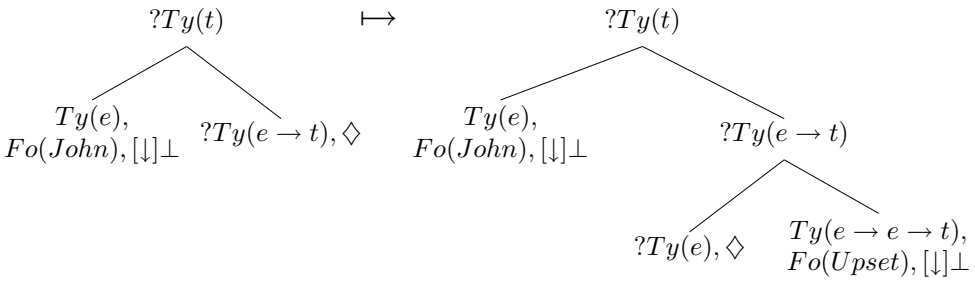

Figure 2: Parsing *John upset*

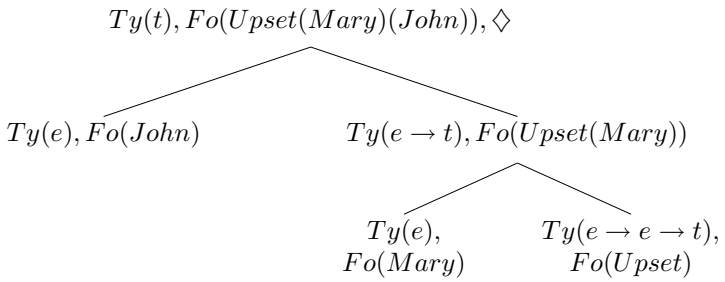

Figure 3: Completing a parse of *John upset Mary*

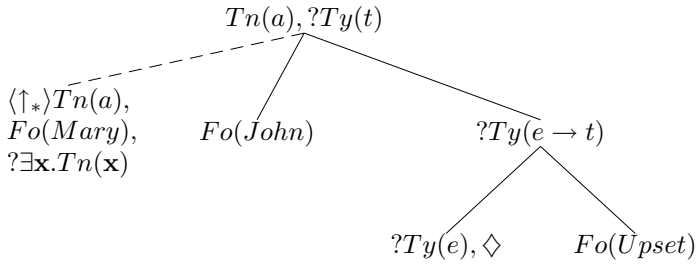

Figure 4: Parsing *Mary, John upset*

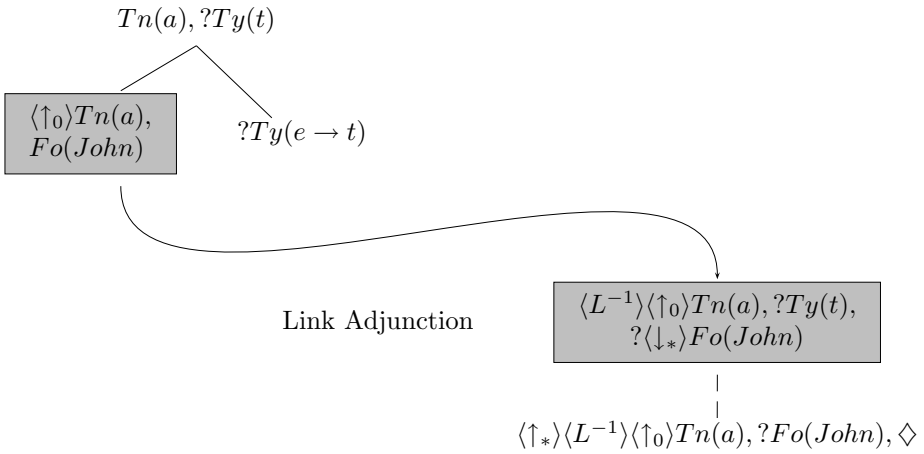

Figure 5: Building a *LINK* transition with *LINK Adjunction* plus **Adjunction*

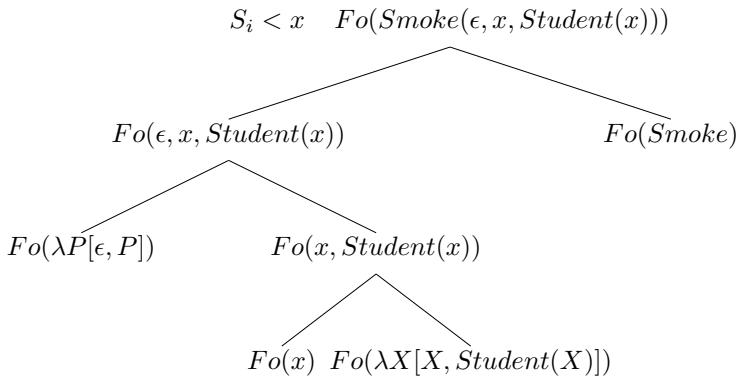

Figure 6: Parsing *A student smokes*

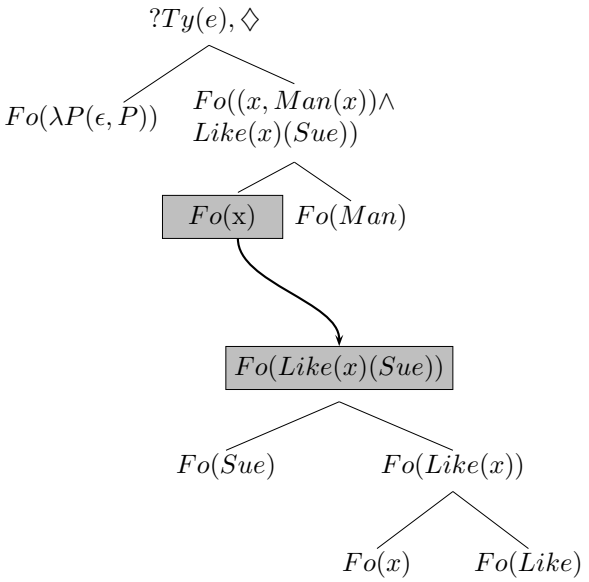

Figure 7: Completing interpretation for a restrictive relative

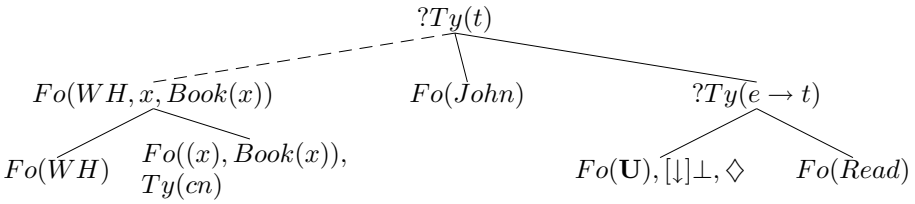

Figure 8: Resumptive pronouns in questions