Gender and the Interpretation of Pronouns in French

A view from Relevance Theory*

Paul Hedley

This paper is primarily intended to consider the role of grammatical gender on French pronouns in the process of their interpretation in utterance contexts. I will first discuss the theoretical context which underlies my general account of pronominal interpretation, the cognitive perspective of Relevance Theory, and sketch the bare bones of that account. I will then move on to a fuller discussion of grammatical gender on pronouns, its effect on interpretation and its representational status, using French as a test-bed, and taking psychological and psycholinguistic data into account. I conclude that in terms of their semantics, French pronouns carry primarily procedural meaning which has a fundamentally pragmatic effect on interpretation, but that gender is conceptual, and as such contributes in a rather different fashion both to the semantics of the pronoun, and to the process of its interpretation.

1. Concepts, Language and the Mind

At some level it seems incontrovertible that linguistic interpretation is a cognitive process, and as such a theory which intends to explain and account for it must have some cognitive component at the very least. For Relevance Theory, the cognitive perspective is fundamental, both in terms of its context, and its application. In the sphere of pronominals and their interpretation in particular, this cognitive view seems a fruitful path to follow: greater generalisations regarding usage and interpretation seem to be accessible, and we find ourselves in a position to ask different, deeper and more interesting questions, as well as receiving somewhat fuller answers. Similarly, when one tries to think about a linguistic feature such as gramatical gender, particularly if that feature has clear surface realisation in a particular language, it seems that to ignore the cognitive side of things is to avoid grasping the central concern of accounting for linguistic interpretation. For a speaker of a language that makes use of grammatical gender, the strategies used for interpretation are likely to make use of that feature to a greater or lesser extent. In considering pronouns, it may be that such information is categorical, directly constraining the process of reference resolution. Alternatively, speakers might use it as a guide, as evidence to weigh in order to sucessfully resolve the speaker's intended reference. Either way, such processes take place against the backdrop of a fundamentally cognitive system, and one in which hearers as well as speakers bear a significant responsibility for successful communication.

Relevance Theory takes the reasonably uncontroversial view of the mind as involving representations of some kind which are manipulated by the mental computational

* This paper draws on a paper I gave at the Durham Postgraduate Conference in July 2004 entitled 'Pronouns, Procedures and Relevance Theory'. Thanks to Stephanie Pourcel, an attendee of the talk, who drew my attention to the work of Boroditsky et al. (2003), which led to this extension and application of the approach put forward in that original paper.

apparatus. If such mental representations do take the form of concepts at some level, as many have argued,[1] our mental computational apparatus must include some system for the manipulation of those concepts. Applying such a picture to the processes of utterance interpretation results in a clear bipartite model, involving two fundamentally different types of process. As Wilson & Sperber (1993 :1) put it:

> a modular decoding phase is seen as providing input to a central inferential phase in which a linguistically encoded logical form is contextually enriched and used to construct a hypothesis about the speaker's informative intention.

Clearly, these two processes must be of a radically different nature, one based on the decoding of the linguistic signal into conceptual representations, and the other appealing to cognitive faculties of inference in order to reason towards a rational hypothesis concerning the intended meaning of that signal, and its import (relevance) to the individual(s) concerned. But what about the relative importance of these two processes in relation to each other? Many have argued, perhaps beginning with Grice (1967), that a significant amount of inferential processing is needed in order to interpret utterances, particularly in terms of the notion of implicature,[2] a position which is now widely accepted by linguists of widely differing outlook. However, one of the key advances of Relevance Theory is the demonstration that such inferential processing is not just a factor in the construction of implicatures, i.e. in the field of the implicit. Carston (2002) shows convincingly that linguistically encoded meaning underdetermines not only 'what is meant' by a speaker in a particular context (a point disputed by very few), but also 'what is said' or explicit,[3] a point of view which she terms "The Underdeterminacy Thesis" (2002: 19). In short, she argues, inferential processing is not confined to implicit content (as Grice had argued) but also has a significant bearing on explicit content.

Consider the semantics of a pronoun, say *il*, for a moment. Apart from some sort of gender information (which I will come to later) and some notion of the type of linguistic element such a lexical item may replace or stand for, we seem to be able to say little about what *il* might **mean** — a situation that is strikingly different from most nouns. All speakers seem able to do is indicate the referent in the particular usage at hand: '*il, ça veut dire Thierry*'.[4]

So, what do the truth conditions for a sentence like (1) look like?

---

[1] Aitchison (1994) to name but one.

[2] In brief, the implicit content of an utterance.

[3] There is substantial debate as to whether notions such as 'what is said' have any real place in theorising of this sort, particularly as such a concept has turned out to be so problematic to define. Grice clearly felt that some notion of 'what is said' was of central importance, and much of the Neo-Gricean camp follows this view (notably Levinson 1995, 2000; Bach 1994a, b also bases his approach on such a construct). Relevance theorists however do not accept that any notion of 'what is said' is needed to interpret utterances, nor that human beings actually have or use such a level of representation (see particularly Carston 2000).

[4] '"*il*" means Thierry'. It might also be interesting to investigate whether native speakers of a language like French whose grammar is comprehensively taught in schools are able to come up with a more satisfactory definition of a word like *il* than speakers of British English, for example, could for the corresponding English pronoun.

(1) *Il porte un pyjama.*

Clearly, this is true iff the person referred to by indexical *il* is indeed wearing pyjamas. It seems that it is the referent of the pronoun which enters into considerations of truth or falsity, and not the pronoun itself.[5] The proposition expressed thus contains the referent of the pronoun, raising the interesting question of whether any of the semantics of the pronoun itself (if it has any actual semantic content) survive, and show up on the surface. We shall return to this later on. For the moment however, the key consequence to be drawn from all this is that pronoun resolution must take place prior to any level of propositional evaluation. We are seeing inferential pragmatic processing here, the output of which enters into the proposition expressed.

What are speakers doing when they interpret a pronoun like this one? Intuitively they seem to be following a heuristic along the following lines:

> Accept the first candidate referent that yields an overall interpretation that is relevant.
>
> (Wilson & Matsui 1998: 188, paraphrasing Erku & Gundel 1987: 541-3)

However, we clearly want to be rather more explicit than this regarding the nature of the sub-processes involved in the assignment of such reference to a pronominal, and the contribution of the pronominal itself. In a system as complicated and sophisticated as that needed for utterance interpretation, it seems likely that any information accessible either from the content of the particular linguistic item, or from the context that is potentially relevant to interpretation will be made use of in the interpretative process. This is particularly clear in cases of reference resolution, where accompanying gesture and physical indication, and the immediate physical and linguistic contexts can all have a significant role in fixing referents for items like pronominals. Sperber & Wilson (1986/95) argue that cognitive processes in general are geared towards maximising relevance,[6] defined thus by Carston (2002: 44):

> relevance is a property of the inputs to cognitive processes (whether perceptual or higher-level conceptual); it is a positive function of cognitive effects and a negative function of the processing effort expended in deriving those effects.

In short, an utterance is relevant if it achieves a cognitive effect, and as it is the speaker's prerogative to make his utterance worth the attention of his intended addressee, the claim is that any ostensive stimulus (a paradigm case being linguistic communication) carries a presumption of its own relevance. This is known in the theory as the (Communicative) Principle of Relevance. The extension of this principle is the assumption that an utterance (and the linguistic items within it) will achieve relevance in certain ways which the speaker

---

[5] I will leave aside the whole debate regarding whether considerations of truth actually have a role to play in interpretation, and refer the reader to Wilson & Sperber (2000) for discussion.

[6] The First (or Cognitive) Principle of Relevance: 'human cognition *tends* to be organised so as to maximise relevance' (Sperber & Wilson 1986/95: 262).

might manifestly have foreseen, and which the hearer can safely assume were so predicted by the speaker. So, it seems reasonable to claim that a speaker's utterance of a pronoun will achieve relevance by uniquely picking out an accessible individual from the context, so satisfying the hearer's expectations of relevance. It also seems reasonable to suggest that features of the pronoun selected will aid the hearer in its resolution in a predicatable way (predictable indeed by the speaker himself). Grammatical gender is, of course, a paradigm case of such a feature which undoubtedly has an effect of some sort on the process of interpretation. It just remains to show how hearers interpret such items, satisfying their own expectations of relevance, what sort of 'sub-elements' exist in pronominal representation which aid the process of interpretation (if any), and the nature and effects of those elements.

2. Content and Character, Concepts and Procedures

In terms of the semantic side of the divide, Kaplan (1989) provides an interesting point of view on these issues: he distinguishes between the 'content' and the 'character' of lexical items. For pronominals, 'content' is the individual, and 'character' refers to a rule for identifying the content of such an expression in any given context. Wilson & Sperber (1993) reformulate this distinction, in terms of a opposition not within some concept of 'meaning', but between two different types of encoded meaning: conceptual meaning and procedural meaning. The crux of this argument is the pronoun *I* in the following sentence:

(2) I do not exist.

Kaplan argues that if *I* means 'the speaker of this utterance', such a sentence would be necessarily false — its truth conditions being that the sentence is true in any situation where the speaker of the sentence does not exist, a distinctly counterfactual circumstance. What we are seeing in this analysis is an instantiation of the direct encoding of the concept of 'the speaker'. If, however, *I* is treated as an instruction to the hearer to identify the referent of the pronoun by first identifying the speaker of the utterance, i.e. a procedure, we do not have such a problem: *I* would be used here to refer to an individual, and the sentence would most likely come out false, but would not be necessarily false.[7] As Wilson & Sperber claim, Kaplan's distinction is a striking forerunner of the conceptual/procedural one in Relevance Theory, and indicates that treating such pronominal elements as encoding procedures rather than concepts looks very much like the way we want to go, particularly as the reformulation accounts straightforwardly for the fact that pronominals do not appear in explicit propositional content: their meaning is computational, not representational, and so is not the sort of meaning that would or indeed should show up.[8]

[7] For Kaplan (1989: 523), 'they [indexicals] "determine" the content (the propositional constituent) for a particular occurrence of an indexical. But they are not "part" of the content (they constitute no part of the propositional consitutent)'.

[8] Kaplan, of course, still needs some sort of ad hoc mechanism to prevent the 'character' from showing up in truth-conditional content — perhaps a feature along the lines of Recanati's (1993) REF feature.

The theoretical notion of procedural meaning has been developed primarily by Diane Blakemore (1987 onwards) in relation to the account of the two-phase process of utterance interpretation discussed above: decoding and inference. In such a model based on a Fodorian representational–computational system and governed by principles of relevance, this idea of another kind of meaning seems to find a natural home. If the inferential phase plays as significant a role, as the evidence seems to suggest, it may not be immediately obvious to a hearer how the speaker intends his utterance to be interpreted, and which contextual assumptions should be used to derive what sorts of effects. Therefore, Blakemore argues, one might expect that languages and human users of those languages would have developed some means by which the hearer might be guided towards the intended context and cognitive effects in the first instance, and thus towards the speaker's intended meaning. On this formulation then, procedural expressions reduce the processing effort required on the part of the hearer by limiting the range of potential hypotheses that must be evaluated concerning the intended meaning, thus contributing to the overall relevance of utterances.

Wilson & Sperber (1993) cite one piece of direct evidence for this conception of linguistically procedural items in discourse. Under this general cognitive view of language and understanding, it seems plausible to claim that human thoughts are structured strings of concepts, and that human beings can typically be conscious of their thoughts. From this perspective, utterance interpretation can sensibly be categorised in terms of the formation and manipulation of conceptual representations, as put forward above. So, if we accept that the semantic distinction between conceptual and procedural meaning does indeed reflect a particular cognitive opposition (representation vs. computation), it follows that the meaning of a linguistic expression that encodes conceptual information should be mentally accessible, in the sense that a speaker should be able to bring it to mind. Native speakers of any particular language generally do have specific ideas about the meanings of lexical items in their language, or the concepts invoked by them. However, there are also computational processes that occur in the mind to which human beings do not seem to have such direct access: namely phonological computations, syntactic computations, or indeed the inferential computations used in the comprehension of utterances. Blakemore's account predicts that the ‘meanings’ of linguistic items which encode procedural information should be very difficult to ‘bring to consiousness’, and this is what we seem to find.

> If ‘now’ or ‘well’ encodes a proposition, why can it not be brought to consciousness? […] The procedural account suggests an answer […]. Conceptual representations can be brought to consciousness: procedures cannot. We have direct access neither to grammatical computations nor to the inferential computations used in comprehension.
>
> (Wilson & Sperber 1993: 16)

While this argument was originally proposed in terms of non-truth-conditional expressions like discourse connectives, it seems that items like pronouns have a very similar status in cognitive terms. In a particular context, speakers will quite happily provide a ‘definition’

of *il,* relating to a specific referent in the discourse or situation, use the verb *mean* to do so. While this is a clearly non-technical use of *mean*, we do not want pronominals to be infinitely linguistically ambiguous. As formal semanticists claim, there is a certain sense in which these expressions are 'variable(s)', though such a formulation tells us little either about the nature of the semantics of pronominals and the actual information encoded or represented, or the processes by which they are resolved. Heim & Kratzer (1998: 274-5 n.) for example have this to say:

> If pronouns are listed in the lexicon at all, they are listed there without an index and as semantically vacuous items.

They, and many others, talk in general terms about assigning to the pronoun 'the most salient individual that allows the hearer to make the most sense of the utterance', but say little about either the processes by which that is accomplished, or how they define salience.[9] It is precisely these underlying processes of reference assignment that concern me here, and what input to those processes linguistic features like gender might have. The idea of individuality turns out to be an important one too when thinking about concepts, as Powell (1998) argues. For him, the crucial question is whether or not we believe that a given concept is a 'representation of an individual' (Powell 1998: 13).[10] He draws a distinction between *individual concepts* (those which we believe correspond to an individual in the world), and *general concepts* (those which we do not believe to uniquely represent such an individual). On this schema, each individual concept will contain one or more general concepts, making up a 'dossier' of information.[11] To illustrate, a speaker might have an individual concept of 'my best friend', which would presumably be made up of a range of different sorts of information gained both by direct contact with that person, and otherwise (reports of other people etc), and contain general concepts such as 'friend' and 'best' (and, most likely, the general concept 'best friend'). This notion will turn out to be both intuitive and useful.

3. Procedures and Pronouns

So, what might these procedural meanings for pronouns look like? I am going to make some suggestions here in meta-linguistic terms, alongside some arguments and examples.

---

[9] Salience is, of course, an extremely complex issue, and one that has puzzled psychologists and linguists alike. Intuitively, the idea is clear, but its definition and integration into any framework has proved extremely problematic (see Ariel 1990, Gundel et al. 1993, Almor 1996). Breheny, for example, in a paper on anaphoric pronouns, argues that pragmatic approaches while being on the right track, 'are of questionable value unless a coherent story about salience or accessibility is provided' (forthcoming: 5). It is interesting to note that many of the critics of Relevance Theory take this lack of an overt formulation for salience as counting against the paradigm, while Deirdre Wilson (personal communication) has indicated that Relevance Theory was neither designed to provide such a formulation, nor does it have need of one. A combination of general relevance theoretic principles and the comprehension procedure should do the job for us anyway.

[10] This notion of individuality is clearly not confined to animate individuals, but is a 'catch-all' term intended to incorporate all items to which a speaker might refer, from plants and CD players to kangaroos and people.

[11] Recanati's conception distinguishes between 'egocentric' concepts ('temporary dossiers dominated by non-descriptive (perceptual) information', Powell 1998: 12), serving to register information gained in a certain way (i.e. primarily perceptually), and 'encyclopaedic concepts', seen as much more 'stable, long term dossiers of predominantly descriptive information' (*ibid.*).

Firstly, let us consider the pronoun *je*. Using the idea of 'individual concepts' just introduced, and taking Wilson & Sperber's initial (1993) formulation ('an instruction to the hearer to identify the referent of the pronoun by first identifying the speaker of the utterance') as a starting point, the encoded procedure for *je* might look something like 'find an individual concept of the speaker'. It would then be up to the pragmatic component to apply general principles of relevance and the application of the standard relevance theoretic comprehension procedure to arrive at the intended referent (most likely).[12] This procedure is characterised by Carston (2002: 143) as follows:

> (a) Consider interpretations (disambiguations, reference assignments, enrichments, contextual assumptions, etc.) in order of accessibility (i.e. follow a path of least effort in computing cognitive effects).
>
> (b) Stop when the expected level of relevance is reached.

Essentially, rather than being direct signals, utterances in general are seen as pieces of evidence about the speaker's meaning, which needs to be inferred by the hearer.

But how does such a conception help us deal with gendered pronouns such as English *he* and *she*, and French *il* and *elle*. Kaplan would probably see the semantics of English *he* as directly constraining reference to a male entity, but not appearing in the proposition expressed. However, Larson & Segal (1995: 214) argue that we should 'treat gender as semantically inert' in English, and consider it as providing only pragmatic guidance to the interpretation of such forms. They cite the situation of a speaker pointing to King's College London and uttering the following sentence:

(3) She is going to be closed over Christmas.

While the utterance is clearly anomalous in some way, it seems to be the case that the speaker has succeeded in fixing the referent of the pronoun through his overt gesticulation, and we do not seem to want to claim that the pronoun *she* could not possibly refer to King's. But, neither is the utterance straightforwardly false. Powell (1998) supports this conclusion using the case of Dr. James Barry, a prominent 19th century doctor, who was discovered after his death to have been a woman. Imagine the situation where Amy knows the truth, but her friend Ollie does not:

(4) Amy: When he was laid out after he died, they discovered that he was actually a woman.

This example shows that the natural way to refer to this individual is using a pronoun with some sort of 'maleness' feature, despite the fact that the referent of that pronoun is actually female. (Substituting *she* gives a rather odd result, making the fixing of the referent on the part of the hearer extremely difficult.) In short, it looks rather like such gender features on

[12] One of the other advantages of this sort of approach is that it can account for instances of misunderstanding, miscommunication, error and sloppiness through considerations of relevance and the application of the comprehension procedure.

English pronouns are not categorical, but rather are interpretative aids, or instructions to the hearer as to the best way to resolve the reference of the pronoun as intended by the speaker, i.e. they look distinctly procedural.[13] Adopting such an approach would seem to give us a very natural way of integrating these ideas into the theory, and of accounting for such problematic data.

So, the procedural formulation for English *he* might look something like this:

find an individual concept with the feature 'male'

Thus, presented with an utterance containing the pronoun *he*, Ollie can assume that Amy intended him to use some property of maleness in his search for the intended referent of the pronoun. He is also justified in narrowing the range of his search to individual concepts, whether pre-existing or formed ad hoc, containing the information '*x* is male' (i.e. the range of concepts he takes to be concepts of male individuals).[14] The oddness of the variant of the above example using *she* is also straightforwardly accounted for in this picture by standard relevance theoretic principles. Such a speaker would be presenting her hearer with the property of femaleness as an interpretative signal, so through the presumption of relevance, justifying a search of individual concepts of whatever sort containing the feature 'female' on the part of the hearer. However, Amy's intended referent for the pronoun contains the information '*x* is male' at this point in the mind of the hearer, resulting in gratuitous processing effort on their part, and so failure of optimal relevance, and probably of the process of pronoun resolution itself.

However, where does such a picture leave us with regard to languages that exhibit phenomena of grammatical gender? English, while exhibiting tripartite differentiation in pronominal gender, at least in the 3rd person singular, does not manifest adjectival gender agreement in any significant way, and inanimate nouns are unmarked for gender.[15] A language like French does not lend itself to the English type analysis very well at all. French gives a choice between only two pronouns (masculine ~ feminine), adjectival agreement for gender is mandatory, and all nouns are overtly gendered.[16]

3.1. Gender Systems

We seem to have need of some background here about gender systems in languages, and about the French gender system in particular, a good starting point being Hockett's (1958: 231) definition of gender:

> Genders are classes of nouns [recognised by being] reflected in the behaviour of associated words

[13] Features like animacy look rather like this too.

[14] Of course, such categories, and any properties that may be within them, are linguistic constructs.

[15] One might argue that they are simply neuter, but this debate does not affect the argument here.

[16] While this gendering does show up on articles, it would also be true to say that overt marking of gender normally only appears in agreement patterns on words other than the noun whose gender is in question. (Thanks to an anonymous reviewer for this observation.)

This underlines the fundamentally classificatory nature of gender systems in language, but, in its quest for generality, it fails to mention semantics. That there is some link, however slight, between grammatical and semantic gender is virtually irrefutable, given that male and female, masculine and feminine go hand in hand across a wide variety of the world's languages. The issue is how far this semantic side of the debate goes, which at one level is primarily a language-specific rather than a general issue. Many languages incorporate systems by which gender is assigned according to the meaning of a noun,[17] either in strict semantic terms, or according to principles of exclusion (where a positive assignment rule is accompanied by an elsewhere condition).[18] The fact that the English word *gender* derives from the Old French *gendre* meaning 'kind' or 'sort' (and from Latin GENVS before that) indicates that while gender classifications often have some similarity to real-world distinctions of sex, such an association, while often important, is rarely sufficient. Indeed, as Corbett (1991) states, French is generally regarded as possessing one of the most opaque gender systems in the world's languages, assignment not being generally predicatable on semantic or morphological grounds.[19] However, Tucker et al. (1977) propose a gender assigment system for French, primarily based on phonological criteria, and more specifically on word-final phones, or pairs of phones. They claim that laying aside the semantic and morphological assignment rules and treating them as exceptions (though exceptions that take precedence), the gender of about 85% of French nouns can be predicted. Interestingly, they also cite experimental data to support their hypothesis, which indicates that deaf children who learn to speak French do not learn to assign nouns to gender, for if the rules are phonological, such data is not available to those language learners (Tucker et al. 1977: 59). This particular area of debate pinpoints the central question of what grammatical gender is, and whether it is calculated by speakers on the hoof, or forms part of lexical representation in some way. It is this cognitive side that I want to look into: the representational status of gender. Whether work is conducted on pronominal forms and extended to other lexical items, or vice versa, the issue is the same.

3.2. The French Situation

Let us take example (3) above, about King's College, and consider its French equivalent:

(5) *Elle sera fermée pour Noël.*

---

[17] Dravidian languages such as Tamil, for example (Corbett 1991: 8).

[18] Diyari is like this, where nouns denoting females are feminine, and nouns in the semantic residue are masculine (Corbett 1991: 13).

[19] Of course, this is not to say that French does not have semantic or morphological assignment rules, characterised thus by Tucker et al. (1977):

Semantic assigment rules

1. Sex-differentiable nouns denoting males are masculine.
2. Sex-differentiable nouns denoting females are feminine.

Morphological assignment

1. Compound nouns formed from a verb plus some other element are masculine.

In such an utterance with the corresponding gesture as described above, the pronominal would most likely be unproblematically interpreted as referring to the university, (*l'université,* f.). However, if we substitute the masculine pronoun *il*, the hearer is left looking for a potential referent. The key difference here seems to be that *il* cannot refer to King's. The hearer could potentially make sense of the utterance by fixing the referent of the pronoun and its accomanying gesticulation as *le batiment/l'immeuble* ('building', m.), and by extension take the utterance describing the closure of the building to entail the closure of the university, but that is not the same situation as the claim we were considering earlier, that the pronoun is not excluded from referring to King's. Of course, this latter case is comparable to the English one in another way: use of the pronoun *il* would result in the forcing of gratuitous processing effort on the part of the hearer, and so fail the test of optimal relevance. However, we should not let that distract us from the conclusion to be drawn here, that we are looking at a fundamentally different situation: it seems that the pronoun cannot be interpreted as referring to King's, and that we have a clear opposition in need of explanation.

There is an increasing body of psychological and psycholinguistic work on the subject of gender that suggests that grammatical gender may actually affect meaning for speakers of a particular language, i.e. that mental representations may be influenced by abstract linguistic (and indeed language-specific) notions such as grammatical gender.[20] Efforts to assess this weaker, less deterministic view of the Sapir–Whorf hypothesis, (the question 'does language shape thought?'), have proved extremely problematic and produced wildly varying conclusions in different studies.[21] In terms of specific work on gender, some early studies in the field such as Jakobson (1966) and Sera et al. (1994), though perhaps methodologically flawed, seem to indicate that there may be some truth in the idea of gender having some semantic reflex. Days of the weeks were consistently personified according to their grammatical gender by Russian speakers (Jakobson 1966), and Spanish speakers similarly classified pictured objects according to whether the word for the object depicted was masculine or feminine (Sera et al. 1994). However, such results from monolingual tests have been questioned in terms of their cross-lingustic applicability, and neither is it clear that the effects of experience with a particular language on thought, can be extended to mental processes in any wider sense.

Boroditsky et al. (2003) aim to address this shortcoming by trying to observe 'a crosslinguistic difference on some more covert measure in a non-language-specific task' (2003: 67). They report a (2002) experiment conducted entirely in English, in which Spanish and German speakers were taught proper names for objects in English, and had their memory for these names tested. (The example they give is of an apple being called *Patrick*.) The objects chosen were selected so as to have opposing grammatical genders in the two languages, and the name assignment controlled so as to be either consistent or inconsistent with the grammatical gender of the object's name in the native language of the participant. As they predicted, memory for gender-consistent pairs was better than for

[20] See Lucy (1992) and Boroditsky (2001).

[21] Compare Slobin (1996) and Li & Gleitman (2002) for example.

inconsistent pairs for both sets of speakers: participants showed opposite memory biases. A control group of native English speakers however showed no such bias, performing as well as the Spanish and German speakers on gender consistent items, and significantly better on inconsistent ones. How are we to interpret such results? Are such interference phenomena indicative of something fundamental going on at the level of mental representation of gender? Boroditsky et al. (2003: 69) certainly think so:

> Since both groups performed the task in English, it appears that the semantic representation of gender (once it has been established) is not language specific. Objects do appear to have conceptual gender, and this gender is consistent with the grammatical gender assigned by language.

The conclusion I particularly want to pick up on here is the idea of conceptual gender. As I stated above, Relevance Theory takes the reasonably uncontroversial view of the mind as involving representations of some kind which are manipulated by the mental computational apparatus:[22] a position often couched in terms of concepts. As Aitchison (1994) states, whether or not we want to accept an abstract layer of concepts as separate from word meaning, it is generally assumed that words are linked to things in the world via concepts, though it is far from clear what might 'count' as such a concept. Most nouns clearly have some sort of concept associated with them, as is evidenced by the accessibility of definitions for such lexical items:[23]

faucon n.m. 1. *Oiseau rapace diurne, au bec court et crochu et aux ailes pointues.*[24]
(*Le Petit Robert* 1993 s.v. faucon)

Speakers of French obviously have a concept FAUCON,[25] but it is much less clear that they have a concept IL in the same way at all. The 'meaning' of *il* in a context depends on who *il* refers to. What I am interested in here is the process by which speakers work out this referent (clearly a cognitive process in some sense), and the role of linguistic and non-linguistic cues in that process.

As I argued above, we seem to want to ascribe procedural meaning to such pronominal elements, meaning which plays the role of aiding resolution by providing evidence to point the hearer towards the intended referent for that pronoun. However, the conclusion that seems to be presenting itself from the evidence discussed is that in languages where grammatical gender exists (and has a surface reflex etc), we seem to have a different situation from that described above for pronouns in English. Gender looks conceptual rather than procedural in these languages, a conclusion which would force us to revise our suggestion for what a French pronominal representation might look like. We have a distinctly more Kaplanian picture emerging here, where gender does seem categorical (unlike in English), with the resulting ungrammaticality rather than pragmatic

[22] An approach developed by Fodor (e.g. 1983) and others.

[23] See also the discussion above regarding mental accessibility and conceptual vs. procedural meaning.

[24] '**Hawk** n.m. **1.** Diurnal bird of prey with a short, hooked beak and pointed wings.'

[25] I adopt the general convention of using block capitals to refer to mental concepts rather than linguistic items.

infelicity in cases like that described above in examples (3) and (5). This is supported by the accessibility facts that I discussed earlier: while we can say little about the 'meaning' of *il*, the one thing we can be clear about is the gender (whether natural or grammatical) of the individual being referred to. The crucial difference between languages like French which have grammatical gender and those like English which do not, is that the linguistic interpretation system does not utilise the gender differentiation that does exist in the latter case precisely because it is not implemented across that language, and has no linguistic reflex. The formulation I would put forward for the procedural meaning of French pronominals is thus 'find an individual concept'. The categorical gender feature thus plays a significantly different role in the process of interpretation in languages exhibiting grammatical gender from those which do not. Thinking in developmental terms for a moment, a child acquiring English will note gender differences and the meaningful distinctions to which they apply in the world (with the odd inanimate exception), but has no reason (or need) to take this any further. By contrast, a child acquiring French has no a priori reason to believe that grammatical gender oppositions do not indicate meaningful distinctions (despite the fact that to a significant extent they do not), given the fact that they have an overt linguistic reflex, just as phenomena like number do. And indeed, this conception of native speakers acquiring some system of conceptual gender which fits their native language is also supported by the claims advanced by Tucker et al. (1977) that there is such a system.

So, how are we to square this circle regarding the notion of pronominal gender, if we accept the conclusions of Boroditsky et al. that gender has some conceptual content, and given the clear difference between pronouns and most other common nouns? Well, it seems to me that we need to look at the representation of such elements in the theory as presented above. One of Powell's claims (2002: 24) is that indexicals 'encode their status as individual concept communicators', as well as some sort of property that plays the pragmatic role described in interpretation. It would seem to me that the most sensible conclusion to draw from this given our current theory and the cross-linguistic data would be to consider pronouns not as empty lexical items, (as Heim & Kratzer 1998 claim), but as pro-concepts, carrying some information to direct the hearer to the intended referent (procedural meaning), but where that referent seems to amount to something akin to Kaplanian 'content'. In that sense, we might see them as the ultimate variable, dependant largely on context and pragmatic utterance interpretation strategies for their instantiation. Indeed, there is a pre-existing relevance theoretic construct which seems to have significant common ground with this idea — that of the concept schema (Carston 2002). The underlying issue that Carston is trying to address in proposing this approach is the nature of word meaning, the prevailing wisdom generally being that concepts encoded by lexical items provide a starting point for a pragmatic process, eventually resulting in a an interpreted concept which differs (to a greater or lesser extent) from the lexical concept.[26] She questions this idea of encoded concepts actually being fully-fledged concepts at all, proposing a picture whereby concept schemas act as pointers to a conceptual space, 'on the

[26] Processes of broadening or narrowing may occur, or something more radical in the case of metaphorical uses, all of which is overseen by the principle of relevance (see Carston 2002).

basis of which, on *every* occasion of their use, an actual concept […] is pragmatically inferred' (Carston 2002: 360).

Intuitively, this is precisely the sort of situation we have been discussing with pronominal elements, their context dependence, and the importance of pragmatic inference in their interpretation. The fact that Carston developed this approach in work on underdetermination, particularly in adjectivals, also provides us with a template for integration of conceptual material into such concept schemas. If the word *heureux* 'happy' encodes a particular concept HEUREUX which should provide communicative access to a wide range of other more specific concepts (relating to varying levels of *bonheur* 'happiness'), that lexically encoded concept HEUREUX will be more general and abstract than any individual use of the word, while providing the bedrock upon which processes of pragmatic enrichment can build in order to create a more specific concept that satisfies expectations of relevance in a particular context, and that can be intergrated into a hearer's representation of a speaker's thoughts. On this account, pronouns would operate in precisely the same way. For native speakers of those languages which exhibit grammatical gender, gender would appear as an abstract conceptual component of the underdetermined underlying representation, or concept schema, alongside the procedural meaning, making a complex semantics, while in those languages which do not exhibit such grammaticalised gender phenomena, pronominal semantics would look somewhat simpler. The role of gender in interpretation in the former cases would therefore be fundamentally different from that of the procedural meaning discussed above, but integrated alongside it in the semantic representation. Such an analysis would account not only for the fact that unlike in a language like English, the French pronominal in example (5) cannot be interpreted as referring to *l'université*, but also the psycholinguistic evidence presented by Boroditsky et al. (2003). In essence, what we are looking at is a linguistic constraint in such languages, and one that often has a surface realisation in things like agreement phenomena; in essesnce, a typological difference between languages like French and those like English. The fact that integration of pronominal elements into a much more general picture of the underlying semantics of lexical items is also straightforward, a claim that could not be made for many (if not most) theories of the semantics of pronominals, seems to be the icing on the cake. Pronouns no longer look like the oddball exceptions they have long been considered, but more like crucial threads in the fabric of an integrated linguistic system.

Conclusion

I have tried to show in this paper that if one takes a cognitive view on matters relating to the interpretation of pronominals, an interesting typological distinction emerges between those languages which have and utilise grammatical gender and those which do not. If pronominals are taken to encode an element of procedural meaning which serves as evidence to guide the hearer towards the speaker's intended referent, the resultant system seems to provide an intuitive and plausible account of the general process of pronominal interpretation, (a fundamentally pragmatic, inferential process), as well giving an insight into the nature of pronominal representation and its sub-parts. The idea of gender as a conceptual element of underlying representation in languages like French also integrates

into the idea of pronouns as concept schemas, needing substantial inferential processing for their resolution, and provides a much more uniform template for lexical representation that integrates pronouns into the larger linguistic system.

References

Aitchison, J. (1994). *Words in the Mind: an introduction to the mental lexicon*. Oxford: Blackwell.

Almor, A. (1996). 'A functional theory of NP anaphor processing', Poster presented at the 9th Annual CUNY Conference on Human Sentence Processing. New York.

———. (1999). 'Noun-phrase anaphora and focus: the informational load hypothesis', *Psychological Review* 106: 748-65.

Ariel, M. (1990). *Accessing Noun-Phrase Antecedents*. London, New York: Routledge.

Bach, K. (1994a). 'Semantic slack: what is said and more' in S. Tsohatzides (ed.), *Foundations of Speech Act Theory: philosophical and linguistic perspectives*. London and New York: Routledge, 267-91.

———. (1994b). 'Conversational implicature', *Mind and Language* 9: 124-62.

Blakemore, D. (1987). *Semantic Constraints on Relevance*. Oxford: Blackwell.

Boroditsky, L. (2001). 'Does language shape thought? Mandarin and English speakers' conceptions of time', *Cognitive Psychology* 43: 1-22.

———, L.A. Schmidt & W. Phillips. (2003). 'Sex, syntax and semantics' in D. Gentner & S. Goldin-Meadow (edd.), *Language in Mind*. Cambridge, Mass., London: MIT Press, 61-78.

Breheny, R. (forthcoming). 'Pragmatic analyses of anaphoric pronouns: do things look better in 2-D?' to appear in M. Garcia-Carpentero & J. Macia (edd.) book on Two-Dimensionalism. Oxford: OUP.

Carston, R. (2000). 'Explicature and semantics', *UCL Working Papers in Linguistics* 12: 1-44.

———. (2002). *Thoughts and Utterances: the pragmatics of explicit communication.* Oxford: Blackwell.

Corbett, G. (1991). *Gender*. Cambridge: CUP.

Erku, F. & J.K. Gundel. (1987). 'The pragmatics of indirect anaphors' in J. Verschueren & M. Bertucelli-Papi (edd.), *The Pragmatic Perspective: selected papers from the 1985 International Pragmatics Conference.* Amsterdam: John Benjamins, 533-46.

Fodor, J. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Grice, H.P. (1967). 'Logic and Conversation', William James Lectures. Harvard University.

Gundel, J.K., N. Hedberg & R. Zacharski. (1993). 'Cognitive status and the form of referring expressions in discourse', *Language* 69: 274-307.

Hall, A. (2004). 'The meaning of *but*: a procedural reanalysis', *UCL Working Papers in Linguistics* 16: 199-236.

Heim, I. & A. Kratzer. (1998). *Semantics in Generative Grammar*. Oxford: Blackwell.

Hockett, C.F. (1958). *A Course in Modern Linguistics*. New York: Macmillan.

Jakobson, R. (1966). 'On linguistic aspects of translation' in R.A. Brower (ed.), *On Translation*. New York: OUP, 232-9.

Kaplan, D. (1989). 'Demonstratives' in J. Almog, J. Perry & H. Wettstein (edd.), *Themes from Kaplan*. Oxford: OUP, 481-563.

Larson, R. & G. Segal. (1995). *Knowledge of Meaning: an introduction to semantic theory*. Cambridge MA: MIT Press.

Levinson, S. (1995). 'Three levels of meaning' in F. Palmer (ed.), *Grammar and Meaning*. Cambridge: CUP, 90-115.

———. (2000). *Presumptive Meanings: the theory of generalized conversational implicature*. Cambridge MA: MIT Press.

Lucy, J. (1992). *Grammatical Categories and Cognition: a case study of the linguistic relativity hypothesis.* Cambridge: CUP.

Li, P. & L. Gleitman. (2002). 'Turning the tables', *Cognition* 83: 265-94.

Phillips, W. & L. Boroditsky. (2003). 'Can quirks of grammar affect the way you think? Grammatical gender and object concepts', *Proceedings from the 25th Annual Meeting of the Cognitive Science Society.* Boston, Mass.

Powell, G. (1998). 'The deferred interpretation of indexicals and proper names', *UCL Working Papers in Linguistics* 10: 143-72.

———. (2002). Chapter 4 of PhD thesis - 'Reference and Relevance: Indexicals' Unpublished manuscript, University College London.

Recanati, F. (1993). *Direct Reference: from language to thought.* Oxford: Blackwell.

Sera, M., C. Berge & J. del Castillo. (1994). 'Grammatical and conceptual forces in the attribution of gender by English and Spanish speakers', *Cognitive Development* 9: 261-92.

Slobin, D. (1996). 'From "thought and language" to "thinking for speaking"' in J.J. Gumperz & S.C. Levinson (edd.), *Rethinking Linguistic Relativity.* Cambridge: CUP, 70-96.

Sperber, D. & D. Wilson. (1986/1995). *Relevance: communication and cognition*. Oxford: Blackwell.

———. (1997) 'The mapping between the mental and the public lexicon', *UCL Working Papers in Linguistics* 9: 107-25.

Tucker, G.R., W.E. Lambert & A.A. Rigault. (1977). *The French Speaker's Skill with Grammatical Gender: an example of rule-governed behaviour*. The Hague: Mouton.

Wilson, D. & T. Matsui. (1998). 'Recent approaches to bridging: truth coherence, relevance', *UCL Working Papers in Linguistics* 10: 173-200.

Wilson, D. & D. Sperber. (1993). 'Linguistic form and relevance', *Lingua* 90: 1-25.

———. (2000). 'Truthfulness and relevance', *UCL Working Papers in Linguistics* 12: 215-54.