Abstract

The paper offers an analysis of pied piping within the theoretical framework of Word Grammar. This framework combines cognitive linguistics with dependency grammar, so it assumes that the full power of domain-general cognition is available for syntax and that phrasal categories are not needed. Instead, syntactic structure is a network of relations between individual words which combines dependencies with the ‘landmark’ relations that determine word order. To handle the special characteristics of pied piping, the analysis includes a single special relation, ‘pipee’, which links the piper (the wh-type word) to the word which replaces it in the landmark structure. The analysis is compared with previous analyses and extended to accommodate both the pied piping with inversion found in Meso-American languages, and the boundary markers found in other languages. The main things missing from the analysis are explanations for the restrictions on pied piping in free relatives, and the social restrictions on ‘quiz interrogatives’.

1. Introduction

As a serious challenge for any theory of syntactic structure, ‘pied piping’ has enjoyed a great deal of attention from theory-builders since it was first identified and named (Ross 1967), so we know how it works both in English and in a number of other languages. A typical example is (1), where which is said to ‘pied pipe’ the preposition phrase in which into a position which is normal for which but not for a preposition phrase.

(1) This is the book in which I found it.

Even the ‘pied piping’ terminology is widely agreed, though not universal. On the other hand, in spite of widespread agreement on the facts, our theories of syntax are as divided on pied piping as on other matters, and the dominant HPSG theory based on feature percolation (Pollard and Sag 1994) has been challenged by both Optimality and Minimalist alternatives (Broadwell 1999, Cable 2012). The main differences between the alternatives lies in their theoretical assumptions about the apparatus needed for syntax, and in the special apparatus that they introduce for pied piping, rather than in their ability to handle the data. But what unites them is the phrase-structure apparatus that they all assume for linguistic structure.

This paper assumes a very different theoretical foundation, Word Grammar, and shows how it can be applied to pied piping. Word Grammar (ANON) is a version of cognitive linguistics, so its most controversial theoretical assumption is that we use the same mental apparatus for learning and processing language structure as for ordinary everyday cognition. Pied piping is a particularly good testing ground for this theory because it is such a peculiarly ‘linguistic’ pattern, apparently as far removed from everyday cognition as syntax can be.

Apart from the cognitive orientation, however Word Grammar also assumes dependency structure, which is just as controversial within cognitive linguistics as it is in mainstream American and European linguistics. For instance, in (1) this theory does not recognise in

1 [Acknowledgements]
2 An earlier theory of pied piping, which the present one supersedes, is presented in (ANON).
which as a preposition phrase; instead, it just recognises a dependency from in to which, and another from found to in. Once again, the challenge of pied piping is special because the apparatus of phrase structure plays such an important part in other analyses.

The paper starts by building on the insights of earlier work (section 2), and then introduces the theoretical apparatus needed for pied piping (section 3). The main body of the paper is a fairly full discussion of English pied piping (sections 4 to 6), but the survey of how pied piping relates to general cognition also considers data from other languages (section 7).

2. Insights from other approaches

Most discussions of pied piping in English limit themselves to the pattern found with wh-words – i.e. relative and interrogative words – and this paper has the same focus. What these examples have in common is that a phrase containing a wh-word as a dependent, rather than as head, occurs in a position where the normal rules allow a wh-word. For instance, take the following examples of interrogative clauses ((Huddleston and Pullum 2002: 910):

(2) In which drawer do you keep the bank statements?
(3) What size shoes do you take?
(4) How big a hole did it make?

Challenges

In (2), the interrogative determiner which stands, as expected, at the start of the clause, but so does the whole of the underlined phrase in which drawer. This is not expected, and the only reason for this phrase being in this position is that it contains the word which; but crucially, which is not the head of the phrase, so exceptionally the whole phrase takes its position from a non-head word inside it. The same is true of the other two examples.

It could be objected that preposition phrases can be freely topicalized, with or without a wh-pronoun:

(5) In this drawer I keep the bank statements.
But this explanation clearly fails for questions, and fails even more spectacularly for relative clauses such as (6).

(6) This is the drawer in which I keep the bank statements.
Here pied-piping is the only possible explanation for the position of in which, there being no grammatical alternative without a wh-pronoun:

(7) *This is the drawer in it I keep the bank statements.

The main theoretical challenge of pied piping, therefore, is the mismatch between linear position and dominance: in terms of dominance, the head of in which drawer is the preposition in; but it is the wh-word which, not in, that determines the linear position of the whole phrase. This conflicts with the very general principle that a phrase’s position is determined by its head word. For instance, the phrases very big and about linguistics can both modify the noun book, but their position before or after book is determined by their head’s word class, giving very big book about linguistics, but not *about linguistics book very big.
Another theoretical challenge is to explain why pied piping applies to words such as our wh-words, and only to these. The obvious peculiarity of these words is their dual linkage. For example, in *the book which I bought*, the relative pronoun *which* has ‘internal’ links within the relative clause (as object of *bought*), but also has ‘external’ links to the antecedent noun *book*, its antecedent. The same is clearly true of interrogative wh-words in subordinate clauses, as in *I wonder what happened*, where *what* is internally linked to *happened* but also externally linked to *wonder*. Arguably, even main-clause interrogatives have a similar combination of internal and external links, though the external link is purely semantic. The second challenge, therefore, is to explain how the dual internal/external linkage of wh-type words relates to pied piping.

**Theory**

Until recently the most popular explanation for the mismatch between linear position and dominance has been ‘feature percolation’, in which the ‘WH’ feature (actually, either QUE or REL) of the wh-word passes up to the node representing the entire phrase. This is an attractively simple explanation for the word-order mismatch, especially given the important role played by feature-passing in the HPSG treatment of other phenomena. On the other hand, this analysis also has major costs, not least of which is the fact that feature percolation from a non-head to its containing phrase exists solely for the sake of pied piping.

Another cost is in the semantics of the percolated feature. REL and QUE are defined not only by their position, but also by their meaning; so REL provides an index for the relative pronoun which binds it to an antecedent noun (Pollard and Sag 1994:208-214). But if this index is percolated up to the pied-piped phrase, the semantics fails. For example, in (6) the drawer is the antecedent of *which*, and not of *in which*. This is even clearer in an example like (8):

(8) This is the drawer *whose key* I lost.

Here the drawer is the antecedent of *whose*, but the pied-piped phrase is *whose key*, referring to the key, not the drawer. Such examples show that the semantics is out of step with the word order: even though *whose* shares its position with *key*, it does not share its semantics. Consequently, pied-piping must find some way to separate the semantic consequences of wh-words from their consequences for word order.

Similar objections apply to a more recent Minimalist analysis which includes extra syntactic structure containing an invisible operator (Cable 2010) which affects word order but has no effect on meaning. The best evidence for this operator comes from a minority of languages in which it is visible, so the analysis depends heavily on controversial assumptions about universal grammar. On the other hand, the analysis also points to the possibility of an abstract structural peculiarity of pied-piped structures which triggers the special word order (as in the analysis to be developed below).

Another recent analysis in the Chomskyan tradition presents pied-piping as strictly a matter of word order, without any other implications for meaning or syntactic structure (Heck 2009). For this analyses, the challenge is to find a way to tweak the structure and the usual movement rules so that they move the entire phrase along with the wh-phrase that it contains. The important insight here is that pied-piping is merely a matter of word order. The strict distinction between word order and the rest of phrase structure is of course reminiscent of the
ID/LP split in the GPSG literature, in which ‘immediate dominance’ was separated from ‘linear precedence’ (Gazdar and Pullum 1981).

Alongside the formal analyses in the PSG or Minimalist traditions, there is an important contribution that applies Optimality Theory (Broadwell 1999; Broadwell 2006). We consider the relevant data in section 7, but the important insight to be noted here is the idea that syntactic patterns, including pied piping, are due to independent but interacting constraints, such as the requirement that wh-words should stand at the start of the clause. It is a short step from this kind of analysis to one which links these constraints to functional benefits for speakers and hearers, and raises questions about why pied piping exists in so many languages.

**Description**

Meanwhile, more descriptive studies have produced statistical evidence to support the impression that pied piping is more academic or formal than its alternative, preposition stranding (Hoffman 2008). As we might expect, private conversations and correspondence strongly favour stranding over pied piping, in contrast with more formal varieties where the balance is reversed. This comparison is possible in English because at least with preposition stranding we have a free choice between the two patterns, as in (9) and (10):

(9) This is the drawer in which I keep the bank statements.
(10) This is the drawer which I keep the bank statements in.

In languages which forbid stranding, of course, there is no choice so it is much harder to assess formality. But why should pied piping be more formal? Is it simply a matter of a historical link between the Romance pied piping and formal learning, or is stranding inherently easier and more appropriate for casual conversation; and if so, is there any counteracting benefit of pied piping? These questions bring us back to the functional motivation for pied piping.

One last descriptive fact to emerge from earlier work is the difference, mentioned earlier, between different kinds of ‘wh’ construction. Preposition pied piping is much freeer in relative clauses than in interrogative clauses. The following examples (from Huddleston and Pullum 2002:1040) illustrate the freedom found in relative clauses, and especially in non-restrictive relative clauses:

(11) the curtain behind which Kim was hiding
(12) She’s just sat her final exam, the result of which we expect next week.
(13) They will be involved in several projects, one of the most important of which will be to find ways to use the new superconductor in chips that can provide the brains of a new generation of supercomputers.
(14) The many varieties of mammalian skin secretions perform a wide range of functions, prominent among which is sexual attraction.
(15) I became disturbed by a ‘higher criticism’ of the Bible, to refute which I felt the need of a better knowledge of Hebrew and archaeology.
(16) They take a rigorous examination, passing which confers on the student a virtual guarantee of a place at the university.

In free relatives, however, preposition pied piping is hardly possible at all:

(17) I enjoyed what I filled my time with.
(18) *I enjoyed with what I filled my time.
(19) *I amused myself with what I filled my time.

These distinctions are normally described in terms of different clause types (interrogative clause, etc), but each clause type allows a different range of wh-pronouns, so the distinctions could be made equally well in terms of word classes: interrogative pronoun, relative pronoun, and so on through finer subdivisions.

To summarise this brief survey, earlier analyses have contributed some important insights to our understanding of pied piping:

- Pied piping is primarily a matter of word order, completely separated from other structural issues such as classification and meaning.
- But there must be some structural connection between the wh-word (the ‘piper’) and the node representing the entire pied-piped phrase (the ‘pipee’), whether this is described in terms of a percolated feature or an invisible operator or by some other means.
- The structural analysis may enable a functional explanation for the existence of pied piping, and also for the choice, in preposition phrases, between pied piping and stranding.
- The analysis must differentiate different types of wh constructions (or different kinds of wh-word), as well as providing a single unifying explanation for their similarities.

3. Towards a cognitive analysis

Suppose our aim is to produce a cognitively plausible account of pied piping, with the minimum of special cognitive apparatus. Clearly the theoretical minimum is none at all, so our aim should be to explain pied piping without any special apparatus, and only to import extra assumptions or apparatus when we’re sure that ordinary cognition does not offer what we need. One of the claims of a cognitive theory such as Word Grammar is that, in fact, no extra apparatus is ever needed; but this claim clearly requires an audit of relevant apparatus which is readily available in ordinary cognition. Three items seem particularly relevant to pied piping, so this section will introduce them, and later sections will explain how they apply to pied piping. They are listed below:

- Networks
- Default inheritance
- Landmarks

Networks

The idea that knowledge forms a network is uncontroversial. The clearest experimental evidence comes from priming experiments, which show that activation (manifested in mental accessibility) spills over to ‘neighbouring’ concepts in a way that can only be explained if the concepts concerned are directly linked in a network. Everyday experience of associations supports the idea that concepts are organised in a network. Consequently, cognitive theories of language often present language as a network of objects such as constructions or attribute-value matrices.

However, the ‘network notion’ takes this idea even further, by explaining all the properties of one concept as nothing but its links to other concepts (Reisberg 2007). In this view,
concepts are atoms without any inherent structure or content, rather than complex ‘boxes’ of information (such as constructions or lexical entries). Without this notion we might think of a concept such as ‘dog’ as a complex collection of information, consisting of special features such as size, function, and so on; and the network idea would be relevant to the relations between these complex boxes. In contrast, the network notion merges the concept’s ‘internal’ features with its external relations to other concepts, so that the atom labelled ‘dog’ is defined by its relation to other atoms such as ‘bark’, ‘mammal’, ‘pet’, ‘kennel’, ‘size’ and so on, and is also part of the definition of each of these other atoms. This merger has the great advantage of removing the need for a distinct set of ‘features’ (which threaten an infinite regress because they too need some kind of definition).

Word Grammar (following Stratificational and Neurocognitive Linguistics - Lamb 1966; Lamb 1998) accepts the network notion, so each unit (e.g. word) is shown simply as a node whose content is provided by its relations to other units in the network. Figure 1 illustrates this idea with a simplified view of the language network round *likes*, the third-person singular of the verb LIKE_v which also shows the homonymy of this verb with the preposition LIKE_p. This diagram also introduces two notational conventions of Word Grammar: the small triangle means ‘isa’ (so LIKE,3sg isa both LIKE_v and 3sg); and the ellipses indicate relational concepts in contrast with the non-relational entities in rectangles. In words, *{like}* is the realization of both LIKE_v and LIKE_p, *{likes}* is the realization of LIKE,3sg, and *x* is the object of LIKE_v.

![Figure 1: A network for likes](image)

The theoretical claim behind this network notation is that, because language is a kind of knowledge and knowledge forms a network, language too is just a network, and nothing else. There are no rules, no processes, no principles, so it’s a single integrated network all the way down. In short, Word Grammar is a constraint-based theory of language structure like HPSG, LFG and others, but with a single subnetwork (within the super-network of cognition).
which shows both the interrelations among ‘linguistic items’ and their internal structures. And of course one aim of this article is to demonstrate that such an impoverished formal apparatus can accommodate the complexities of pied piping.

Default inheritance

Like many other cognitive theories, Word Grammar assumes that the underlying logic of thought is default inheritance – a process whereby stored patterns are inherited by tokens of experience, and which allows defaults to be overridden by exceptions. For instance, if we hear the form {likes}, we inherit the fact that it is the realization of LIKE,3sg, which in turn isa 3sg and LIKE_v, so it has an object, and so on. Word Grammar avoids the familiar problems of non-monotonicity by a simple assumption: that inheritance only happens when new nodes are first created during processing, and it works strictly bottom-up, so the first value inherited is always the most specific and can never be overridden by an even more specific value (ANON). Thus inheritance is part of the process of enriching newly created nodes for utterance tokens, rather than a freely available process for enriching the established network.

One immediate consequence of this assumption, when applied to language, is that because the structure of an utterance is inherited from a network, it too must be a network (albeit, of course, a mere fragment of the total stored network), with the full formal freedom of a network. In particular, individuals nodes can freely relate to one another (as they do outside language, for instance, in a social network); so there is no formal reason for assuming the very limited geometry of phrase structure which only recognises part-whole relations. Consider the two consequences of applying this principle to a sentence such as (20):

(20) Small children were playing outside.

First, there is no need for an NP node for small children to show that they belong together semantically, because general cognition allows individual items (such as people) to be related directly to one another. Applied to this sentence, this means that small can be related directly to children by a dependency which shows that small modifies the meaning of children. Moreover, since general cognition allows us to create concepts as needed, this modification creates a distinct node for the modified children, as opposed to the children that inherits directly from the grammar. So although we seem to have just one node for children, we actually have two: the token children as first inherited, meaning ‘children’, and the ‘sub-token’ children’, resulting from the modification by small and meaning ‘small children’. This sub-token is very similar to an NP node, in that it carries the meaning of the two words combined and reflects a direct dependency relation between them, but crucially children' has exactly the same syntactic classification as children, rather than a distinct ‘phrasal’ classification. In short, dependency structure provides all the structure needed, without invoking phrase nodes as well. In concrete terms, the subject of the verb is not the phrase small children, but the sub-token children'.

And second, children' is not only the subject of were, but (thanks to raising) also of playing. This double dependency is like the HPSG notion of re-entrance, where a single item is the value for two attributes. But once again, there are reasons for thinking that it is not a single item that is involved in two dependencies, but a different one in each. This is because each dependency provides a different set of inheritable properties, including word order: as subject of were, children' should stand just before were, but as subject of playing, it should be
next to playing. To explain why this conflict has to be resolved in favour of were (and more generally, in favour of the higher verb), we can once again invoke two distinct nodes, the sub-token children' and the sub-sub-token children'', with an isa link between them:

- children'' is subject of were, and isa children'
- children' is subject of playing.

The properties of children'' override those of children', just as the latter’s override those of children. Similar arguments apply to the other words, so the full analysis for this sentence is as shown at the top of Figure 2, where the lower diagram is a simplified version which ignores extra word tokens.

Figure 2: The syntax of modification and raising.

In prose, Figure 2 says that small modifies children, turning it into a distinct token concept children' with the meaning ‘small children’ and acting as the subject of playing''.

Similarly, outside turns playing into playing', which children' turns into playing'' so that it means ‘children playing outside’. As for were, it combines with playing'' to produce were', which in turn is modified into were" by children'' to carry the meaning of the entire sentence, ‘Small children were playing outside’. The similarities to Categorial Grammar are obvious, but instead of composing words to produce phrases, we compose them to produce extra word tokens. The difference is subtle, but important because it involves default inheritance rather than composition, and therefore allows overriding.

Landmarks

Like networks and default inheritance, landmarks are an obvious part of ordinary cognition. They play an important part in the semantics of Cognitive Grammar (Langacker 2007), where location is described in terms of a landmark and a trajector (the thing located); for instance, the book on the table locates the book (the trajector) in relation to the table (the landmark).
However Word Grammar extends their use into syntax, where they are the main mechanism for handling word order. After all, the mechanism for expressing the meaning of *Monday is before Tuesday* could well be relevant too for describing the order of words in *Monday, Tuesday, Wednesday*.

Landmarks are an important part of any network model for the simple reason that a network has no left-right dimension. Linguistics has always rested heavily on our writing system which uses this dimension for linear order, but the fact is that linear order is a relation, just like the other relations of syntax. A diagram such as Figure 2 implies this relation, but does not show it explicitly. We know from experimental psychology that the linear order of a string of words (or other objects) is separate from their individual identities, so we can treat a word’s position as one of its properties, along with its class membership, its realization, its dependents, its meaning, and so on.

Linear order in a network seems to require two different kinds of relation. At the most basic level, we need a ‘next’ relation between adjacent words in a string: in the string *Monday, Tuesday, Wednesday*, the word *Tuesday* is the ‘next’ of *Monday*, and in *Small children were playing outside*, each word is the ‘next’ of the word before it. This relation is certainly needed in morphology and phonology in order to explain, for example, that the word *AN* is realized as {a} or {an} according to whether its ‘next’ begins in a consonant or a vowel, as in *a pear* and *an apple*.

But alongside this low-level relation we also need the more abstract ‘landmark’ relation in order to allow general ordering constraints, because these rarely apply just to adjacent items. Instead, we need generalizations such as “A verb’s subject stands before it” or even the very general “A word’s dependent stands near to it”. In each of these generalizations, one word is treated as the landmark, the fixed point, and the other as the trajector, the item that takes its position from the landmark; so the relation is asymmetrical, as in a syntactic dependency. Indeed, many dependencies are defined in part by a landmark relation in which the ‘parent’ word (the word on which it depends) is the landmark and the dependent is the trajector.

The decision to treat word order in terms of ‘next’ and ‘landmark’ has important theoretical consequences. For one thing, it means that the syntactic structure of a sentence can be a true network which does not rely on the left-right dimension – though, of course, for purely practical reasons of presentation the networks are easier to read if we can read them from left to right. More importantly, if a word’s position in linear order is a property of that word, the position can be inherited by the usual mechanism of default inheritance. By default, a word’s parent is also its landmark, so landmark relations can be inherited from dependency relations; but like any other default, this can be overridden – exactly what we need in order to explain pied piping.

A further consequence of recognising ‘landmark’ as a relation is that relations are a type of concept, so they can be subclassified; so the general ‘landmark’ relation can be subdivided into ‘before’ and ‘after’. For example, although a verb is the landmark for all its dependents, it is the ‘before’ for its subject and the ‘after’ for its object; e.g. in *John loves Mary, loves* is both the landmark and the ‘before’ of *John*. 
It will be helpful to have a notation for syntax which allows us to combine landmark relations with dependency relations. There is no established convention for this in Word Grammar so we now introduce the convention of writing all dependencies above the words and all landmark relations below them. The convention is illustrated (with simplified dependencies) in Figure 3.

![Figure 3: Landmarks combined with dependencies](image)

Notice that in general, each arrow above the words in Figure 3 is paired with an arrow below the words that points in the opposite direction; this is because the former point towards the dependent, but landmark arrows point towards the dependent’s parent. But, exceptionally, the dependency arrow pointing from *playing* to *children* (or more accurately, from *playing*’ to *children*’) has no paired landmark relation because, by default inheritance, the landmark link from *children*” to *were*” overrides it. This possibility of overriding default landmarks will provide the mechanism needed for pied piping.

The distinction between landmarks and dependencies is important because it is the landmarks, not the dependencies, that define the phrases that are continuous (or, in dependency terminology, ‘projective’). In Figure 3, *small children playing outside* is a dependency-based phrase because both *children* and *outside* depend on *playing*; but it is not a landmark-based phrase because it is interrupted by *were*. In formal terms, landmark arrows never intersect each other, in contrast with dependencies, which often do (as we shall see in later diagrams). This ban on intersecting landmarks has an easy non-linguistic explanation: the same is true of landmarks in general cognition. According to the Best Landmark Principle (ANON), the best landmark for an object is the nearest object which is more prominent. Given a row of objects A B C, if the best landmark for A is B, then A cannot be the best landmark for C; after all, B is both more prominent than A, and closer to C. For example, if my socks are by the chair and my shirt is on the chair, we would not use the shirt as the landmark for saying where the socks were. And given a longer row of objects A B C D, there is no situation in which landmark links could intersect because that would mean that some object was preferring a more distant landmark to a nearer one.

4. English preposition pied piping
We are now ready to explain the pied piping found in examples like (1), *This is the book in which I found it.*
Mutual dependency

On the way to this explanation we must review the Word Grammar analysis of wh-constructions such as wh-interrogatives and relative clauses, so we start with a very simple example:

(21) What happened?
The dependency structure is unexpectedly complicated. One thing is clear: what is the subject of happened, so what depends on happened. However the converse is almost equally clear: that we expect what to introduce a clause, or (in dependency terminology) its valency includes a finite verb. This means that happened must be the complement of what – in short, this is a clear case of mutual dependency, which modern grammatical theories, whether based on phrase structure or on dependency structure, assume is impossible, but which is accepted in Word Grammar because it clearly exists outside language (ANON).

However, there may be an intermediate position between the rejection of mutual dependency and its acceptance. This is because of the extra flexibility provided by sub-tokens. Given this assumption, what is the subject of happened', but when the latter depends on what it creates a sub-token what', so it is what' rather than what itself that happened' depends on. If this is right, we seem to need a structure like Figure 4. Notice that the landmark links (below the words) are also mutual, and both require the same order, with what before happened. Strictly speaking, these links apply to the sub-tokens, but our notation prevents us from showing this.

![Figure 4: Mutual dependency](image)

Extraction

Next we consider extraction, where the wh-word is moved out of its default position by being attached as a preceding ‘extractee’ to the finite verb, as in (22):

(22) What did you find?
In the Word Grammar analysis, what depends (as object) on find, but it is also the ‘extractee’ of did; each of these relations defines a different token of what: the object what and the extractee what'. The conflict between their properties, including their expected positional properties, is resolved by the fact that what’isa what in the inheritance hierarchy. The structure, ignoring all other sub-tokens, is shown in Figure 5. The main point of this example is that the default position of what after find is overridden by that of what' before did, giving What did ... rather than ... find what. In all other respects, however, the dependency of what on find is exactly the same as it would have been in a non-extracted example such as I found.
it. (We might also note that *you* takes its position from *did* rather than *find*, although thanks to raising it depends on both; the explanation would involve the same overriding mechanism, but isn’t directly relevant here.)

![Diagram of dependency structure]

*Figure 5: A simple wh extraction*

**Pied piping**

Returning to pied piping, we start with a pied-piped interrogative, (23):

(23) **With whom** do you live?

This is just like the previous example except that it is *with* rather than *whom* that takes its position from *did*, thereby overriding its default position after *live* (as in *I live with Jane*). And yet it is clearly *whom*, rather than *with whom*, that is extracted: after all, a wh-question needs to start with an interrogative pronoun, not a preposition phrase, and semantically the question is about the referent of *whom* (rather than of *with whom*), so a good answer would be simply *Jane*. In short, the dependency structure is, in fact, exactly the same as in the stranded equivalent (24):

(24) **Who** do you live **with**?

As noted in section 2, pied piping is strictly a matter of word order, and nothing else. When reading the two dependency structures shown in Figure 6 it is important to bear in mind that, in a network, these two structures are topologically identical and indistinguishable.
The question, therefore, is how to explain the different positions of \textit{with}. What seems to be happening in the pied-piped example is that \textit{with} replaces \textit{whom} in the landmark structure; so \textit{with}, instead of \textit{whom}, takes \textit{do} as its landmark, and \textit{with}, instead of \textit{whom}, also acts as landmark for \textit{do}. More generally, this is possible because \textit{with} is the parent of \textit{whom}, but of course this is equally true whether or not pied piping applies, so we need an extra trigger for pied piping which is not part of the regular dependency structure; and since it is \textit{whom} that permits pied piping and controls it, the trigger must be a special relation between \textit{whom} and \textit{with}. The obvious name for this relation is the term introduced earlier, \textbf{`pipee’} (meaning ‘what is pied-piped’); so \textit{with} is the pipee of \textit{whom}. Although this is neither a dependency or a landmark, we can freely add it to the analysis because the latter is just a small part of the total cognitive network which already includes many different kinds of relations. The ‘pipee’ relation is not the only such additional relation; as Rosta argues (Rosta 2005), a more general relation which he calls ‘surrogate’ may be involved not only in pied piping but also in a range of constructions such as the optional \textit{that} which introduces subordinate clauses and degree modifiers such as \textit{so} in \textit{so big a house}.

With this addition, then, the complete analyses for our two interrogative examples are shown in Figure 7\footnote{Error! Reference source not found.}, where the crucial relations in the pied-piped structure are shown as dotted lines.
We now have the full apparatus needed for pied piping:

- the ordinary dependency structure, in which dependency relations are properties of word tokens.
- the landmark structure which treats linear order as a property of word tokens.
- the isa relations between tokens and sub-tokens which resolve conflicts, including conflicts of linear order.
- the special ‘pipee’ relation from a wh-word to the word which replaces it in the landmark structure.

This analysis extends easily to other kinds of interrogative such as our original example, (2): *In which drawer do you keep the bank statements?* The only relevant difference is that the wh-word is the determiner *which*, rather than the pronoun. As it happens, Word Grammar treats determiners as heads (ANON) so *which* is directly linked to the preposition, just as in the *with whom* example; but we shall consider examples below in which this is not so.

**Pied piping in subordinate interrogatives**

Subordinate interrogatives such as (25) are slightly more complicated because the interrogative pronoun has an extra dependency.

(25) I wonder what she wants.

In this case, *what*, or rather its sub-token *what'*, must depend on *wonder* as well as on *wants*, because the latter needs an interrogative word such as *what*. Moreover, *wonder* must be the landmark of *what'*, because the latter, and the clause it introduces, takes its place among the higher verb’s dependents, as in (26) where *what she wants* separates *wonder* from its adjunct *because she looks hungry* (the cause of the wondering, not of the wanting).
(26) I wonder what she wants because she looks hungry. This landmark link from what to wonder explains the order wonder ... what. But it also removes the existing explanation for what ... wants. If wonder is the landmark of what, and a word may only have one landmark, this relation must override the other potential landmark of what, the lower verb wants. In that case, why does wants still have to follow what? Fortunately, the mutual dependency mentioned earlier offers an explanation: as can be seen in Figure 8, wh-words are mutually dependent with their complement verb, so (in a main-clause interrogative) what must be the landmark of wants as well as the other way round. Both landmark relations guarantee the same order (what ... wants), so the order persists even without one of the landmarks.

Figure 8: A subordinate interrogative without pied piping

This analysis is confirmed by a curious difference of word order between main-clause and subordinate interrogatives: an adjunct may be fronted before the wh-word in a main clause, but not in a subordinate clause (ANON). For example, compare the following:

(27) Tomorrow where shall we meet?
(28) *I wonder tomorrow where we shall meet.
In (27), tomorrow is a semantically coherent adjunct of shall, but this interpretation is not possible in (28). The difference can be explained in terms of the Best Landmark Principle explained above. In the main clause, tomorrow takes shall as its landmark, and so does when. But in the subordinate clause, when has wonder as its landmark, so tomorrow can’t take shall because the two landmark links would intersect.

The same kind of analysis applies to a pied-piped example such as (29) giving the structure in Figure 9, where once again the heavy dotted lines show the pied piping of with and its effect on word order.

(29) I wonder with whom she lives.
Pied piping in relative clauses
Subordinate interrogatives lead comfortably into relative clauses, which have very similar syntactic structures except for the much greater freedom in pied piping discussed in more detail below. Here too the wh-word has an extra dependency external to the subordinate clause; and here too, the wh-word has the external parent as its landmark. Take, for example, the stranded equivalent of our first example:

(30) This is the book which I found it in.
Here which is an adjunct of the antecedent, book, but it is also the complement of in and the extractee of found. As far as landmarks are concerned, book is the landmark of which, overriding both in and found; but the order which ... found is still guaranteed because which is the landmark of found (see Figure 10).

As with interrogatives, pied piping is permitted by the ‘pipee’ relation, so we start with a simple example where a single preposition is pied-piped, our (1): This is the book in which I found it. Here the pipee is in, so it is in, rather than which, that has book as its landmark and provides the landmark for found. Figure 11 gives the structure for this example.
Figure 11: A relative clause with simple preposition pied piping

The main point of this section has been to introduce the relation ‘pipee’ and to show how it extends the basic analyses of interrogative and relative clauses to include elementary prepositional pied piping. The next section considers other and more complicated constructions.

5. The limits of pied piping
Recursive pied piping in relative clauses
The flexibility for prepositional pipees in relative clauses was illustrated earlier in (12) to (16), where we find the following pied-piped phrases:

(31) ... her final exam, the result of which we expect next week.
(32) ... several projects, one of the most important of which will be to ....
(33) ... a wide range of functions, prominent among which is sexual attraction.
(34) ... a ‘higher criticism’ of the Bible, to refute which I felt the need of ...
(35) ... a rigorous examination, passing which confers on the student ...

Alongside these real examples we can put the noun phrases in Ross’s famous made-up examples (Ross 1967):

(36) reports the covers of which the government prescribes the height of the lettering on
(37) reports on the covers of which the government prescribes the height of the lettering
(38) reports the lettering on the covers of which the government prescribes the height of
(39) reports of the lettering on the covers of which the government prescribes the height
(40) reports the height of the lettering on the covers of which the government prescribes

Admittedly Ross himself claims that (37) and (39) are ungrammatical, but the differences seem too subtle to rely on.

The easiest generalisation for relative clauses is that, as far as the grammar is concerned, the pipee may be almost any ‘ancestor’ of the piper, where ‘ancestor’ is a recursive generalisation across ‘parent’ (just as in ordinary kinship terminology). The exception is finite verbs, which seem to be absolutely impossible as pipees. A plausible example would be (41).

(41) *She’s just taken her exams, that she’s passed which I very much doubt.

It is true that even some of the earlier examples are hard to process, and might be considered impossible because of that; but (41) feels different, and its badness can’t be explained just in terms of processing difficulty because it is probably easier than some of the earlier examples.
It seems, therefore, that the grammar allows pied piping to apply recursively across nouns and prepositions until it is blocked by a finite verb.

**Restrictions on pied piping in relative clauses**

On the other hand, alongside this great potential freedom for recursive pied piping, there are severe lexical limitations (Johansson and Geisler 1998). In principle, pied piping is an alternative to stranding, but this choice is actually only available in half the cases studied by Johansson and Geisler. In some of the other cases, the only permitted option was stranding, while in others it was pied piping. They report that pied piping is obligatory with some antecedent nouns (e.g. *way*, *extent*, *point*, *sense*, *degree*, *time*, *moment*) and some prepositions (e.g. *beyond*, *during*, *underneath*):

(42) I love the way in which you casually mentioned her death.
(43) *I love the way which you casually mentioned her death in.
(44) The trees stretched half-way up the mountain, beyond which there was no cover at all.
(45) *The trees stretched half-way up the mountain, which there was no cover at all beyond.

In contrast, stranding is obligatory with phrasal verbs (e.g. *look after*) and other idioms (e.g. *get rid of*) where the preposition does not have its normal meaning:

(46) This is the book which I’m looking after for her.
(47) *This is the book after which I’m looking for her.
(48) Where’s the body which we have to get rid of?
(49) *Where’s the body of which we have to get rid?
(50) *Where’s the body _ of which we have to get?

These two sets of constraints are rather different. The stranding-only cases are easy to explain in terms of processing, because they inevitably mislead the hearer/reader by inviting a default interpretation of the preposition which later has to be abandoned, producing a ‘garden-path’ effect; for instance, an incremental reading of *this is the book after which* invites the reader to construe *after which* as a time adverbial. In contrast, the piping-only cases look like constructional patterns which are learned item by item, and hardly generalise beyond these items: so we learn *way in which* and *beyond which* as fixed phrases. Moreover, the possibility of stranding must be controlled lexically by the preposition rather than the wh-word because BEYOND or DURING cannot be stranded with any relativizer, including THAT or zero:

(51) That was the concert during which I went to sleep.
(52) *That was the concert which I went to sleep during.
(53) *That was the concert that I went to sleep during.
(54) *That was the concert _ I went to sleep during.

It may be significant that these prepositions are also unstrandalbe in passives:

(55) *This line has never been walked beyond.
(56) *His lectures are never slept during.

These facts suggest that the grammatical entry for each preposition may show both whether it can be pied piped and also whether it can be stranded – two grammatical configurations that can easily be accommodated in the local network for the preposition concerned.
Another kind of limitation on pied piping in relative clauses is social: it is far more common in formal and academic writing than in conversation (Biber et al. 1999: 625). But this may follow from the social limitations on the wh-words which pied piping requires, and which also have the same distribution (Biber et al. 1999: 610). Other research confirms that, when the relativizer is a wh-word rather than that or zero, pied piping is actually more common than stranding (Johansson and Geisler 1998). Consequently, it may be better to assume that pied piping in relative clauses actually has no social limits either, so it is limited only by the lexical restrictions mentioned above.

But prepositional piping in relative clauses is not the only kind of pied piping, and different kinds have different limits. A full account of pied piping needs to accommodate the following:

- prepositional pied piping in interrogatives
- *how* as degree modifier (e.g. *how big a car, how fast*)
- *wh* determiners: *what, which, whose*
- complex pied piping combining a mix of types
- free relatives

**Prepositional pied piping in interrogatives**

One restriction on interrogative pied piping is a matter of sociolinguistics: in contrast with relative clauses, prepositional pied piping is much more formal and academic than its alternative, stranding. This generalisation emerges clearly from corpus studies; for example, “full wh-questions with a preceding preposition chiefly occur in academic prose and news. They are rare in fiction and virtually non-existent in conversation” (Biber et al. 1999: 106). Unlike relative clauses, wh-interrogatives have no alternative without a wh-word, and of course there are no social restrictions on the use of common interrogative words such as *who, what or how*. The only possible explanation for the social limitation on pied piping in interrogatives, therefore, is that the pied piping pattern itself is socially restricted.

But even when pied piping does occur in interrogatives, the syntactic structure itself also seems much more limited than for relatives (ANON) Horvath 2006: 579). It is doubtful whether recursion is possible at all in interrogatives:

(57) *In which films have you ever fallen asleep?*

(58) ??*By the end of which films* have you ever fallen asleep?

The simplest general conclusion about interrogative clauses – i.e., in dependency terms, interrogative wh-words – is that they only allow one-step pied piping, rather than the recursive pied piping found in relative clauses; in other words, the pipee of the wh-word is always also the parent preposition.

However, this claim misses an important exception: what could be called ‘quiz-show questions’ such as (59) which are syntactically unlimited.

(59) *To the daughter of which famous statesman was he engaged?* (Huddleston and Pullum 2002: 912)

Huddleston and Pullum accept this example without comment, but it seems very unlikely in ordinary conversation, while being very easy to imagine in a quiz show. Similar examples are easy to find on the internet (e.g. on the ‘pubquizreference’ site):
(60) The name of which plant comes from the Greek meaning 'Earth Apple'? 
(61) In the grounds of which house is the largest private tomb/mausoleum in England? 
(62) Officers in which army were given copies of 'Les Misérables'? 
(63) On the banks of which river is the Taj Mahal? 
(64) Six ounces of what contains the minimum daily requirement for vitamin c? 
(65) The assassination of what country's Archduke led to World War I? 
(66) The closure of which British nuclear re-processing plant was announced in 1998? 

However even in this specialized genre the syntax is clearly not easy to control, as can be seen in the following internet examples where pied piping is combined with stranding of the same preposition:

(67) At what approximately rounded figure does the Earth revolve at? 
(68) In the world of communications for what do the letters U R L stand for? 
(69) Of what are corolla, filament and stigma a part of? 

Alongside the general syntactic restrictions on pied piping in interrogatives, we must also recognise that even this limited pied piping is lexically restricted, just like the pied piping found in relative clauses (Huddleston and Pullum 2002: 913). An obvious example is the combination what ... for, meaning 'why' (or more accurately, 'with what purpose in mind'):

(70) What did you do that for? 
Here pied piping is absolutely impossible, even in quiz shows:

(71) *For what did you do that? 
The same ban on pied piping applies to a few other idiomatic combinations such as what ... like; and as with relative clauses, some preposition + wh + noun combinations (e.g. in what way) have to be pied piped:

(72) In what way can I help you? 
(73) *What way can I help you in? 
Once again, these lexical restrictions suggest an analysis which includes a large number of lexically specific constructions alongside some rather weak generalizations. The next section will suggest how this analysis can be modelled in a network grammar. 

This completes the discussion of prepositional pied piping, so it will be helpful to stand back from the detail to review this construction. The most obvious point to make is that this kind of pied piping has the effect of separating the wh-word from its preferred position at the start of the clause; so instead of starting with, say, what the clause starts with in what. Another effect is to shift material from the end of the clause to the beginning. We shall consider possible functional explanations for the pied piping in section 7.

How as degree modifier
We now consider examples in which the wh-word comes first in its phrase, starting with the ‘wh-word’ how:

(74) How fast did she run? 
(75) How big a house do you want? 
These examples count as pied piping because the wh-word determines the position of its phrase without being the head of this phrase. But unlike prepositional pied piping, the pied piping is part of a more general pattern which has nothing to do with relative or interrogative
clauses. As a degree modifier, how behaves like so, too, as, this and that because, when modifying an adjective the latter must modify a following a (which, in turn, needs its usual singular noun):

(76) They bought so big a house that we always get lost there.
(77) They bought too big a house to heat properly.
(78) They bought as big a house as they needed.

As these examples show, most of these modifiers have the additional property of projecting their valency upwards; so in (76) for example it is so that lexically sanctions the that clause, but this depends on house, rather than on so. (In Rosta’s terms, either a or house is the ‘surrogate’ of the degree modifier.) This isn’t just a matter of word order, but of meaning. For example, in as big as X, the parameter of comparison is simply size, whereas as big a girl as X builds ‘girl’ into the parameter, so that any standard of comparison must also be a girl.

This is why (79) is sensible but (80) is anomalous.

(79) She is as big as her brother.
(80) #She is as big a girl as her brother.

In short, degree modifiers sanction a dependency for a noun higher up the dependency chain.

This analysis provides an explanation for the pied piping found with the degree modifier how: how sanctions a finite verb complement, but this depends on the higher noun, so it takes the higher noun as its landmark. This explains why how big or how big a house are clause-initial in (74) and (75), but it does not explain why these phrases can take their position in a higher clause, as in (81) and (82):

(81) I wonder how big a house they want.
(82) How big a house they want is a mystery.

This external positioning of the subordinate clause is handled in our analysis by the ‘pipee’ relation, so this relation must also be available for how. We shall see below that this will help to explain why how can be combined with prepositional pied piping, as in In how big a house do they live?

Wh determiners
Like how, the wh determiners what, which and whose occur on the left of their phrase, but unlike how, they qualify as head of this phrase; or at least they do if we accept the DP-like analysis favoured by Word Grammar (ANON). Given this analysis, there is no more need for pied piping in (83) than in (84).

(83) What clothes shall we take?
(84) What shall we take?

But whose is rather different from the other wh determiners because a very persuasive analysis (Rosta 1997) takes whose as two syntactic words, who + Z, realized by a single word-form in the morphology. In that analysis, Z is the determiner and the head, just as it is in at least one analysis of John’s or everyone else’s (ANON), so even a simple example such as (85) involves pied piping, with who as a dependent and Z as the pipee.

(85) Whose car shall we use?
(86) Who Z car shall we use?

Consequently, we must recognise the who of who Z as a pied piper.
Both *how* and the *wh* determiners stand firmly at the start of their phrase, so we can call them ‘phrase-initial pied piping’. In contrast with prepositional pied piping, the result of phrase-initial pied piping is that the *wh*-word stands at the start of the clause, rather than being separated from it by the pied-piped material, so we should not expect the same functional explanations to apply to both.

Complex pied piping
The two kinds of pied piping, prepositional and phrase-initial, can combine freely, as in the following examples.

(87) *In how big a house* do they live?
(88) *In whose car* did you come?
(89) I’ve found the book *about whose cover’s colour* we were arguing.

Moreover, phrase-initial pied piping is indefinitely recursive when combined with the possessive *Z* discussed earlier:

(90) How many people’s neighbours’ dogs dig up their gardens, I wonder.
(91) The man whose mother’s father’s brother’s house is up for sale has just died.

In all these cases, the pied-piped phrase follows ordinary syntactic patterns except for its position, which is determined exceptionally by the *wh*-word.

Once again the simplest conclusion is that the pipee relation can be passed recursively up the dependency chain without limitation. On the other hand, we still find the limitations on prepositional pied piping noted previously: an absolute prohibition of finite verbs as the pipee, and a ban on recursive prepositional pied piping in normal interrogative clauses.

Free relatives
In contrast with all other kinds of *wh*-word, those that introduce free relative clauses don’t allow any pied piping (Heck 2008: 342). For example, whereas *with which* is an alternative to *which ... with* in (92) and (93), *what ... with* in (94) cannot be changed into the *with what* of (95).

(92) This is the thing *which* I blocked the hole with.
(93) This is the thing *with which* I blocked the hole.
(94) This is *what* I blocked the hole *with*.
(95) *This is *with what* I blocked the hole.*

This restriction cannot be explained in terms of dependency structure, because pied piping does not affect dependency structure so (94) and (95) have just the same dependency structure. But a functional explanation may be possible in terms of the processing difficulties of *this is with what*, where the normal procedures try to treat *with* as a dependent of *is*. On the other hand, some languages do allow free relatives with pied piping. For example, German examples like (96) have been found in published writing (Müller 1999:61, quoted in Hoffman 2011:58):

(96) Aus wem noch etwas herausgequetscht werden kann, ist sozial dazu verpflichtet, es abzuliefern.

‘Those who have not yet been bled dry are socially obliged to hand it over.’
This example is remarkable for the fact that *wem*, ‘who’, is in the dative, as required by the preposition *aus*, but is also functioning as the subject of the higher clause, which would normally require an accusative *wer*. If German speakers can process (96), why can’t English speakers process (95)? In short, the proposed analysis fails to explain the restrictions on pied piping in free relatives.

The next section explains how these English patterns can be formalized in Word Grammar, and then section 7 will explore the cognitive background for pied piping.

6. Formalizing the pipee

Section 4 presented a Word Grammar analysis of single-step prepositional pied piping, so we now need to consider how this grammar needs to be expanded in order to accommodate the more complicated examples reviewed in section 5. The earlier analyses all relied on the special relation ‘pipee’, which connected the wh-word to the word which replaced it in the landmark structure (responsible for word order). We shall now see that this relation is all we need to accommodate the more complex examples as well.

The main complexities are as follows:

- different provisions for relative, interrogative and free relative clauses
- recursive pied piping in relative clauses and some interrogatives, with an absolute ban on finite verbs as pipees
- lexical restrictions on pipers (the wh-words), pipees and antecedents
- social restrictions on some kinds of pied piping.

Arguably, all these possibilities and limitations are part of our ‘competence’, i.e. our knowledge of English, so the challenge is to incorporate them all into a formal model of that knowledge.

Pipers

We start with the wh-words themselves. Word Grammar provides a hierarchical classification of words, so we can recognise a hierarchy containing ‘pronoun’ and ‘relative pronoun’ as well as the individual lexemes such as WHO. This inheritance hierarchy, based on the ‘isa’ relation introduced earlier, replaces the syntactic features favoured by most other theories, and there is certainly no suggestion that the hierarchy might be part of ‘universal grammar’; it is learned rather than innate, and can vary freely between languages. One of the advantages of a hierarchical classification is the prevalence of this kind of classification outside language, in contrast with feature-structures, which are much more limited outside language. Another is the unified treatment that it allows for general categories and lexical items, contrasting with the fundamental distinction forced by feature structures.

One example of this unified treatment is the possibility of distinguishing ‘*sub-lexemes*’ – distinct uses of what is clearly a single item, comparable to the ‘sub-tokens’ discussed earlier. For example, it allows us to recognise a single lexeme WHO which requires a human referent, while also distinguishing the relative pronoun WHOrel from the interrogative WHOint. These sub-lexemes can inherit their distinct properties from the more general categories ‘relative pronoun’ and ‘interrogative pronoun’ because the hierarchy allows multiple inheritance; so WHOrel isa both WHO and ‘relative pronoun’. Similarly, ‘interrogative pronoun’ inherits from ‘interrogative word’ and ‘pronoun’, so HOW can inherit from ‘adverb’ as well as from ‘interrogative word’, and assuming that there are no other such
words, the resulting intersection could equally well be called either ‘HOW’ or ‘interrogative adverb’. (Arguably, both WHEN and WHERE are pronouns, since they can both be used as relative pronouns.) The network in Figure 12 shows some of the relevant nodes.

This classification of words will allow us to identify words as ‘pipers’ in pied piping. For instance, we can say that any interrogative word is a potential piper; or we can say more specifically that WHOint is a potential piper. Since pipers are lexically restricted, it seems likely that we shall need to include both general and specific information. How this can be done will be explained below.

Constructions
The subclassification system of Word Grammar accommodates constructions in terms of sublexemes. For example, an idiom such as KICK THE BUCKET is analysed as containing a special sub-lexeme of each of the words concerned: KICKbucket, THEbucket and BUCKETkick. (The choice of labels is, of course, trivial; indeed, one of the basic assumptions of network analysis is that labels are merely an analytical convenience, without any material content.) Each of these words has some unique properties; for example, KICKbucket is the only example of KICK that means ‘die’ and requires THEbucket, and the latter is the only example of THE that occurs with KICKbucket and BUCKETkick. But at the same time, these words are also linked to the default lexemes KICK, THE and BUCKET, as has been shown through priming experiments (Sprenger, Levelt, and Kempen 2006). The proposed analysis captures the similarities to the default entries as well as the differences.

Returning to pied piping, the same principles apply. For example, WHATfor is the sublexeme of WHAT which occurs in the what ... for construction, along with FORwhat, the

Figure 12: Classification of words in a network
special sub-lexeme of FOR. This system allows both pipers (e.g. *what*) and potential pipees (e.g. *for*) to be specified either crudely, in terms of general word classes, or more specifically in terms of sub-classes or even specific lexical items. Moreover, in this system we can tie lexical constraints to the pied-piping or stranding constructions: it is just *WHATfor*, rather than every *WHAT*, that cannot have a pipee, and it is *WHICHway* (the *which* found in *the way in which ...*) that must have a pipee. The first restriction rules out *for what* (in *For what did you do that?*) while the second rules out *the way which ... in* (in *I liked the way which he did it in*). In view of the very construction-specific constraints on pied piping reviewed in section 4, this degree of specificity is essential.

**The optionality of prepositional pied piping**

Assuming, then, that we can distinguish the constructions which require pied piping from those which forbid it or merely allow it, we now need a mechanism for expressing these differences in a network. Optionality is handled in Word Grammar by means of an elementary ‘quantity’ relation, whose values range over numbers. By default the quantity of any entity is 1, meaning that when it is possible structurally, we expect to find one exemplar in our experience; but the default can be overridden by values such as ‘0-1’ (either 0 or 1) or ‘0’. In network diagrams, the quantity relation is labelled ‘#’, so the total network would include the sub-network shown in Figure 13 for *WHATfor* and *WHICHway*. The diagram shows that, by default, a pipee is optional, but it is impossible with *WHATfor* and obligatory with *WHICHway*.

![Network Diagram](https://example.com/network_diagram.png)

*Figure 13: WHATfor cannot have a pipee but WHICHway must have one*

**Pipees and landmarks**

Now that we have a mechanism for adding the pipee relation as required, we can show how this relation produces the word order of pied piping. As explained earlier, a word’s parent is normally also its landmark; this is shown in the top section of Figure 14. This is the default arrangement, but pied piping is exceptional because the ‘piper’, defined as a particular kind of word (e.g. a relative pronoun) has a ‘pipee’ which takes over the piper’s role in landmark
structure: any parent of the piper acts as the landmark (‘lm’) of the pipee, and any dependent of the piper takes the pipee as its landmark.

Figure 14: Pipees as exceptional landmark bearers

The sub-networks presented above combine as a formal grammar which generates the structures shown in the earlier syntactic diagrams. For example, Figure 7 shows that exactly the same dependency structure is mapped onto two different word orders according to whether the pipee relation is included. In the default case, with stranding and no pied piping, the order is: *Who do you live with?* In this sentence, *do* is the landmark of *who* and *who* is also the landmark of *do*. But when *with* is the pipee of *who*, these landmark relations transfer from *who* to *with*.

Recursive pied piping

Finally, the grammar needs to allow recursive pied piping in some cases but not in others. Recursion is found in both prepositional and phrase-initial pied piping, but it involves different mechanisms so we consider the two types separately.
By default, prepositional pied piping is freely recursive and allows complicated structures like (31) to (40). The only general constraint is that a finite verb cannot act as a pipee (as in the ungrammatical (41)). The sub-network in Figure 15 shows how this recursion can be expressed in a network. In prose, if word B is the pipee of word A, and B depends on C, then C may also be the pipee of a new token of A, A’. (The extra token A’ is needed in order to allow A to have more than one pipee.) But in the even more specific case Y where C is finite, this extra pipee relation is impossible. The optionality of the pipee in one case and its impossibility in the other are shown by the ‘quantity’ relations (#).

Figure 15: Defining ‘pipee’ recursively

One possible objection to this analysis is that it identifies too many pipees, each of which could have the piper’s parent as its landmark. For example, in the book on the cover of which I wrote, the piper which creates four different pipees: on, the, cover and of; but only on takes over the piper’s landmark role. Why isn’t this possible for the other three pipees? The answer is that the pipee relations form an isa hierarchy: which: on (the pipee relation from which to on) isa which:the, which isa which:cover, which isa which:of. By default inheritance, the winning value is the one furthest from the default, at the bottom of the hierarchy; so which: on automatically beats all the other values for ‘pipee’ when applying the rule for landmarks.

The discussion so far has focussed on prepositional pied piping, and has almost finished the formal analysis of this kind of pied piping in English. The only fact not yet covered is the fact that ordinary interrogatives are different from quiz interrogatives; since this involves a complex sociolinguistic category, we must leave this for future research.

We now turn to the phrase-initial pied piping found with interrogative how, what and which, and either interrogative or relative whose, as in how big a hole, what size or what size shoes, which book and whose book. These need not delay us, because they are already covered by the analysis. As we saw in relation to examples (87) to (91), the wh-word can
pied-pipe recursively, so the pipee may be either the head of the local phrase (e.g. *shoes* in *what size shoes*) or a more remote ancestor, whether a preposition, as in a sentence such as *In how big a house do they live?*, or a possessive *Z*, as in *How many people’s neighbours’ dogs dig up their gardens, I wonder*.

7. Cognitive ends and means
The previous sections have all concentrated on showing how pied piping can be accommodated in a theory such as Word Grammar which makes all the machinery of ordinary cognition, and only this machinery, available to grammatical analysis. Because of this focus, the discussion was all about how pied piping works in Modern English. However, one of the basic assumptions of Word Grammar is that the challenge for linguistics is not to explain the similarities among languages, but rather to explain their differences (Evans and Levinson 2009). Given the problems of communication, each language evolves its own ‘engineering solutions’ which strike some kind of balance between competing pressures.

What, then, are the communication problems for which pied piping is a solution? What relevant cognitive apparatus is available? And what solutions are found in other languages?

Communicative ends
The communication problem addressed by pied piping is information-packaging: how to achieve a user-friendly distribution of the information in a sentence. The same could be said, of course, about any part of language, because every word has evolved as a way to package information in a user-friendly way; but, in contrast with most words and constructions, pied piping does not increase the expressive power of a language in the technical sense of allowing more meanings to be expressed. Instead, it is just a matter of word order (section 2), so it applies to dependency structures that are licensed independently by the grammar, and allows a word order which would not otherwise be possible but which brings some identifiable functional benefits.

The functional benefits of pied piping have to be explored in the context of a number of potentially competing functional pressures:

1. endocentricity: a phrase’s head determines all its properties relevant to the rest of the sentence, including its position in word order. This principle disfavours all pied piping.
2. boundary linkers: a word which indicates how a whole phrase is linked to other words in the sentence is located on the phrase’s boundary between that phrase and those words. This principle disfavours prepositional pied piping, because the linking word (the *wh*-word) is not at the clause’s edge, but it allows phrase-initial pied piping.
3. anti-stranding: a word should not be separated from its complement. This principle favours pied piping as an alternative to stranding.
4. memory protection: structures should be selected so as to minimize the burden on the working memory of both speaker and hearer. This principle favours any structure which allows information to be distributed evenly, including pied piping.
5. grammar simplification: grammars should be as simple as possible. This principle disfavours the extra complexity of pied piping.

The first four pressures apply to pied piped and stranded pairs such as (97) and (97).
This is the book which I found it in.
This is the book in which I found it.
The stranded (97) wins on two of the pressures and the pied-piped (97) on just one, with almost equal scores on the remaining one:

- **Endocentricity** is satisfied by *which ... in*, because the phrase’s head (*in*) is in its expected position (even if its complement is not), whereas *in which* defies endocentricity by taking its position from a non-head (*which*).

- **Boundary linkers** are as expected in *which ... in* because *which* (the word which links its clause to the antecedent book) is the first word in its clause, in contrast with *in which*. (This constraint is similar to Broadwell’s ‘align’ which places the interrogative phrase at the left edge of the clause — Broadwell 1999).

- **Anti-stranding** disfavours structures such as *which ... in*, including those which we call ‘stranding’, in which a word’s complement precedes it and is separated from it. It may be that this constraint could be constrained so as to apply only to function words (such as prepositions) and their complements; this would explain why prepositions cannot be stranded in languages where verbs can, and might explain why languages that forbid preposition stranding also tend to allow fusion of a preposition with a definite article (e.g. French *au* for *à le* and German *zum* for *zu dem*, both meaning ‘to the’), showing a specially tight connection between prepositions and their complements.

- For memory protection, the two versions make equal memory demands when measured in terms of dependency distance, the number of words separating each word from its parent (ANON): *which ... in* scores one more on *in*, while *in which ...* scores one more on *which*. However the pied-piped version also includes the pipee relation from *which* to *in*, so its structure is slightly richer and therefore makes more memory demands.

Grammar simplification is a different kind of pressure because it applies to the diachronic development of a grammar by the language community concerned rather than to the way in which that grammar is applied; so given that English already allows both pied piping and stranding, this criterion does not distinguish *which ... in* from *in which*. The fact is that pied piping requires special apparatus, as outlined above, and however simple that apparatus is, it still needs to be learned. Similarly, a grammatical ban on stranding needs to be learned (as a special constraint on recursive extraction), so the simplest grammar is one which allows stranding but not pied piping. For any language community, therefore, the challenge is to evolve a grammar which strikes an acceptable balance among these five competing pressures.

For linguists, the challenge is to explain why any community chooses pied piping as a solution, given the balance of advantage that appears to favour stranding. One benefit of pied piping has already been noted: anti-stranding forbids the obvious alternative, which is stranding. But this doesn’t in itself explain why pied piping develops; after all, French-speaking children apparently manage without it (Guasti and Shlonsky 1995), just as English-speaking children do (McDaniel, McKee, and Bernstein 1998), even though they have no stranding alternative. If children can tolerate the lack of pied piping, why not adults too?

Another benefit involves phrase-initial pied piping, in which a non-head such as *whose* or *how* pipes its phrase to the front of the clause. Any language that has such words needs the equivalent of our ‘pipee’ relation, even if it does not allow prepositional pied piping; so at
least that much of the extra grammatical apparatus is already available. However, not every language has such words: for instance, examples such as (99), a word-for-word translation of English *How big is it?*, are possible in German but not in French, which demands (101) instead of (100).

(99) Wie gross ist es?
    ‘How big is it?’
(100) *Combien grand est-il?
(101) Combien est-il grand?
    how is-it big?

And although determiners such as quel, ‘which’, exist in French as well as in German, these are heads according to the DP analysis favoured in Word Grammar, so they don’t require pied piping.

The most convincing functional explanation for the existence of pied piping is that it helps when the clause introduced is complex, by distributing information more evenly than the stranded alternative (Johansson and Geisler 1998). Johansson and Geisler quote the following examples from their corpus of spoken language:

(102) I mean there are some subjects in which I get the impress you know that there is a pump going and the stuff is pumped from one mind into another mind.
(103) The system architecture which is the focus of today's seminar is really an er a complete environment within which we can deal with the challenges of this new generation of enterprise client server applications.

In both of these examples, the stranded alternative would place a heavy strain on working memory because of the many words separating which from the stranded preposition:

(104) I mean there are some subjects which I get the impress you know that there is a pump going in and the stuff is pumped from one mind into another mind.
(105) The system architecture which is the focus of today's seminar is really an er a complete environment which we can deal with the challenges of this new generation of enterprise client server applications within.

This benefit is similar to one of the benefits of topicalization, which allows information to be redistributed within a complex clause. For instance, we might compare (103) and (105) with the following pair:

(106) Within this environment, we can deal with the challenges of this new generation of enterprise client server applications.
(107) We can deal with the challenges of this new generation of enterprise client server applications within this environment.

Topicalizing the adjunct allows it to be processed separately from the rest of the clause, and allows it to be linked directly to *can*, whereas the end position in (107) leaves it separated by a lot of other words from *can*.

If this explanation is correct, and pied piping arises in order to facilitate the formation of complex clauses, it could explain a number of other things. Maybe this is why pied piping is so strongly associated in English with formal (and therefore complex) writing. And maybe it explains some of the strong lexical restrictions noted earlier, if we can also assume that the
constructions concerned, such as *way in which* or *beyond which*, are typically used in complex structures.

**Cognitive means**

Suppose, then, that the extra machinery of pied piping would help communication if it was introduced. And suppose the ordinary machinery of grammar is supported mentally, as claimed in section 3, by ordinary cognitive abilities such as networks, default inheritance and landmarks. Would the extra machinery for pied piping require special cognition unique to language, or could we invoke other bits of ordinary cognition here too? Since the extra machinery amounts to just one relation, ‘pipee’, the question is whether this has analogues outside language.

What we are looking for, therefore, is a non-linguistic entity A which humans locate (by landmark) not according to its own inherent classification, but according to that of some other entity B which is subordinate to it to the extent of taking A as its own landmark. This would be exactly equivalent to a preposition A taking as its landmark the parent of a wh-word B for which A is the landmark (as in *the house in which he lives*). And if we do find a non-linguistic analog for pied piping, then we can assume that our minds establish a relation between A and B which provides a non-linguistic analog for the ‘pipee’ relation – and evidence that the cognitive apparatus for pied piping is not unique to language.

For a simple example, consider supermarkets, where all the goods are classified and located according to some system which was devised by the shop owner so as to be meaningful to shoppers – a clear case of human decisions about location, defined in terms of landmarks such as ‘aisle 24’ or ‘the back left corner of the shop’, but also based on how the items concerned are typically classified by members of the relevant culture: for example, as food, drink or household goods and more precisely as salad, fruit, meat, dressings and so on. The crucial point of the example is that many items are sold in some kind of packaging or container such as cans, bottles, boxes and bags, each of which acts as landmark for the thing contained; so the beans are in the can, rather than the can being round the beans. But the location of the containers depends on their contents, with cans of beans and peas located near to each other, together with other kinds of container full of prepared vegetables. In short, the contents could be described as ‘pied piping’ the container. (In contrast, we might imagine a shop in which all the cans are on one shelf, all the bottles on another, and so on, all regardless of what they contain.)

Since speech is linear action, it may be helpful to find another non-linguistic example that involves the ordering of actions in time rather than physical location. In this case we need to consider a structured activity such as baking a cake that follows a ‘script’ (in the sense of Schank and Abelson 1977) of steps to be taken. Each step defines a goal and the sub-goals that lead to this goal, and each sub-goal may define further sub-goals; for example, the goal of mixing the ingredients includes the sub-goal of getting the ingredients, and the goal of decorating the cake includes the sub-goal of preparing the icing, which in turn has the sub-goal of getting the sugar. In terms of ordering, the script defines the order of the sub-goals, with sub-goals typically timed to be achieved just before their goal. But this ordering can be disrupted by various constraints which can require advanced action – i.e. preparation – such as shopping for missing ingredients, washing unwashed bowls or warming butter. In such cases, the preparation, such as warming butter for the icing, takes place before the main
baking activity starts and may lead to a re-ordering of the higher sub-goal (in this case, with the icing made before the rest of the cake). In formal terms, the hard butter ‘pied pipes’ the making of the icing.

The point of these examples is to show that the special characteristics of pied piping, which centre on the ‘pipee’ relation, are by no means special to language. Instead, we find similar formal patterns in everyday non-linguistic thinking and behaviour. This is not to say that we necessarily use the same mental relation in all these cases (though it is possible that we do); and it is certainly not to say that we are born with the ‘pipee’ relation hard-wired. All it means is that we can give ourselves extra mental flexibility by creating ad hoc relations such as ‘pipee’ which produce the same effect outside language as we find in syntactic pied piping.

Different languages, different solutions
It is clear that there is nothing ‘natural’ about the engineering solutions offered by English to the fundamental problem of wh-type words, which combine two conflicting sets of properties. On the one hand, a word such as what in (108) is an ordinary pronoun, just like it in (109), with the same relations to wants.

(108) I wonder what she wants.
(109) She wants it.

On the other hand, what is also different from it in facing outwards as well as inwards – in this case, in providing the interrogative word that wonder needs as its complement. These properties are associated with conflicting word-order demands, which is why what is extracted from its expected default position after wants. Other languages offer different ways of resolving this conflict, such as ‘wh in situ’ which favours the clause-internal relations.

The same conflict leads to pied piping, where the wh-word combines its normal grammatical function (e.g. as a pronoun) with a function that requires it to stand at the start of its clause; but in pied piping, this conflict goes beyond the mere reordering of words that we find in extraction, as it threatens the fundamental principle of endocentricity because the wh-word is not the head of the displaced phrase. English-style pied piping is just one engineering solution to the problem, with stranding of a preposition as an alternative in some cases. But although stranding is common in Modern English, it was rare in Middle English, when (unlike Modern English) prepositional pied piping was almost obligatory with wh-words (Johansson 2002). And of course, in many European languages stranding is absolutely forbidden.

English-style pied piping and stranding are not the only possible solutions, and, as we might expect, other solutions can be found in other languages. We note here two particular variations on pied piping: inversion and boundary marking.

Inversion solves a problem of English pied-piped sentences such as (110).

(110) Until when shall we wait?
The problem here is that the wh-word when is not clause-initial, in spite of the strong pressure to put wh-words first. As a solution, inversion is very simple: it changes the default order within preposition phrases so that the wh-word does come first, giving the equivalent of ‘when until shall we wait?’ Although this is not possible in Modern English, it was common in Middle English with purely pronominal wh-words, which were realised as where:
wherefore, whereof, whereon, wheresoever, wherethrough, whereto, wherewith. Cognates are still normal in German and Dutch (e.g. German worin, ‘in what’). However, these may not be clear examples of syntactic inversion because the inversion may be just in the morphology, on the assumption that the pronoun is cliticized to the preposition (as indicated by standard orthography). In other words, Middle English wherefore may have the normal order for what in syntactic structure but the reverse order only in the fused morphological realisation.

A much more productive kind of inversion, called either ‘pied piping with inversion’ or ‘secondary wh-movement’, is found in some head-initial Meso-American languages (Broadwell 1999; Broadwell 2006). This phenomenon is more relevant to the present discussion because it clearly involves syntax rather than morphology. For instance, (111) to (114) from San Dionicio Ocotepec Zapotec show the effect of pied piping with inversion when the piper is a possessor (Broadwell 2006). Normally the dependent possessor follows the possessed, as expected in a head-initial language; this is shown in the first two examples, where the possessor is Màríí. But when, as in the second pair of examples, this is replaced by an interrogative, the latter triggers both pied piping and inversion of the possessor and possessed so that the interrogative word is clause-initial.

(111) Cù’á Juààny x-pèh’cw Màríí.
  com:grab Juan  p-dog Mary
  ‘Juan grabbed Mary’s dog.’

(112) *Cù’á Juààny Màríí x-pèh’cw].
  com:grab Juan Mary p-dog

(113) Túú x-pèh’cw cù’á Juààny?
  who p-dog com:grab Juan
  ‘Whose dog did Juan grab?’

(114) *X-pèh’cw túú cù’á Juààny?
  p-dog who com:grab Juan

Such examples are easy to accommodate in the same kind of analysis as for English, based on the pipee relation. Once again, the pipee takes over the piper’s role in landmark structure, but this time it also affects this structure by reversing the normal landmark relation between the piper and pipee. Since the piper (here túú, ‘who’) depends on the pipee, it would normally follow the pipee; but this default, giving the order possessed-possessor, is overridden by a special inversion rule giving possessor-possessed. Given the detailed lexical variability found in English prepositional pied piping, it is not surprising that similar variability is found among prepositions in Zapotec, or indeed that the scope of the rule, in terms of word classes to which it applies, varies from language to language.

In short, pied piping with inversion requires a very small extension of the analysis proposed for English, but (as Broadwell points out) the restricted variability gives no reason to assume that either pied piping or inversion is in any sense part of universal grammar. Given the functional benefits and costs of pied piping, inversion is a rather obvious way to reduce the costs.
The other variation on pied piping is boundary marking, an overt word which in some languages marks the end of a pied-piped phrase and which Heck calls ‘Q’ (Heck 2008; Heck 2009). The languages include Tlingit, for which Heck reports the following examples:

(115) Daa sā aawaxaa i éesh?  
what Q he.ate.it your father  
‘What did your father eat?’

(116) Goodēi sā has oowajée wugootx i shagónich?  
where.to Q they think he.went your parents.ERG  
‘Where do your parents think he went?’

(117) Wāa kwligeyi xáat sā i twúa sigóo?  
how it.is.big.REL fish Q your spirit.at it.is.happy  
‘How big a fish do you want?’ (Lit. ‘A fish that is how big do you want?’)

(118) Aadóo teen sā yeegoot?  
who with Q you.went  
‘Who did you go with?’

(119) Aadóo yaagú sā ysitone?  
who boat Q you.saw  
‘Whose boat did you see?’

(120) X’oon keiit sā ysitone?  
how many dog Q you.saw  
‘How many dogs did you see?’

The boundary-marking word is sā, which is obligatory after an interrogative word, and can only occur in the position indicated in the examples. With both the interrogative word and sā underlined, it is easy to see how they provide a clear structural guide to hearers (not dissimilar in function to the English subordinator that, which signals the start of a clause). Although functional explanations are inevitably speculative, it is easy to see this strategy as yet another engineering solution to the problems of wh-type words. And in terms of the structural analysis, it is easy to see sā as a post-dependent of the pipee – a case where the pipee is more intimately related to the general dependency structure than in languages like English.

8. Conclusions
The main conclusion is that even pied piping, for all its apparent peculiarities, is just part of general cognition, along with the rest of language. The analysis presented is at least compatible with all the known facts about pied piping not only in English but in other languages; but it also assumes no theoretical apparatus beyond what we know is already available in general cognition for dealing with space, society and so on. This cognition includes networks (with rich hierarchies of relationships), default inheritance and landmarks, which are enough not only for ordinary syntax, but even for such an abstruse part of syntax as pied piping.
But the ‘pipee’ relation which is the defining characteristic of pied piping is peculiar in that it interrupts the normally direct relation between dependency structure and landmark structure. Similar relations are found outside language, so the ‘pipee’ relation does not support the idea that language is conceptually unique; on the contrary, if anything it shows that language is subject to much the same functional pressures as the rest of cognition, and that our minds are so flexible that we can invent new relations as needed.

Another general conclusion is that the observed patterns of pied piping cover the expected range of engineering solutions to the problems posed by wh-words – stranding, pied piping, inversion and boundary marking. If our aim is to understand the diversity of human language, then our explanations must start with these functional pressures, plus the cognitive resources that are available for dealing with them.

References


