Journal of Linguistics (2026), 1–30
doi:10.1017/S002222672610125X

Journal of Linguistics
CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Mutual dependency, English *wh*-clauses and Word Grammar

Nikolas Gisborne

University of Edinburgh
Email: n.gisborne@ed.ac.uk

Received: 28 June 2024; **Revised:** 26 November 2025; **Accepted:** 16 March 2026

Keywords: extraction; landmarks; mutual dependency

Abstract

Word Grammar (WG), a dependency grammar, has historically permitted mutual dependency, a structure where two words depend on each other. This paper re-evaluates mutual dependency within WG, focusing on contradictions arising when dependency relations are inherently tied to word-order via WG's 'landmark' relation. It is argued that mutual dependency which involves conflicting landmark relations is untenable. I examine earlier WG analyses of English dependent interrogatives and free relatives, which were the principal motivators for mutual dependency analyses. It is proposed that dependent interrogatives do not require mutual dependency and are better analysed with the verb as the head, with selection being semantic. Free relatives, however, with their matching effects where the *wh*-word is head, necessitate a form of mutual constraint. A solution is offered by refining WG's theory of word-order and its relationship to extraction, by distinguishing between the general relationship of word-order, 'landmark' and a distinct relation of word-order that is part of the extraction start-rule. This approach allows for mutually constraining relations in free relatives without contradictory landmark assignments, thereby preserving a coherent theory of word-order while accommodating mutual dependency.

1. Introduction

Word Grammar (WG) is a dependency grammar (Hudson 1984, 1990, 2007). That is to say it is a syntactic theory where the ontology is restricted to nodes, which are words, and arcs, which are the pairwise dependency relations between words. Mutual dependency involves a pair of words where each word depends on the other: many dependency grammars reject mutual dependency, but Word Grammar allows it. In this it is not unique: in a sense, so do the models developed by Tesnière (1959) and Gerdes & Kahane (2011). That said, dependency grammars typically reject mutual dependency, usually by being axiomatised to exclude cyclic structures because dependency is usually defined to be antisymmetric and irreflexive. Moreover, a dependency grammar needs working definitions of 'head' and 'dependent' and the challenge in mutual dependency is in avoiding contradictions or incoherence. In a theory

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

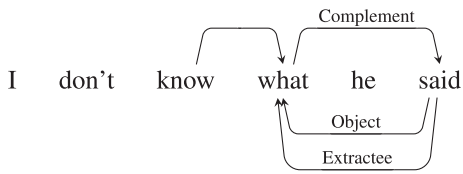
that includes word-order in its syntax, without further theorising, a mutual dependency pair involves the contradiction that each word takes its position from the other.

It is possible to deal with such issues. Word Grammar has always treated word-order as part of dependency, but it has always banned mutual word-order relations, allowing there to be dependencies that are related to word-order and dependencies that are not. There have been three successive theories specifically addressing word-order in WG and its relationship to dependency structure (Hudson 2003, 2007, 2018).[1] In this paper, I argue that there is an omission in WG's most recent theories (Hudson 2007, 2018), to do with the syntax of extraction, which requires a slight reworking of the theory of word-order in Hudson (2007).

The alternative I present has two parts. First, given Hudson's (1987) arguments about the properties of heads in dependency grammars, mutual dependency must be exceptional: Hudson argues that the criteria for heads typically converge on a single word in a two-word syntagm. WG has a default inheritance architecture with relational categories being typed in an inheritance hierarchy; therefore, the convergence of head properties on a single word is the (proto-)typical default case but not an absolute rule. Given this, mutual dependency overrides the default. If mutual dependency is exceptional, then it must be restricted to constructions that have exceptional properties. In consequence, analyses involving mutual dependency should be dispreferred if a non-mutual account is possible and offers a plausible and coherent analysis of the data. Second, I revise the theory of word-order and the start-rule of extraction to allow for a non-contradictory account accommodating mutual dependency in extraction.

Hudson (1984: 84–5) originally motivated mutual dependency for WG in an account of dependent interrogatives in structures such as (1).[2]

(1)



In Hudson's analysis *what* is dependent on *said* and *said* is dependent on *what.* He argues that *what* must be dependent on *said* because it is the (extracted) direct object of *said.* But at the same time, informally speaking, *said* must be dependent on *what* because *know* selects for an interrogative complement, and the only candidate is the interrogative word, *what.* In representations such as (1) the formalism itself does not identify the external 'distributional' head of *what she said*, which could be either *what* or *said.* Hudson (1984) argues that it must be *what* because *know* selects an interrogative word. Hudson (1984: 85) defends mutual dependency by arguing that it is well-formed as long as different dependencies are involved, because then there will be no conflicts in the syntactic information: 'there are interdependence relations in sentence structure, but these always involve a pair of separate entities in the grammar'. In WG, each dependency is intransitive and typed, therefore as long as each dependency is different, there will be no contradictions.

[1] Hudson (1990: 114–120) presents an outline of a theory.

[2] Hudson doesn't give the dependency annotation; I limit myself to the key dependencies here.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

We can see how this works in (1). Let us assume that, as a head, *what* selects a finite verb and therefore governs *said.* At the same time, *said* selects a nominal element as its direct object: *said* governs *what.* There need not be a contradiction: the relation 'governs' is a linguists' generalisation. The relevant constraints don't involve an abstract 'governs' relation: they involve the requirement for a finite verb as the complement of *what* and for an object-form nominal element as the direct object of *said.* In that pair of dependency-specific requirements, there is no contradiction. But we can also see some challenges in (1). The direct object in English usually occurs immediately after the verb. In (1) it is initial, because it is extracted. The theory captures non-default word-order with the extractee relation. But extraction is a word-order relation: *said* takes its position from *what* because *what* is the only word that links it to the rest of the sentence, and *what* takes its position from *said* because it is extracted. This is a problem for current WG, because there is a transitive relation of word-order, which entails that in (1) each of *what* and *said* would take their position from themselves by transitivity, which is ill-formed. This is the problem I set out to resolve in this paper.

Hudson (1990: 385) extends the mutual dependency analysis of dependent interrogatives to free relatives, arguing that the syntactic structures of free relatives and dependent interrogatives are indistinguishable. Free relatives require a mutual dependency analysis because of their matching structure (Bresnan & Grimshaw 1978), in which the lexical category of the *wh* word determines their distribution, so in Section 5 I provide a revision to the theory of word-order in mutual dependency presented in Hudson (2007: 146). Although the paper is focused on English data, the theoretical claims and analyses are relevant to extraction data in other languages and there is a key argument from Farsi sluicing. My arguments in part draw on Hudson's (1990) original analysis of extraction, which in turn drew on ideas originally developed in Generalised Phrase Structure Grammar (GPSG; Gazdar et al. 1985). My analysis of free relatives is a dependency analogue of the 'fused relative' analysis in Huddleston & Pullum (2002: 1073–1074). On the other hand, in Section 4 I argue that dependent interrogatives do not require mutual dependency because the verb is the head and the interrogative properties of the clause can be accounted for by semantic selection.

In Section 2, I begin with a brief introduction to Dependency Grammar in order to introduce the formal constraints on word-order dependencies in WG; Section 3 discusses Word Grammar, the role of mutual dependency in WG and issues to do with word-order. As noted, in Section 4, I explore the *wh* constructions which motivate mutual dependency in the theory, concluding that dependent interrogatives do not require mutual dependency, but free relatives do. Then, Section 5 sets out a revision of WG theory which captures the necessity of mutually constraining structures while ensuring that word-order is negotiated without the introduction of conflicting information. Finally, Section 6 presents the conclusions.

2. Dependency Grammar and constraints on representations

Dependency grammar (DG) has a long history. Modern works include Anderson (1971), Debusmann (2006), Hudson (1984, 1990), Gerdes & Kahane (2001), de Marneffe et al. (2021), de Marneffe & Nivre (2019), Mel'čuk (1988), Osborne (2019), Osborne & Gerdes (2019), Sgall et al. (1986) and Tesnière (1953, 1959, 2015). The chapters in Ágel et al. (2003, 2006) present the state of the art at the time of their publication. As I noted above,

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

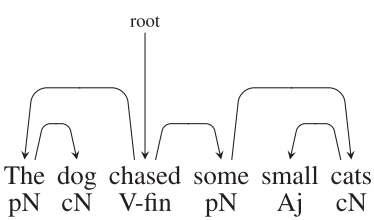
the main representational claim of DGs is that syntax consists of pairwise relations between words. One key question concerns how the formal constraints on the representation interact with the syntactic analyses. In some theories, the representation constrains the grammar directly, which is the case for GPSG. Head-driven Phrase Structure Grammar (HPSG; Pollard & Sag 1987, 1994), on the other hand, is Turing complete, which is to say that it has unrestricted computational power (Przepiórkowski 2023). This is not seen as a problem by HPSG theorists, who argue that the theorems of the grammar have to be restricted, not the formalism in which those theorems are stated (Pollard 1997, Müller 2023). As noted in Section 1, like HPSG, WG does not assume that the representation should constrain the theory, except in the case of word-order.

In WG, the dependencies associated with word-order are represented in a projective tree, i.e. a tree with no crossing branches (Hudson 1990: 114–120, Hudson 2003, 2007, 2018). The early formal literature on Dependency Grammars formalised the representations as rooted trees, with a single root node and with only one head permitted per node; through the axioms of the formal representation, these constraints enforce acyclicity, they enforce each word having only one head (i.e. only one incoming dependency), and they enforce connectedness, a necessary property of such graphs. This treatment is found in Hays (1964), Gaifman (1965), and Robinson (1970); see also Miller (1999: 20-22). Robinson (1970: 260) argues for a dependency representation where the dependency relation is 'transitive, irreflexive, and anti-symmetric'. She presents the axioms in (2) as the 'axioms of the theory which was […] formalized by Hays (1964) and by Gaifman (1965)'.

(2) (a) One and only one element is independent;
(b) all others depend directly on some element;
(c) no element depends directly on more than one other; and
(d) if *A* depends directly on *B* and some element *C* intervenes between them (in the linear order of the string), then *C* depends directly on *A* or *B* or some other intervening element.

Robinson's axioms define a DG as a rooted tree, with the constraint in (2d) also enforcing projectivity, otherwise known in the WG literature as the no-tangling constraint, which was originally introduced in Lecerf (1960). WG therefore argues that syntax is a network, with reticulations (i.e. mutual dependency), but that there is a projective sub-network that accounts for word-order. A tree as defined as in (2) is weakly equivalent to a Context-free Phrase Structure Grammar. (3) shows such a graph.[3]

(3)

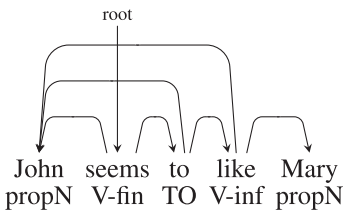


[3] pN = pronoun; cN = common noun; V-fin = finite verb; Aj = adjective; in the subsequent diagram, V-inf = infinitive form of the verb; propN = proper noun; TO = infinitival *to*. The classification of *the* shows WG's analysis of 'determiners' as transitive pronouns, see Hudson (1984: 90–92). The analysis of 'determiner' as head is the WG tradition so I retain it here for consistency's sake, although I note the arguments against the analysis in Pullum & Miller (2022).

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

Therefore, it is not the case that the WG network allows unrestricted networks as in (1). However, it is not only mutual dependency where the relationship between the larger network and the projective tree of word-order dependencies has to be worked out. The WG analysis of raising, control and other types of non-finite predication result in a directed acyclic graph, as in (4); here the theory permits nodes to have more than one incoming dependency relation. (4) presents the representation of subject-raising of Hudson (1984: 83), and assumes that *to* is a non-finite auxiliary verb.

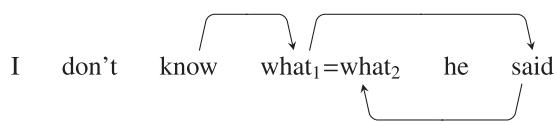
(4)



In the analysis in (4) *John* is the subject of *seem*, the subject of *to*, and the subject of *like*; for this reason, *John* has three incoming dependencies. The analysis still shows an acyclic graph, but it is more complex than the simple tree structure of (3) because it permits a single word to have more than one head, and because the arrows from *to* and *like* cross the 'root' arrow, which violates projectivity. Given that the representation involves *John* having more than one head, graphs such as (4) also require the theory to factor out the projective tree of word-order dependencies from the more complex structure.

There is an existing DG theory that solves some of the problems with mutual dependency identified in Section 1. However, for reasons that become clear below, I reject this approach. Tesnière (2015: 570–572) developed an analysis of relative clauses where he argued that relative pronouns had a double function, occupying two places in the graph. One way of thinking about this is to distinguish between morpho-syntax and morpho-phonology. If we extend Tesnière's view to subordinate interrogatives, we could argue that (1) has the syntactic structure in (5), where the formulation $what_1$=$what_2$ shows two syntactic words $what_1$, a complementiser, and $what_2$, an extracted pronoun, which are realised by a single morpho-phonological word pronounced /wɒt/.[4]

(5)



According to the analysis in (5), the interrogative complementiser, $what_1$ is selected by *know* and takes *said* as its complement. The pronoun, $what_2$ is the extracted direct object of *said.* This approach potentially solves the word-order/projectivity problem, inasmuch as it could be argued that *said* takes its position from $what_1$ with $what_2$ (not $what_1$) taking its position from *said* (as long as word-order can be analysed in such an abstract syntax).

[4] Tesnière treats the different syntactic parts as different morphological elements in French *qui*, with *qu* and *i* in different syntactic relationships; WG, however, has a strictly word-based syntax.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

The analysis in (5) is a 'fused word' analysis, where a single morpho-phonological word realises two syntactic words. There are contexts where such analyses are required. For example, the English form *we're* rhymes with *weir* for me and *where* for others but involves two syntactic words *we* and *are.* French *du* is the morpho-phonological realisation of the syntactic word *de* plus the syntactic word *le.* Such analyses are justified either by commutation or by direct replacement of the fused form with two full syntactic words. French *du* commutes with *de la* in the context of feminine nouns. English *we're* can be replaced by *we are* in all contexts. In the case of relative pronouns, Tesnière (2015: 571) justifies his analysis with a speculative, erroneous etymology: 'This analysis leads to the insight that the personal relative pronoun of the European languages is probably the result of a prehistoric agglutination between the invariable translative element and a variable anaphoric element'. In fact, the Latin relative pronoun descends from the use of indefinite–interrogative pronouns (Haspelmath 1997) in correlative clauses (Belyaev & Haug 2020) which are ambiguous between conditional and relative interpretations (Andrews 1975); see Pompei (2011) for discussion. Furthermore, there are much more recent Indo-European varieties that have relative pronouns co-occurring with complementisers; see (6), from Chaucer's *Boethius* (from Gisborne & Truswell 2017: 27, with their structural annotations).

(6) blisfulnesse is [$_{NP}$ that thing [$_{CP}$ [for whiche] that [$_{IP}$ alle thise othere thinges ben desired]]] (cmboeth,434.C1.226)

The example shows that relative pronouns and complementisers co-occur in English as recently as the late fourteenth century, which is incompatible with the claim that there was a prehistoric agglutination of them. Without independent linguistic justification for a fused word analysis, Tesnière's strategy solves a theoretical problem by recourse to an unwarranted abstraction which is not justified with linguistic facts; indeed, the linguistic facts undermine it.

In Section 3, I discuss the theory of word-order and how word-order dependencies relate to the wider syntactic network in earlier work on WG.

3. Word Grammar

As a dependency grammar, Word Grammar limits its formal syntactic ontology to nodes (words) and asymmetrical pairwise relations between them. However, the axioms of WG are not Robinson's in (2). First, the dependency relations of Word Grammar are typed in an inheritance hierarchy. This ensures that they are intransitive, not transitive: in *she likes brown dogs*, *brown* is the adjunct of *dogs*, not the direct object of *likes* by transitivity; otherwise the dependencies of Word Grammar are asymmetrical and irreflexive. As well as the basic intransitive dependency relation, it has proven necessary to define two antisymmetrical transitive relations for WG: a relation 'subordinate' (Hudson 1987) and the projective relation of word-order, the landmark relation (Hudson 2007), discussed below in Section 3.1.[5] As noted in §2, the two main departures from Robinson are that it is possible for a word to have more than one head, and that mutual dependency is permitted.

[5] These relations are antisymmetrical because the word at the head of the chain of these transitive relations is its own landmark or its own subordinate.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

Hudson (1984: 77) asserts that WG's dependency relations are asymmetrical because 'it is the head that provides the link between the modifier and the rest of the sentence'. Hudson (1984: 78) gives the criteria in (7) for heads, which I have rephrased in the first case.

(7) (a) word-order – the head determines the relative position of the dependent;
(b) possibility of occurrence – whether the modifier may, or must, occur at all rests with the head, and not vice versa;
(c) inflectional selection (traditional 'government') – the head decides which inflectional form of the modifier occurs;
(d) lexical selection (alias 'collocation') – the head selects a particular, formally specified, word (e.g. *decide* selects the preposition *on*);
(e) semantic structure – the head provides a structure in which the modifier fits (e.g. *brown* defines the colour of the eyes, rather than *eyes* defining the application-range of *brown*).

Hudson (1990: 106–107) presents a slightly different list of the properties of the head–dependent relation, presented here in (8) which I have rephrased and summarised.

(8) (a) The construct refers to a hyponym of the head;
(b) the semantic relation between a head and a dependent involves the head's sense and the dependent's (intra-mental) referent;
(c) dependents take their position from their head, particularly in terms of rules of adjacency – *in the autumn* cannot be **the in autumn* because *autumn* must take its position from *the* not from *in*;
(d) any rule that restricts the relative position of any dependent in terms of whether it occurs before or after the head refers to the head, including in complex cases such as the relative ordering of indirect objects and direct objects in English;
(e) the external syntactic relations of a construct are due to the head: the head is the distributional head;
(f) the set of possible complements of any given head is determined by the head, i.e. the head is the subcategorisand and the governor in Zwicky's (1985) terms. (Adjuncts are different.)

Criterion (8b) is a refinement of (7e); criteria (8c) and (8d) are refinements of (7a). The distributional criterion (8e) is new in 1990 as is the criterion in (8f). There is an extended discussion of headedness in Hudson (1987), a response to Zwicky (1985), see also Zwicky (1993) and Hudson (1993). In those papers, Zwicky argued that the properties of heads did not converge on a single word and Hudson argued that they prototypically did. The two key facts for WG theorising are: (i) head properties prototypically converge, but can be dissociated (Gisborne 2006); and (ii) given the definitions in (7) and (8), mutual dependency must be exceptional.

As the representation in (1) does not identify the distributional head, criterion (8e) cannot apply in the case of mutual dependency, unless there is an extension to basic dependency theory. However, because in WG dependencies are labelled and intransitive, not all cases of mutual dependency necessarily introduce contradictions. Hudson argues that the mutual dependency involves different relations, and as a result there are no contradictions. In his analysis of (1), *said* is the complement of *what* and *what* requires a finite verb, whereas *what* is the direct object of *said* and has to satisfy the constraints on direct objects. Mutual

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

government as mentioned in (7c) does not introduce contradictions at this level of specificity. Furthermore, WG's relations are defined to be intransitive, so even if (7c) applied in the same way in both cases, it is not the case that each word would govern each other and in turn by transitivity govern itself, as I noted in Section 1.

These criteria do not apply universally. Criterion (7d) cannot apply in cases such as the direct object of most transitive verbs, where the selection is typically just one of lexical category. The semantic criteria (7e) and (8b–c) also cannot apply in all cases: in the raising example in (4), *John* is not an argument of *seems* despite being the subject of *seems*. Without developing a theory of it, it is not clear whether an open question is a hyponym of the interrogative semantics of its *wh* word, or of its verb. Criterion (8b) cannot apply in the case of questions if the *wh* word is the head: what is the sense of *what*? The criteria for heads in (7) and (8) are not hard-and-fast rules; instead, they present a set of properties that allow us to define a number of asymmetrical intransitive relations which can apply mutually as long as the mutual relations are different relations. Note that Kahane (2025) reports that Word Grammar research has consistently applied distributional criteria for heads. He argues that there are three (heuristic) distributional criteria which establish heads, called the 'Positive distributional criterion with removal'; the 'Negative distributional criterion with removal'; and the 'Distributional criterion without removal'. I return to his first criterion when I discuss sluicing below. Meanwhile, we can use the criteria in (7) and (8) to help determine which word in a mutual dependency pair is the best candidate for the head when it could be either. In Section 4, I argue that in the case of (1), the best candidate is the finite verb. However, word-order is defined by a transitive relation, the landmark relation, which I return to in Section 3.1.

Hudson's (1984) treatment of mutual dependency relies on a restricted notion of dependency, where the words involved may be selected or licensed by each other and governed by each other as long as different relations are involved. However, we need to examine the syntax of word-order in these mutually constraining structures, which obliges us to examine how the theory of word-order relates to extraction.

3.1. Word-order, landmarks and dependency transitivity

Not all dependency theories include word-order within syntax: for example, Mel'čuk (2014: 21) argues that word-order should be factored out from dependency in Meaning-Text Theory. However, WG always has. Hudson (1990: 208) divides dependency relations into 'pre-dependents' and 'post-dependents': this classification captures the fact that English is a mixed word-order language. Later work (Rosta 1994, 1997, Hudson 2003) distinguishes between those dependencies that are implicated in word-order and those that are not. The approach is updated in Hudson (2007) and again in Hudson (2018).

The treatment of word-order in Hudson (2007: 146) partially dissociates word-order from dependencies. Rather than refining the theory of dependencies as positional or not, he introduces a new relation of word-order, the 'landmark' relation, which is then associated with the dependencies. He sets out to exclude the landmark relation from dependencies where it would cause contradictions. I quote the relevant material here.

- ***Ordinary dependencies****, which typically (but not always) carry meaning and other inter-word relations such as selection and agreement.*

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

- *A **'landmark'** relation, which is subdivided into 'before' and 'after' and is transitive; it is this transitivity that keeps all the words in a phrase together.*[6]

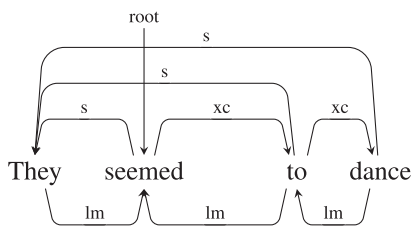
  *– A **'Parents are landmarks'** rule (51): a word's parent (the word on which it depends) is its landmark.*[7]

  *– A **'Subordinate parents are not landmarks'** rule […] which handles most cases where a word has two or more parents by blocking the landmark role on lower parents–i.e. which favours raising rather than lowering. This rule takes priority over the first rule, but may itself be overridden by specific rules for lowering.*
- *Some **extra dependencies** such as Extractee and Extraposee which always combine with ordinary dependencies and have the effect of 'raising' a word so that it depends on its grandparent (or higher ancestor) as well as its own parent.*

This theory works in terms of WG's theory of default inheritance. By default, heads are landmarks, but subordinate heads override that rule: they are not landmarks. In WG, overriding is bottom-up, so the more specific fact overrides the more general. Both raising and extraction invoke the Subordinate Head Rule which establishes the non-subordinate head as the landmark.

We can see how the landmark relation works in a simple case of raising by looking at the analysis of *they seemed to dance* in (9).

(9)



The diagram shows that *seemed* is the landmark of *They* & *to*, and *to* is the landmark of *dance*, but there are no landmark relations associated with the subject relations between *to* & *They* or *dance* & *They*, which are subject to the 'subordinate parents' rule. Although these subject relations violate projectivity by crossing the root vertex, as they are not associated with landmarks, this is not a violation of the premises of the theory. As the theory is developed in Hudson (2007: 130–151) landmarks are a property of certain dependencies in the following way:

- A subset of head–dependent relations is relevant to word-order. These head–dependent relations are associated with a landmark relation, where the landmark of a word is its head.
- The landmark relations make up a projective tree, as defined in (2) above.
- The landmark relation is transitive, so as to ensure contiguity of phrases.[8]

[6] Hudson defines two classes of landmark: the landmarks of head–dependent relation, and the landmark of sister ordering.

[7] Hudson has renamed heads with the less familiar term 'parent'. Rule (51) mentioned here is the rule that by default heads are landmarks (Hudson 2007: 132).

[8] The theory of landmarks is reminiscent of Gazdar & Pullum's (1981) factorisation of Immediate Dominance and Linear Precedence rules. Gazdar & Pullum (1981) are concerned with the expansion rules for phrasal categories and their linear precedence rules concern the relative ordering of phrasal categories (relative to the lexical category

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

This means that the theory of Hudson (2007) has mutual dependency in structures where mutual dependency does not involve any conflicting properties, but there is a projective sub-tree of word-order relations which does not permit mutuality: the theory of landmarks makes up a projective tree of transitive antisymmetrical relations as defined by Robinson (1970) quoted in Section 1 above (see also Hudson 2018: 110).

Here is the nub of the research challenge. The subordinate landmark rule asserts that there cannot be a landmark relation associated with the extractee relation. But the landmark relation is a word-order relation, and the extractee relation is an exceptional dependency which is only associated with word-order, offering the apparent paradox that in free relatives, the relation of word-order is not involved in the projective sub-tree of word-order relations. See (10).

(10) She wondered whom$_i$ he saw __$_i$.

In (10), *whom* must be extracted given that its form marks it as the direct object of *saw.* The Subordinate Head Rule says that it cannot take its word-order from the lower head, *saw*, but because it is extracted, *saw* is responsible for its position: this is a contradiction. I offer a solution in Section 5, where I argue that the Subordinate Head Rule applies, and that the specific nature of the extractee relation allows it (in exceptional constructions) to provide additional, compatible, word-order information without being the landmark.

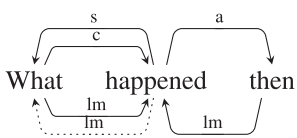
However, there is an alternative WG theory of word-order which we should consider. Hudson (2018: 109–113) rejects the 2007 theory and updates it with a number of theoretical innovations. Like the 2007 theory, landmarks define word-order (Hudson 2018: 109) but Hudson also introduces optionality so that in mutual dependency the choice of landmark is either optional or determined by the wider construction; this argument looks at data like those in (11).

(11) (a) What happened?
(b) What happened then?
(c) Then what happened?
(d) I wonder what happened then.
(e) *I wonder then what happened.

For Hudson, *What* and *happened* in (11a) are in a mutual dependency structure. He argues that either word can be the landmark of the other, so the possible landmark relations are alternatives.

The optionality of the landmark relation in this case is developed in the 'Best Landmark Principle' which claims 'the best landmark for an object is the nearest object that is more prominent' (see also Hudson 2010: 53). The dotted landmark relation in (11) is a rejected alternative.

(12)

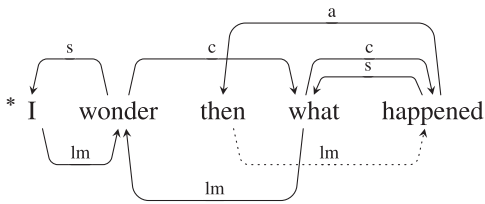


of the head) within a phrase. Word Grammar does not refer to phrasal categories, and the theory of landmarks is more concerned with linearising structures than linearising rules, which is the objective of Gazdar & Pullum (1981). That said, there is no scope within a word-based dependency grammar for linearising anything other than word–word pairings, as defined by a given dependency relation.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

The Best Landmark Principle accounts for (11b), diagrammed in (12), because *then* depends on *happened*, making *happened* more prominent than *What* and so the best potential landmark within the syntagm; the same explanation handles (11c); and it explains (11d) in similar terms, except that *wonder* is the best potential landmark; but, Hudson argues, it makes (11e) ungrammatical, because of a clash in potential landmarks as in (13).

(13)



Hudson (2018: 112) argues that in (13), '*what* mediates between *wonder* and *happened* … so it must be the landmark for *happened*'. Each word can have only one landmark. Hudson argues 'the fronted *then* excludes *what* as landmark of *happened*' by the same logic that applies to (12). We can see that *wonder* being the landmark of *what* excludes *happened* from being the landmark of *what* and the crossing landmark relations (one solid, one dotted in the diagram) violate projectivity, and so this structure is also ruled out.

This theory changes the nature of landmarks. In the 2007 theory, they are associated by default with ordinary dependencies, but delinked in specific syntactic structures. In the later theory, landmarks are assigned in an online operation, and not just inherited by default. The new theory does not provide a solution for a two-word structure such as *What happened?*: in this case, it just leaves the choice of landmark arbitrary. However, both 'complement' and 'subject' are defined as being associated with a landmark relation, so what brings about the cancellation of the landmark of one or other of those dependencies?

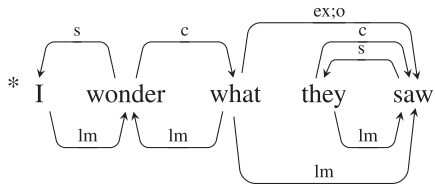
There is another problem. In English, some landmarks are clearly grammaticised: subjects take their heads as their landmarks, and must occur to the left of their heads. Verb-object bonding means that objects must occur immediately to the right of their heads with no intervening words. And extraction is positionally determined. An extractee does not occur in its canonical position.

While Hudson's (2018) Best Landmark Principle introduces variability to landmark assignment, its application to complex cases involving multiple potential heads and grammaticised positional facts, like extraction, remains challenging. An alternative approach, which I develop in Section 5, focuses on a more constrained system, where conflicts are resolved not by optionality in landmark choice, but by identifying an exceptional context which includes an exceptional local statement of word-order. The challenge for the theory of the Best Landmark Principle lies in working out how it interacts with these grammaticised facts. How can we ensure that the relevant head–dependent pairings are associated with landmarks, and that others are not? Let us take an example that is similar to *I wonder what happened* but which includes an unambiguous case of extraction.[9] The asterisk reflects the ungrammaticality of the proposed analysis in (14).

[9] s = subject; c = complement; o = object; ex=extractee; lm = landmark. The same arc is labelled as both extractee and object to reduce visual clutter.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

(14)



As the analysis in (14) shows, *what* is the extractee and direct object of *saw.* Given that the extractee relation is a word-order relation, the expectation might be that the landmark associated with the direct object has been overridden by the landmark associated with the extractee relation. However, following Hudson's analysis of (13), it must be *wonder* not *saw* that is the landmark of *what* in (14). The problem with that analysis is that the word-order facts of extraction entail that an extracted word must always take its position from the last head it is extracted from, and there is no place in the representation for that analysis. Given that each word can have only one landmark by projectivity, it is not possible for there to be a landmark relation between *what* and *saw* as well as *what* and *wonder*; but if *saw* were the best landmark, there would be no word-order rules connecting the subordinate clause with the main clause.

Therefore, the syntax of extraction is a challenge for the workings of the Best Landmark Principle. Given the problems I have noted, in the rest of this paper I set out to provide an alternative that is compatible with the Subordinate Head Rule. I argue that the verb is the head of dependent interrogatives, removing that construction from the set of problems to do with word-order, mutual dependency and extraction, and I accommodate the facts of extraction to the Subordinate Head Rule in free relatives, within an analysis that treats free relatives as exceptional.

4. Mutually constraining structures and *wh*-words

In this section, I investigate free relatives and dependent interrogatives. I conclude that it is possible to analyse dependent interrogatives without mutual dependency but that free relatives require a mutual dependency analysis, which itself requires a rethinking of the relationship between the landmark relation and extraction.

As the data below shows, free relatives have a different internal structure from dependent interrogatives. The head of a free relative, responsible for its external distribution and its internal structure, is the *wh* word. Dependent interrogatives, on the other hand, behave like clauses and require a different structural analysis.[10] In the case of an interrogative clause, the issue is whether the head is the *wh* word or the verb. I argue that it is the verb. The facts about free relatives mean that it is necessary to find a syntactic/theoretical solution to the problem of mutual dependency because the *wh* word is both the head of the whole structure and a dependent of the verb within the structure; it also shows the signature of extraction (Jacobson 1995). Interrogatives present a different analytical challenge: unlike free relatives they do not have the external distribution of the lexical category of their *wh* word, and therefore it is possible to analyse dependent interrogatives without mutual dependency. The rest of this

[10] I use the term 'clause' informally. There are no clauses in the WG ontology, because the theory is limited to word–word pairings.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

section presents the syntactic differences between free relatives and dependent interrogatives before turning to the structure of free relatives in Section 5.

The starting point is Bresnan & Grimshaw's (1978) presentation of the structure of free relatives, following Baker (1968), which showed that they have a 'matching' structure, where their distribution matches the distribution of the word class of the *wh* item. In (15a), *whatever you want to sell* has the distribution of a Noun Phrase, like *the turkey* in (15b) and likewise for the examples in (16); in (17b), *however tall his father was* has the same distribution as *six feet tall* and so forth for the remaining pairs of examples. The examples in (15)–(20) are taken from Bresnan & Grimshaw (1978: 335).

(15) (a) I'll buy *whatever you want to sell.*
(b) I'll buy *the turkey.*

(16) (a) I'll read *whichever book you give me.*
(b) I'll read *your book.*

(17) (a) John will be *however tall his father was.*
(b) John will be *six feet tall.*

(18) (a) She vowed to become *however rich you need to be to get into that club.*
(b) She vowed to become *very rich.*

(19) (a) I'll word my letter *however you word yours.*
(b) I'll word my letter *carefully.*

(20) (a) I'll put my books *wherever you put yours.*
(b) I'll put my books *anywhere.*

On the other hand, as the examples in (21)–(23) show, dependent interrogatives have the same distributions as dependent declarative clauses (modulo the semantic selection restrictions placed by verbs such as *wonder*).

(21) (a) I know *who won/that she won.*
(b) *Who won/That she won* is immaterial.
(c) It is immaterial *who won/that she won.*

The examples in (21a) show clausal complementation; those in (21b) show that the clauses are subjects; and the final examples show clausal extraposition. In (22), the clauses are complements in equative or specifying constructions and the examples in (23) make the point that both the *who* and the *that* clauses are the subjects of the following adjectives.

(22) (a) The main point is *who we have sufficient evidence to convict.*
(b) The main point is *that we have sufficient evidence to convict.*

(23) (a) I consider *who cheated* less important than why they did.
(b) I consider *that he cheated* less important than why he did.

Not all of these distributions are limited to clauses, but *it*-extraposition is a clausal property as the examples in (24) from Bresnan & Grimshaw (1978: 332) show.

(24) (a) *It* is not obvious *that you are six feet tall.*
(b) *It* is not obvious *whether you are tall enough*

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

(c) **It* is not obvious to me *six feet tall/a tall man/all the facts.*

It is possible to extrapose a dependent open interrogative as in (25).

(25) *It* is not obvious *how they got the answer.*

In contrast, like ordinary nominal structures (or Noun Phrases), free relatives cannot be extraposed in this way (Huddleston & Pullum 2002: 1069).

(26) (a) *What she suggests* is unreasonable.
(b) **It* is unreasonable *what she suggests.*

Further evidence of the difference between dependent interrogatives and free relatives is that the bound form *-ever* is restricted to free relatives (27);[11] it is possible for there to be *ever* in interrogatives, but this is a floating form (28). *Buy* selects a free relative; *inquire* selects an interrogative. The examples are taken from Bresnan & Grimshaw (1978: 334–5)

(27) (a) I'll buy what he's selling.
(b) I'll inquire what he's selling.
(c) I'll buy whatever he's selling.
(d) *I'll inquire whatever he's selling.

(28) (a) Who did he EVER kiss?
(b) What EVER is the matter with him now?
(c) What is EVER the matter with him now?

As (28b,c) show, this *ever* floats.

There is a further complication: Bresnan & Grimshaw (1978: 333) show that dependent interrogative clauses do not just behave like clauses; they also have the distributional properties of NPs. This is a property of dependent clauses; the lexical category of their *wh*-words is not relevant: they can have the word class of an adverb, as *how* does in (29d), but the clause still has the distribution of an NP.[12]

(29) (a) We didn't talk about *the muddy shoes* with Fred.
(b) We didn't talk about *whose shoes were muddy* with Fred.
(c) We didn't talk about *whether or not our shoes were muddy* with Fred.
(d) We didn't talk about *how muddy our shoes were* with Fred.

This distributional pattern is in addition to the clausal distribution of subordinate interrogatives.

The examples in (30) are taken from Huddleston & Pullum (2002: 1072) and introduce some further contrasts between free relatives and interrogatives.

(30) (a) *What ideas he has to offer* are likely to be half-baked. [FR]

[11] It is possible to have *whatever* in main clause interrogatives as in *Whatever are you doing?* and in dependent clauses in direct speech as in *I asked, 'Whatever are you doing?'* but **I asked whatever you were doing* is not well formed.

[12] This property is also arguably found with *that* clauses, at least vestigially. In the example *John knew Mr Colson but not that he worked at the White House* (from McCloskey 1979) a *that* clause is coordinated with a proper name. In the WG analysis, coordination requires sentence elements to be in like relationships to their head, not to be in like syntactic categories. From this, we conclude that *Mr Colson* and *that…* are both direct objects of *know.*

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

(b) *What ideas he has to offer* remains to be seen. [I]
(c) *What he's referring to/*To what he's referring* is Riga. [FR]
(d) I can't imagine *what he's referring to / to what he's referring.* [I]

(30a) and (30b) show that subject free relatives trigger different agreement patterns on the verb from subject dependent interrogatives. The free relative in (30a) triggers plural agreement, which indicates that it is plural. The dependent interrogative triggers singular agreement, indicating that it has a different internal structure from the free relative. This is evidence that the *wh* word is not the head in (30b). In theories where 'clause' belongs within the theoretical ontology, this would be a case of clauses triggering singular agreement. In a theory like WG, the only other candidate head is the verb, which triggers singular agreement. (30c) shows that free relatives do not permit pied-piping, whereas (30d) shows that dependent interrogatives allow it.

The pied-piping distinction addressed in Huddleston & Pullum's examples in (30c,d) was previously discussed in more detail by Bresnan & Grimshaw (1978: 342–3). Bresnan & Grimshaw observe that there are prepositions that cannot be stranded, as in (31), where *in* cannot be stranded when it is dependent on the noun *manner.*

(31) (a) John described the manner in which Dickens died.
(b) *John described the manner which Dickens died in.[13]

With a dependent interrogative, pied piping is acceptable, as in (32).

(32) (a) I'd like to know in what manner Dickens died.
(b) *I'd like to know what manner Dickens died in.

But with a free relative, neither the stranded nor the unstranded variant is possible.

(33) (a) *John will describe in whatever manner Dickens died.
(b) *John will describe whatever manner Dickens died in.

Bresnan & Grimshaw (1978: 344) argue that the apparent pied-piping in (34) is a case of the free relative occurring in a PP. They argue that the apparent extraction site is occupied by a Pro-PP which is deleted by an operation called 'Controlled Pro Deletion'.

(34) (a) I'll live in whatever town you live.
(b) I'll put my books in whatever cupboard you put yours.
(c) I'll open an account at whatever bank you open an account.
(d) I'll move to whatever town you move.
(e) I'll send my books to whatever bookstore John sends his.

Müller (1999) sees similar patterns in German as evidence of pied-piping. Such structures are part of the matching criterion of free relatives. Van Riemsdijk (2006: 352) gives pairs such as (35).

(35) (a) He'll remain in whatever town he has been living all his life.
(b) He'll remain in whatever town he has been living in all his life.

[13] However, note that Jim Donaldson (pc) says that this is grammatical for him.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

The implication of such examples is that pied-piping is possible when the pied-piped preposition is compatible with the head that licences or selects the free relative, but this is not an uncontroversial conclusion: Van Riemsdijk (2006: §4.1.2) points to the debate between Larson (1987) and Grosu (1996) about the correct analysis.

The conclusion from Bresnan & Grimshaw's (1978) evidence is that the head of the free relative is the *wh* word, because a free relative has the distribution of the word-class of its *wh* word. However, Jacobson (1995) argues that free relatives require a 'CP' analysis, which involves a treatment where the *wh* word is extracted, because of their structural similarities with other *wh* constructions, and I agree that there is extraction: the issue is how to resolve these apparently conflicting facts. Culicover & Jackendoff (2005: 323) review these two analyses as shown in (36), their example (37). In (36b), the head is empty, whereas in (36c), the extractee is a 'covert operator'.

(36) (a) What you see is what you get.
(b) *Wh moves; empty NP head:* $[_{NP}$ e $[_{CP}$ what$_i$ [you see t$_i$]]]
(c) *Wh is head of NP; covert operator OP moves:* $[_{NP}$ what$_i$ $[_{CP}$ OP$_i$ [you see t$_i$]]]

These analyses are forced by 'Interface and Derivational Uniformity'. On the other hand, in Simpler Syntax, 'there is nothing in principle to prohibit the trace from being bound to a *wh*-word that serves as head of the NP'. As I develop the analysis in Section 4.2 and Section 5 below, in WG there is nothing to prevent the *wh* word from being both the head of the free relative construction and the extractee. The analysis I adopt here, which is the dependency variant of the 'fused relative' analysis in Huddleston & Pullum (2002: 1073–1074), is that the *wh* word is both the head and extracted. In a declarative framework such as WG, for an item to be extracted is a statement of its relationship with other elements in the clause, and not an assertion that there has been movement.

A dependent interrogative, on the other hand, has a mixed pattern. The syntactic evidence reviewed in this section militates against the *wh* element being the head, because dependent interrogatives don't have the distributional properties associated with the lexical category of the *wh* word; nor do they trigger the agreement patterns that would be consistent with it being the head. But it is the interrogative *wh* word that carries the interrogative features of the clause. I argue below that dependent interrogatives are not syntactically interrogative, but that they are semantically selected. By taking the verb to be the head and the interrogative-selection pattern to be semantic, there is no mutual dependency and therefore dependent interrogatives can be removed from the discussion.

The free relative pattern requires some kind of mutually constraining structure: in (30a) *What* is not only the head of the free relative but is also the direct object of *offer* and extracted. I return to this problem in Section 5. For dependent interrogatives, we have to explore the remaining arguments that the *wh* word is the head. This is addressed in subsection 4.1.

4.1. Previous Word Grammar analyses

As I've noted, previous WG analyses (Hudson 1990, 2010) treat free relatives and dependent interrogatives in the same way. In both cases, it is argued that the *wh* word is the head of the construction, and also that it is the *wh* word that is responsible for the distribution of the construction. On the evidence just reviewed, this is the correct analysis for free relatives. The differences between free relatives and dependent interrogatives imply, however, that they should have a different treatment. If free relatives have the distribution of their *wh* word's

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

lexical category (as pronoun, or as adverb) and yet dependent interrogatives do not have such a distribution, the arguments for treating the *wh* word as the head of dependent interrogatives are weakened. In the case of a dependent interrogative, the *wh* word is a pronoun, not a complementiser, which is witnessed by its dependency relations with the dependent verb. A theorist who wants to develop the argument that the *wh* word is both head of the dependent verb and extracted from it has to develop a theory of the lexical category of the *wh* word that is compatible with the distributional facts. For these reasons, I take the view that a dependent interrogative is headed by the verb. Selection of an interrogative clause is semantic, not syntactic, and the verb-as-head analysis is supported by the agreement facts in the case of subject interrogatives – see the examples in (30) – and by the fact that the verb is the head of the other elements in the interrogative clause, including the *wh* word.

However, although the evidence just reviewed implies that it is not the correct analysis for a dependent interrogative, Hudson (1990: 362–365) has a number of arguments for treating *wh* as head in these constructions too, so I review these arguments here. He presents four main arguments that *wh* is the head in dependent interrogatives. One of these arguments is the similarity to free relatives, which I shall say no more about given the evidence in the previous section. The others are: sluicing (37a); the fact that *why* clauses cannot be non-finite (37b); and the selection of dependent interrogatives (37c).[14]

(37) (a) He was angry about something but I don't know what.
(b) I asked how/why I should do it *vs.* I asked how/(*why) to do it.
(c) He wondered if/whether he should eat/whether to eat/what to eat/(*that he should eat).

Hudson's sluicing argument involves Kahane's (2025) 'Positive distributional criterion with removal', according to which it is possible to identify the head–dependent relationship in a pair of items, A and B, distributionally. In Kahane's formulæ, U is a syntagm, and A and B are the atoms of the syntagm, so U=AB should be read as 'a syntagm consisting of A and B'. Given U=AB, if both A and B can occur independently, the head of AB is A if AB has the distribution of A. The head is B if it has the distribution of B. For example, in *I haven't seen him since we had that argument* both *since* and the clause *we had that argument* can occur independently of each other, but *I haven't seen him since* is grammatical, whereas **I haven't seen him we had that argument* is not. Therefore, *since* meets the requirements of this diagnostic heuristic for heads. Sluicing appears to fit the criterion; however, I explain below that the sluicing facts of Farsi show that sluicing is not adequate evidence that the *wh* word is the head.

As Culicover & Jackendoff (2005: 266) put it, 'Sluicing is typified by a bare interrogative phrase in a context where an indirect question would be expected'. As an argument for *wh* as head, it follows from a general property of heads and ellipsis: the claim is that if every dependent of *x* to the right of *x* can be elided, then *x* is the head. See (38), where *has*, *hasn't* and *to* are the heads of their respective phrases (in PSG terms).

(38) (a) Jane has written a book but Peter hasn't.
(b) I want to climb Kilimanjaro, do you want to?

[14] There is a further semantic argument; however, as Hudson himself describes this argument as 'relatively weak' and presents a counter-argument, I don't consider it here.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

For sluicing to show that *wh* is the parent of the dependent interrogative, the phenomenon of sluicing would itself need to be syntactic and it would need to be clear that the ellipsis worked in the same way as in (38). In some theories sluicing is syntactic. Such theories claim that it involves the deletion of syntactic material at the P(honetic) F(orm) level of representation in the Y-model of syntactic architecture. But the empirical facts do not force a syntactic analysis of sluicing: deletion is not the only possible approach in parallel architecture models such as WG (and HPSG and Simpler Syntax). Culicover & Jackendoff (2005) discuss sluicing in the larger context of other kinds of ellipsis, as well as contextualising it in terms of *wh*-movement. They evaluate a number of arguments for a syntactic view of sluicing (such as those found in Ross 1967, 1969, Merchant 2001, 2005) and conclude that the syntactic explanation runs into a number of difficulties which are avoided by the account that they develop in terms of Simpler Syntax.[15]

For example, it is possible for sluicing to be grammatical when a full version of the sentence would involve an island violation – an observation that goes back to Ross (1967). The example in (39), and its structural analysis, is taken from Rottman & Yoshida (2013: 651).

(39) They want to hire someone who speaks a Balkan language, but I don't remember which (Balkan language (*[$_{TP}$ they want to hire [$_{NP}$ someone [$_{CP}$ who speaks __]]])).

The fully spelled out version of the structure violates the Complex NP Constraint. From this, as well as from other examples that cause problems for a syntactic-deletion account such as the ungrammaticality of (40b) compared with (40a), Culicover & Jackendoff (2005: 267—268) conclude that sluicing is not only a syntactic phenomenon.

(40) (a) She went to the ball with a friend, but I don't know who with.
(b) *She put the crown on someone but I don't know who on.

Culicover and Jackendoff argue that sluicing must be partly syntactic, in that sluiced structures have the distribution of clauses, but there are also semantic constraints, and given the island violations, purely syntactic accounts will come unstuck.

There is evidence from Farsi, an SOV language with *wh*-in situ, which shows that sluicing does not demonstrate that the *wh* word is selected by the matrix verb and that it does not select the subordinate verb. See (41). Toosarvandani (2008) shows that these constructions are not pseudo-sluices of the kind discussed in Merchant's (2001) analysis of Japanese cleft-based sluicing-like structures.

(41) (a) sohrāb moz-o xord
Sohrab banana-OBJ eat.3SG
'Sohrab ate the banana'
(b) sohrāb chi-o xord
Sohrab what-OBJ eat.3SG
'What did Sohrab eat?' Farsi (Toosarvandani 2008: 692)

Farsi preserves the *wh*-in-situ structure in embedded questions, as in (42).

[15] See Nykiel & Kim (2021) for further arguments for treating ellipsis, including sluicing, without deletion.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

(42) hads bezan rāmin chi xarid
guess hit.2SG Ramin what bought.3SG
'Guess what Ramin bought.' Farsi (Toosarvandani 2008: 679)

Farsi permits sluicing, (43).[16]

(43) mahin ye chiz-i xaride vali be sohrāb ne-mige chi.
Mahin one thing-IND bought.3SG but to Sohrab NEG-say.3SG what
'Mahin bought something, but she didn't tell Sohrab what.' Farsi
(Toosarvandani 2008: 680)

It is not possible for the (in situ) *wh* word to be the head in a dependent interrogative construction in Farsi: it is too embedded and it would result in tangled landmark relations if it were. Given that Farsi has sluicing, sluicing is not evidence of head-hood. Toosarvandani (2008) shows that although some kind of *wh* fronting is possible in Farsi, it actually involves Focus movement, not movement to Spec,CP and can co-occur with a complementiser to its left, in which case it is not possible that the *wh* word is selected by the matrix verb or selects the dependent verb. Nouns can be focus-fronted as well as *wh* words. (44) shows that even if Farsi sluicing involves fronting, it does not involve classic *wh* movement. In the example, *ke* is a complementiser, and *chi* is the *wh*-word remnant in the sluice. Given this, *ke* must be the head.

(44) mahin mixād ye chiz-i bexare valie yād-esh ne-miyād ke
Mahin want.3SG one thing-IND buy.3SG but memory-her NEG-come.3SG that
chi
what
'Mahin wants to buy something but she doesn't remember what.' Farsi
(Toosarvandani 2008: 701)

The conclusion must be that sluicing is not evidence that the *wh* word is the head; anyone who wants to argue that sluicing is evidence that *wh* is the head in English is obliged to explain why it's not evidence in Farsi.

The second data point Hudson offers in favour of *wh* being the head of the dependent interrogative clause is found in (37b). The example shows that *ask* can be complemented by interrogative manner and purpose clauses, with *how* and *why* respectively. It also shows that *ask* can be complemented by both finite and non-finite clauses, but it cannot be complemented by a non-finite *why* clause. The argument that *wh* is the head assumes that it is the *to* infinitive clause that is not licensed by *why.* That is to say that it claims that *why* selects a finite clause complement and excludes a non-finite clause complement. But the facts are equally consistent with a different restriction: there are constraints on extracting *why* from lower clauses.[17]

Example (45) shows this. I've included *that* to force the correct structural interpretation. As the example shows, it is hard to get an interpretation where *why* is a modifier of *did.*

[16] The examples in (41) and (42) are Toosarvandani's. The variation between (41b), where *chi* has the object marker, and (42) & (44), where it doesn't, is not explained.

[17] There is a literature on the place of *why* in cartographic theories; for example Rizzi (2001) argues that Italian *perche* 'why' is base-generated high. See Shlonsky & Soare (2011) and the references therein for further discussion.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

(45) ?/*Why do you think that he did it __?

For me, at least, (45) is ungrammatical and therefore implies that *why* cannot be extracted from the lower clause.[18] Also note that *why* has a strong preference for being the first *wh* item in a multiple *wh* question.

(46) (a) ?/*What did you say why?
(b) Why did you say what?

This also suggests that *why* cannot be extracted. An alternative explanation of the absence of non-finite *why* clauses involves finding out the constraints on the extractability of *why*, including inter-speaker variation. It's not enough to take the facts about *why* as evidence for *wh* as head without further discussion of alternative analyses.

This leaves the argument from the selection of dependent interrogatives in which Hudson (1990: 363) argues from the selection facts of *wonder*, which selects for *whether*, *if* or a clause with an initial *wh*-proform. The issue of distribution is central, given the arguments about free relatives in Bresnan & Grimshaw (1978) reviewed above. Hudson's argument in full is given here.

> *It is the interrogative pronoun that links the rest of the clause to a verb such as WONDER, whose valency requires a subordinate interrogative clause; what is crucial to the definition of a clause as 'subordinate interrogative' is the presence of either an interrogative pronoun, or WHETHER or IF, rather than the presence of some particular kind of root verb. Thus in (20) the link between* wondered *and* could *is mediated by* what*, so the* what *must depend on* wondered *and be the head of* could.
>
> *(20) He wondered what he could do to help.*
>
> *It seems reasonable to generalize from such subordinate examples to all wh-interrogatives.*

This argument is plausible, especially considering a syntactic theory which is limited to word–word relations, but if we contextualise the data, the picture starts to look more complex. First of all, selection need not entail that the *wh* word is the head, because semantic selection is a possible alternative to syntactic selection. Secondly, syntactic selection need not entail that the *wh* word is the head, because an analysis is possible in terms of feature percolation.

The theory has to account for feature percolation because of examples like those in (47).

(47) (a) I wondered *how quickly* they ran.
(b) the arras *behind which* Polonius hid

In (47a), *how* is dependent on *quickly* but the interrogative feature on *how* percolates to *quickly*, which licenses the extraction. In (47b), pied-piping of *behind* is also licensed by feature percolation. In Word Grammar, there are various approaches to feature percolation, represented in (Hudson 1990: 367–375, 2007, 2018) and Rosta (1997: 68–77). We can set the details of the mechanism aside; what is at issue is that the theory has to accommodate the

[18] Cattell (1978) finds (45) grammatical, but its negated counterpart ungrammatical, and presents an analysis in terms of 'stance' verbs, which eschews a movement-based account. Jim Donaldson (pc) tells me that he agrees with Cattell's judgment. Dick Hudson (pc) tells me that *Why do you think he did it?* is ambiguous for him between a high and a low reading of *why*.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

phenomenon. Given the existence of the phenomenon, there is an open invitation to analyse dependent interrogatives as constructions where the interrogative feature has percolated from the *wh* word to the verb which is the clausal root: Hudson (2003: 638–639) offers this analysis, analogous to non-local feature percolation in HPSG (Pollard & Sag 1994: 164), of dependent interrogatives in Italian. He argues that there is feature percolation of the interrogative feature from the *wh* word to the verb, so a fully syntactic analysis with the verb as head is possible. This approach takes the verb as the head of the clause, and the *wh* item as the source of the interrogative feature: it is no more the head of the clause than relative *which* is head of the phrase *behind which* in (47b). Either way, a structure with an extracted *wh* word is different from a structure with *whether* or *if*, which are complementisers, as they are invariant and not associated with the gap.[19]

An alternative approach is to treat selection as semantic. There has to be semantic selection: compare *wonder*, *ask*, and *know* in (48).

(48) (a) I wondered {if it was time to go / whether it was time to go / what the time was / *(that) he ran away / *the time}.
(b) I asked {if it was time to go / whether it was time to go / what the time was / *(that) he ran away/ the time}.
(c) She knows {if he ran away / whether he ran away / when he ran away / (that) he ran away / the time}.

Syntactically, *wonder* is like *ask* except that it cannot take a direct object.[20] *Know* is the most permissive of the three verbs because semantically it permits both a declarative and an interrogative complement, whereas *wonder* and *ask* only permit interrogatives. An example which helps unpick semantic selection is *I asked the time.* Such constructions are known as 'concealed questions' (Frana 2017, 2021, Huddleston & Pullum 2002: 976). As Frana (2021) puts it, Grimshaw (1979) argues that concealed questions require there to be semantic selection as well as syntactic selection: *ask* selects for a question semantically and not a proposition; syntactically, unlike *wonder*, it permits a direct object; it is different from *know* which selects for both a proposition and a question. The three verbs oblige us to factor out semantic selection from syntactic selection. Chomsky (1995: 32–33), following Pesetsky (1982), runs a similar argument. Given that a direct object phrase such as *the time* is formally indistinguishable from any other noun phrase, examples such as *I asked the time* invite us to develop theories where the selection of interrogatives reduces to semantic selection.

In support of the semantic account, we should note that English Subject-Auxiliary Inversion isn't interrogative. Although polar interrogatives have this word-order, the same word-order can be emphatic as in (49a). The examples in (49b–d) show a clausal pattern where subject-auxiliary inversion is preceded by an adjunct which is a candidate for an extraction analysis. (49c) exemplifies rude negation and (49d) exemplifies a frequency construction with negative semantic prosody.

[19] A referee asks for an analysis of pied-piping. I refer the reader to the WG works mentioned in this paragraph. The earlier work assumes feature percolation; the most recent WG theory in Hudson (2018) replaces the feature percolation theory with a construction-specific relation that reassigns landmarks.

[20] A referee points out that *ask* is restricted in the direct objects it can take: you can *ask the time/the way/the reason* but not **ask the decision.* However, Frana (2021: 1) gives examples such as *Sam asked me the outcome of the election*, as well as several other examples involving other matrix predicates: the pattern is general.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

(49) (a) (And) did he (ever) win!
(b) Never did Dobby dream that Harry Potter would get to school another way!
(c) Bollocks did he.
(d) Hardly ever have I read beyond p.2 of *The Lord of the Rings.*

Furthermore, subject questions do not involve inversion and the properties of extraction including the island constraints are shared between open matrix interrogatives and topics. For these reasons, as Culicover & Jackendoff (2005: 309) also argue, rather than taking interrogative *wh*-clauses as syntactically interrogative, it makes sense to state that there are various overlapping constructional patterns, which are associated with the semantics of *wh* words when they are focused. After all, there is no sentence pattern that includes all open interrogative clauses and (49b–d) provide some evidence that auxiliary inversion and extraction are compatible with non-interrogatives.

Given that it is not possible to find a clause type that subsumes all open matrix interrogatives, we might wonder whether dependent open interrogatives have to be analysed as having a special interrogative structure and whether that requires them to be uniformly headed by their *wh* word. Like matrix interrogatives, dependent interrogatives share their surface structure with free relatives as we saw at the beginning of Section 4, as well as sharing structural patterns with headed relatives and indirect exclamatives. Consider (50).

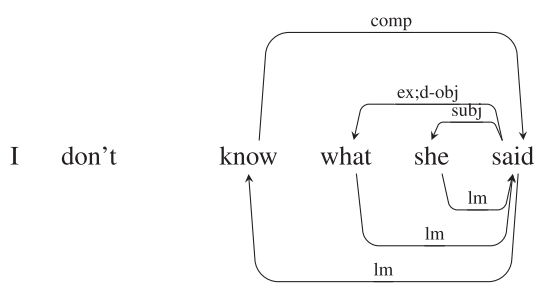
(50) (a) I saw *who she saw.*
(b) I asked *who she saw.*
(c) The person *who she saw* was eating a pear.
(d) She told them *what bores they were.*[21]

As it is not possible to find a clause type that is confined to interrogatives, there is no argument for asserting that selection of an interrogative is syntactic when an interrogative clause is selected by an external head: the head of an interrogative clause is its verb and the selection is semantic.

4.2. Summary

The discussion in this section has two main consequences: free relatives are headed by their *wh* word, but interrogatives are not. This conclusion means that the appropriate representation for dependent interrogatives has the analysis in (51) with the verb as head, not the one in (52), which is the traditional WG analysis.

(51)

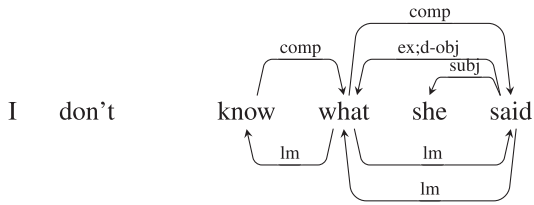


[21] Huddleston & Pullum (2002: 993). The particular constraints on indirect exclamatives require the use of a different *wh* word.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

(51) assumes that the interrogative properties of the clause are semantic not syntactic. The head is the verb *said.* With this decision made, the rest follows and there is no need for a mutual dependency analysis. The arc labelled *ex;d-obj* shows both the 'extractee' relation – the relation that governs extraction – and the direct object.[22] The diagram in (52) shows the traditional WG analysis in contrast.

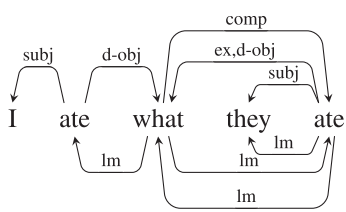
(52)



The example in (52) shows how the traditional WG analysis of *wh* dependent interrogatives with the *wh* word as head requires mutual dependency because of the loop between *what* and *said.*[23] It causes there to be mutual landmarks, which are ruled out by the theory's axioms, and which would cause a logical problem if they were not ruled out: by transitivity each word would be its own landmark, which is ill-formed. In the diagram, I have shown the landmark relation associated with the extractee relation and the landmark associated with the complement relation between *what* and *said.* By the Subordinate Head Rule, *said* would not be the landmark of *what*, but then there would be no grammar of extraction. However, the analysis in (51), unlike (52), is consistent with the syntactic facts reviewed in this section and it also extends to dependent subject interrogatives so we can set the analysis in (52) aside.

The outstanding issue concerns the structure of free relatives and the displaced object relation and the extractee relation, which form loops as in (53). The diagram in (53) shows a free relative where the *wh* word is the direct object of the verb within the free relative, and the subsequent analysis in (54) shows the situation where the *wh* word is the subject of the verb within the free relative. The analyses in (53) and (54) are impossible because of the mutual landmark relations.

(53)

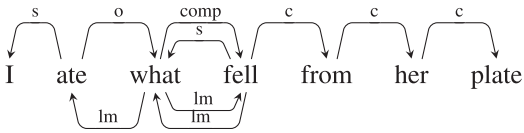


[22] A referee asks how I rule out, **I know she said what.* The answer is that *wh* words have a lexical specification that says that by default they're extracted. In English, this is a default because it is overridden in the case of multiple wh questions such as *I know what she said to whom*; in other languages, such as Bulgarian, multiple *wh* movement is required (Rudin 1988).

[23] Note that this is not the analysis for dependent clauses with *whether* or *if*, because those words are not extracted.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

(54)



From (53) and (54) we can see that free relatives raise two problems. In (53), there is a mutual dependency loop between *what* and *ate*, which involves an extractee relation and a displaced direct object. In (54), the mutual dependency loop between *what* and *fell* involves the adjunct dependency and the subject dependency. On the face of it, (54) is a significant challenge because it involves two positional dependencies involving landmarks. Although the discussion in this section has shown that there is no problem in the case of dependent interrogatives if we accept that they are semantically selected and that their head is the verb, free relatives continue to pose a research problem which I explore in the next section.

5. Free relatives: breaking the landmark cycle

If we set Hudson (2018) aside, WG's theory of extraction is most completely presented in Hudson (1990: 355–357), which essentially follows Gazdar et al. (1985).[24] The extractee relation is controlled by 3 rules: the start-rule, accounting for extractees being sentence initial; the end-rule, which matches the extractee with some postdependent such as the direct object; and the continuation-rule which negotiates the island constraints. It is the start-rule that is responsible for word-order. As has been widely pointed out, extraction does not require a *wh* item. As Culicover & Jackendoff (2005: 331) say, extraction requires the 'signature' of a long-distance dependency construction: it is also witnessed in both topicalisation and some analyses of *tough*-movement. Hudson (1990: 355) has different start-rules: his start-rule for relative pronouns states that a relative pronoun has a complement; that its complement is a tensed verb; that it occurs before its tensed verb. The start-rule for a topic says that a finite verb has optionally many extractees. The continuation- and end-rules are independent of the start rules.

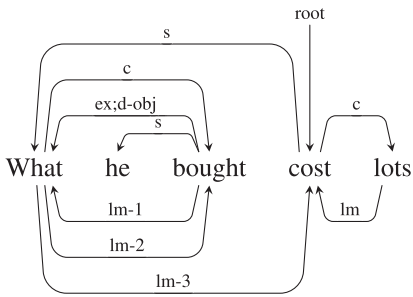
The solution to the mutual dependency problem in the case of free relatives involves defining dependencies in such a way that they do not conflict. (53) and (54) show that it is necessary to make a minor revision to the theory of word-order. In *I saw what they ate*, the Subordinate Head Rule entails that neither the direct object nor the extractee is responsible for the position of *what* but extractee is a positional dependency so there is a rule conflict; the complement relation is also positional, and without further theoretical reasoning, extractee and complement involve conflicting landmarks. The solution is to define the start-rule of extraction as a non-landmark positional dependency.

The analysis in (55) makes the problem more specific. It represents the analysis of free relatives in Section 4.2 above, where the *wh* word is identified as the head of the construction. As we can see, the initial problem with the analysis is that there is a mutual landmark relationship between *What* and *bought.* A secondary problem is that *What* has a second landmark, because it is the distributional head of the free relative.

[24] Hudson (1990), the extractee relation is called a 'visitor' because unbounded dependencies 'visit' successive heads; 'extractee' is the more usual later name.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

(55)



The problem presented by (55) is that the landmarks labelled *lm-1* and *lm-2* are in a mutual relationship, and because landmarks are transitive, by transitivity this makes each word its own landmark. The fact of *What* having two landmarks is a problem for reasons to do with the theory's axioms: landmarks are defined in such a way that they make up a projective tree. However, we know that *What* must have two heads: on the one hand, its position within the free relative is governed by its being extracted; on the other, the whole of the free relative clause has the distribution of the lexical category of *What*, which is a pronoun, so the free relative has the distribution of a noun phrase. The landmark relation between *What* and *bought* controls the internal structure of the free relative (i.e. the positional syntax within the clause) whereas the landmark relation between *What* and *cost* controls the distribution of the free relative.

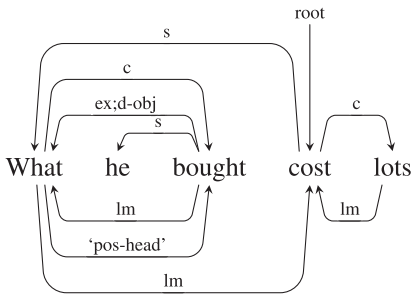
There is a straightforward solution which is to do with the extractee start-rule. As noted, this relation is exceptional in that it overrides normal word-order facts and in that it is not associated with any other morpho-syntactic or semantic properties. Unlike the subject relation, it doesn't trigger subject morphology on pronouns. Unlike the subject relation, it is not associated with an agent prototype in the semantics. Whereas by default we would expect the extractee relation to be associated with the landmark relation because it is only related to word-order, its exceptional properties imply that if there is an exceptional construction like the free relative construction, then it is the extractee relation that should be dissociated from the landmark relation: the landmark should be identified with the distributional head given Hudson's criteria for heads in (7) and (8) above, and the extractee relation is not associated with distributional syntax, i.e. the syntax of syntactic selection. This allows us to define an extraction-specific word-order relation, outside of the landmark system and with different properties from the landmark relation, which allows there to be two discrete relations of word-order between two words in a pair.

This move allows us to maintain the Subordinate Head Rule of Hudson (2007), while allowing extractee to be a word-order relation. The new positional relation is part of the start-rule of extraction in all cases. Let us call this relation the 'positional head'. By default, in the start-rule the positional head of an extracted word will be its landmark. But as long as the word-order properties of the landmark relation and the positional head relation are compatible, there won't be any contradictions if they are in a mutual relationship. With this relation in place, extraction can be subject to the Subordinate Head Rule, because the extracted word will still take its position from the nearest verbal head. But the positional head relation is outside the system of landmarks: like the start-rule of extraction, it is an intransitive word-order relation that relates the extracted word to its nearest head. Indeed, the start-rule of extraction is defined by this relation.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

This approach is represented in (56).

(56)



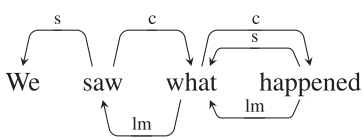
The word-order facts associated with the landmark relation and the positional head relation are mutually compatible, because complements follow their heads and extractees, by the positional head rule, precede theirs, and 'before' and 'after' are converses. The following propositions are all consistent with one another.

(57) (a) *bought* is the complement of *what*;
(b) therefore *What* is the landmark of *bought*;
(c) the complement landmark relation ensures that *bought* follows *what*;
(d) *What* is the extractee of *bought*;
(e) therefore, *bought* is the positional head of *What*;
(f) the positional head relation ensures that *What* precedes *bought.*

The problem caused by the Subordinate Head Rule in the case of extraction is solved by defining a locally specific, intransitive positional dependency which is part of the start-rule of extraction. There is an open research question that I return to briefly in the conclusion about whether this treatment is limited to the free relative construction, with its exceptional properties of matching, or whether it is also found in headed relatives.

There is, however, an outstanding question. Are subject free relatives accounted for by the mechanism of the Subordinate Head Rule or are they also extracted? This is a question about regularity: do subject free relatives align with all other free relatives, or do they align with other cases subject to the Subordinate Head Rule? The diagram in (58) shows a subject free relative without an extraction analysis.

(58)



However, from the point of view of resolving word-order conflicts in mutual dependency, there is no discussion to be had: neither analysis brings about mutual landmark relations, so I leave this question for further work.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

6. Conclusions

In Section 1, I observed that the issue of whether there could be mutual dependency or not depends on how heads and dependents are defined in the theory and on whether mutual dependency introduces any logical incoherence, or causes a contradiction. In Word Grammar it has always been held that mutual dependency is possible as long as word-order information is not involved. It is not possible to have mutual landmarks, and in earlier versions of the theory it was not possible to have mutual positional dependencies. However, the WG tradition has not taken into account the fact that the start-rule of extraction is only ever positional, so the fact that an extractee takes its position from its nearest head has to be taken into account in the theory of mutual dependency. In this paper I have set out to address the problem these facts cause for the subordinate subject rule of Hudson (2007) by exploiting Hudson's (2018) distinction between landmarks and the positional relations 'before' and 'after'. This was necessary for the analysis of free relatives. However, it is possible to treat dependent interrogatives as headed by their verb, given which there is no occurrence of mutual dependency in their syntax. I have argued that dependent interrogatives are selected semantically, not syntactically. I didn't pursue an alternative account where the dependent interrogative was headed by its verb, with feature percolation of the interrogative feature from the *wh*-word to the verb.

Given the treatment of dependent interrogatives, the locus of the problem moves to free relatives. Because of the matching constraint, free relatives have a number of exceptional properties. The key challenge of free relatives is that the *wh* word is both the distributional head and the extractee of the main verb. These facts require an analysis which factors out the positional information of the extractee start-rule from the projective tree of landmark relations. Such an approach treats the start-rule as exceptional but it does not involve an excessive burden of exceptionality given the exceptional structure of free relatives and it is consistent with their mutual dependency structure.

This, however, leaves headed relatives. If we take an adjunction approach to headed relatives and set the raising and matching theories aside, there are two possible analyses: either (i) headed *wh* relatives have the structure of free relatives, with the *wh* word being the dependent of the modified noun, or (ii) headed *wh* relatives have the structure of dependent interrogatives, with the verb being the dependent of the modified noun. I leave this to further research.

In this paper, I have made relatively small tweaks to Word Grammar. My main innovations have involved following the logic that it is not possible for there to be mutual landmarks and the argument that the start-rule of extraction must always be positional. But theories are not just their formalisms, they are also their canonical analyses, and I have therefore also revised the theory by arguing that the *wh* word is never the head of a question, and by arguing that all *wh* words on the left of their sentence have been extracted, including when they are subjects.

Acknowledgements. Thanks to James Donaldson and Brett Reynolds for running the seminar in honour of Geoff Pullum in August 2023 and to them and the other editors of this special issue for including this paper. Thanks to the audience at the seminar, three anonymous reviewers for *JL*, and Jim and Brett for their comments. Thanks to Dick Hudson for extensive discussion of earlier versions. None of these colleagues is responsible for any remaining errors. For many years, I taught syntax with Geoff whose influence is all over this paper: I am very pleased to be able to contribute to this volume. This work was supported by a Major Research Fellowship from the Leverhulme Trust, grant number MRF-2021-139.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

References

Ágel, Vilmos, Ludwig M. Eichinger, Hans Werner Eroms, Peter Hellwig, Hans Jürgen Heringer & Henning Lobin (eds.). 2003. *Dependency and valency: An international handbook of contemporary research. Volume* 1. Berlin: de Gruyter.

Ágel, Vilmos, Ludwig M. Eichinger, Hans Werner Eroms, Peter Hellwig, Hans Jürgen Heringer & Henning Lobin (eds.). 2006. *Dependency and Valency: An international handbook of contemporary research. Volume* 2. Berlin: De Gruyter.

Anderson, John M. 1971. *The grammar of case: Towards a localist theory*. Cambridge: Cambridge University Press.

Andrews, Avery. 1975. *Studies in the syntax of relative and comparative clauses*: Massachusetts Institute of Technology dissertation.

Baker, Carl LeRoy. 1968. *Indirect questions in English*. Urbana, Illinois: University of Illinois dissertation.

Belyaev, Oleg & Dag Haug. 2020. The genesis and typology of correlatives. *Language* 96(4). 874–907. doi:10.1353/lan.2020.0065.

Bresnan, Joan & Jane Grimshaw. 1978. The syntax of free relatives in English. *Linguistic Inquiry* 9. 331–391.

Cattell, Ray. 1978. On the Source of Interrogative Adverbs. *Language* 54(1). 61. doi:10.2307/412999.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Culicover, Peter & Ray Jackendoff. 2005. *Simpler syntax*. Oxford: Oxford University Press.

de Marneffe, Marie-Catherine, Christopher D. Manning, Joakim Nivre & Daniel Zeman. 2021. Universal Dependencies. *Computational Linguistics* 47(2). 255–308. doi:10.1162/coli_a_00402.

de Marneffe, Marie-Catherine & Joakim Nivre. 2019. Dependency Grammar. *Annual Review of Linguistics* 5. 197–218. doi:10.1146/annurev-linguistics-011718–011842.

Debusmann, Ralph. 2006. *Extensible dependency grammar: A modular grammar formalism based on multigraph description*: Universität des Saarlandes dissertation.

Frana, Ilaria. 2017. *Concealed Questions*. Oxford: Oxford University Press.

Frana, Ilaria. 2021. Concealed Questions. In Daniel Gutzmann, Lisa Matthewson, Cécile Meier, Hotze Rullmann & Thomas Zimmermann (eds.), *The Wiley Blackwell Companion to Semantics*, 1–24. Wiley 1st edn.

Gaifman, Haim. 1965. Dependency systems and phrase-structure systems. *Information and Control* 8. 304–337.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan Sag. 1985. *Generalized phrase structure grammar*. Oxford: Blackwell.

Gazdar, Gerald & Geoffrey K. Pullum. 1981. Subcategorization, constituent order, and the notion 'head'. In Michael Moortgat, Harry van der Hulst & Teun Hoekstra (eds.), *The Scope of Lexical Rules*, 107–123. Dordrecht: Foris.

Gerdes, Kim & Sylvain Kahane. 2001. Word order in German: A formal dependency grammar using a topological hierarchy. In *Proceedings of the 39th annual meeting of the association for computational linguistics*, .

Gerdes, Kim & Sylvain Kahane. 2011. Defining dependencies (and constituents). *International Conference on Dependency linguistics (Depling 2011), September 2011, Barcelona, Spain* 17–27.

Gisborne, Nikolas. 2006. Factoring out the subject dependency. In *Word Grammar: New perspectives on a theory of language structure*, 204–224. London: Continuum.

Gisborne, Nikolas & Robert Truswell. 2017. Where do relative specifiers come from? In Éric Mathieu & Robert Truswell (eds.), *Micro-change and macro-change in diachronic syntax*, 25–42. Oxford: Oxford University Press.

Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10. 279–326.

Grosu, Alexander. 1996. The Proper Analysis of "Missing-P" Free Relative Constructions. *Linguistic Inquiry* 27(27). 257–293.

Haspelmath, Martin. 1997. *Indefinite pronouns*. Oxford: Oxford University Press.

Hays, David G. 1964. Dependency theory: A formalism and some observations. *Language* 40(4). 511–525.

Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Hudson, Richard. 1984. *Word Grammar*. Blackwell Oxford.

Hudson, Richard. 1987. Zwicky on Heads. *Journal of Linguistics* 23(1). 109–132.

Hudson, Richard. 1990. *English Word Grammar*. Oxford: Blackwell.

Hudson, Richard. 1993. Do we have heads in our minds? *Heads in grammatical theory*. 266–291.

Hudson, Richard. 2003. Trouble on the left periphery. *Lingua* 113(7). 607–642. doi:10.1016/S0024-3841(02)00086-4.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

Hudson, Richard. 2007. *Language networks: the new Word Grammar*. Oxford: Oxford University Press.
Hudson, Richard. 2010. *An introduction to Word Grammar*. Cambridge: Cambridge University Press.
Hudson, Richard. 2018. Pied-piping in cognition. *Journal of Linguistics* 54(1). 85–138. doi:10.1017/S0022226717000056.
Jacobson, Pauline. 1995. On the quantificational force of English free relatives. In Emmon Bach, Eloise Jelinek, Angelika Kratzer & Barbara Partee (eds.), *Quantification in natural languages*, 451–486. Dordrecht: Kluwer.
Kahane, Sylvain. 2025. Hudson on Heads: About Distributional Criteria. In Eva Duran Eppler, Nikolas Gisborne & Andrew Rosta (eds.), *Word Grammar, Cognition and Dependency*, 163–177. Cambridge University Press 1st edn.
Larson, Richard. 1987. 'Missing prepositions' and the analysis of English free relative clauses. *Linguistic Inquiry* 18. 239–266.
Lecerf, Yves. 1960. Programme des conflits, modèle des conflits. *Bulletin bimestriel de l'ATALA* 4(5).
McCloskey, Jim. 1979. *Transformational syntax and model theoretic semantics: A case study in modern irish*. Dordrecht: Reidel.
Mel'čuk, Igor A. 1988. *Dependency syntax: Theory and practice*. Albany, New York: SUNY press.
Mel'čuk, Igor A. 2014. Dependency in language. In Kim Gerdes, Eva Hajičová & Leo Wanner (eds.), *Dependency Linguistics: Recent advances in linguistic theory using dependency structures*, 1–32. Amsterdam/Philadelphia: Benjamins.
Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.
Merchant, Jason. 2005. Fragments and ellipsis. *Linguistics and Philosophy* 27. 661–738.
Miller, Philip H. 1999. *Strong Generative Capacity: The Semantics of Linguistic Formalism*. Stanford, CA: CSLI Publications.
Müller, Stefan. 1999. An HPSG-Analysis for Free Relative Clauses in German. *Grammars* 2. 23–105.
Müller, Stefan. 2023. *Grammatical theory: From Transformational Grammar to constraint-based approaches*, vol. 1. Berlin: Language Science Press 5th edn.
Nykiel, Joanna & Jong-Bok Kim. 2021. Ellipsis. In Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The Handbook*, 905–950. Berlin: Language Science Press.
Osborne, Timothy. 2019. *A dependency grammar of English: An introduction and beyond*. Amsterdam: Benjamins.
Osborne, Timothy & Kim Gerdes. 2019. The status of function words in dependency grammar: A critique of Universal Dependencies (UD). *Glossa: a journal of general linguistics* 4(1). doi:10.5334/gjgl.537.
Pesetsky, David. 1982. *Paths and categories*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
Pollard, Carl. 1997. The nature of constraint-based grammar. *Linguistic Research* 15. 1–18.
Pollard, Carl & Ivan Sag. 1994. *Head-driven phrase structure grammar*. Stanford, CA: CSLI.
Pollard, Carl & Ivan A Sag. 1987. *Information-Based Syntax and Semantics, Volume 1: Fundamentals*. CSLI Lecture Notes. Stanford, CA: CSLI Publications.
Pompei, Anna. 2011. Relative clauses. In Philip Baldi & Pierluigi Cuzzolin (eds.), *New Perspectives on Latin Historical Syntax, Volume 4: Complex Sentences, Grammaticalization, Typology*, 427–547. Berlin, Boston: De Gruyter Mouton.
Przepiórkowski, Adam. 2023. LFG and HPSG. In Mary Dalrymple (ed.), *The Handbook of Lexical Functional Grammar Empirically Oriented Theoretical Morphology and Syntax*, Berlin: Language Science Press.
Pullum, Geoffrey K. & Philip Miller. 2022. NPs versus DPs: Why Chomsky was right. *LingBuzz archive* Paper no. 6845.
Rizzi, Luigi. 2001. On the position "Int(errogative)" in the left periphery of the clause. In Guglielmo Cinque & Giampaolo Salvi (eds.), *Current issues in Italian syntax: Essays offered to Lorenzo Renzi*, 287–297. Amsterdam: Elsevier.
Robinson, Jane J. 1970. Dependency structure and transformational rules. *Language* 46(2). 259–285.
Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
Ross, John Robert. 1969. Guess who? In Robert Binnick, Alice Davison, Georgia Green & Jerry Morgan (eds.), *Papers from the fifth regional meeting of the Chicago Linguistic Society*, 252–286. Chicago: Chicago Linguistic Society.
Rosta, Andrew. 1994. Dependency and grammatical relations. *Working papers in Linguistics* 64. 219–258.
Rosta, Andrew. 1997. *English syntax and Word Grammar theory*: University College London (University of London) dissertation.

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press

Rottman, Isaac & Masaya Yoshida. 2013. Sluicing, Idioms, and Island Repair. *Linguistic Inquiry* 44(4). 651–668. doi:10.1162/LING_a_00142.
Rudin, Catherine. 1988. On multiple questions and multiple WH fronting. *Natural Language and Linguistic Theory* 6(4). 445–501. doi:10.1007/BF00134489.
Sgall, Petr, Eva Hajicová & Jarmila Panevová. 1986. *The meaning of the sentence in its semantic and pragmatic aspects*. Dordrecht: Reidel.
Shlonsky, Ur & Gabriela Soare. 2011. Where's 'Why'? *Linguistic Inquiry* 42(4). 651–669. doi:10.1162/LING_a_00064.
Tesnière, Lucien. 1953. *Esquisse d'une syntaxe structurale*. Paris: Librarie C. Klinksieck.
Tesnière, Lucien. 1959. *Eléments de syntaxe structurale*. Paris: Librarie C. Klincksieck.
Tesnière, Lucien. 2015. *Elements of structural syntax*. 2015: John Benjamins.
Toosarvandani, Maziar. 2008. *Wh*-movement and the syntax of sluicing. *Journal of Linguistics* 44(3). 677–722. doi:10.1017/S0022226708005367.
Van Riemsdijk, Henk. 2006. Free Relatives. In Martin Everaert & Henk Van Riemsdijk (eds.), *The Blackwell Companion to Syntax*, 338–382. Wiley 1st edn.
Zwicky, Arnold. 1985. Heads. *Journal of Linguistics* 21. 1–29.
Zwicky, Arnold M. 1993. Heads, bases and functors. In Greville G Corbett, Norman M Fraser & Scott McGlashan (eds.), *Heads in grammatical theory*, 292–315. Cambridge: Cambridge University Press.

Cite this article: Gisborne, Nikolas. 2026. "Mutual dependency, English *wh*-clauses and Word Grammar." *Journal of Linguistics*. doi:10.1017/S002222672610125X

https://doi.org/10.1017/S002222672610125X Published online by Cambridge University Press