

What's what?

Nomi Erteschik-Shir

Ben-Gurion University

1 What is the status of gradience?¹

My purpose in this paper is to demonstrate that the source of graded acceptability judgments cannot be purely syntactic. Instead, such data is predicted by information structure (IS) constraints.

Since the early days of generative grammar, it was observed that acceptability often patterns as squishes (e.g. Ross 1971). Ross' explanation was couched in terms of the strength of the transformation, the strength of the construction (island) and the strength of the language. Danish was therefore considered to be a 'strong' language because it allowed extraction out of relative clauses, which were graded as 'strong' islands, by a 'weak' transformation such as *wh*-movement:

- (1) Hvad for en slags is er der mange børn der kan li?
what kind of ice cream are there many children who like

The idea that gradience is the result of the 'strength' of the processes (or constraints) involved has survived till today, particularly in some versions of optimality theory. Less sophisticated attempts at this type of theory were made in the 1990s: for example, the ECP was considered to be a stronger constraint and subjacency was considered to be a weak constraint. The violation of both these constraints together was predicted to result in a stronger grammaticality infraction than the violation of just one. This provided an explanation for the distinction between the examples in (2) and (3):

- (2) a ?This is the guy that I don't know whether to invite t
 b ?This is the guy that I don't know whether I should ask t to come to the party.
 c ?This is the guy that I asked whether Peter had seen t at the party.
- (3) a *This is the guy that I don't know whether t should be asked to come to the party.
 b *This is the guy that I asked whether t had seen Peter.

The hierarchy in (4) (Lasnik and Saito 1992, 88) illustrates that the strength of subjacency can be seen as depending on the number of barriers crossed. In the last example of the three subjacency is doubly violated, in the others it is only singly violated.

- (4) a ??What did you wonder whether John bought?
 b ?*Where did you wonder what John put?
 c *Where did you see the book which John put?

This shows that in principle it is possible to have, as the output of the syntax, sentences of various levels of acceptability. The number of constraints violated, as well as the strength of the relevant constraint/s will render different outputs.

However, as observed already in Ross (1971), the extraction data are complicated by the fact that not all processes of extraction render the same results. The sentences in (4) and (4b) are worse than the ones in (2). An attempt to explain why *wh*-movement can render worse outputs than relativization is to be found in Cinque 1990. Cinque demonstrates that extracted phrases which can be interpreted as being d-linked render superior results to those in which the extracted phrases cannot be interpreted in this way. The type of extraction

illustrated in (2) is more readily interpreted as being d-linked than simple *wh*-movement. D-linked *wh*-movement also improves the examples in (4) and (4b), as shown in (5):

- (5) a ?Which book did you wonder whether John bought?
b ??Which place did you wonder what John put?

D-linking depends on whether a contextual referent for the *wh*-phrase is available. Hence in a context in which a set of relevant books ((5a)) or a set of relevant places ((5b)) is available (4a) and (4b) should be as good as the examples in (5). Cinque builds this notion of referentiality into his syntax and predicts that when the context provides the required referent, the extraction should be perfectly acceptable. What is missing in Cinque's approach is an explanation of why referentiality should interact with syntactic constraints, such as subjacency, in this way.

On the basis of squishy data of this type I argued in Erteschik-Shir 1973 that extraction is completely determined by IS constraints, in particular that only focus domains are transparent for purposes of extraction. The intuition behind this idea was that potential focus domains are processed differently from non-focus domains in that gaps are only visible in the former.² In view of the fact that the availability of focus domains depend on context, the results will be graded according to the discourse into which the target sentence is embedded.

(6) provides an illustration:

- (6) a Who did John say that he had seen?
b ?Who did John mumble that he had seen?
c *Who did John lisp that he had seen?

(6b) is improved in a context in which 'mumbling' has been mentioned (e.g. following 'At our meetings everyone always mumbles'). (6c) is acceptable in a context in which it is known that John lisps. This is because such a context enables the main verb to be defocused

and consequently enables the subordinate *that*-clause to be focused. If intuitions are elicited out of context, judgments for sentences of this kind will depend on whatever context the informant happens to come up with. The examples in (7) illustrate other kinds of contextual factors that interact with focus-assignment and the concomitant acceptability judgments.

- (7) a ??What did the paper editorialize that the minister had done?
b *What did you animadvert that he had done?

(7a) would sound much better if uttered by a member of an editorial board, and (7b) probably can't be contextually improved due to the fact that highly infrequent items such as *animadvert* are necessarily focused. Contrastive contexts also interact with extraction judgments:

- (8) a ?Who did John SAY that he had seen? [=contrastive]
b Who did JOHN say that he had seen? [=contrastive]

Contrastive focus on the main verb, as in (8a), or on another constituent of the main clause, as in (8b), does not preclude focus on the subordinate clause. Therefore these sentences are fine with contrastive interpretations. The reason (8a) is slightly more degraded than (8b) is because it is harder to construe a likely context for it.

These examples illustrate that the positive response of informants is conditional on their ability to contextualize in such a way that the clause from which extraction has occurred is interpreted as a focus domain. In view of the fact that informants differ with respect to the contexts they are able to construct, the results across informants are predicted to be non-uniform.

A number of different syntactic solutions have been suggested over the years to account for such squishes in grammaticality. This type of solution does not, however, explain the gradience of the output, nor does it explain the contextual effects. Speakers'

judgments with respect to data of the kind illustrated in (6)-(8) is rarely stable: Differences are found across speakers and sometimes the responses of the same speaker change. This type of instability occurs whenever grammaticality is context dependent because the judgments in such cases are also context dependent. Therefore, if a sentence of this type is presented to an informant out of context, it is judged good to the extent that the speaker can imagine a context in which the verb is defocused. The lowest grade will be assigned to a sentence for which the particular informant does not come up with a context which improves it. No syntactic account of data of the type in (6)-(8), even if it can predict gradience, will be able to predict the contexts which improve acceptability. Syntactic constraints will therefore always fail empirically.

Let us examine if this is indeed the case: It has been suggested that the *that*-clauses following verbs of manner-of-speaking, other than *say*, are adjuncts rather than complements (e.g. Baltin 1982) and adjuncts are, of course, islands. Extraction is therefore predicted to be blocked. Formulating the constraint on extraction out of *that*-clauses as a syntactic constraint on extraction out of adjuncts predicts that with a particular verb extraction will either be perfectly good, or totally bad. Moreover, it will not allow for the influence of context in improving extractability.

Supporting the ‘adjunct’ analysis is the fact that the *that*-clauses which are argued to be adjuncts are optional:

(9) John mumbled/lisped/*said.

This correlation, however, also follows from an analysis in terms of IS: The verbs that require a complement are light verbs which do not provide a focus. Sentences with such verbs without a complement are ruled out because they do not contain an informational focus, a minimal requirement for any sentence.

An other reason that the oddness of the sentences in (6)-(8) cannot be due to a syntactic constraint is that not only must a level of acceptability be assigned, but an account of the context dependency of the grammaticality judgments must also be given. Such an account is provided by a theory of IS which accounts for the contextual properties of sentences.

I conclude that any phenomenon which varies with context among and across speakers cannot receive a syntactic account. An account in terms of IS is geared to predict this type of variation. It follows that syntactic constraints will always render ungraded results. A violation of a syntactic constraint will therefore be ungrammatical, a violation of an IS constraint will be open to contextual variation and will therefore result in gradience. There will be no weak syntactic constraints, only strong ones. Examples of violations of such 'real' syntactic constraints are shown in (10):

- (10) a *John eat soup. (agreement)³
 b *Eats John soup? (do)
 c *John likes he. (case)

If we adopt the proposal that violations of syntactic constraints cannot be graded, whereas violations of IS constraints can be, we can employ the presence of context-sensitive grammaticality squishes as a diagnostic for whether a syntactic or an IS constraint is involved: Whenever context interacts with acceptability, the constraint cannot be syntactic. The answer to the question posed in the title is therefore that IS constraints can generate graded output whereas syntactic constraints cannot.

It is often assumed that IS constraints must be universal since they are based on universal concepts such as topic and focus. This is only partially true and depends on how the particular language codes these basic concepts. Danish is a language in which topics are fronted whenever possible. In English, however, topics are generally interpreted **in situ**. This difference triggers a difference in the application of island constraints in the two languages.

Compare, for example, the following case of extraction out of a relative clause in the two languages (both are licensed by the IS constraint in (16) below):

- (11) a Den slags is er der mange der kan li.
 This kind icecream are there many who like
 ‘This kind of icecream there are many who like.’
- b ?This is the kind of icecream that there are many people who like.

(11)a is perfect in Danish and sentences of this sort are common. (11b) is surprisingly good in English in view of the fact that it violates the complex NP constraint, yet it is not considered perfect by speakers of English. In Erteschik-Shir 1982 I offer more comparative data and illustrate that the acceptability squish in English is exactly the same as in Danish, yet all the examples in English are judged to be somewhat worse than their Danish counterparts.

In Erteschik-Shir 1997, I introduce a theory of IS, f(ocus)-structure theory. F-structure is geared to interact with syntax, phonology and semantics and is therefore viewed as an integral part of grammar. Here I argue that this approach predicts gradience effects of various kinds. In section 2, I map out the theory of f-structure. Section 3 demonstrates the f-structure constraint on extraction. In section 4, I show that the same constraint which accounts for extraction also accounts for Superiority in English and the concomitant gradience effects. In section 5, I extend this account to explain different superiority effects in Hebrew, German and Danish. Section 6 provides a conclusion.

2 Introduction to f(ocus)-structure theory

The primitives of f-structure are topic and focus.⁴ These features are legible to both interfaces: In PF this shows up across languages as intonational marking as well as f-structure motivated displacements. At the interpretative interface it allows for the calculation of truth values.

Following Strawson 1964 and Reinhart 1981, I define topics as the ‘address’ in a file system

under which sentences are evaluated (Erteschik-Shir 1997). If truth values are calculated with respect to the topic, it also follows that every sentence must have a topic. Topics are selected from the set of referents previously introduced in the discourse, which correspond to cards in the common ground. Topics are therefore necessarily specific: they identify an element in the common ground that the sentence is about.

Whereas different types of focus have been defined in the literatures (e.g., information focus, contrastive focus, broad focus, narrow focus), I propose only one type of focus which functions to introduce or activate discourse referents. The different types of focus are derived in this framework by allowing for multiple topic-focus assignments. As an illustration, examine the contrastive f-structure in (12):

$$(12) \quad I_{\text{top}} \text{ read } \left[\left\{ \begin{array}{l} \text{a BOOK}_{\text{foc}} \\ \text{a magazine} \end{array} \right\}_{\text{top}} \right]_{\text{foc}}, \text{ (not a magazine)}$$

In cases of contrast, a contrast set with two members is either discursively available or else it is accommodated. In (12) ‘a book and a magazine’ form such a set. In view of the fact that this set is discursively available, it provides a topic (as indicated by the top marking on the curly brackets). One of the members of this set is focused (in this case ‘a book’) and in this way, the set is partitioned, excluding the nonfocused member of the set from the assertion. Since foci are stressed in English, stress is assigned to the contrasted element.^{5,6} Contrastive elements are thus marked as both topic (the discourse-available pair) and as focus (the selected element). Such an element can play the role of a focus in the sentence as a whole as it does in (12), and it can also function as a main topic, forming a contrastive topic; this is illustrated in (13):

$$(13) \quad \left[\left\{ \begin{array}{l} \text{TOM}_{\text{foc}} \\ \text{Bill} \end{array} \right\}_{\text{top}} \right]_{\text{top}} \text{ is handsome.}$$

(13) asserts that *Tom* (and not *Bill*) is handsome.

Not all f-structure assignments are equally good. (14) illustrates a well-known asymmetry: objects are harder to interpret as topics than subjects (in languages with fixed word order and no morphological marking of top/foc):⁷

(14) Tell me about John:

- a He is in love with Mary.
- b ??Mary is in love with him.

In view of the fact that this constraint figures prominently in languages such as English which have fixed word order, I propose that the reason for this asymmetry is that there is a preference for aligning f-structure with syntactic structure. The alignment is shown in (15):

(15) **Canonical** f-structure:

SUBJECT_{top} [... X ...]_{foc}

In other words, an unmarked f-structure is one in which syntactic structure is isomorphic with f-structure: Either the subject is the topic and the VP is the focus or there is a stage topic and the remaining sentence is the focus.⁸ It follows that a marked f-structure is one in which an object is the topic.

This section provided a brief outline of the discourse properties of f-structure.⁹ I now turn to f-structure constraints on syntax, constraints which provide a graded output.

3 F-structure constraints

As pointed out above, only focus domains are transparent for purposes of extraction. In Erteschik-Shir 1997 I argue that this constraint on extraction falls under a more general constraint on ‘I-(identificational) dependencies’ which include anaphora, *wh*-trace

dependencies, multiple *wh*-dependencies, negation and focus of negation, and copular sentences. What all these dependencies have in common is that the dependent is identified in the construction, either by its antecedent or by an operator. The constraint which governs I-dependencies is the **Subject Constraint**:

(16) An I-dependency can occur only in a **canonical** f-structure:

$$\begin{array}{c} \text{SUBJECT}_{\text{top}} \quad [\dots X \dots]_{\text{foc}} \\ | \\ \text{I-dependency} \end{array}$$

I-dependencies are thus restricted to f-structures in which the subject is the topic and the dependent is contained in the focus. The intuition behind this constraint is that dependents must be identified and that a canonical f-structure, in which f-structure and syntactic structure are aligned, enables the processing of this identification. In the case of *wh*-traces, for example, the trace must be identified with the fronted *wh*-phrase. The proposed constraint restricts such identification to canonical f-structures. The constraint is thus couched in processing terms in which f-structure plays a critical role.

Let us first examine how the constraint applies to the graded extraction facts in (6)-(8). In Erteschik-Shir and Rapoport in preparation, we offer a lexical analysis of verbs in terms of meaning components. We claim that verbs of speaking have a Manner (**M**) meaning-component. **M**-components are interpreted as adverbial modifiers, which normally attract focus. The **M**-component of ‘light’ manner-of-speaking verbs such as *say* is light, i.e., there is no adverbial modification, and the verb cannot be focused. **M**-components can be defocused contextually, enabling focus on the subordinate clause, which then meets the requirement on extraction, since, according to the subject constraint, the dependent (the trace) must be contained in the focus domain. It follows that, out of context, only *that*-clauses under *say* allow extraction. All the other manner-of-speaking verbs require some sort of

contextualization in order for the adverbial element of the verb to be defocused, thus allowing the subordinate clause to be focused. Extraction is judged acceptable in these cases to the extent that the context enables such a focus assignment.

The subject constraint, which constrains dependencies according to whether the syntactic structure and the f-structure are aligned in a certain way can, in cases such as this one, generate graded results. This is not always the case. Extraction out of sentential subjects is always ungrammatical and cannot be contextually ameliorated. (17) gives the f-structure assigned to such a case:

(17) *Who is [that John likes t]_{top} [interesting]_{foc}

In order to comply with the subject constraint, the subject, in this case a sentential one, must be assigned topic. Since dependents must be in the focus domain, they can not be identified within topics and extraction will always be blocked. Although the subject constraint involves f-structure, it does not necessarily render graded results. This is because the constraint involves not only f-structure but also the alignment of f-structure with syntactic structure. Sentential subjects are absolute islands because they are both IS topics **and** syntactic subjects.

4 Superiority

Superiority effects are graded as the examples in (18) show:

- (18) a Who ate what?
b *What did who eat?
c Which boy read which of the books?
d Which of the books did which boy read?
e ?What did which boy read?
f ?*Which of the books did who read?

Superiority therefore provides a good test case to demonstrate how gradience is predicted by f-structure theoretical constraints.

The answer to a multiple *wh*-question forms a paired list, as demonstrated in (19):

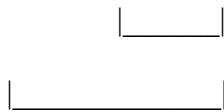
(19) Q: Who read what?

A: John read the Odyssey and Peter read Daniel Deronda.

Such an answer can be viewed as ‘identifying’ each object (answer to *what*) with one of the subjects (answer to *who*). In this sense the multiple *wh*-question itself forms an I-dependency in which one *wh*-phrase is dependent on the other.

Superiority effects are the result of two I-dependences in the same structure:

(20) *What did who read t



One I-dependency is between the fronted *wh*-phrase and its trace. The other one is between the two *wh*-phrases. As (20) illustrates, the dependent is identified in two different dependencies at once. This results in an interpretative clash, thus blocking the processing of the sentence.

The subject constraint is not violated, however, since the subject *wh*-phrase can be assigned topic (the question ranges over a discourse specified set; it is d-linked) and the trace can be analyzed as within the focus domain. Since it is not f-structure assignment which rules out the sentence, context should not have an effect on cases of superiority. This prediction is false, however, as shown by the following well-known example:

(21) I know that we need to install transistor A, transistor B, and transistor C, and I know that these three holes are for transistors, but I'll be damned if I can figure out from the

instructions **where what** goes! (Pesetsky 1987, from Bolinger 1978).

The answer to this puzzle lies in a proper understanding of the distinction between d-linked and non-d-linked questions, however not the one proposed in Pesetsky 1987 (see Erteschik-Shir 1986; Erteschik-Shir 1997). (22a) and (22b) illustrate a non-d-linked and a d-linked question, respectively, and the f-structure of each one:

- (22) a What did you choose?
 [What] did you_{top} [choose t]_{foc}
- b Which book did you choose?
 [Which book]_{top} [did you choose t]_{foc}

In the non-d-linked question in (22a), the fronted *wh*-phrase and its trace form an I-dependency and the trace is interpreted as an anaphor. Such a question must therefore conform to the Subject Constraint. In (22b), however, the fronted *wh*-phrase functions as a topic in that it ranges over a contextually available set (of books). The trace can therefore be interpreted on a par with a coreferent pronoun, since the set over which it ranges is discursively available. Since no I-dependency is defined, the subject constraint is not invoked, hence no superiority effects are predicted with *which*-phrases, which must be interpreted as d-linked. Questions with simple *wh*-phrases can be interpreted as being d-linked if the context provides a set over which they must range. That is why superiority violations such as (21) can be contextually ameliorated. They are always degraded, however. The reason is that both *wh*-phrases have to be interpreted as topics as shown in (23):

- (23) [Where]_{top} [[what]_{top} [goes t]_{foc}]_{foc}
- |_____|
- I-dependency

The subject *wh*-phrase forms an I-dependency with the trace in order to render the pair-list reading. The subject constraint on I-dependencies requires the subject to be a topic. The fronted *wh*-phrase must be interpreted as a topic because otherwise it will form an I-dependency with the trace which will then be doubly identified as in (21). Bolinger's detailed context allows for such an interpretation. The question will be viewed as degraded relative to whether the context forces a topic reading on both *wh*-phrases or not. Note that both *wh*-phrases must be interpreted as d-linked. *Which*-phrases are necessarily d-linked. Therefore, multiple *wh*-questions involving only *which*-phrases are perfect, as shown in (18d). When only one of the *wh*-phrases is a *which*-phrase, the other depends on context to receive a d-linked interpretation. This is why (18e) and (18f) are degraded.

The examples in (24) provide further evidence for the analysis of superiority effects proposed here. They illustrate that superiority effects also arise in single *wh*-questions when the subject is a nonspecific indefinite, i.e., a subject which cannot be interpreted as a topic:

- (24) a *What did a boy find?
 b (?)Which book did a boy find?
 c What did a certain boy find? ≈ (18e)
 d What did a BOY find?
 e What do boys like?

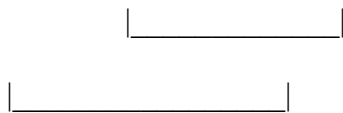
(24a) violates the Subject Constraint because the subject cannot be interpreted as a topic. In (24b), the fronted *wh*-phrase is d-linked and therefore does not form an I-dependency with its trace. It is degraded on a par with a sentence with an indefinite subject and a definite object as in (25):

- (25) a (?)A boy found the book.
 b A BOY found the book.

(25) is degraded because it is a non-canonical f-structure (cf. (15)) in which the object is the topic. Note that contrastive stress on the subject as in (25b) enables its interpretation as a topic, rendering a canonical f-structure. (24c), (24d) and (24e) do not violate the subject constraint because specific, contrastive and generic indefinite subjects are interpretable as topics.

Kayne's (1984) facts in (26) and (27) show that, surprisingly, an extra *wh*-improves superiority violations:

(26) What did [who]_{top} [hide t where]_{loc}



(27) Who knows what who_{top} [saw t]_{loc}



This is because the extra *wh*-phrase makes it possible to circumvent doubly-identifying the trace. In (26), for example, the fronted *wh*-phrase forms an I-dependency with the trace. This dependency is licensed by the subject constraint since the subject is interpreted as a topic and the trace is embedded in the focus. Another I-dependency is formed between the two remaining *wh*-phrases. This I-dependency is also licensed by the subject constraint since the dependent is embedded in the focus. The presence of the extra *wh*-phrase enables the formation of two separate I-dependencies without forcing a double identification of the trace as in the classic case in (20). This is how the extra *wh*-phrase saves the construction.

Although Kayne-type questions are an improvement on the classical case, they are still quite degraded. There are two reasons for this. First, the subject *wh*-phrase has to be

contextualized as ranging over a topic set (due to the subject constraint). Second, the integration of the two separate dependencies poses a heavy processing load: One I-dependency in (26) is between *who* and *where* allowing for the pair-list interpretation of these two *wh*-phrases. However, in order to process the question the fronted *wh*-phrase *what* must also be accommodated so that the interpretation of the question is that it asks for a ‘triple’ list-reading.¹⁰

The account of superiority effects proposed here thus affords an explanation of when context can improve acceptability and when it cannot and predicts the fine distinctions in acceptability evident in the English data.

5 Superiority in other languages

The account of the observed gradience in the English superiority data extends to other languages, once the nature of their canonical f-structures is determined. This section discusses Hebrew, German and Danish data and shows that superiority effects are determined by the same considerations as in English. Differences are due to variation in the application of the subject constraint which is in turn determined by the particular canonical focus structure of the language in question.

5.1 Hebrew

The first observation concerning Hebrew is that although topicalization may result in OSV, superiority violations are licensed only in the order OVS, as shown in (28a) and (28b) from Fanselow 2004:

- (28) a ma kana mi?
 what bought who
- b *ma mi kana?
- c mi kana ma?

(28a) is only licensed in a d-linked context in which a set of goods are contextually specified and (28c) requires a d-linked context in which a set of buyers are contextually specified. D-linking is not employed in Hebrew as a way to avoid double ID as it is in English.¹¹ The fronted *wh*-phrase therefore does not form an I-dependency with its trace. It follows that only one I-dependency is at work in Hebrew multiple *wh*-questions, namely the one that renders the paired reading:

(29) a mi kana ma

 |_____|

I-dependency

b ma kana mi

 |_____|

I-dependency

I conclude that the subject constraint is not operative in Hebrew as it is in English. This conclusion is also supported by the fact that adding a third *wh*-phrase not only does not help, as it does in English, but is blocked in all cases:

(30) a *mi kana ma eifo?

b *ma kana mi eifo?

c *ma mi kana eifo?

The subject constraint constrains I-dependencies to the canonical f-structure of a particular language. In English, the canonical f-structure is one in which syntactic structure and f-structure are aligned. The fact that the OVS and SVO orders of (28a) and c are equally good in Hebrew and that the OSV order of (28b) is ruled out may mean that it is the OSV word order which is the culprit. The difference between OSV and OVS in Hebrew is

associated with the function of the subject when the object is fronted. When it is interpreted as a topic, it is placed preverbally and when it is focused, it is placed after the verb. The examples in (31)-(33) demonstrate that this is the case:

(31) a et hasefer moshe kana.¹²
 the-book Moshe bought

b et hasefer kana moshe.

(32) a *et hasefer yeled exad kana
 the-book boy one bought

‘Some boy bought the book.’

b et hasefer kana yeled exad.

(33) a et hasefer hu kana.
 the-book he bought

b *et hasefer kana hu

(31) shows that a definite subject which can function as both a topic and a focus can occur both preverbally and postverbally. (32) shows that an indefinite subject which cannot be interpreted as a topic is restricted to the postverbal position. (33), in turn, shows that a subject pronoun, which must be interpreted as a topic, can only occur preverbally. (31)a and (33)a also require contextualization in view of the fact that both the topicalized object and the preverbal subject are interpreted as topics. Since every sentence requires a focus, this forces the verb to be focused or else one of the arguments must be interpreted contrastively. In either case the f-structure is marked. To complete our investigation of the unmarked f-structure in Hebrew, we must also examine the untopicalized cases:

(34) a moshe kana et hasefer/sefer

Moshe bought the-book/(a) book

b ?yeled exad kana et hasefer

boy one ought the book

The most natural f-structure of (34a) is one in which the subject is the topic and the VP or object is focused. (34b) with the definite object interpreted as a topic is marked.¹³ The results of both orders are schematized in (35):

(35) a *O_{top} S_{foc} V

b ?O_{top} S_{top} V

c O_{top} V S_{foc}

d *O_{top} V S_{top}

e S_{top} V O_{foc}

f ?S_{foc} V O_{top}

(35c) and (35e) are the only unmarked cases. I conclude that the unmarked focus structure in Hebrew is one in which the topic precedes the verb and the focus follows it. Hebrew dependencies therefore do not depend on the syntactic structure of the sentence, but only on the linear order of topic and focus with respect to the verb. The (subject) constraint on I-dependencies which applies in Hebrew is shown in (36):

(36) An I-dependency can occur only in a **canonical** f-structure:

$$X_{\text{top}} \text{ V } [\dots Y \dots]_{\text{foc}}$$

(36) correctly rules out (28)b and predicts that both (28)a and c are restricted to d-linked contexts (the initial *wh*-phrase must be a topic).

I conclude that multiple *wh*-questions in Hebrew are governed by the same considerations as they are in English. Differences between the two languages follow from

their different canonical f-structures.

5.2 German

According to many authors, German lacks superiority effects. Wiltschko 1998 not only argues that this is not the case, but also explains why German superiority effects have been overlooked. One of the reasons she offers is that controlling for d-linking is difficult since ‘discourse-related contrasts are often rather subtle’ (443). Along these lines, Featherston 2003 performed an experiment in which informants were asked to grade the data according to an open-ended scale. His results showed that superiority effects are ‘robustly active’ in German. It turns out, then, that German does not differ significantly from English in this respect. Fanselow 2004, although aware of Featherston's results, still distinguishes the status of English and German with respect to superiority effects. Fanselow points out that in German the superiority effect does appear when the subject *wh*-phrase is in Spec, IP: (his (35))

- (37) a wann hat's wer gesehen
 when has it who seen
- b ?*wann hat wer's gesehen
 ‘who saw it when?’

In (37a) the subject follows the object clitic, indicating its VP-internal position. In (37b), it precedes the object clitic and so must be outside the VP. This data is reminiscent of the Hebrew facts just discussed: German subjects in Spec,IP must be interpreted as topics, whereas VP-internal subjects are interpreted as foci.

D-linking is also required, as noted by Wiltschko. Fanselow 2004 gives the following illustration (his (42)):

- (38) wir haben bereits herausgefunden
 we have already found out

a wer jemanden gestern anrief, und wer nicht
who.nom someone.acc yesterday called and who.nom not

b wen jemand gestern anrief, und wen nicht
who.acc someone.nom yesterday called and who.acc not

Aber wir sind nicht eher zufrieden, bis wir auch wissen

But we are not earlier content until we also know

a' wer WEN angerufen hat
who.nom who.acc called has

b' wen WER angerufen hat

According to Fanselow, OSV order is licensed only if the object is discourse linked, but SOV order is also allowed in an out-of-the-blue multiple *wh*-question (his (43)):

(39) Erzähl mir was über die Party.

‘Tell me something about the party.’

a Wer hat wen getroffen?
who.nom has who.acc met?

b ??Wen hat wer getroffen

Fanselow's example cannot, however, be considered out-of-the-blue. A party necessarily involves a set of participants. These are what the *wh*-phrases range over in the questions following the initial sentence. Since both *wh*-phrases range over the same set of party-participants, they are equivalent. (39a), in which no reordering has occurred, is therefore preferred.

From this data I gather that the German canonical f-structure is similar to the one proposed for Hebrew, with only one small difference: German, too, requires that the first argument be the topic and the second be the focus, yet the status of the subject is determined

differently: German subjects are interpreted as foci when they are VP-internal, and as topics when they're not, as shown in (37). The position of the subject is transparent only in the presence of adverbials or other elements that mark the VP boundary.¹⁴ In many of the examples in which such elements are absent, the linear position of the subject *wh*-phrase gives no clue as to its syntactic position. In those cases, the subject will be interpreted according to contextual clues. An I-dependency is licensed between two *wh*-phrases in German when the first one is interpreted as a topic and the second as a focus.

5.3 Danish

According to Fanselow 2004, Swedish does not exhibit superiority effects (his (12)):

- (40) Vad köpte vem
what bought who

In Danish, the same question is degraded:

- (41) a Hvem købte hvad?
who bought what
b ?Hvad købte hvem?
what bought who

Overt d-linking significantly improves the question:

- (42) Hvilken bog købte hvilken pige?
Which book bought which girl?

Danish may have a preference for overtly marking d-linked *wh*-phrases instead of just depending on contextual clues. Danish is like English in this respect, except that the preference in English is even stronger. Danish differs from English in that superiority effects

in subordinate clauses are not ameliorated by overtly d-linked *wh*-phrases:

- (43) a *Jeg ved ikke hvad hvem købte
I know not what who bought
- b *Jeg ved ikke hvilken bog hvilken pige købte
I know not which book which girl bought

Danish generally marks the topic by fronting it to sentence-initial position. This is also the case if the topic is located in the subordinate clause. Topicalization within a subordinate clause is therefore excluded.¹⁵ It follows that whereas word order may signal the f-structure of the main clause, the order within subordinate clauses does not. This is the explanation I propose for the different behavior of Danish main and subordinate clauses with respect to superiority effects. Scrambling languages such as German differ: scrambling positions topics outside the VP in subordinate clauses as well as main clauses. No difference between main and subordinate clauses is predicted in scrambling languages. This prediction is borne out for German.¹⁶ Fanselow 2004 rejects the idea that the availability of scrambling is what explains the lack of superiority effects because there are non-scrambling languages which also lack superiority effects. I would not be surprised if non-scrambling languages exhibit the same difference between main and subordinate clauses for Danish.

Since the verb in Danish main clauses must appear in second position, the canonical f-structure is identical to the one proposed for Hebrew. The only difference between Hebrew and Danish is the preference for overtly d-linked *wh*-phrases.

What is common to the languages examined here is the need for d-linking at least one of the *wh*-phrases in multiple *wh*-questions. That is why such questions are always sensitive to context and therefore exhibit gradience. Variation among languages follows from three parameters: the canonical f-structure, the availability of topicalization and scrambling processes, and the array of *wh*-phrases available in a particular language. As I have shown

here, all three must be taken into account in order to predict the crosslinguistic distribution of superiority effects.

6 What's what?

In this paper, I have shown that f-structure constraints, which are sensitive to context, generally result in gradient output. Speaker judgments, which are generally solicited out of context, depend on how likely it is for a given informant to contextualize the test sentence appropriately. This will be hard if the required f-structure is marked or if accommodation is necessary. The ability of speakers to contextualize appropriately will also vary. It follows that gradience within and across speakers is to be found whenever grammaticality is constrained by f-structure principles.

I expect that the (subject) constraint on I-dependencies is universal and that its *raison d'être* is to enable the processing of the dependency. Sentences which exhibit a canonical f-structure are easy to process because they do not require complex contextualization. Dependencies also impose a processing burden. They are therefore restricted to structures which impose only a minimal processing burden themselves. Language variation follows from differences in canonical f-structure.

The answer to my question 'What's what?' is that gradience can only result when f-structure is involved. Violations of syntactic constraints necessarily cause strong grammaticality infractions, thus resulting in ungrammatical sentences. It follows that context-sensitive grammaticality squishes provide a diagnostic for whether a syntactic or a focus-structure constraint is involved: Whenever context interacts with acceptability, the constraint cannot be syntactic. There is therefore no need for 'weighted constraints' in syntactic theory.

¹ Thanks to Gisbert Fanselow and an anonymous reviewer for their comments and to Tova Rapoport, Sofie Raviv and the audience of the ‘Conference on Gradedness’ at the University of Potsdam for their feedback. This work is partially supported by Israel Science foundation Grant #1012/03.

² See Erteschik-Shir and Lappin 1983, p. 87 for this proposal which also provides an explanation for why resumptive pronouns salvage islands.

³ Timmermans, Schriefers, Dijkstra, and Haverkort 2004 argue that agreement involves both a syntactic procedure and a conceptual-semantic procedure which affects person agreement with Dutch and German coordinated elements which differ in person features. The former, according to these authors, ‘hardly ever derails’. This is what I have in mind here. The fact that nonsyntactic procedures are also involved in certain agreement configurations is irrelevant to the point I'm making here.

⁴ In Erteschik-Shir 1997 I assume that the output of syntax is freely annotated for topic and focus features. In Erteschik-Shir 2003, I introduce topic-focus features at initial merge on a par with ϕ -features in order to abide by the inclusiveness principle. The issue of how top/foc features are introduced into the grammar is immaterial to the topic of this paper.

⁵ See Erteschik-Shir 1997; Erteschik-Shir 1999 for an account of intonation in which f-structure provides the input to a stress rule which assigns stress to foci.

⁶ (12) also illustrates that not all constituents of a sentence need be assigned either top or foc features. Here the verb is assigned neither and the sentence could provide an answer to a question such as ‘Did you read a book or a magazine?’

⁷ See, among others, Li and Thompson 1976; Reinhart 1981; Andersen 1991; and Lambrecht 1994.

⁸ Sentences uttered out-of-the-blue are contextually linked to the here-and-now of the discourse. I argue in Erteschik-Shir 1997 that such sentences are to be analyzed as all-focus predicated of a ‘stage’ topic. The sentence *It is raining*, for example, has such a stage topic

and is therefore evaluated with respect to the here-and-now. All-focus sentences also have a canonical f-structure in which the (covert) topic precedes the focus.

⁹ I have included only those aspects of f-structure strictly needed for the discussion in this paper. See Erteschik-Shir 1997 for a more complete introduction to f-structure theory.

¹⁰ Triple dependences are not derivable in this framework, a desirable result since they do not render an optimal output.

¹¹ There is no parallel to a ‘*which*-phrase’ in Hebrew. ‘eize X’ is best paraphrased as ‘what X’.

¹² ‘et’ marks definite objects. *mi* (= ‘who’) in object position is most naturally marked with ‘et’ whereas *ma* (= ‘what’) is not. I do not have an explanation for this distinction.

¹³ (i), in which both arguments are indefinite, is interpreted as all-focus:

- (i) yeled exad kana sefer
boy one bought book
‘Some boy bought a book.’

In this paper all-focus sentences are ignored. For an account of such sentences within f-structure theory, see Erteschik-Shir 1997.

¹⁴ See Diesing 1992 for this effect.

¹⁵ Topicalization is licensed in subordinate clauses under a few bridge-verbs such as *think*. In such cases the syntactically subordinate clause functions as a main clause.

¹⁶ Hebrew is like Danish in this respect. Since Hebrew is not a scrambling language, this is what is predicted. Since English is not a scrambling language, English should also exhibit a difference between main and subordinate clauses. This is not the case:

- i) Which book did which boy buy?
- ii) I don't know which book which boy bought.

The difference between main and subordinate clauses in Danish arises because only in the former is f-structure marked by word order. English main clauses do not differ from

subordinate clauses in this way. This may explain why no difference in superiority effects between main and subordinate clauses can be detected.