The Empirical Challenge

Say we take the following empirical observations to have been established:

- extraction from surface subject position (specifier of TP) yields a measurable level of degradation in acceptability
- the level of degradation is not constant but rather seems to depend on the origin-site of the subject (along with other factors, such as D-linking and so on). Specifically:
  1. extraction from the subject of a transitive or unergative verb yields the most severe violation
  2. extraction from the subject of an unaccusative or passive verb yields an output which, while not fully well-formed, is not as ill-formed as the cases in (i) (though the amelioration effect is relatively small, judging by the presentation by Sturgeon, Polinsky, Kravtchenko et al in SynLab on March 5th last).

These observations/claims are illustrated by the examples in (1) (unannotated). The first example of each pair involves extraction of a DP, so runs the risk of the Cinque-Postal Confound.

(1)

a. this is a song that many versions of have caused offense
   this is a song of which many versions have caused offense
b. this is a song that many versions of have been recorded
   this is a song of which many versions have been recorded
c. this is a song that many versions of exist
   this is a song of which many versions exist

If these observations are confirmable, then we have a classic problem to address: how to understand gradient acceptability. Notice also that, if they are confirmable, they clearly suggest that in assessing deviance we must look, to some degree at least, at the derivational history of phrases which host extractions.
Analytical Background

The system of On Phases: does not allow a full understanding of why (1b) and (1c) are as degraded as they in fact are.

In that system, c interacts with the host phrase many versions of which only in its base position.

GOALS TODAY:

(i) exploit the fact that certain aspects of recent minimalism provide the kind of access to derivational history that seems to be needed for a full understanding of the observations in (1) to:

(ii) explore a simple-minded modification of that system which might allow us to capture the gradience seen in (1).

More specifically: take from much recent work the idea that movement (Internal Merge) results in the creation of structures of multi-dominance, structures in which a DP, for example, might simultaneously enter into the daughter-of relation with VP and with TP.

Consider (2):

(2) They were arrested

and its partial derivation in (3):

(3) a. [VP arrested [DP they]]
b. [VP arrested [DP they]]
c. [VP be [VP arrested [DP they]]]
d. [TP [past] [VP be [VP arrested [DP they]]]]

At that point:

(i) T enters into the case-agreement interaction with the DP they (rendering they nominative and T 3rd plural) and

(ii) assuming that T bears the EPP property (either inherently or by inheritance from C), the DP they will become a specifier of T, satisfying the EPP requirement of T.

CRUCIALLY: the DP they is both a specifier of T and a sister of the passive verb arrested. That is, there is a single syntactic object they which has two occurrences—one the specifier of T and the other the complement of V.

John Frampton suggests visual representations like that in (4) for such structures:
The derivation then proceeds to the point where a C which will trigger WH-movement is introduced (the next phase-defining head). It in turn will probe its command-domain (TP) seeking a WH-phrase to interact with. So let’s adjust (3) accordingly:

(4) \[ [TP \quad T_{\text{[PAST]}} \quad [vP \quad be \quad [vP \quad v \quad [vP \quad arrested \quad [DP \quad how \quad many \quad demonstrators \quad ]]]]] \]

After the derivation reaches state (4), C is introduced bearing the EPP property and whatever other features are implicated in WH-movement. This C probes the contents of TP, finds relevant features in the label of the DP how many protestors, enters into an Agree relation with those features and attracts a phrase whose label contains those features into its specifier position, yielding ultimately (5):

(5) How many protestors were arrested

CRUCIAL FOR OUR PURPOSES: When C probes, it interacts with a single syntactic object that is both a complement and a specifier.

It is almost unavoidable in this framework to consider the ‘derivational history’ of a host-phrase in assessing potential island violations. To avoid this conclusion, one would have to add a stipulation (only look at the highest occurrence, say, of a host phrase in assessing possible island violations).

With that in mind …
A Simple-Minded Proposal

Make the obvious moves:

**OBSOURED MOVE 1:** Take from 'On Phases' the intuition that probing **into** certain kinds of constituents (for Chomsky phase-edge material) 'comes at a cost' and cash it out in the following simple terms:

**OBSOURED MOVE 2:** Assign a mark of deviance on each occasion on which a probe interacts with a goal which is a proper subpart of a specifier or of an adjunct (the **CED**).

**OBSOURED MOVE 3:** Assume the stupid theory (since we don't know any better): The overall deviance of a structure is simply the sum of the marks of deviance on its constituent phases.

A QUESTION: what is the mark of deviance assigned to?

(i) to the entire derivation?
(ii) to a chain?
(iii) to one of the positions in a chain?
(iv) to the relevant phase?

Look again at our initial examples, repeated here as (6):

(6) a. this is a song that many versions of have caused offense
   this is a song of which many versions have caused offense
 b. this is a song that many versions of have been recorded
   this is a song of which many versions have been recorded
 c. this is a song that many versions of exist
   this is a song of which many versions exist

Then:

(i) (6a) will then be assigned two marks of deviance, since the raised element (a deleted **wh**-phrase or silent operator or whatever) is a proper subpart of the specifier of **TP** and a property subpart of the specifier of **VP**, since **cause** is presumably a transitive verb whose surface subject (external argument) originates in the specifier of **VP**.

(ii) (6b) will be assigned one mark of deviance since the raised element is a proper subpart of the specifier of **TP**, but of no other specifier or adjunct (given the assumption that the surface subject of a passive originates as a complement of the passive verb).

(iii) (6c) will similarly be assigned one mark of deviance on the assumption that **exist** is unaccusative.
A CONSEQUENCE: The system sketched will make (at least) a four-way distinction among:
- structures which are not computable at all
- structures which bear two marks of deviance
- structures which bear one mark of deviance
- structures which bear no marks of deviance

NOTE: 1 Of the four outcomes above, presumably the final three (but not the first) are assigned interpretations by the compositional process in the normal way (i.e. they should be meaningful)

NOTE: 2 The intuitive content of OBVIOUS MOVE 2 is the same as the intuitive content of Kayne’s (1983) Connectedness Condition—that in the search for a gap to connect with a filler, search flows most easily down head paths and into complements.

NOTE: 3 The general approach partially incorporates the intuition behind the Freezing Principle, in the sense that moved elements have the potential to produce cumulative deviance-scores (as happens in (6a)) more than un-moved phrases do. Movement at least typically targets specifier and adjunct positions. Further movements out of such positions, then, will likely invoke the mechanisms just described.

NOTE: 4 Pied Piping will frequently emerge in this system as a means of avoiding the cost (the ‘cedtax’) associated with OBVIOUS MOVE 2.

NOTE: 5 These distinctions might or might not correspond to what we actually observe when we measure carefully, but at least we have the beginnings of a way of determining degree of fit between model and data.

Extensions

One: Raising Predicates

(7)

TP

T

vP

v
[UNACC]

VP

PP
[EXP]

V

CP
[FIN]

(8) a. It seems to me that the crisis is easing.

b. The crisis seems to me to be easing.
What do we expect?

Subextraction from the experiencer argument should give rise to at least one deviance mark (with the possibility of tax-avoidance by way of Pied Piping).

(9)  a. Which of us does it seem to that the crisis is easing?
    b. To which of us does it seem that the crisis is easing?
    c. These are the companies of which it seems to employees that the crisis is easing.
    d. These are the companies to employees of which it seems that the crisis is easing.

(10) should be one step worse:

(10)  This is a club of which members seem to me to be dangerous.

And (11) one step worse again:

(11)  This is a club of which members seem to me to have caused a lot of harm.

because it involves extraction out of two specifier of TP positions. (This hardly seems believable.)

Double Object Verbs

(12)  \[ \begin{array}{c}
T \\
\Downarrow
\end{array} \quad \begin{array}{c}
TP \\
\Downarrow
\end{array} \quad \begin{array}{c}
T \\
\Downarrow
\end{array} \quad \begin{array}{c}
vP \\
\Downarrow
\end{array} \quad \begin{array}{c}
DP \\
\Downarrow
\end{array} \quad \begin{array}{c}
u \\
[ACC]
\end{array} \quad \begin{array}{c}
VP \\
\Downarrow
\end{array} \quad \begin{array}{c}
DP \\
[GOAL]
\end{array} \quad \begin{array}{c}
V \\
\Downarrow
\end{array} \quad \begin{array}{c}
DP \\
[THEME]
\end{array} \]

(13)  I gave the dog a bone.

(14)  a. I gave a supporter of this ballot-measure a large donation.
    b. this is a ballot-measure of which I gave a supporter a large donation
    c. they gave the author of this book a prestigious prize
    d. this is the book of which they gave the author a prestigious prize
Subjects of Small Clauses

Examples like those in (15), which involve extraction out of small clause subjects, should be similar to ?? and ?? in involving a one-strike level of deviance:

(15)  a. This is a song of which there are many versions available.
    b. This is a song of which I’ve heard many versions being performed.

Adjunct island condition violations, given this simple-minded theory, should also be on a par with ?? and ??, and they should also be on a par with (15).

Direct Objects—An Obvious Difficulty

There is a tension between what’s been said so far and the theory of objecthood that we find in *On Phases* and that we found some support for in our class discussion. Here the relevant probe is $v$ rather than $c$, and the object moves from a complement position (in the simplest cases) to the specifier position of $v$. Given that starting point, extraction out of a direct object should accrue a single mark of deviance at the $vP$-phase level (whose source is the movement out of the object’s second position, a specifier). This is probably not right. Examples like those in (16) have usually been reported as fully well-formed:

(16)  a. this is a song of which they’ve recorded many versions
    b. this is an office to which I own many keys

The system does draw a distinction between (16) and (17), the ECM or Raising to Object cases:

(17)  a. This is a song of which I believe many versions to be available.
    b. This is a politician of whom I consider supporters to be fascists

There is a long tradition (Kayne especially in ‘ECM Extensions’) of judging such cases to be more deviant than simple object subextractions. The system under discussion here does in fact make a distinction between (17) and (16). Pied Piping may once again avoid the relevant cost:

(18)  This is a politician supporters of whom I consider to be fascists.

We can, for better or for worse, go further. (19) should be yet worse than (17):

(19)  a. This is a politician of whom I consider supporters to have caused great harm.
    b. This is a club of which I consider members to have broken the law.

Here, the raised $wh$-$pp$ originates in three specifiers (in the case of (19b): Spec,$vP$ (*break*), Spec,$tp$ (*to*), and Spec,$vP$ (*consider*)) and so would accrue three marks of deviance at the $vP$-phase level.
Other Left Branch Extractions

(20)  
  a. Is he the guy that you stayed in a friend of’s apartment?
  b. This is a song that I’m not sure how many versions of have been recorded.

As things stand, these should be assessed one strike each (given the idea that the assessment of deviance is phase-bound). This is surely wrong, once we see past the Cinque-Postal confound.

(21)  
  a. Is he the guy of whom you stayed in a friend’s apartment?
  b. This is a song of which I’m not sure how many versions have been recorded.

Possible Responses to the Direct Object Problem

(i) Return to an earlier and more conventional theory of objecthood?
(ii) Add: Probe Goal interactions in which the goal is a proper subpart of the edge material of a completed phase ‘come at a cost’ (essentially the idea of On Phases)?
(iii) See how the data actually turn out when measured more carefully?

Lifting the Gaze

This is a simple-minded system, which surely needs significant refinement (at least). But it does follow out the implications of the initial observations about the Subject Condition (which seem sort of unavoidable). Maybe more important, it holds out the promise of a theoretical world in which we could begin to calibrate relative island-strength within a theoretical framework and then measure carefully how close the fit is between the theoretical model and our empirical measurements.
I. Revisiting the relationship between acceptability and grammar

1. The ‘inputs’ to acceptability are numerous. These may be plausibly distinct, though overlap in the expressions they target. E.g.:
   a. The status assigned an expression by the grammar
   b. The difficulty of parsing the string re: architectural constraints—issues of memory and control for parsers
   c. The difficulty of parsing the string re: available and intended interpretations
   d. Similarity of other expressions, edit distance
   e. Conditions of usage
   f. Conditions of elicitation—task instructions, task design

2. How do we reason backwards from an acceptability rating to grammatical status?
   a. Probably we will need to start conjointly measuring the relevant ‘sensations’ (e.g., acceptability and complexity, etc).
   b. Or manipulate what we might reasonably expect the contributions of n-factors are, but get something like a perception of similarity
   c. We can view the recent contributions of Hofmeister & Sag, & Sprouse + friends, to be pushing us in this direction

3. Notice there are a number of separate problems here
   a. How independent is (a)(i)-(vi)?
   b. How are their contributions weighted and ‘squashed’ onto the scales we use for measurement?
   c. Supposing we held (ii)-(iv) constant, now how do we interpret the magnitude of differences on the scales we use?

II. Example. Data from Anne Sturgeon’s presentation (Kravtchenko, Polinsky, Xiang, CUNY2009 poster):

4. What is the significance of the advantage unaccusatives enjoy?
   a. Note: we’re not talking about the reliability of a statistical test.
   Given that a reliable difference exists between unaccusatives and [transitive/unergatives], but that the difference is small (0.12 “ratings points”), what do we then say about the grammatical status of unaccusatives?

(10) English Judgment Task: (1-5)

(12) Russian Judgment Task (1-5):
5. A simple demonstration: difference on ratings scale does not map directly onto ‘mental distance’
   a. Suppose the grammar assigns status continuously, as below, with some noise. The ‘solid’ condition differs from the ‘dashed’ condition below by 2 ‘mental units’.
   b. Suppose participants map this variable onto a 5-point scale depending on how likely it is to be drawn from one distribution (‘solid’) versus the other (‘dashed’).
      In this example, the mean difference between conditions is ~3.

c. Now suppose we hold the grammatical status values constant, but add some noise.
    Unsurprisingly, responses become more spread out.
    The mean difference is reduced to ~2.
d. There’s nothing deep here, except that it can be difficult to reason about the relationship between an absolute difference on the ratings scale in the absence of better knowledge about the function that takes ‘mental distance’ and maps it to the scale.

e. What we seem to want is an ordering of types, and distance along some scale.

f. Why/how?

g. Explicit measures of contrast. Measures of similarity are one direction we could take our investigations.

h. This is the typical means of discovering the structure of representations [see Suppes, Krantz, Luce & Tversky, 1989; Levelt, 1974]. How could we do this?


(3) Soft constraints on extraction

a. DEFINITENESS (Def): a picture NP has to be marked [−DEFINITE].

b. VERBCLASS (Verb): a verb subcategorizing for a picture NP has to be marked [−EXISTENCE].

c. REFERENTIALITY (Ref): an NP extracted from a picture NP has to be marked [+REFERENTIAL].

(4) Hard constraints on extraction

a. INVERSION (Inv): subject and auxiliary have to be inverted.

b. AGREEMENT (Agr): subject and verb have to agree in number.

c. RESUMPTIVE (Res): resumptive pronouns are disallowed.

![Graph showing mean acceptability for soft and hard constraints](image-url)
III. Taking a step backward to go forward?

1. We remain relatively ignorant about how ratings relate to grammatical status (see above).

2. There have been a couple of ‘surprising’ results in the experimental syntax literature.
   + Featherston and Superiority in German
   + Jurka and CED in German
   + Lack of resumptive rescue (modulo the ECP cases)
   + The ‘slightness’ of the unaccusativity rescue
   + Non-robustness of satiation

3. A hunch/sense/intuition that the way rating tasks work is not the way linguists have traditionally (often) arrived at their data; which is by way of contrast and comparison.
   [With recognition of a potential feedback/bias problem here: how is the relevant contrast set selected? With the aid of the theory…]

4. Magnitude Estimation was, in a sense, advertised to work better because it is based on a comparison. But we know it doesn’t really work out that way. It seems to function covertly as a plain ratings scale.

5. Yet it seems odd to ask participants to judge the similarity of a pair of sentences.

6. The system we’ve sketched out orders sentences according to how many specifier mothers they have.

7. Conclusion
   Let’s try some potentially simpler for participants and which attempts to build in the benefits we’d like from MagE:
   Let’s ask our participants to order pairs/tuples of sentences.

8. Given sentence types \( a, b \), we ask which is the best statement:

\[
\begin{array}{ccc}
 a > b & a < b & a \sim b \\
 1 & -1 & 0 \\
\end{array}
\]

9. What would our data look like?
   a. How consistent are the rankings for a pair?
   b. Or rankings within a type? \( ? a_1 \sim a_2 \)
   c. What are the properties of rankings? Transitivity?
      \( a > b, \ b > c, \ a > c \)

10. \[
\begin{array}{cccc}
 a > b & a < b & a \sim b \\
 1 & -1 & 0 \\
\end{array}
\]

\[
\begin{array}{cccc}
 a & b & c & d \\
 0 & -1 & 0 & 0 \\
 +1 & +0.7 & 0 & 0 \\
 -0.7 & -0.3 & -0.3 & 0 \\
\end{array}
\]

Could this constitute the relevant measure of similarity/dissimilarity?
Cases to test

a. Basic cases: [Judith]

<table>
<thead>
<tr>
<th>T</th>
<th>UnE</th>
<th>P</th>
<th>UnA</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>transitive.subj</td>
<td>~</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unergative.subj</td>
<td>~</td>
<td>~</td>
<td></td>
<td></td>
</tr>
<tr>
<td>passivized.subj</td>
<td>&gt;</td>
<td>&gt;</td>
<td>~</td>
<td></td>
</tr>
<tr>
<td>unaccusative.subj</td>
<td>&gt;</td>
<td>&gt;</td>
<td>~</td>
<td>~</td>
</tr>
<tr>
<td>objects\text{predicted}</td>
<td>&gt;</td>
<td>&gt;</td>
<td>~</td>
<td>~</td>
</tr>
<tr>
<td>objects\text{observed}</td>
<td>&gt;</td>
<td>&gt;</td>
<td>&gt;</td>
<td>&gt;</td>
</tr>
</tbody>
</table>

b. Raising-to-subject

c. Experiencer objects of raising verbs

d. ECM/Raising-to-object [Mark]

e. Left-branch extraction [Sam]

Spec,CP extraction [Mark]

f. Internal VP arguments

[Goal [V Theme]] order

[Theme [V [P Goal]]] order

g. Small clause subject, predicate

IV. Some design concerns

An exhaustive experiment may not be possible. How many possible comparisons for \( n \) sentence types? [handshake problem]

\[(n-1) + (n-2) + ... + 1\]

4: 6 comparisons
5: 10 comparisons
6: 15 comparisons!!
7: 21 comparisons!!