Processing Ellipsis: A Processing Solution to the Undergeneration Problem?

Lyn Frazier
University of Massachusetts, Amherst

The study of ellipsis has exploded in recent years, especially the study of Verb Phrase Ellipsis (VPE) in (1) and Sluicing in (2).

(1) Josh laughed and Bella did __ too. (Empty VP following aux., *not,* infinitival *to)*

(2) Lucy bought something but I don't know what [ ].

Since the pioneering work of Sag (1976) and Williams (1977), syntactic approaches to the grammar of these structures have typically assumed that, morphological details aside, the clause containing the elided constituent and the clause containing the antecedent must match at LF, modulo any focused constituents in them (Rooth, 1992). Hence, examples like (3) are unambiguous, exhibiting only the parallel reading in (3a) not the nonparallel reading in (3b), and examples like (4) are unambiguous due to the lack of syntactic matching.

(3) The pig was ready to eat and George was __ too.
   a. 'The pig was ready to eat something and George was ready to eat something too.'
   b. 'The pig was ready to be eaten and George was read to eat something too.'

(4) *The pig was caught and John did.
   'The pig was caught and John caught it.'

Assuming that VPE targets a constituent above the Voice Phrase (See Merchant ms. for an alternative view, and Appendix 1 for discussion.) the matching condition requires the clause containing the antecedent and the clause containing the ellipsis to show the same voice feature. Both must be active or both passive. However, many attested examples of VPE exhibit syntactic mismatches, such as the voice mismatch in (5), as has been emphasized by Webber (1978), Hardt (1993), and Kehler (2002).

(5) This information could have been released by Gorbachov, but he chose not to. (Daniel Shorr, NPR, 10/17/92, reported by D. Hardt)

Based on such examples, some investigators have urged a semantic approach to ellipsis where an elided constituent only needs a salient semantic antecedent (Dalrymple et al., 1991; Hardt, 1993; Hardt and Romero, 2004). Others, in particular Kehler (2002) have proposed an elegant solution based on the idea that the syntactic matching requirement between antecedent and elided constituent comes into play only for examples with a particular (Resemblance) discourse coherence relation but not for examples exhibiting other (e.g., Cause-Effect) discourse-coherence relations (but see too Frazier and Clifton, 2006, who report experimental studies where manipulating discourse coherence relations fails to have the expected effect, but directly manipulating parallelism does influence judgments).

* Acknowledgements. I am very grateful to Chuck Clifton, Kyle Johnson, Angelika Kratzer, Jason Merchant, Chris Potts and Ivan Sag for discussion of the ideas presented here.

 Proceedings of the 26th West Coast Conference on Formal Linguistics, ed. Charles B. Chang and Hannah J. Haynie, 21-32. Somerville, MA: Cascadilla Proceedings Project.

Here I will pursue a syntactic approach to ellipsis (see Johnson, 1996, for overview). I will assume both that there exists syntactic structure in the ellipsis site (Merchant 2005, to appear) and that (morphology aside) the syntax of the antecedent and the syntax of the elided constituent must match at LF. The assumption that the elided constituent is fully articulated syntactically explains, for example, why in (6a) the goal of *buy* cannot be expressed by a *to*- phrase even though goals often are expressed in this manner. It explains the correlation between preposition stranding in ordinary questions and in sluicing (Merchant, 1999), and readily explains generalizations about examples with traces (Fiengo and May, 1992; Kennedy, 1997), scope (Lasnik, 2001) and examples with missing antecedents where the syntax is needed in order to identify the intended interpretation (Grinder and Postal, 1971). See too Frazier and Clifton, 2005, and Martin and McElree, 2007 for psycholinguistic evidence.

(6) a. *Max bought some presents and Fred did __ to Mary.
b. Max bought some presents and Fred did __ for Mary.

Despite the evidence favoring a syntactic approach to ellipsis, a syntactic approach requiring syntactic matching between the antecedent clause and the ellipsis clause clearly undergenerates, as shown by Webber (1978), Dalrymple et al. (1991), Hardt (1993), Kehler (2002) and others. In this paper, I will pursue a processing approach to the undergeneration problem. The proposal is based on the assumption that it is the most explanatory grammatical theory of ellipsis, together with the most explanatory theory of processing, that must account for actual data. Intuitions of native speakers reveal whether sentences are 'good' or 'bad.' But these pre-theoretic judgments do not themselves determine whether a sentence is judged 'bad' because it is ungrammatical or whether it is deviant in some other way, e.g., it is impossible to parse. It is the best overall theory of language that will determine the status of sentences. I will argue here that some sentences judged 'good' are strictly speaking ungrammatical but nevertheless useable and relatively acceptable (for early discussions of acceptable ungrammaticality see Otero (1972) and Langendoen and Bever (1973)). Before turning to the evidence for this account, I will discuss the outlines of the processing account of VPE that has emerged from recent work (see Tanenhaus and Carlson, 1990, Mauner et al., 1995, Carlson, 2002 and references therein for earlier work).

1. Processing assumptions

Various studies indicate that at the elided verb phrase, there is cost-free copying of the antecedent or, equivalently, structure sharing. For example, Frazier and Clifton (2001) report a self-paced reading study where sentence like (7) were presented frame by frame (slashes in (7) indicate presentation frames). Elided verb phrases with bigger antecedents (7b) took no longer to read on the final frame (Tina did too) than elided verb phrases with a smaller antecedent (7a) (despite significant differences in reading times elsewhere in the experiment, which thus show the sensitivity of the experiment). See also Frazier and Clifton, 2001, 2005, and Martin and McElree, submitted for further evidence.

(7) a. Sarah left her boyfriend last May. / Tina did too.
b. Sarah got up the courage / to leave her boyfriend last May. / Tina did too.

Frazier and Clifton (2005) argue that antecedents which are part of the main assertion of an utterance are preferred, especially across sentence boundaries. Thus, in a written paraphase selection study, the probability of participants reporting the matrix antecedent (said Fred...) for the elided verb phrase in the final clause was higher when the antecedent was in a separate sentence from the elided constituent, as in (8b), than when it was in the same sentence (8a).

(8) a. John said Fred went to Europe and Mary did too. 40% matrix VP antecedents
b. John said Fred went to Europe. Mary did too. 55% matrix VP antecedents

In a self paced reading study, the answers to questions indicated that participants chose the main clause verb phrase as antecedent three quarters of the time, regardless of whether the main clause occurred

before (9a) or after (9b) a subordinate clause. This is expected if listeners and readers prefer antecedents that are part of the main assertion.

(9) a. If Mary laughed after she made a joke about the supervisor. Then Tina did too.
b. After Mary laughed she made a joke about the supervisor. Then Tina did too.

To be sure that the preference is due to information structure (the main assertion) rather than tree geometry, Frazier and Clifton (2005) also tested sentences like (10a), where the matrix clause is open to an 'epistemic' interpretation where it simply expresses the degree of certainty with which the speaker makes the assertion conveyed by the complement clause. The prediction is that (10a), where the matrix clause is open to an epistemic interpretation, should show fewer main clause interpretations than (10b), which is not open to an epistemic interpretation. The prediction was confirmed in a self-paced reading study by the answers participants gave to questions about the sentences they had just read. A control study tested P1 vs P3 subjects with verbs like *announce*, which are not open to an epistemic interpretation. In the control study, there was no difference in antecedent choice due to the person marking on the subject of the antecedent.

(10) a. I think Mary smokes. Sam does too. 32% matrix VP antecedents
b. The teacher thinks Mary smokes. Sam does too. 50% matrix VP antecedents

With this introduction to the processing of VPE, we now return to the main plot, namely, how to solve the undergeneration problem, given a syntactic matching approach to ellipsis.

2. Processing sentences with flawed antecedents

2.1 Processing elided constituents without syntactically matching antecedents

Arregui, Clifton, Frazier and Moulton (2006) proposed that when an elided VP has a flawed antecedent, the processor repairs the antecedent at LF. If this can be done easily, using only grammatical operations for which there is evidence, then the ellipsis will be relatively acceptable (on a par with syntactic reanalysis of garden path structures). They presented several studies in support of this 'Recycling' hypothesis. In a visual speeded acceptability judgment study, they presented sentences like those in (11), where the slash indicates frame breaks. In (11a), the elided verb phrase has a matching antecedent "see(+PAST) the comet". In (11b) the elided verb phrase has a matching antecedent, but it is not in the prototypical verb phrase position. Instead it is but embedded under a gerundive DP. In (11c) an antecedent must be formed by replacing the empty object of "see" with the controller of its binder "the comet." Finally, in (11d) the processor would have to create a verb phrase from scratch since it does not even have access to the head verb of the needed verb phrase without going inside a word. According to the recycling hypothesis, acceptability should degrade systematically from a- through d-. In the speeded acceptability judgment study, the prediction was confirmed: there was a significant drop in acceptability for each pairwise comparison: a > b > c > d.

(11) a. None of the astronomers saw the comet, /but John did. (Available verb phrase)
b. Seeing the comet was nearly impossible, /but John did. (Embedded verb phrase)
c. The comet was nearly impossible to see, /but John did. (Verb phrase with trace)
d. The comet was nearly unseeable, /but John did. (Negative adjective)

The recycling hypothesis was further tested by comparing nominal and verbal gerunds. Nominal gerunds (12a) display the expected properties of DPs (12b): they take a determiner, the object of the head nominal is case-marked by the preposition *of*, and they are modified by an adjective. Verbal gerunds (12c) display the expected properties of verb phrases (12d): they do not take a determiner, the object of the verb is case-marked by the verb, and they are modified by adverbs.

(12) a. The (loud) singing of the opera — nominal gerund
b. The lyrics of the opera — noun phrase
c. (*The) Singing the opera loudly — verbal gerund
d. John sang the opera. — verb phrase

According to the recycling hypothesis, verbal gerunds should be more acceptable antecedents for VPE than nominal gerunds because various operations must apply in the nominal gerund to create a verb phrase antecedent. This prediction was tested in a visual speeded acceptability judgment study, using materials like those in (13). The forms with modifiers were included to determine whether adding a modifier to a verbal gerundive actually increases acceptability by making it easier to identify the antecedent whereas including a modifier in the nominal gerund might decrease acceptability because altering the modifier to the appropriate form might take an extra operation.

(13) a. Singing the arias tomorrow night will be difficult / but Maria will. (Verbal, -mod)
b. Singing the arias slowly tomorrow night will be difficult / but Maria will. (Verbal, mod)
c. Tomorrow night's singing of the arias will be difficult / but Maria will. (Nominal, -mod)
d. Tomorrow night's slow singing of the arias will be difficult / but Maria will. (Nominal, +mod)

As predicted, the verbal gerunds were judged more acceptable than the nominal gerunds. Numerically, the presence of a modifier did slightly increase acceptability for the verbal gerunds but decrease acceptability for the nominal gerunds, but this tendency was not reliable.

In this section, evidence has been presented that when the processor doesn't have an appropriate antecedent for an elided constituent, it simply creates one at LF by repairing whatever antecedent is available. When that takes more operations, acceptability drops more than when it takes fewer operations. There are complexity limits on how far the processor is willing to go in recycling an imperfect antecedent. For example, the processor at least requires access to the head of the desired verb phrase without stripping off derivational affixes.

Although recycling of available material provides the beginning of an account of how listeners and readers deal with flawed antecedents, it does not address the question of why speakers would produce such examples in the first place. If they are ungrammatical, why are they produced?

2.2 Why would speakers utter ungrammatical sentences?

If VPE is ungrammatical when there is no syntactically matching antecedent, then why are such examples attested rather often in speech corpora? Presumably speakers should not opt for ungrammatical structures to convey their messages. However, speakers often do make speech errors during sentence production. One relatively common form of speech error is a 'syntactic blend' where the speaker mixes two common ways of formulating the same message (see Coppock, 2006 for a recent discussion). Speakers might have several options concerning the syntactic structure they might use for an antecedent clause. They might forget which form they chose for the antecedent clause when, at a later point, they plan the clause containing the ellipsis. In short, examples of VPE with flawed antecedents may be instances of syntactic blends.

The hypothesis that imperfect antecedents result from syntactic blends makes an interesting prediction. Mehler (1963) has shown that passives are misrecalled as actives more often than actives are misrecalled as passives. This leads to the expectation that mismatching VPE examples will be more acceptable when they involve a passive antecedent clause and an active ellipsis clause than an active antecedent clause and a passive ellipsis clause. This prediction was tested by Arregui et al. in a written acceptability judgment task (where 1 was unacceptable and 5 was perfect).

The study also tested the hypothesis that mismatched forms of ellipsis would be more acceptable when the elided clause contained a presupposition trigger (*already*, or *too*) because the presupposition trigger might imply that an antecedent was available in context. The study tested examples like those in (14) and (15).

(14) a. The dessert was praised by the customer after the critic did already.
b. The dessert was praised by the customer and the critic did.
c. The customer praised the dessert after the appetizer was already.
d. The customer praised the dessert and the appetizer was.

(15) a. The student was praised by the old schoolmaster, and the advisor did too.
b. The student was praised by the old schoolmaster, and the advisor did.
c. The advisor praised the student, and the old schoolmaster was too.
d. The advisor praised the student, and the old school master was.

The results confirmed both predictions: passive-active examples were rated as more acceptable than active-passive examples, and examples containing a presupposition trigger were rated as more acceptable than sentences lacking a presupposition trigger (though the effect was largely due to the examples with *already*.) The results of the study thus encourage the view that antecedent clause - elided clause mismatches are due to syntactic blends.

2.3 Non-actuality implicatures

Many naturally occurring examples of elided constituents with flawed antecedents involve salient goal states or non-actuality implicatures. Consider the semantic contrast between (16) and (17).

(16) A trip is planned for August.
(17) A trip should be planned for August.

Sentence (16), with an indicative tense, entails that a trip is planned for August. It has an actuality/veridicality entailment. Sentence (17), which contains a modal, does not have either an actuality entailment or a non-actuality entailment (A trip should be planned for August and it has been). However, pragmatically, (17) implies that a trip has not already been planned. (I'll dub this a non-actuality implicature.) I'll assume that this non-actuality implicature is represented as in (18), where Wo stands for the actual world.[1]

(18) Wo: NOT (a trip is planned for August)

Non-actuality implicatures may implicitly focus the content of the clause which triggers them due to the implied contrast between the actual world and the asserted content of the clause. This focusing may influence processing of the ellipsis clause, as stated in (19).

(19) Non-actuality implicature hypothesis: The implied contrast between the actual world and the non-actuality implicature serves to focus the content of the antecedent clause and thus makes it a salient antecedent for elided constituents.

---

[1] Given this representation, a sentence with a nonactuality implicature should be representationally complex and might give rise to the exponential increase in complexity characteristic of double (John didn't neglect it) and triple negation sentences when it is found in sentences with negation(s). Intuitively, this seems to be the case, as indicated in the examples below. It is puzzling to me, however, that examples like "wanted to go to Paris" also give rise to a non-actuality implicature but not to the same level of complexity.

(i) If only John had gone to Paris. (Wo: NOT(John went to Paris))-1 neg
If only nobody had gone to Paris.-2negs
If only nobody had ignored Paris.-3negs

(ii) John is too lazy to go to Paris. (Wo: NOT(John went to Paris)
Nobody is too lazy to go to Paris.
Nobody is too lazy to ignore Paris.

Note the comparison to the classic example "No head injury is too trivial to ignore."

The non-actuality implicature hypothesis (19) makes several predictions.

One prediction concerns preferred interpretations of elided constituents. An antecedent from a clause with a nonactuality implicature should be relatively good even if the potential antecedent is in a syntactically nonprominent position. For example, in (20a) either the matrix verb phrase or the relative clause verb phrase is in principle available as the antecedent for the elided verb phrase in the second sentence. However, in (20b) with a modal and a non-actuality implicature, the relative clause antecedent is predicted to be more available than it is in (20a).

(20) a. The supervisor chastised a worker who brought production to a halt. Fred did recently.
b. The supervisor chastised a worker who would have brought production to a halt. Fred did recently.

The prediction that more relative clause antecedent interpretations should occur in sentences like (20b) (containing "would have, should have, could have, wanted to, almost, eager to") than in counterparts lacking the implicature as in (20a) was tested and confirmed in a written paraphrase selection study conducted with Chuck Clifton. Sentences with an indicative past tense received 53% relative clause verb phrase responses whereas those with a non-actuality implicature received 62% ($p < .024$).

A second prediction of (19) is that non-actuality implicatures should make flawed antecedents better than they would be otherwise. In a written acceptability judgment study with Chuck Clifton, 'experimental' sentences like (21)-(23) were tested to determine whether indeed the b-forms, with non-actuality implicatures, are considered more acceptable than their a-counterparts.

(21) a. This information was released but Gorbachev didn't.
b. This information needed to be released but Gorbachev didn't. (Adapted from Hardt's example in (7) above)

(22) a. A trip is planned for August, but Anne didn't.
b. A trip must be planned for August, but Anne didn't. Acc judgments: 1-5

(23) a. A private firm will be hired, but the Chancellor can't. 1.6
b. A private firm should be hired, but the Chancellor can't. 2.2

'Control' sentences such as (24a) with no actuality/veridicality entailment and no non-actuality implicature were also tested.

(24) a. The cookies may have been made, but the babysitter didn't. 1.9
b. The cookies needed to be made, but the babysitter didn't. 2.2

The prediction is that non-actuality implicatures should increase acceptability, especially in the experimental items. The prediction was confirmed. The 'experimental' sentences were rated less acceptable than the 'control' sentences [$F1(1,47) = 7.603$, $p < .01$]; crucially, the non-implicature a-forms were rated as less acceptable than the b-form (containing the non-actuality implicature) [$F1(1,47) = 41.141$, $p < .001$]. There was also an interaction of these two factors due to a larger effect of the non-actuality implicature in the experimental sentences than in the control sentences [$F1(1,47) = 5.45$, $p < .03$].

To summarize, non-actuality implicatures serve as a focusing device, guiding the processor to, seemingly effortlessly, build just the structure/interpretation required for elided constituents with flawed antecedents. In effect, the processor is set up for a contrast in advance of encountering the elided constituent.

3. Acceptable ungrammaticality: can it be explanatory?

An account of ellipsis relying in part on the notion of acceptable ungrammaticality might seem outlandish. But the notion of acceptable ungrammaticality is independently required. For example,

Staum and Sag (2007) investigated examples like those in (25). Consider first the critical example (25d). Despite being ungrammatical, it is more acceptable than its grammatical counterpart (25c). To convince yourself that (25d) actually is ungrammatical, compare it with (25b) where there is only a short constituent intervening between the two complementizers. Sentence (25b) is sharply ungrammatical for all speakers.

(25) a. John reminded Mary **that** soon his brother would be ready to leave.
b. John reminded Mary **that** soon **that** his brother would be ready to leave. (Short intervener)
c. John reminded Mary **that** after he was finished with his meeting his brother would be ready to leave.
d. John reminded Mary **that** after he was finished with his meeting **that** his brother would be ready to leave. (Long intervener)

Staum and Sag (2007) claim sentences like (25d) occur frequently, and show that they are read FASTER than their one-*that* (c-) counterparts; as expected, sentences like (25b) with short dependencies and an extra-that are read slower than their one-*that* (a-) counterparts. In short, independent of ellipsis, linguistic theory must recognize the notion of acceptable ungrammaticality.

One notion of 'explanatoriness' is restrictiveness: the overall theory of language is explanatory if it limits the principles of grammar and the principles of processing to just a small number for which we find evidence in many different domains. Predicting and explaining actual empirical generalizations is due to the interaction of these principles. This contrasts with a theory consisting of a very large number of surface descriptions which primarily restate the data rather than providing insight into them. From this perspective, it is important that the principles implicated in a theory of acceptable ungrammaticality are the same as those implicated in the theory of unacceptable grammaticality and the same as those generally required to give a theory of processing acceptable grammatical sentences.

The recycling hypothesis exploits mechanisms already required in the theory of language. Consider sluicing on the account of Chung, Ladusaw and McCloskey (1995). An antecedent clause, italicized in (26), is copied into the elided clause, as illustrated in (26a). An appropriate (usually indefinite) phrase, *something* in (26a), is then replaced with a variable that is bound by the interrogative operator. But in examples like (26b) (or "Ian mowed the lawn but I don't know for whom") there will be no syntactic position in the antecedent to house the variable. On Chung et al.'s account, in such cases the processor simply assumes the existence of the necessary phrase providing it doesn't violate any grammatical constraint. Notice that this process of 'sprouting' structure already contains the seeds of recycling: a process of structure building (or repair of the antecedent/elided constituent) is triggered by evidence that the structure is needed (so that the interrogative operator will have a variable to bind).

(26) a. *Josh ate something* but I don't remember what [*Josh ate vbl.*]
b. *Josh ate* but I don't remember what [*Josh ate vbl.*]

According to the recycling hypothesis, the processor is repairing antecedents at LF using the same mechanisms required for syntactic reanalysis. Studies of syntactic reanalysis (e.g., Frazier and Clifton, 1998, Fodor and Ferreira, 1998) reveal that the ease of reanalysis depends on the nature of the error signal that indicates reanalysis of the current analysis is necessary. When the error signal is informative about the nature of the required repair, and few changes are needed, reanalysis is easy, as in (27) where the tensed verb phrase *was correct* indicates the need for a DP subject, thereby triggering reanalysis of the preceding DP as subject of the embedded clause rather than as the simple direct object of *know*. By contrast, in the classic conscious garden path in (28), the error signal is the same, a tensed verb phrase (*fell*) is missing its subject. But taking the preceding DP, *the barn*, to be the subject will not repair the sentence. Indeed, the very problem is that what should be the DP subject of *fell* has not been identified as a DP at all.

(27) Celia knew the answer to the difficult problem was correct.
First analysis: [knew DP]
Error signal: [was correct] requires a subject
The preceding DP can be the required subject.
Correct analysis: [knew CP]

(28) The horse raced past the barn fell.
First analysis: [The horse raced past the barn fell] = IP
Error signal: [fell] requires a subject
The preceding DP can NOT be the required subject.
Correct analysis: [The horse raced past the barn fell] = DP

In approaches to ellipsis where only a salient semantic antecedent is required, it has been emphasized that inference is required for an account of ellipsis (Webber, 1978, Hardt, 1993). Hardt (2005, p.3) in particular cites examples like (29) to show that "the interpretation without inference is unacceptable, involving as it does an agreement violation between the singular subject .. and the VP ... This fact is the key to solving the problem posed by Webber about how to delimit the class of available inferences in ellipsis: only inferences triggered by violations are possible."

(29) Martha and Irv had planned to nominated each other, but Martha couldn't __, because of her political obligations.

On the present account, it is expected that inferences must be triggered by "violations" since in general reanalysis is triggered by "violations" or "evidence". Typically the processor does not reanalyze its current analysis when there is no need.

Another characteristic of reanalysis is that it's easier when the first (incorrect) analysis is implausible or ungrammatical, and it's easier when the final analysis is plausible and is or appears to be grammatical. This applies to ellipsis processing too. For example, Garnham and Oakhill (1987) investigated the processing of sentences like those in (30). With the misleading continuation "the nurse had too", people often gave incorrect answers to questions, suggesting that they interpreted "The nurse had too" as "The nurse had examined the child too."

(30) It had been a busy morning in the hospital. The elderly patient had been examined by the doctor. The child/nurse had too.

When the material needed for reanalysis is focused or highly accessible, reanalysis is easier than when it is not. In the discussion of non-actuality implicatures in Section 2.3 above, it was argued that this too applies in the case of ellipsis. In short, repair of an antecedent (recycling) shows the same properties as syntactic reanalysis in ordinary language comprehension. No special principles are needed to account for the recycling process.

We turn now to sentence production, and to its effect on acceptability judgments. Memory errors, the presence of presupposition triggers, and the availability of systematic alternative formulations of utterances (active/passive, *is sufficient enough for*, *pales next to comparison of*) should all have an effect on the acceptability of mismatched antecedent-ellipsis examples because they should increase the likelihood of producing syntactic blends, such as VPE in an active clause with a passive antecedent. These and other ungrammatical sentences are expected to sound more acceptable if they have been heard repeatedly or sound like something likely to have been heard/produced.

In the experimental literature, it has been shown that acceptability judgments are higher if even one example of the structure has been read or heard before rating a novel example. In five experiments, Luka and Barsalou (2005) had participants read grammatical sentences first and then rate novel sentences. Mere exposure to a sentence, or to a sentence structure with different words, resulted in higher ratings. For example, "Egor lugged Dr. Frankenstein the corpse" was rated higher after reading unrelated sentences with double object structures; "What the pharmacist recommended is to read the directions" was rated higher after reading sentences with a pseudo-cleft structure.

Perhaps even more striking are indications that acceptability judgments are influenced by how 'natural' an error seems. E.g., errors that would avoid an unwanted implicature are judged more acceptable than those that don't. In a small (two-item) experiment with Chuck Clifton we investigated sentences like those in (31). Sentence (31a) is an actual speech error, heard on National Public Radio. It was clear that the interviewer who said the sentence switched from impersonal "you" in clause one, to imperson "we" in the main clause as a matter of politeness: otherwise the person interviewed might have interpreted the main clause as an insult, with the reporter accusing the person interviewed of deluding himself.

(31) a. If you think this is going to solve the terrible problems in Najaf, we're deluding ourselves.
b. If we think this is going to solve the terrible problems in Najaf, you're deluding yourself.

By contrast, in (31b) the switch from impersonal "we" to impersonal "you" would not avoid an unwanted implicature, but instead give rise to the possibility of one. In an acceptability judgment study (using a scale from 1-5, where 5 was perfectly acceptable) the two examples like (31a) were rated 4.05 which was significantly higher than the b-forms (which received a mean rating of 3.37). This result suggests that the acceptability of sentences depends not just on what sentences have been heard but on the probability that a human would produce the sentence as a speech error/syntactic blend.

Other cases argued to be instances of acceptable ungrammaticality also are plausibly analyzed as syntactic blends. Otero (1972) argued that Spanish (and also other Romance speaking) language users get confused about which SE appears in an ungrammatical sentence like (32b). The reflexive/reciprocal SE in (33) is grammatical. But in (32b) the same sentence form is attributed the meaning of impersonal SE (appropriate in (32a)) even though that would be ungrammatical. Nevertheless, native Spanish speakers consider (32b) acceptable. Otero argued that Spanish speakers essentially get confused about which SE they have used. Notice that this blend of syntactic structures must affect both speakers and listeners just like in the ellipsis examples discussed in Section 2.2.

(32) a. SE alquila apartamentos. ('Impersonal SE')
'PRO rent-sg. apartments.'
b. *SE alquilan apartamentos.
'PRO rent-pl. apartments.'
'PRO rents apartments'

(33) SE alquilan apartamentos. (Reflexive, Reciprocal SE)
'They rent apartments to themselves/each other'

Elided constituents with flawed antecedents are sometimes easy to repair and they are judged acceptable when they correspond to natural speech errors. If they involve a highly salient contrast, as with non-actuality implicatures, they are especially acceptable. Does this view suggest that we want a theory of acceptable ungrammaticality in which the various factors affecting acceptable ungrammatical sentences are assigned a weight and summed to determine some specific numerical level of acceptability?[2]

My view is that a more explanatory approach is to understand the principles involved in creating acceptable ungrammaticality. Consider the slightly adapted form of (5/21) in (34).

(34) This information needed to be released, but Gorbachov chose not to.

In addition to being a likely syntactic blend and being easy to repair, this sentence in effect also satisfies the likely rhetorical needs of the speaker. Assuming that the relevant information needed to be released by someone (not necessarily by Gorbachov) but that Gorbachov was the person to get the

[2] Thanks to Line Mikkelson for raising this question.

blame because he was the person with the authority to release the information, the ungrammatical (34) may actually convey the intended message better than any grammatical alternative.

Register also seems to influence acceptability. In casual speech, sentence (35) seems to be reasonably acceptable on a negative concord reading, where it means "I miss going to the gym."

(35) I miss not going to the gym.
(Paired with the interpretation: I miss going to the gym.)

It seems to be a natural speech error and it is easy to repair by treating "not" as an instance of negative concord. Though listeners might well accept (35) with a negative concord reading in casual speech, (35) seems far less likely to be acceptable with a negative concord interpretation when written. Assigning weights to factors, even if they are adjusted for rhetorical purpose, register, etc. might bring some insight to the study of acceptable ungrammaticality, but it's just as likely to obscure it by becoming an end in itself rather than a means to a more adequate and explanatory theory. Instead, developing a theory of processing and a theory of grammar and carefully investigating how they interact seems more promising to my mind.

Presumably the reason why some easy-to-repair sentences are not acceptable and others are is because of the nature of grammatical conditions. In "I miss not going to the gym," the operation required for a negative concord reading is available in UG and in some dialects of English (AAE), but not generally available in standard English. If some instances of grammatical negative concord structures were to be incorporated into standard English, one could well imagine the acceptable ungrammatical (35) becoming grammatical.

4. Conclusions

A restrictive syntactic approach to the grammar of VPE requires the existence of articulated syntactic structure in the ellipsis site and also the existence of a syntactically-matching antecedent. Linguistic and psycholinguistic evidence (alluded to in the introduction) support the claim that the elided verb phrase contains articulated syntactic structure. The claim that a matching antecedent is required appears on the face of it to be too strong. However, it has been argued here that the 'parallelism/matching antecedent' condition can and should be maintained. Once an adequate processing theory is combined with a restrictive grammar of ellipsis, the theory of language (grammar plus processing theory) can explain why ungrammatical examples with flawed antecedents can be easily understood and judged relatively acceptable.

When an appropriate antecedent is not available, the processor may create one at LF. It is easy to recycle in the same circumstances where it is easy to make a syntactic repair (e.g., the first analysis is not semantically and pragmatically confirmed, there's clear evidence of an error and the error signal is informative about the nature of the correction required, the repair produces a plausible interpretation).

The grammar may define ideal sentences. When speakers often fail to fully conform with the ideal in a particular (systematic) way, and hearers easily and correctly identify the intended structure/meaning of the sentence, this may produce ungrammatical but relatively acceptable sentences, especially when supported by a clear contrast implied by a non-actuality implicature.

5. Appendix

Merchant ms. suggests that VPE targets the sister of Voice, which is sufficiently low in the tree that it is neutral between a passive or an active clause and thus may serve as the antecedent of either clause-type. On this analysis, voice-mismatched VPE is grammatical, though there may be a preference for matching overall structure. The asymmetry established experimentally, favoring passive antecedents and active ellipsis clauses, is attributed to the poorer memory for passive clauses. The central argument given for the approach is the observation that it is only VPE, not sluicing, that permits voice mismatches, which is expected on this account (since sluicing clearly targets a constituent higher in the tree than Voice Phrase).

(1) The filthy kitchen was cleaned. The janitor did.

(2) The filthy kitchen was cleaned but I don't know which janitor.

In general, it may be true that sluicing is worse with voice mismatches than VPE is. The question is why. Is this because there's a categorical difference in grammaticality, as Merchant's analysis implies, with the sentence in (1) being grammatical but the sentence in (2) being ungrammatical? Alternatively, are the sentences in (1) and (2) ungrammatical but for some reason (2) is more difficult to rescue?

Several considerations appear to favor the ungrammaticality (recycling) account. If (1) is already grammatical, it's not clear why it needs adornment or pragmatic help. Further, Sluicing requires an elided syntactic structure with a variable in it. Perhaps this makes sluicing 'fussier' than VPE. In other words, the matching condition is not simply a formal condition but determines the interpretation of the sluiced clause. This line of reasoning predicts that VPE with a variable inside the elided VP should also be fussy. This prediction seems to be borne out. A voice mismatch in VPE when the elided VP contains a variable seems rather unacceptable, as in "Every room will have been cleaned that the janitor will". Finally, on the recycling account, it should matter how much evidence the processor has concerning the LF operation needed to create an antecedent that satisfies grammatical conditions. The passive-active mismatched sluice in (2) doesn't provide the processor with evidence that it is the implicit *by*-phrase agent that is replaced by the interrogative phrase. However, in active-passive mismatched sluices, "by which janitor" does provide exactly this evidence. My intuitions do indeed suggest that (3) is better than (2), and that (4a) is no worse than (4b), even though the former is ungrammatical and the latter grammatical.

(3) Someone cleaned the filthy kitchen, but I don't know by which janitor.

(4) a. A famous politician told an obvious lie last night on national tv. I bet you can guess by which scumbag.
    b. An obvious lie was told by a famous politician last night on national tv. I bet you can guess which scumbag.

In short, although more evidence and more systematic evidence is clearly needed, some mismatched-voice sluicing examples don't seem completely unacceptable, at least not to all speakers.

References

Arregui, Ana; Clifton, Charles; Frazier, Lyn & Moulton, Keir. 2006. Processing elided VPs with flawed antecedents. *Journal of Memory & Language*, 55: 232-246.

Carlson, Katy. 2002. *Parallelism and prosody in the processing of ellipsis sentences*. NY: Routledge.

Chung, Sandra: Ladusaw, William: & McCloskey, James. 1995. Sluicing, and Logical Form. *Natural Language Semantics*, 3:239-282.

Coppock, Elizabeth. 2006. Alignment in syntactic blending. *MIT Working Papers in Linguistics*, 13: 239-255

Dalrymple, Mary; Sheiber, Stuart; & Pereira, Fernando C. N. 1991. Ellipsis and Higher-Order Unification. *Linguistics and Philosophy* 14: 399-452.

Fiengo, Robert & May, Robert 1992. The eliminative puzzles of ellipsis. In *Proceedings of the Stuttgart Ellipsis Workshop* (reported in Johnson, 1996).

Fiengo, Robert & May, Robert. 1994. *Indices and Identity*. Cambridge, Mass: MIT Press.

Fodor, Janet D. & Ferreira, Fernanda 1998. *Reanalysis in Sentence Processing*. Dordrecht: Kluwer.

Frazier, Lyn & Clifton, Charles E. 1998a. Comprehension of sluiced sentences. *Language and Cognitive Processes*, 13:499-520.

Frazier, Lyn & Clifton, Charles E. 1998b. Sentence reanalysis, and visibility. In J.D. Fodor & F. Ferreira (eds.) *Reanalysis in Sentence Processing*. Dordrecht: Kluwer Academic Publishers.

Frazier, Lyn & Clifton, Charles E. 2000. On bound variable interpretations: The LF-only hypothesis. *Journal of Psycholinguistic Research*, 29, 125-139.

Frazier, Lyn & Clifton, Charles E. 2001. Processing coordinates and ellipsis: Copy ä. *Syntax* 4:1-22.

Frazier, Lyn & Clifton, Charles E. 2005. The syntax-discourse divide: Processing ellipsis. *Syntax*, 8:121-174.

Frazier, Lyn & Clifton, Charles E. 2006. Ellipsis and discourse coherence. *Linguistics and Philosophy*,29:315-46.

Garnham, Alan & Oakhill, Jane. 1987. Interpreting elliptical VPs. *Q. J. of Experimental Psych*, 39A:611-627.

Grinder, John & Postal, Paul M. 1971. Missing antecedents. *Linguistic Inquiry*, 2:269-312.

Hardt, Daniel. 1993. Verb Phrase Ellipsis: Form Meaning, and Processing. Ph.D. dissertation, University of Pennsylvania.

Hardt, Daniel. 2005. Inference, ellipsis and deaccenting. Amsterdam Colloquium 2005.

Hardt, Daniel & Romero, Maribel. 2004. Ellipsis and discourse structure. *Journal of Semantics*, 21:375-414.

Johnson, Kyle. 1996. When verb phrases go missing. *Glot International*, 2:3-9.

Kehler, Andrew. 2002. *Coherence, Reference and the Theory of Grammar*, CSLI Publications. Stanford University.

Kennedy, Chris. 1997. Antecedent contained deletion and the syntax of quantification. *Linguistic Inquiry*, 28:662-688.

Langendoen, D. Terence, & Bever, Thomas G. 1973. Can a not unhappy man be called a not sad one? In S.R. Anderson & P. Kiparsky (eds.), *A festschrift for Morris Halle*, (392-409). New York: Holt, Rinehart and Winston, Inc.

Lasnik, Howard. 2001. When can you save a structre by destroying it? In M. Kim and U Strauss (eds.) *Proceedings of the North Eastern Linguistic Society* 31 Volume 2, 301-320, GLSA.

Luka, Barbara & Barsalou, Lawrence. 2005. Structural facilitation: Mere exposure effects for grammatical acceptability as evidence for syntactic priming in comprehension. *Journal of Memory and Language*, 52:436-459.

Martin, Andrea & McElree, Brian. 2007. Retrieval and nonadjacent dependencies: Speed-accuracy tradeoff evidence from sluicing. Poster presented at 20th annual CUNY conference on human sentence processing. La Jolla, Ca, March 29-31, 2007.

Martin, Andrea & McElree, Brian. Submitted. A content-addressable pointer mechanism underlies comprehension of verb phrase ellipsis. NYU manuscript.

Mauner, Gail; Tanenhaus, Michael; & Carlson, Gregory. 1995. A note on parallelism effects in processing deep and surface verb-phrase anaphora. *Language and Cognitive Processes*, 10:1-12.

Mehler, Jacques. 1963 . Some effects of grammatical transformations on the recall of English sentences. *Journal of Verbal Learning and Verbal Behavior*, 2:346-351.

Merchant, Jason. 1999. The syntax of silence: Sluicing, islands and identity in ellipsis. Ph.D. dissertation, University of California, Santa Cruz.

Merchant, Jason. 2005. Fragments and ellipsis. *Linguistics & Philosophy*, 27:661-738.

Merchant, Jason. To appear. Variable island repair under ellipsis. To appear in K. Johnson (ed.) *Topics in ellipsis*. Cambridge: Cambridge University Press.

Merchant, Jason. Ms. Voice and ellipsis. University of Chicago manuscript.

Otero, Carlos. 1972. Acceptable ungrammatical sentences in Spanish. *Linguistic Inquiry*, 3:233-242.

Rooth, Mats. 1992. A Theory of Focus Interpretation. *Natural Language Semantics* 1:117-121.

Sag, Ivan. 1976. Deletion and Logical Form. Doctoral Dissertation, Massachusetts Institute of Technology. [Published in 1980 by Garland Press, New York].

Sag, Ivan & Hankamer, Jorge. 1984. Towards a theory of anaphoric processing. *Linguistics & Philosophy*, 7:325-345.

Staum, Laura & Sag, Ivan. 2007. Multiple *That*: A strategy for reducing integration costs. Poster presented at CUNY 2007, La Jolla, California.

Tanenhaus, Michael K.& Carlson, Greg. 1990. Comprehension of deep and surface verb phrase anaphors. *Language and Cognitive Processes*, 5:257-352.

Webber, Bonnie. 1978. A Formal Approach to Discourse Anaphora. Harvard University doctoral thesis.

Williams, Edwin. 1977. Discourse and Logical Form. *Linguistic Inquiry* 8:101-139.

Proceedings of the 26th West Coast Conference on Formal Linguistics

edited by Charles B. Chang
and Hannah J. Haynie

Cascadilla Proceedings Project Somerville, MA 2008

ISBN 1-57473-423-2 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Frazier, Lyn. 2008. Processing Ellipsis: A Processing Solution to the Undergeneration Problem? In *Proceedings of the 26th West Coast Conference on Formal Linguistics,* ed. Charles B. Chang and Hannah J. Haynie, 21-32. Somerville, MA: Cascadilla Proceedings Project.

or:

Frazier, Lyn. 2008. Processing Ellipsis: A Processing Solution to the Undergeneration Problem? In *Proceedings of the 26th West Coast Conference on Formal Linguistics,* ed. Charles B. Chang and Hannah J. Haynie, 21-32. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1652.