**The Production and Comprehension of Resumptive Pronouns
in Relative Clause "Island" Contexts**

**Fernanda Ferreira and Benjamin Swets
Michigan State University**

Correspondence to:

Fernanda Ferreira
Cognitive Science Program and Department of Psychology
Michigan State University
East Lansing, MI 48824-1118
USA

Email: Fernanda@eyelab.msu.edu

Draft: Please do not distribute or quote.

INTRODUCTION

Many investigators working on language processing are coming around to the idea that the field should adopt a more naturalistic and ecological approach to psycholinguistics (Pickering & Garrod, in press; Trueswell & Tanenhaus, in press). The trick is to hold on to the insights and rigorous methodology associated with the formal, representation-based psycholinguistic tradition while expanding the range of phenomena to include those found particularly in naturalistic speech and human conversation. For instance, people gesture while they talk, but there is little work on the way that linguistic representations are built and coordinated with the production of gestures (for an important exception, see Levelt, Richardson, & La Heij, 1985). Similarly, when people talk, they are often disfluent (Clark, 1996): They produce fillers such as *uh* and *um*, and they repeat words, backtrack, abandon utterances, start them over, and so on. In our recent work, we have begun to investigate how disfluencies affect the parser's structure-building operations (Bailey & Ferreira, 2003; Ferreira, Lau, & Bailey, in press). We have not only demonstrated that disfluencies have a systematic effect on parsing, but we have also shown that it is possible to study a phenomenon of the "wild" like disfluencies in a way that retains the assumption that formal linguistic structures are built during processing, and in real time.

How might this more ecological approach be applied to language production? One idea is to relax the assumption that people produce sentences that conform perfectly to the grammar of their language. Of course, we know that speakers of English will virtually never say something like *the cat black is asleep*, while speakers of Portuguese will say exactly the Portuguese equivalent, because of the fundamental word order constraints given by the grammars of the two languages. At the same time, we also know that people often produce utterances that contain "mistakes" such as subject-verb agreement errors. More interestingly, because of the demanding nature of conversation, people sometimes produce a certain sentence type that has a number of interesting properties. Here is an example, heard on the United States' National Public Radio on September 16, 2001:

> (1) We're afraid of things that we don't know what they are.
>
> [matrix clause *We're afraid of things* [relative clause *that we don't know*[complement clause *what they are*]]]

Sentences of this type occur reasonably often in conversation (Tony Kroch estimates that one may expect to hear one or two of them every day – personal communication). Thus, although people do not ignore strict grammatical rules such as those that describe sequencing within noun phrases, they do appear to relax other sorts of constraints, such as those that determine the kinds of long-distance dependencies that are legal (Bock & Miller, 1991)[1].

It could be very useful to study what takes place in the language production system when it creates utterances such as (1). One reason these structures are potentially illuminating is that they are ungrammatical or at least marginal in English, but they are nevertheless produced

---

[1] It is important to note that this structure is grammatical in some languages – for example, in Hebrew and Arabic (Prince, 1990). Thus, this constraint on long-distance dependencies is language specific.

on a fairly frequent basis. What makes these sentences illicit is that they violate the subjacency constraint on wh-movement. The position in (1) occupied by the pronoun *they* is the original source for the wh-element that moves to form the relative clause, but the clause in which it is housed is an "island" (Ross, 1967), meaning that no subconstituent is supposed to leave it. This illicit movement of the wh-element leaves behind a gap (between *what* and *are*), resulting in *we're afraid of things that we don't know what <gap> are*. Typically, when utterances like this are said, the speaker plugs up the gap with what is termed a "resumptive pronoun", resulting in something that is clearly still marginal but much better than the version with just a gap. We will refer to sentences like (1) as "islands plus resumptives", to capture their two key properties: They violate an island constraint, and they contain a resumptive pronoun.

Another reason that it might be productive to investigate the island + resumptive structure is that it is much more complex than the forms that are typically investigated in studies of language production. Almost all work up to now has been done on single clause sentences, but (1) contains three: the matrix clause (the entire sentence), the relative clause (*that we don't know what they are*), and the complement clause of *know* (*what they are*). It seems likely that considering structurally more complex forms might shed greater light on how grammatical encoding works; at the very least, more aspects of the grammatical encoder's mechanisms will be on display in a more demanding structure. But the island + resumptive form is not just quantitatively more challenging; it also has a property that allows us to address a critical issue in theories of language production, and that is incrementality. The question of incrementality concerns how much look-ahead the production system requires during utterance encoding (Levelt, 1989). Applying it specifically to the case of syntactic processing, we can ask whether the system plans several words in advance before passing the syntactic representation to the phonological encoder, or is it more of a cascading architecture, where a minimal syntactic unit such as a word is sent off for further processing? The reason the island + resumptive structure is potentially useful for examining this question is that the locus of the ultimate flaw in the sentence is several words downstream and buried in the lowest of the three clauses. A sentence such as (1) does not go wrong until *they*, as revealed by the perfect acceptability of *We're afraid of things that we don't know very much about*. Therefore, if the system is highly incremental, it should not "know" about the island and the gap/resumptive pronoun until it is essentially at that point in the sentence or one word before. Even a less incremental model which assumes that processing is essentially clause-based (Garrett, 1988; Bock & Cutting, 1992) would predict that any effects of the gap/resumptive pronoun would be seen fairly late during articulation – specifically, inside the complement (lowest) clause.

What we can do, then, is see whether speakers give off any clues to suggest that the system is aware of the funny property of island + resumptive structures before the most embedded clause. For example, are initiation times for such sentences longer than for comparable controls? Or is the duration of the first word or two longer? Before turning to the details of an experiment to examine these predictions, we should address the issue of incrementality in a bit more detail. Space limitations do not allow us to review here the entire literature relevant to this question (but see Ferreira, 2000; Griffin & Bock, 2000), but we will mention that at the moment the evidence is mixed. Some experiments seem to suggest that the system plans no more than one phrase at a time (for example, Smith & Wheeldon, 2001), but other work implies that there is a fair bit of look-ahead (such as Ford & Holmes, 1978). We will describe in some detail one study that provided support for only limited incrementality, in part because the methodology that was used is similar to what we will present here (Ferreira & Swets, 2002). Participants were

undergraduates at Michigan State University who were asked to answer arithmetic problems and who gave their answers in the form of an utterance (*The answer is fifty-eight*). The problems always included at least one two-digit addend (e.g., 53+5), so participants were highly unlikely to be able to simply retrieve the sum. The logic of the study was that if the production system is incremental, then the difficulty of the sum should show up while it is being articulated, but difficulty should not influence initiation times for the utterance as a whole. In one experiment, participants were allowed to begin to speak whenever they felt ready, and what we found was no evidence for incremental production: Only initiation times were affected by problem difficulty, not the duration of any part of the utterance. In a follow-up experiment, speakers were required to begin to speak quickly to avoid hearing the sound of a "beep". That sound was their punishment for waiting too long to begin talking. This manipulation reduced initiation times overall from over 2 seconds in the first experiment to about 700 ms in the one with the deadline. But even with this pressure to begin to speak quickly, initiation times still reflected the difficulty of computing the sum. But at the same time, the duration of the earlier part of the utterance was similarly affected, suggesting that speakers did plan the sum to some extent during articulation. Thus, this evidence suggests that the degree to which the system is incremental depends on the speaker's goals and strategies for managing the communicative situation; if a premium is placed on beginning to speak quickly, then the production system does indeed become more incremental, but if there is time to plan, speakers seem to prefer to do so. Note that this tradeoff does not affect the quality of the utterance: We were surprised to find that the accuracy of the sums was just as high when speakers were under time pressure as it was when they controlled when they began to speak. This finding suggests that there is a dimension to human information processing related to intensity of effort (Just & Carpenter, 1993), and clearly participants in experiments only want to work so hard; but this is a topic for another paper.

One limitation of this study is that most of the utterances we say in our day to day lives can be produced without the need for arithmetic computations. The task is not entirely unnatural, of course (consider a situation in which you're telling your dinner companion what his contribution to the bill is), but it is important to see if the same pattern holds with an entirely different kind of utterance. Moreover, as mentioned earlier, it would be useful to assess how incremental the system is when it produces not just simple one-clause sentences (which the arithmetic utterances were) but also multi-clausal utterances with some syntactic complexity. In addition, because the island + resumptive structure is multi-word and multi-clausal, speakers have the opportunity to revise it while they are talking. This mechanism is implicit in the idea of incrementality; speakers do not plan very far ahead, so if they find themselves in the middle of an utterance that looks like it will turn out odd or wrong, they use their knowledge of the grammar online to try to come up with something better. In the case of (1), the speaker could begin with *We're afraid of things that we don't…* and, realizing that an illicit structure is about to emerge, change the utterance to something like (2):

(2) We're afraid of things that we don't understand

The experiments we report here allowed us to investigate these major questions: First, how incremental is the production system? Second, and related to the first question, do speakers reformulate online as a consequence of making early syntactic commitments that will lead to ungrammaticality? And third, how does the production system balance its desire to encode the true, communicative intentions of the speaker with the requirements and constraints imposed by the grammar of the language? We turn now to a description of the experiments designed to

address these issues. Later we will discuss theoretical implications, focusing particularly on the production – comprehension relationship.

EXPERIMENTS TO ELICIT ISLAND + RESUMPTIVE SENTENCES

Consider a situation in which someone sees a certain species of dog and describes it to another person in this way: *That's a dog that comes from California*. Then imagine she sees another breed of dog (perhaps the interlocutors are at a dog show) and she says (somewhat contrastively), *That's a dog that comes from Brazil*. Now imagine that a dog appears which the speaker is not familiar with. Because of the utterances she produced earlier (i.e., possibly because of structural priming; Bock, 1986; Pickering & Branigan, 1998), she might find herself saying *Geez, that's a dog that I don't know where it comes from*. This general scenario is the basic idea behind our paradigm. Speakers see pictures of people, animals, or cartoon figures, and the pictures are grouped so that three examples of the same category (e.g., three brides, three dogs, three muppets) appear together, in sequence. The first two instances have a descriptor; the last picture of the three is specified as lacking information. The pattern set in motion by the first two utterances is expected to lure our participants into producing the desired island + resumptive structure on the majority of trials. The initiation times for the utterances and the durations of the words that make them up can then be measured and compared to appropriate controls (to be described shortly).

Two experiments will be reported, as well as the results of two follow-up grammaticality judgment studies. In the first production experiment, speakers produced utterances as soon as they felt they were ready. Results showed that the earliest words were longer in the island + resumptive condition, suggesting that the production system uses a great deal of lookahead. In the second production experiment, speakers were given a deadline to respond. This change in procedure reduced response latencies and yielded evidence for more incremental production (because effects on word duration were found in a sentential location closer to the resumptive pronoun). The grammaticality judgment experiments show that comprehenders indeed find these sentences unacceptable. The implications of the experiments as a whole is that incrementality is a parameter of the production system that changes depending on speakers' goals. In addition, the production system appears to be unaware of grammatical constraints to which the comprehension system is quite clearly sensitive, suggesting a production – comprehension asymmetry.

Method

Both experiments consisted of just two conditions: the island + resumptive condition and what we will term the surface-structure control. The same stimuli were used for both experiments. Two different stimulus lists were created; a given participant saw only one of those lists. In each list there were 48 arrays, and each array consisted of three pictures next to three short descriptions of the pictures. The description for the first picture of the trio was a short verb phrase such as *lives in California*. The description for the second was identical, except that the location was changed (e.g., *lives in Brazil*). For the third picture of the trio, the descriptor was either *I don't know* (for the island + resumptive condition) or *doesn't know* (for the surface-structure control condition).

Participants were tested individually in a sound-attenuated room. They were told that on a computer monitor they would see three pictures next to three descriptors, and their task was to

learn the pictures and brief descriptions one by one and then hit the space bar to begin each trial. After they studied the trio, they struck the space bar again, and the question "What is this?" replaced the first descriptor. Figure 1 shows this stage in an experimental trial. Participants were instructed to answer the question with a full-sentence response that included at least a noun, a verb, and the content of the descriptor. Although they were instructed to begin speaking as quickly as possible, it was emphasized that it was important to use good sentences that appropriately answered the questions. An example was provided (e.g., *This is a cat that likes fish*). An example of an inappropriate response given the question was also given: (*This cat likes fish*). After describing the first picture of the trio (e.g., *This is a donkey that lives in Brazil*), participants answered the same question for the next two pictures in each slide. After they described the third item from the trio, the participant terminated the screen by pressing the space bar to initiate the next trial.

Figure 1: Paradigm to Elicit Island + Resumptive Sentences		
1st Question	**2nd Question**	**3rd Question (response analyzed)**
←**What is This?**		
-**lives in Brazil**	←**What is this?**	
-**I don't know**	-**I don't know**	←**What is this?**
Target sentence: "This is a donkey that lives In California."	**Target sentence: "This is a donkey that lives In Brazil."**	**Target sentence: "This is a donkey that I don't know where it lives."**

The island + resumptive and the surface-structure control conditions were created as follows. The trials were identical except for the descriptor associated with the third picture of the trio. For the island + resumptive condition, it was the sequence *I don't know*, which entices the participant to say something like *This is a donkey that I don't know where it lives*. For the surface-structure control, the descriptor was *doesn't know*, and so the analogous utterance was expected to be *This is a donkey that doesn't know where it lives.* Notice that the utterances are almost the same except for the words between *that* and *where it lives*, and even that material has the same number of syllables. This feature of the design means that the utterances to be compared are similar in most essential ways.

Each list consisted of 48 trios; half of those were experimental items, and so half of those (12) were in the island + resumptive condition and half (12) in the surface-structure control condition. The remaining 24 trios were filler trials. They were similar to the experimental trios, except that the third picture of each trio was another verb phrase that would likely not elicit an island + resumptive utterance or an utterance like the surface control.

For the first experiment, 35 participants were tested, but data from only 30 were included in the analyses. For the second, 34 were tested and the data from 30 were analyzed. The people omitted from the analyses did not make enough utterances of the expected type to allow means to be computed. Participants proceeded through the trials at their own pace. In the first experiment with no deadline, they were told to speak as soon as they were ready, but they were also encouraged to make sure their utterances were "good". In the second experiment, which employed a deadline procedure, participants were told that as soon as they struck the spacebar to describe the picture, three exclamation points would appear to the right of the screen. The exclamation points would begin disappearing rapidly until, after the third exclamation point disappeared, a "beep" sounded. At any point during the countdown, the participants' voice would terminate the countdown procedure until they had completed their utterance. The countdown was programmed to beep 1750 ms after the onset of the question "What is this?". We pilot-tested various times that were less than the mean initiation time in Experiment 2 (2216 ms) in order to arrive at a timing deadline that would make speakers feel rushed without adversely disrupting performance. As soon as the participant answered the first question, the experimenter pressed a button to start the deadline timer for the answer to the second question (about the second pictured item). The same countdown bars appeared to the right until the participant deactivated them by triggering a voicekey. After the same procedure took place for the third, target sentence description, the participant terminated the trial by pressing the space bar.

Analyses

Participants' utterances were transcribed and then each one was coded into categories, including the two that were of specific interest (island + resumptive, surface-length control). To be included in the initiation time and durational analyses, target utterances of both types had to meet these criteria: First, the sentences had to begin with a deictic subject followed by *is* (*This is…*). Then a determiner/noun phrase had to follow, and then a complementizer or wh-phrase that initiated the relative clause (*…a donkey who/that…*).

Only utterances that described the third picture of the trio and that met the target criteria described above were included in waveform analyses. Initiation time was defined as the time between the second utterance for each array and the onset of the third and final utterance. Duration was defined as the total time taken to complete the utterance once it has been initiated. To investigate where in the utterances speakers slowed down, we analyzed utterance durations into subparts as follows:

(3)
    a. [This is a] [donkey] [that] [I don't know] [where it lives].
    b. [This is a] [donkey] [that] [doesn't know] [where it lives].

At the most simple-minded level, we predicted that if the system is incremental, then durations in the two conditions would not diverge until late in the last clause *where it lives*. In contrast, the more non-incremental the production system is, the earlier in the utterance the durations should differ. In addition, if speakers' subjective sense that they are under time pressure affects the extent to which the production system uses look-ahead, then we also predicted that durations would diverge later in the utterance in the deadline condition compared to the no-deadline condition.

In addition, the kinds of utterances that were produced in the two conditions and the two experiments were also examined. We were interested to see whether participants in the resumptive + island condition might find creative ways to prevent themselves from producing these marginal structures, particularly after the first couple of trials. We also speculated that participants would be more successful at reformulating when they were not under time pressure, because it likely takes time and effort to come up with an alternative structure.

Results and Discussion

This discussion of the findings will be organized as follows. First, we will consider the initiation time and duration results for the no-deadline experiment, followed by those same results in the deadline experiment. We will compare performance in the island + resumptive and the surface-structure control conditions, but including only those trials on which the target sentence was actually produced (that is, (3a) in the island + resumptive condition, (3b) in the other). Also, trials were excluded if the initiation time was too short (< 100 ms), too long (> 15 seconds), or if the overall duration of the sentence was longer than 15 seconds. Second, we will describe the kinds of utterances that were produced in the experimental conditions in each experiment. Finally, we will compare the results for the two experiments, focusing especially on utterance types.

In the no-deadline experiment, processing times were longer across the board when people produced the island + resumptive sentences compared to the surface-structure controls. Initiation times were 2328 and 2104 respectively. The means for utterance duration are shown in Figure 2. The part of the sentence where differences in duration reached conventional levels of statistical significance was at the head noun of the relative clause (i.e., immediately after *This is a*) and on the complementizer (e.g., *that*). In addition, the overall duration of the utterances was longer for the island + resumptive sentences, consistent with the idea that they are harder to produce than are the surface-structure controls.

Figure 2: Durations of Island + Resumptive Sentences and Controls, in Regions
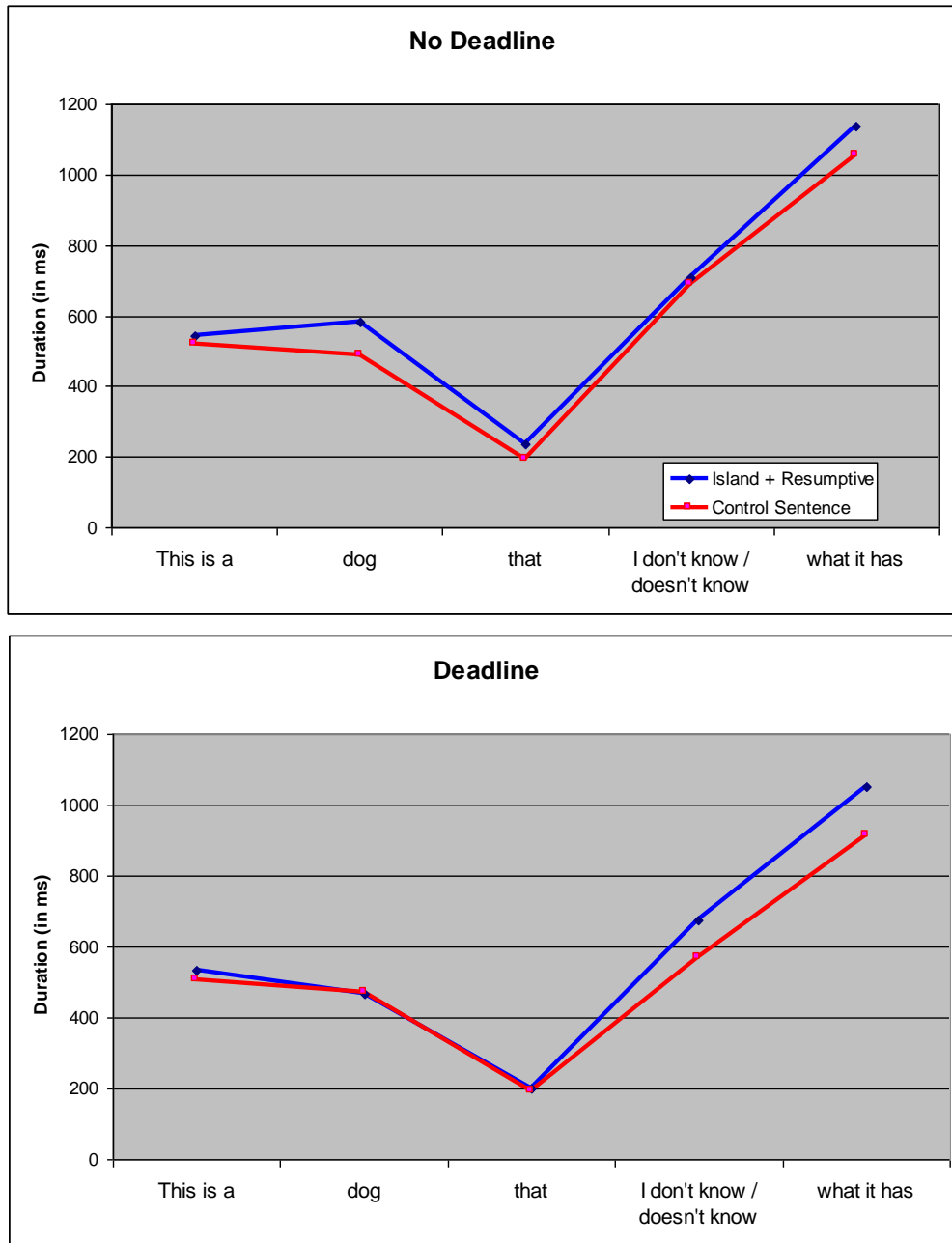
Results for the second experiment, which required participants to respond before a deadline, are also shown in Figure 2. Clearly, the deadline caused people to begin to speak more quickly than they did in the first experiment: Overall initiation times dropped from over two seconds to less than one. As in the no-deadline experiment, those initiation times did not differ significantly between conditions (853 and 797 ms in the island + resumptive and control conditions, respectively). Total utterance duration was longer for the island + resumptive sentences. There also were significant effects in the analyses in which the utterances were broken

down into parts. In the island + resumptive condition, it took people longer to say *I don't know* than it took people in the surface-structure control condition to say *doesn't know* (674 v. 570 ms). Given that these two phrases have the same number of syllables and stress pattern, it is likely that the difference is due to the greater difficulty of producing the island + resumptive sentences. The duration of *what it has* was also longer for the island + resumptive sentences than the controls.

Comparing the findings across conditions and experiments, two points clearly emerge. One is that the island + resumptive sentences take more processing resources to produce than do the acceptable controls. But more interestingly, it also appears that the difficulty of producing these sentences shows up fairly early in their production, especially when the speaker is not under pressure to respond quickly. In the no-deadline experiment, latencies diverged right at the beginning of the relative clause containing the island, which is one clause before the part of the sentence that contains the resumptive pronoun. And even in the deadline experiment, durations were longer a little further into that relative clause (at the *I don't know* portion). At the same time, it also appears that speakers are more incremental in their production when there is a deadline to begin to speak, consistent with what we observed in our arithmetic study (Ferreira & Swets, 2002). The locus of the difference moved further out into the utterance, suggesting that the production system was using less look-ahead. These results overall suggest that (a) the production system is only somewhat incremental, and (b) the extent to which incrementality is observed depends on the speakers' strategies for managing the communicative situation (in this case, the relatively impoverished one associated with participating in psychological experiments).

Let us now turn to the types of utterances that speakers produced, beginning with the no-deadline experiment. In the island + resumptive condition that was designed to elicit island + resumptive sentences, about 67% of all utterances were of the desired type. This finding is actually quite striking considering that the form is not very acceptable (more on this point later). In the deadline experiment, the percentage of sentences of this type dropped to 56%. Surprisingly, then, people are **less** likely to produce this marginal structure when they are under time pressure, a finding which goes against the general belief that the island + resumptive structure is produced when people do not plan properly and so essentially paint themselves into a syntactic corner (Creswell, 2002). Consistent with this idea are the findings for alternative structures. These other forms can be viewed as ways that speakers avoided producing a marginal form like the island + resumptive. The strategies included coordination (e.g., *This is a donkey and I don't know where it lives*) and left-dislocation (e.g., *This donkey I don't know where it lives*). In the no-deadline experiment, these alternative, licit structures occurred on 18% of trials; but in the experiment in which people spoke under time pressure, alternatives were actually produced **more often**: 21% of the time. We conclude that the island + resumptive form is not a mistake; it is a structure that the production system intends to produce. Moreover, its generation clearly requires significant processing resources. Under time pressure, the grammatical encoder opts not to create this form, perhaps because it is a hard structure.

The unexpected finding that the island + resumptive structure is intentionally produced has two further interesting theoretical implications. One is that it seems to support a modular architecture for the production system. The argument goes like this: The reason the grammatical encoder ends up stuck with the job of producing a marginal form is that the message level does not know about grammatical constraints (especially language-specific ones) and so it gives the

grammatical encoder a communicative intention that is not easily accommodated syntactically. The second implication is that the production system is somewhat egocentric, in the sense that it does not take into account the needs of the listener when it formulates utterances. For this point to be more convincing, we need to describe the results of our grammaticality judgment experiments. We turn to those now.

GRAMMATICALITY JUDGMENT EXPERIMENTS

We wanted to obtain clear-cut evidence that the island + resumptive structure is indeed fairly marginal. After all, it is possible that we obtained the results we did in the production experiments because in the sociolinguistic community to which our participants belong, the form is gaining acceptance. Accordingly, in one experiment, people were visually presented sentences like (3) (repeated below as (4)), and their task was to indicate on a scale from 1 (perfect) to 5 (awful) whether the sentences were acceptable in English.

(4)
        a. [This is a] [donkey] [that] [I don't know] [where it lives].
        b. [This is a] [donkey] [that] [doesn't know] [where it lives].

Consistent with the idea that the comprehension system does not like the island + resumptive structure, sentences such as (4a) were given average ratings of 3.3, and controls like (4b) were given average ratings of 1.9. But perhaps sentences like (4a) are from a spoken register; that is, they are not sentences that tend to occur in written English, so an auditory judgment task might yield more meaningful data. Accordingly, this experiment was re-run, but we used spoken materials with normal prosody. The values for (4a) and (4b) were 3.0 and 1.7 – essentially the same. (Both pairs of means are significantly different.)

        Based on these findings, we have begun conducting experiments to examine in more detail the way island + resumptive sentences are comprehended, using eye movement monitoring and detailed follow-up questions to obtain information about how they are interpreted online and how they are ultimately understood. This work is currently in progress, but we expect it to provide results consistent with those of the grammaticality judgment studies: Participants find the sentences harder to understand than controls.

CONCLUSIONS

This research on the production and comprehension of the marginal island + resumptive structure reveals another important properties of both systems. First, it appears that the production system is only moderately incremental, and the amount of look-ahead that the system needs is determined in part by the speaker's strategy for managing the communicative situation. Second, the more marginal structure takes more processing resources to produce, but nevertheless it is a form the production system chooses; it is not the result of poor planning. This finding leads to our third conclusion, which is that the production system has a modular architecture, because the message level system does not consider grammatical constraints when it creates a message level representation for the grammatical encoder. Fourth, the language production system appears to be somewhat egocentric; it considers mainly **its** needs when it creates utterances, and it does not try to avoid producing a form that the comprehension system will dislike (as revealed by our grammaticality judgment data).

What are some of the more general implications? We have seen that the research strongly suggests non-parity between the production and comprehension systems (contrary to what Pickering & Garrod, in press, advocate). The two systems do not consult the exact same database of grammatical rules, as indicated by the finding that the production system allows the island + resumptive structure to leak through, but the comprehension system tends to reject them. Assuming that the two systems are distinguishable, we can ask which one seems to have priority? Given that the production system seems to force the comprehension system to handle a structure the latter does not like, it can be argued that it is the production system that has priority. At the same time, clearly the two systems interact; if they did not, conversations would be impossible. Moreover, there are theoretical approaches to both domains which assume a certain amount of architectural integration: Models of language production use the comprehension system to filter out bad utterances (Levelt, 1989), and there are models of language comprehension which use analysis by synthesis (comprehension system tries to generate the input utterance) to accommodate semantic effects on parsing and other syntactic decisions (Townsend & Bever, 2001).

What all of this means is that although the production and comprehension systems can certainly be studied independently, real progress in understanding the architecture used for language processing will take place when we compare the two and determine the way they work together. This approach will allow us to learn more about humans' remarkable ability to communicate efficiently. And, of course, these investigations will inform us about the nature of the mind, which is the central question in cognitive science.

REFERENCES

Bailey, K. G. B., & Ferreira, F. (2003). Disfluencies influence syntactic parsing. *Journal of Memory and Language*, *49*, 183-200.

Bailey, K. G. B., & Ferreira, F. (in press). The disfluent hairy dog: Can syntactic parsing be affected by non-word disfluencies? In J. Trueswell & M.K. Tanenhaus (Eds.), *World situated language use: Psycholinguistic, linguistic, and computational perspectives on bridging the product and action traditions*. Cambridge, MA: MIT Press.

Bock, J.K. (1986b). Syntactic persistence in language production. *Cognitive Psychology*, *18*, 355-387.

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, *23*, 45-93.

Bock, K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, *31*, 99-127.

Clark, H. H. (1996). *Using language*. Cambridge University Press.

Creswell, C. (2002). Resumptive pronouns, wh-island violations, and sentence production. *Proceedings of the Sixth International Workshop on Tree Adjoining Grammar and Related Frameworks (TAG+6)*, 101-109. Universita di Venezia.

Ferreira, F. (2000). Syntax in language production: An approach using tree-adjoining grammars. In L. Wheeldon (Ed.), *Aspects of language production* (pp. 291-330). Philadelphia, PA: Psychology Press.

Ferreira, F., Lau, E.F., & Bailey, K.G.B. (in press). Disfluencies, parsing, and Tree Adjoining Grammars. *Cognitive Science*.

Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, *46*, 57-84.

Ford, M., & Holmes, V. M. (1978). Planning units in sentence production. *Cognition*, *6*, 35-53.

Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Language: Psychological and biological aspects*. (pp. 69-96). New York: Cambridge University Press.

Griffin, Z., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, 11(4), 274-279.

Just, M. A., & Carpenter, P. A. (1993). The intensity of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology*, *47*, 310-339.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W.J.M., Richardson, G., & La Heij, W. (1985). Pointing and voicing in deictic expressions. *Journal of Memory and Language, 24*, 133-164.

Pickering, M.J., & Branigan, H.P. The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*, 633-651.

Pickering, M. J., & Garrod, S. (in press). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*.

Prince, E. (1990). Syntax and discourse: A look at resumptive pronouns. *Proceedings of the Berkeley Linguistics Society*, *16*, 482-497.

Ross, J. (1967). Constraints on variables in syntax. Unpublished doctoral dissertation, MIT.

Smith, M., & Wheeldon, L. (2001). Syntactic priming in spoken sentence production—an online study. *Cognition*, 78, 123-164.

Townsend, D.J., & Bever, T.G. (2001). *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.

Trueswell, J., & Tanenhaus, M. K. (Eds.) (in press). *World situated language use: Psycholinguistic, linguistic, and computational perspectives on bridging the product and action traditions*. Cambridge, MA: MIT Press.