Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

HHS Public Access
Author manuscript
Lang Speech. Author manuscript; available in PMC 2019 September 01.

Published in final edited form as:
Lang Speech. 2018 September ; 61(3): 480–512. doi:10.1177/0023830917737110.

INFORMATION STRUCTURE PREFERENCES IN FOCUS-SENSITIVE ELLIPSIS: HOW DEFAULTS PERSIST

Jesse A. Harris and
UCLA

Katy Carlson
Morehead State University

Abstract

We compare the roles of overt accent and default focus marking in processing ellipsis structures headed by focus-sensitive coordinators (such as *Danielle couldn't pass the quiz, let alone the final/ Kayla*). In a small auditory corpus study of radio transcripts, we establish that such structures overwhelmingly occur with contrastive pitch accents on the correlate and remnant (*the quiz* and *the final*, or *Danielle* and *Kayla*), and that there is a strong bias to pair the remnant with the most local plausible correlate in production. In two auditory naturalness ratings experiments, we observe that marking a non-local correlate with contrastive pitch accent moderates, but does not fully overturn, the bias for local correlates in comprehension. We propose that the locality preference is due to a sentence-final default position for sentence accent, and that auditory processing is subject to 'enduring focus,' in which default positions for focus continue to influence the focus structure of the sentence even in the presence of overt accents. The importance of these results for models of auditory processing and of the processing of remnant ellipsis structures is discussed.

Keywords

Information structure; default focus; contrastive pitch accent; ellipsis; corpus

1. Introduction

Sentence comprehension is clearly a multifaceted affair, in which a great deal of information must be evaluated in real time against a diverse collection of requirements and constraints. In addition to establishing lexical category information, constructing basic syntactic dependencies, and resolving semantic relationships, language users intuitively relate sentential and sub-sentential meanings to a larger discourse by attending to the *information structural status* of constituents and their configurations within the sentence. The theory of information structure traditionally employs a collection of concepts that identify elements along dimensions like topic, focus, given-new, contrast, and so on, which control how information is entered into, and maintained within, common ground (e.g., Chafe, 1976;

Corresponding author: Jesse A. Harris, Department of Linguistics, 3125 Campbell Hall, University of California, Los Angeles, CA 90095-1543, USA., jharris@humnet.ucla.edu.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Lambrecht, 1994; Krifka, 2008). Although focus can be marked explicitly in English, using such means as pitch accent placement or a cleft structure, such markings are not obligatory. Instead, English often relies on powerful defaults, and deviations therefrom, in order to signal information structure status. In this study, we are concerned with how those defaults interact with explicit focus marking in the interpretation of ellipsis structures.

We concentrate on a particular form of ellipsis known as *focus-sensitive coordination*, a construction which involves coordinators like *let alone* and *much less*, as in *John doesn't like coffee, let alone tea*. We do so for several reasons. First, interpreting focus-sensitive coordination requires two or more elements to be placed in contrastive focus (Fillmore et al., 1988; Hulsey, 2008; Toosarvandani, 2010), making the focus structure of the sentence instrumental for its interpretation. Second, the construction exhibits a strong preference for the contrastive elements to be in a local relationship (Harris & Carlson, 2016), similar to other well-studied ellipsis structures, such as sluicing (Frazier & Clifton, 1998; Carlson et al., 2009; Dickey & Bunger, 2011; Harris, 2015). Though the origin of this preference is not fully settled, it has been attributed to a default sentence-final position for sentence accent (Carlson et al., 2009), which allows us to explore the extent to which default focus interacts with markers of overt focus. Finally, focus-sensitive coordination structures (FSCs) have not been the subject of experimental manipulation until recently (Harris, 2016; Harris & Carlson, 2016), and they exhibit a wide range of strong grammatical and pragmatic requirements that make them ideal for studying the relationship between structure and discourse in a controlled fashion. These considerations provide a valuable case study for exploring the role of default focus structures in auditory language processing.

In one auditory corpus study and two auditory questionnaires, we present evidence that the overt placement of focus marking, i.e. pitch accent position, affects the acceptability and eases the understanding of FSCs, but default expectations about focus position remain active in processing, as well. We are led to reject the idea that the Locality bias can be completely overturned by a pitch accent on a non-local correlate, which in turn lends support to the continued influence of default focus even in the context of overtly focus-marked speech. Further, we replicate the central finding in a follow up experiment, showing that the Locality bias is not limited to simple SVO sentences, and that the bias dissipates when syntactic disambiguation removes impossible correlate-remnant pairings.

1.1 Information structure background

Informally, information structure serves to organize content by its role in the discourse. This organization may be achieved in many ways. Typically, discourse-old/given information precedes discourse-new information (Firbas, 1966; Chafe, 1970; Prince, 1992), a trend that governs how discourse participants expect a sentence to unfold (Clark & Haviland, 1977). In English, and many other languages, special or non-canonical organizations are indicated by special or non-canonical markings (e.g., Horn, 1984; Ward & Birner, 2004). For example, the cleft structure in (1b) focuses *Bob* in a way that a canonical subject-verb-object structure (1a) does not.

(1) **a.** Bob ate the cake.

b. It was Bob who ate the cake.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Similarly, constituents may be marked as prominent or focused via pitch accents: suprasegmental targets that signal the information structural status of an element in the current discourse. In the ToBI (Tones and Breaks Indices) framework, for example, pitch accents include high or low pitch targets, H* and L*, and their combination, e.g., L+H*, with the tone that associates with the stressed syllable of an accented word marked with an asterisk (e.g., Pierrehumbert & Hirschberg, 1990; Beckman & Elam 1997). Accented words are also often longer in duration and spoken with more intensity than unaccented ones. There is debate within the prosodic community in whether H* and L+H* accents are different categories or merely different points on a continuum of more or less high, more or less steep H* accents (e.g, Bartels & Kingston 1994; Ladd, 2008; Ladd & Morton, 1997; Ladd & Schepman, 2003). We will use the L+H* category and describe accents of this type as contrastive, but those who do not believe in the category may assume that the accents referred to are on the higher, steeper end of the H* continuum. Non-contrastive H* accents will be lower, more rounded, and less clearly preceded by a low F0 target.

Most of the semantic literature on focus assumes that a focused constituent contains a pitch accent, though focus may project beyond the accented constituent (e.g., Jackendoff, 1972; Selkirk, 1984; Büring, 2012, inter alia). Although such accents do not always affect the truth-conditional meaning of a sentence, especially in the absence of focus-sensitive operators like *only* (Horn, 1969; Beaver & Clark, 2008), they can strongly affect the appropriateness of an utterance. The most prominent and focused element in (2a) is *Bob*, which corresponds to the information sought by the *wh*-question *Who ate the cake*? Here, *Bob* is appropriately marked as being new (or non-presupposed) information by accenting. However, the same string of words becomes infelicitous as a response to (2) if any of the discourse old / given information (*ate the cake*) receives pitch accent (2b).

(2) Who ate the cake?

a. BOB ate the cake.

b. # Bob ate the CAKE.

We observe the opposite pattern in (3), where the object is now queried by the *wh*-question. Together with (2), the example shows that what should be prosodically marked as new information, and hence what appears in focus, depends on the context and the discourse status of the element (Jackendoff, 1972, among many others).

(3) What did Bob eat?

a. # BOB ate the cake.

b. Bob ate the CAKE.

In case context is absent or weak, prominence in English is usually assumed in the default position for primary (nuclear) sentence accent, on the most deeply embedded constituent at the right edge of a sentence (Chomsky & Halle, 1968; Halle & Vergnaud, 1987; Selkirk, 1984; Cinque, 1993). For instance, the most natural response to the question in (4) is (4b), as the entire sentence is new information in the discourse. Although one might reply with (4a), it would require some degree of accommodation on the part of the addressee, e.g., the speaker believes that eating the cake was or should have been available in common ground.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(4) What happened?

a. ?? BOB ate the cake.

b. Bob ate the CAKE.

Of course, the tendency to put prominence late in the sentence is not absolute, as pitch accent placement is governed by many factors, including the phonological requirement to place an accent within each phonological phrase, or the desirability of alternating rhythm (Ladd, 2008; Selkirk, 1995, among others). Thus, the presence of a pitch accent does not necessarily signal that an element is in focus.[1] As a consequence, the auditory processor may not be certain whether prominence in the acoustic input transparently marks the location of focus within an utterance, and may thus consider default locations for foci during interpretation.

In this project, we propose that default locations for focus interfere with the identification of focus in non-default positions from overtly marked constituents (5).

(5) **Enduring focus:** Locations that typically bear default focus continue to provide potential locations for focus, regardless of overt markers of focus.

In other words, we expect that hearers may be tempted to default to standard locations of focus, even when there are overt cues, like pitch accent, to the contrary. This is not to say that hearers are completely insensitive to the pitch accent, or that they wholly rely on information structural defaults. The enduring focus principle may instead reflect an efficient processing strategy which attempts to reduce uncertainty in the message by relying in part on conventional mappings between given and new information. We expect that this principle is particularly robust in cases of weak or impoverished context, where the hearer has little independent reason for assigning given and new status besides the sentence itself. We speculate on the origins and mechanism of enduring focus further in the General Discussion section below.

There is some prior research compatible with the enduring focus principle. Stolterfoht, Friederici, Alter, & Steube (2007) showed ERP evidence of an expectation for focus on the object in German, with markers of revision of structure when the focus particle *nur* 'only' or a later contrastive remnant indicated a different focus position. Carlson et al. (2009) found a lasting bias for focus on the object or the latest argument evident in studies of the processing of sluicing ellipsis sentences. In a series of acceptability rating studies on corrections in spoken dialog, Clifton and Frazier (2016) manipulated whether a correction to a statement was syntactically parallel with the corrected statement (*Speaker A: Mary brought the pie; Speaker B: No, Susie brought the pie / No, the pie was brought by Susie*). In addition to an advantage for parallelism, they found an interaction in which non-parallel corrections became more acceptable when the contrastive term (*Susie*) was situated in a position bearing default focus, e.g., the object position. The effect of default focus position was present even when given information (*the pie was brought*) carried the most prominent pitch accent,

[1]A reviewer claims that this uncertainty should not pertain to nuclear accents. We suggest that even nuclear accents might, in the midst of sentence processing, be forgotten, misheard, or misinterpreted (a la the noisy channel theory of processing: Gibson, Bergen, & Piantadosi 2013).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

suggesting that default syntactic positions for focus continue to influence which constituent is in focus. All of these results are highly compatible with the enduring focus principle, and suggest that focus marking in English is determined by both syntactic position and pitch accent. We now turn to the use of information structure in resolving ellipsis, and the effect that enduring focus plays in interpreting focus-sensitive constructions.

1.2 Focus-sensitive coordination

Constructions with *let alone, much less*, and certain uses of *never mind*, exemplify focus-sensitive coordination structures. At a descriptive level, these coordinators appear to coordinate two elements of nearly any syntactic type, except fully inflected clauses (Hulsey, 2008). They are typically produced with contrastive focus on the second conjunct (the remnant: *a contractor* in (6a)), as well as the element in the matrix clause with which it contrasts (the correlate: *a carpenter* in (6a)), as shown in (6).

(6) **a.** John didn't call a CARPENTER, let alone a CONTRACTOR.

b. John didn't CALL a carpenter, let alone HIRE one.

c. John didn't CALL a CARPENTER, let alone HIRE a CONTRACTOR.

In mainstream dialects of English, the coordinator is licensed in negative environments of various sorts, as well as pragmatically adverse situations, but not in positive environments.[2]

Besides their grammatical properties, focus-sensitive coordination structures require comparison along a contextually salient scale for the correlate and the remnant, such that the negation of the correlate strongly implies the negation of the remnant – e.g., that John didn't call a carpenter strongly implies that he did not hire one (6b). These scales are typically *ad hoc* and rely heavily on dimensions that can be implicitly ascertained from the context, even when such dimensions are complex (7).

(7) **a.** I can't drink TEA, let alone COFFEE, past 5 PM.

i. *Implication:* If I can't drink tea, then I can't drink coffee.

ii. *Scale ordered by caffeine strength:* Tea has less caffeine than coffee.

b. I didn't have time to FEED the CHILDREN, let alone PREPARE my LECTURE. (Fillmore et al., 1988: 531)

i. *Implication:* If I don't have time to feed the children, I don't have time to prepare my lecture.

ii. *Scale ordered by time required to complete tasks:* Feeding children takes less time than preparing a lecture.

[2]Some dialects support a 'positive' *let alone* use, which permits the coordinator without negation, as in *John called a carpenter, let alone a contractor* (http://itre.cis.upenn.edu/~myl/languagelog/archives/005142.html). These dialects are thought to treat the coordinator as introducing an afterthought, similar to *not to mention*, rather than a scalar entailment between items in contrastive focus (Toosarvandani, 2010; Cappelle, Dugas, & Tobin, 2015).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

That the entailment pattern is strongly encoded in the ordering of the correlate and the remnant can be seen more directly in the infelicitousness of (8b), which violates the logical entailment between *two cookies* and *five cookies*, compared to the minimally different variant (8a).

(8) **a.** I didn't eat TWO of your cookies, let alone FIVE of them.

b. # I didn't eat FIVE of your cookies, let alone TWO of them.

While Fillmore et al. (1988) used the seemingly idiosyncractic grammatical properties of *let alone* and related coordinators to argue for a construction-based grammar, recent studies have converged on an ellipsis account of their syntax (Hulsey, 2008; Toosarvandani, 2010; Harris, 2016). For instance, Harris (2016) proposes that the syntax of *let alone* coordination introduces stripping ellipsis (Frazier, et al., 2012; Sailor & Thoms, 2014), in which the second conjunct is a coordinated *v*P (a VP with an external argument, in the sense of Marantz, 1984; Chomsky, 1995; Kratzer, 1996) or CP (a complementizer phrase) having undergone mandatory ellipsis of all but the focused elements. We assume that there are several sizes of a clause, and refer to the kind of ellipsis in *let alone* constructions as clausal ellipsis, as in the case of fragment ellipsis (Merchant, 2005; Weir, 2014). In this analysis, example (7a) is the result of moving the remnant *coffee* to a focused position (FocP) and eliding the remainder of the clause; a simplified bracketed structure is provided in (9).

(9) I can't drink TEA, let alone $[_{\text{FocP}}$ a COFFEE$]_1$ <I drink t_1 >

Since evidence for the idea that *let alone* structures are distinct from ordinary coordination structures has been documented elsewhere, we simply illustrate the point with two examples from Hulsey (2008) and Harris (2016), respectively. First, in gapping ellipsis under ordinary coordination (*and, or*), an active verb may be omitted or expressed in its uninflected form (10a). The uninflected form (*Sue eat veal*) takes low scope under negation, consistent with the gapping interpretation (*Sue -- veal*), whereas the inflected form (*Sue eats veal*) coordinates above the negation (Siegel, 1984). In contrast, *let alone* cannot coordinate a clause with an active inflected verb (10b), suggesting that it coordinates material below negation or at least up to the vP, assuming the subject-internal hypothesis (e.g., Kuroda, 1988; Koopman & Sportiche, 1988).

(10) *Constraints on inflection*

a. JOHN didn't eat CAVIAR, and SUE (-- / eat / eats) VEAL.

b. JOHN didn't eat CAVIAR, let alone SUE (-- / ?eat / *eats) VEAL.

The second example comes from sprouting, in which the remnant of the ellipsis, e.g., *caviar* in the replacive ellipsis example (11a) or *what* in the sluicing ellipsis example (11b) lacks an overt correlate in the host clause (Chung et al., 1995). *Let alone* structures (11c) pattern with cases of clausal ellipsis (11a–b) in allowing sprouting, and not with ordinary coordination, in which sprouting is not permitted (11d).

(11) *Sprouting*

a. John ate, but not caviar.

b. John ate, but I don't know what.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

c. John didn't eat, let alone caviar.

d. * John ate and/or caviar.

The purpose of the present paper is not to defend any particular syntactic analysis of this construction, but we believe the close correspondence in its processing preferences to other kinds of ellipsis structures, particularly stripping and sluicing, is important. In particular, we believe that it sheds light on the character of the relations that must be established by the processor when interpreting ellipsis in general, briefly reviewed below.

1.3 Processing ellipsis

Although much can be said about approaches to processing ellipsis structures (e.g., Phillips & Parker, 2014), we assume that there are three basic tasks that the processor must solve in order to interpret any ellipsis involving a remnant (12). For concreteness, we adopt the view that a basic covert syntactic / logical structure exists at the ellipsis site at some level of representation, in keeping with a great deal of syntactic literature (e.g., Merchant, 2001) and experimental work (e.g., Frazier, 2008, Poirier et al., 2010).

(12) *Basic tasks of the processor in ellipsis processing:*

1. Parse the remnant by constructing the appropriate phrase structure for the remnant given the input.
2. Locate the correlate, if any, from the antecedent clause.
3. Construct the elided phrase by regenerating or copying a structure at Logical Form.

We use example (13) to illustrate the three tasks in more depth:

(13) I can't drink tea, let alone coffee

1. *Parse the remnant:* Assign the appropriate phrase structure for *coffee*.

   I can't drink tea, let alone [$_{\text{DP=Remnant}}$ coffee]

2. *Locate the correlate:* Retrieve an appropriate correlate that provides a suitable contrast to the remnant *coffee*, using various processing strategies.

   I can't drink [$_{\text{DP=Correlate}}$ tea], let alone [$_{\text{DP=Remnant}}$ coffee]

3. *Construct the elided phrase*: Build the ellipsis structure after the remnant.

   I can't drink [$_{\text{DP=Correlate}}$ tea], let alone [$_{\text{DP=Remnant}}$ coffee]$_1$ < I drink t_1 >

Conceptually, steps 2 and 3 each depend on the step that precedes it. For step 2, assuming that remnants must be of the same syntactic type as the correlate, the remnant needs to be parsed before the processor can retrieve an appropriate correlate.[3] Similarly, for step 3, the

[3]The remnant can also be a larger syntactic unit which contains the contrastive element, so that the correlate contrasts with only a sub-part of the remnant. See Harris & Carlson (2016) for examples and discussion.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

basic argument position for the remnant trace must be determined before the elided structure can be formulated, and we assume that the processor prefers analyses with parallel structures between clauses, as has been observed for basic conjunction with *and* (e.g., Frazier, Taft, Roeper, Clifton, & Ehrlich, 1984) and gapping ellipsis (e.g., Carlson, 2001, 2002).

We are particularly interested in step 2, which perhaps relies the most heavily on default resolution strategies. One such strategy was discussed in Frazier & Clifton's (1998) study of sluicing ellipsis, in which the processor appears to prefer the closest correlate, which was later generalized as the Locality bias (14).

(14) **Locality bias:** Contrast the remnant with the nearest constituent (of the appropriate type) in the preceding clause (Harris, 2015; Harris & Carlson, 2016)

We have remained agnostic about whether the nearest relation should be explicated in structural position, temporal precedence, or linear distance, as these co-vary in our experimental manipulations below. Given the relationship between focus and structural embedding, we suspect that the relationship is structural, although we leave open whether other factors, such as temporal precedence, may independently influence which correlate is paired with the remnant. The experiments discussed below were not designed to tease apart these factors (see comments in the General Discussion section).

In support of the Locality Bias, Frazier & Clifton (1998) observed a reading time advantage for sluices with a viable correlate for the remnant *who* in object position (*someone*), even when there was another possibility in subject position (*somebody*); see also Carlson et al. (2009) and Harris (2015) for congruent results.

(15) *Somebody* claimed that the president fired *someone*, but nobody knows who.

In a follow up paper, Carlson et al. (2009) found that the closest noun phrase, the object *the co-pilot*, was the preferred correlate for ambiguous sluices (16). This was true whenever the prosodic marking was consistent with the possibility, as with contrastive accent on the object in (16a), on the verb (*talked*) in (16b), or on both the subject (*captain*) and the object noun in (16c). Only when the subject alone received contrastive accent (16d) was the subject noun phrase consistently understood as the correlate.

(16)

a. The captain talked with the CO-PILOT, but we couldn't find out who else.

b. The captain TALKED with the co-pilot, but we couldn't find out who else.

c. The CAPTAIN talked with the CO-PILOT, but we couldn't find out who else.

d. The CAPTAIN talked with the co-pilot, but we couldn't find out who else.

However, the bias towards subject correlates in subject accent cases (16d) was markedly lower than the bias toward the object in all other cases (about 60% compared to 80%). Carlson et al. (2009) suggest that the default syntactic position for focus continues to exert an influence on interpretation even when overt marking of focus does not support that

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

interpretation. When the overt cues and structural defaults for focus assignment converge, the interpretation is robust. When they diverge, however, the consistency of judgments is weakened considerably, as comprehenders may have to choose between what their ears perceive and their expectations about the information structure of the sentence. In other words, their results strongly support the idea that pitch accent position guides correlate resolution in sluicing, but that the processes that identify prominence are subject to enduring focus (5).

The bias for more local correlates has been corroborated in studies of other types of ellipsis, including focus-sensitive ellipsis (Harris & Carlson, 2016), as well as the closely related case of stripping / replacive ellipsis (Carlson, 2013). In a corpus study of *let alone* in British and American English, Harris & Carlson found that approximately 84% of remnants of *let alone* ellipsis in their sample were associated with local correlates, with a somewhat weaker bias for NP remnants. In two self-paced reading experiments, they found evidence that the Locality bias is also observed during silent online sentence processing. They manipulated the location of a contrastive adjective like *nicest* to modify the subject (non-local) noun (17b), or the object (local) noun (17a), thereby varying which noun phrase formed a likely semantic/discourse contrast with the remnant containing an adjective like *meanest*. As expected, reading times on the remnant (*the meanest one*) and the spill over region (*and no one at the hospital*) after the *let alone* coordinator were increased following a non-local contrast compared to the expected local contrast.

(17) **a.** The nurse couldn’t stand the *nicest* patient, let alone the meanest one, and no one at the hospital was happy at all. (Object / local correlate)

b. The *nicest* nurse couldn't stand the patient, let alone the meanest one, and no one at the hospital was happy at all. (Subject / non-local correlate)

In all, a range of experimental studies provides compelling evidence that sentence processing is facilitated when there is a local correlate for the remnant of clausal ellipsis. In the case of *let alone*, this preference has been confirmed across a wide range of written material, both in corpora and in silent reading. However, previous studies have not explored how the Locality bias might interact with overt focus marking in focus-sensitive coordination structures. This study aims to distinguish between two possibilities regarding the relationship between focus (as realized through pitch accent) and Locality.

The first possibility identifies focus and information structure as the primary source for Locality preferences in clausal ellipsis, and allows a significant role for enduring focus in the processing of contrastive material. This possibility makes two predictions. First, constituents that are accompanied by contrastive pitch accents should be preferred as correlates to the remnant. Second, assuming that default focus positions can intercede in the assessment of focus structure, accented constituents in non-local positions should be understood as correlates less readily and less easily than those in local positions. We assume that the syntactic structure of the antecedent clause indicates the location of default accent, in that the most deeply embedded constituent will receive nuclear pitch accent (Cinque, 1993).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The second possibility is that focus-sensitive coordination structures require contrastive accents for interpretation, because they are explicitly contrastive. Thus, listeners might well be attuned to explicit pitch accents to the exclusion of default preferences. On this view, the placement of pitch accents within the first clause should fully determine which noun phrase serves as the correlate for the remnant. This possibility would require us to either abandon the enduring focus principle, despite the arguments made above, or limit its domain to cases in which prosodic marking is optional or otherwise less central to the task of interpretation. Assuming a relatively direct pathway from a contrastive pitch accent to a contrastive focus structure, this possibility has a clear appeal for the design of the auditory processing system, as listeners would have little cause to doubt which noun phrase provides the intended correlate.

To investigate these possibilities, we present results from a corpus study of spoken radio transcripts, and two auditory rating studies. The results together support the central predictions of our first information structure account of enduring focus, in which explicit contrastive accent strongly influences, but does not wholly override, default focus position.

2. Auditory corpus study

Over 20,000 links to radio interview transcripts available on the National Public Radio website (npr.org) were collected through the BeautifulSoup and Scrapy web-scraping modules in Python. Transcripts were downloaded into an XML (Extensible Markup Language) format and searched for the coordinators *let alone, much less*, and *never mind*. 227 distinct cases were uncovered. Sixteen examples that included speech errors, irrelevant uses, or clearly non-native speakers were removed, leaving 211 instances (59 female / 152 male) of focus-sensitive coordination (141 cases of *let alone*, 65 of *much less*, and 5 of *never mind*). Audio portions containing the target sentence were excerpted from radio segments provided by NPR. Although this is an interesting set of data, it is still a small one, and given the restriction to NPR, contains several utterances by the same speakers. Thus we take the findings below to be useful but not definitive indicators of how these structures are pronounced.

The authors annotated the dataset along multiple dimensions, including (i) syntax of remnant, (ii) locality of correlate, and (iii) prosodic contour, including the accent / type of accent on correlate, the accent / type of accent on remnant, the presence/type of prosodic boundaries before the coordinator, and the accent on coordinator (*let alone, much less*, or *never mind*).

2.1 Syntax of remnant

The most common remnant types were NP (48%) and VP (32%) remnants, with PPs, sentence complements, bare Vs, Adverbs, Adjectives, and Determiners making up the remaining 20% of cases, each category comprising less than 8% of the data. The syntactic distribution of the remnant categories closely follows the distribution found in the British National Corpus and the Contemporary Corpus of American English (COCA) for both *let alone* and *much less* constructions (Harris & Carlson 2016; Davies 2008), which suggests

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

that this NPR sample, though small, is representative of speech and text recorded in other, previously explored, corpora.

2.2 Locality of correlate

As in Harris & Carlson (2016), the majority of correlates appeared in the most Local position with respect to the remnant for all remnant categories (88% of total cases). NP remnants accounted for the greatest deviance from the Locality bias with 16% of cases being Non Local, followed by PP remnants at 14% Non Local. Examples of Non Local correlates for each of the three coordinator types is provided below, where the correlate and remnant are marked in **bold**, and the more local, non-correlate is presented in *italics*.

(18) What hurts right now is that there are millions of people who cannot put **new clothes** on *their kids' backs*, let alone **designer duds**, or **keep a roof over their heads**, no matter how hard they *try*, let alone **jet off to pricey vacation spots**.

(19) Well, if so, then that's a very interesting chapter in human history, but I'm not going to stake **my life** on *it*, much less **my hope for eternity** …

(20) You know, I don't know **adults** who drive *that car*, never mind **freshmen in college**.

2.3 Prosodic contour

Prosodic analysis of the auditory files revealed that every correlate was produced with an accent, and that the pitch accent was judged to be contrastive 79% of the time (specifically, a steep, high L+H* accent). Similarly, every remnant bore a pitch accent, and most had contrastive accents (73% of cases). For the two major syntactic categories of remnants, NPs and VPs, a majority of examples had contrastive accents on both the correlate and the remnant (56% for NPs, 67% for VPs). Only a small number of examples had non-contrastive accents on both correlate and remnant (9% for NPs, 7% for VPs) and the remaining examples had contrastive accents on either the correlate or the remnant and non-contrastive accents on the other. A prosodic contour with contrastive accents on correlate and remnant was clearly the most common rendition of the construction. Non-local and local examples had similar percentages of accented correlates and remnants to the whole data set, though the number of non-local examples was small (just 25 of 211 examples). Sample pitch tracks for cases with local and non-local correlates are provided in Figure 1.

Interestingly, most examples (85%) had pitch accents on the coordinator phrase itself (*let alone, much less* or *never mind*). This had not been expected but emerged as a phenomenon to analyze through observation of the sound files. These were usually non-contrastive accents (27% were contrastive). However, there were occasional examples where the accent on the coordinator was the most prominent element in the phrase.

Another common pattern was a prosodic boundary before the coordinator, appearing in 97% of the examples. Close to half of these prosodic boundaries (42%) were intermediate phrase boundaries, the smaller of the two prosodic phrase types in the ToBI system (94% of them ending with L− phrase tones). Another 40% of examples with boundaries had Intonational

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Phrase (IPh) boundaries with L−L% boundary tones, 14% had IPh boundaries with continuation rises (L−H%), and 4% had IPhs ending in H−L% plateaus.

2.4 Discussion

Although the relatively small size of the dataset makes it less than conclusive, the auditory corpus data validates the intuitive role of contrastive accent in focus-sensitive coordination structures. Correlate-remnant pairs almost always bore pitch accents of some kind, usually contrastive accents. This finding supports the use of L+H* accent on the correlates and remnants in the experiments that follow as a natural and consistent way to produce FSC sentences, as well as indicating the intended position of focus and marking the contrast between correlate and remnant. Furthermore, we find that the Locality bias is observed in speech recorded for radio interviews, although there are a healthy number of exceptions, indicating that the preference is not absolute.

The fact that correlates and remnants were uniformly marked with clear pitch accents in this dataset provides some naturalistic support for the idea that accents are nearly obligatory in focus-sensitive coordination structures. In addition, it lends some plausibility to the alternative possibility that listeners could in principle rely solely on overt accent placement to determine focus structure, at least in the case of focus-sensitive coordination. In that case, there would be no clear utility to enduring focus, so that comprehenders would resist assigning focus based on its default position.

We explore the relationship between pitch accent marking and the Locality bias in two controlled experiments that follow. As a preview, we find that explicit pitch accent fails to override the preference for the most local correlate, but that this preference is active only when the remnant is compatible with the most local correlate. We interpret these results in terms of the continued, and powerful, role of default focus in sentence comprehension even when prominence is marked overtly.

3. Experiment 1

3.1 Participants

Fifty-six subjects were recruited on Amazon's Mechanical Turk (AMT). Subjects self-reported as native speakers of English, and were compensated with $2 for completing the experiment, which typically took approximately 25 minutes to complete. Although subjects were distributed evenly across four counterbalanced lists in a within-subjects Latin Square design, the data for one subject was corrupted and not replaced.

3.2 Materials

Twenty quartets like (21) were constructed, crossing *Accent location* of a contrastive L+H* accent (Subject vs. Object) and *Remnant type* (Subject vs. Object). Remnant types were disambiguated by the animacy of the second conjunct, in that remnants with Subject correlates were names of people like *Kayla* in (21b,d), or definite descriptions like *the coach*, whereas remnants with Object correlates were inanimate nouns, e.g., *the final* (21a,c). Each Remnant type followed a matrix clause (*Danielle didn't pass the quiz*) with a

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

contrastive pitch accent located either on the Subject or on the Object noun phrase. Although differences in animacy might create differences between the saliency of the subject and object, the animacy manipulation does allow us to unambiguously assess which correlate the participant took as the remnant. The design minimizes the effect of animacy as a confounding factor, in that pitch accent location was crossed with animacy of the remnant, which allows us to test for an interaction between conditions. In addition, we accented the *let alone* coordinator with an H* accent, and placed a L−H% continuation rise at the end of the matrix clause and the end of the remnant, as was seen in the majority of examples from the auditory corpus study. All items are provided in Appendix A.

(21) **a.** Danielle didn't pass the QUIZ, let alone the FINAL.

b. Danielle didn't pass the QUIZ, let alone KAYLA.

c. DANIELLE didn't pass the quiz, let alone the FINAL.

d. DANIELLE didn't pass the quiz, let alone KAYLA.

All sentences were recorded by the second author and analyzed for adherence to these prosodic specifications. Any sentences that did not have the intended contour were re-recorded. Acoustic measurements to substantiate the listed features are in Tables 1–2. Specifically, words with contrastive pitch accents were significantly higher and longer on average than those words in the other conditions, and all had the characteristic narrow peak shape of this type of accent, as illustrated in Figure 2. Objects in subject accent conditions were usually deaccented, while subjects in object accent conditions bore mild H* accents, as recorded sentences beginning with deaccented material sound quite odd.

The materials additionally varied whether the animate remnant could plausibly be understood as the object of the matrix verb. For example, it might be plausible for Ryan to lift either Mary or the barrel in (22a), but it is implausible, not to mention gruesome, for the patient to eat his family in (22b).

(22) **a.** Ryan couldn't lift the jug, let alone the barrel / Cindy. (Plausible)

b. The patient didn't eat dinner, let alone dessert / # his family. (Implausible)

In order to be able to examine this factor, we balanced the materials to make one half of the items plausible and the other half implausible, as judged by the authors. If comprehenders attempt to initially pair the remnant with the closest, i.e., object, correlate no matter the meaning of the nouns and regardless of pitch accent location, then the interaction between Accent location and Remnant type could be driven by items with implausible object interpretations for animate nouns (22b).

A separate post-hoc norming study tested whether participants perceived the location of prominence in the experimental sentences for Experiments 1 and 2. The audio files for conditions (22a) and (22c) of each item were truncated in the pause before *let alone*, so that the remnant and contrastive construction would not influence perceptions of the first clause (e.g., *Danielle didn't pass the quiz.*). Thirty participants listened to these files (along with

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

several other sentence types in one of two counterbalanced lists) and indicated where they thought the most prominent phrase in the sentence was by answering questions like (23) for the item in (21):

(23) Which phrase did the speaker emphasize?

a. just Danielle

b. just the quiz

c. both Danielle and the quiz

d. other (please enter the phrase)

The response options included sole prominence on the subject (23a) or object (23b), dual prominence on both (23c), and other (23d), with a text box for participants to indicate another phrase. As only one response was provided in the text box (23d), it was removed from the data. One item was coded incorrectly and subsequently removed. Participants identified the subject as the sole location of prominence in 69% of subject accent cases, whereas they selected the object as the sole location of prominence for object accent sentences in 76% of cases. Subjects selected the dual prominence response in 29% of subject accent cases, and in 22% of object accent cases, a difference which is not significant in a binomial test, $p = 0.13$. Further inspection revealed that these responses were limited to a handful of listeners: five respondents chose dual prominence for almost every example (regardless of condition), whereas the rest chose it only rarely. The sentences which actually contained multiple accents, the object accent conditions, did not receive significantly more dual prominence responses overall. The norming results suggest that for most listeners, the initial clauses had clear pitch accents of highest prominence on the intended NPs.

The primary goal of Experiment 1 was to determine whether the Locality bias would persist in sentences bearing clear, contrastive pitch accents. If correlate resolution is guided primarily by overt pitch accent placement, there should be no preference for local correlates. If, however, Locality is at least partially independent of explicit pitch accent, as predicted by a theory of auditory processing subject to enduring focus, then the bias for local correlates should be retained. In addition, the design creates two conditions with a potential mismatch: the Subject accent paired with the Object remnant (21c), and the Object accent paired with the Subject remnant (21b). If perceivers' assessment of focus marking is influenced by default focus in addition to explicit accent, then mismatching sentences should not be penalized equally; mismatching Object remnants (21c) should show higher acceptability ratings than mismatching Subject remnants (21b), since the object position is the default focus position.

3.3 Methods

Experimental items were presented in counterbalanced and individually randomized order, interspersed with 24 items from unrelated experiments and 16 non-experimental filler sentences, for a total of 60 items per participant in each session. Each participant heard only one condition of each item and heard equal numbers of items in each condition over the experiment, following a Latin Square design. Sentences were rated on 7-point Likert scale with labeled endpoints (1 = Totally unnatural; 7 = Totally natural).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

3.4 Results

The data were collected and analyzed in R 3.2.3 (R Core Development Team, 2015). Means and standard errors are summarized in Table 3. A bar plot of *z*-score transformed scores, showing the difference of each condition from the grand mean (M = 5.12) is provided in Figure 3, along with the results from Experiment 2.

The ratings were analyzed in a linear mixed effect regression model, treating Accent location, Remnant type and their interaction as fixed effects, with maximal random effect structures, i.e., by-subject and by-item random slopes and intercepts for the predictor variables. Levels of the predictor variables were given deviation coding (testing each group for significant deviations from the grand mean, as in traditional ANOVA analyses) in which the hypothetically most natural condition (Object correlate – Object accent) was treated as the statistical baseline. As there is some debate regarding how best to compute *p*-values from fixed-effect parameters in these models (Bates, 2006), we adopted the convention that *t*-values over |2| are to be considered significant. The model was computed on the raw acceptability rating scores, although qualitatively identical results were obtained in a model fit to centered *z*-score values.

The mixed effects model yielded several effects; see Table 4. First, there was a main effect of Locality: Non Local (Subject) correlates were rated as less natural than Local (Object) correlates, $t = -6.73$. Second, we observed an effect of Accent location, in which Object focus elicited lower naturalness ratings than Subject focus, $t = 2.49$. Crucially, there was a differential effect of Accent location. While Non Local accent degraded Local correlates ($d = -0.67$), it improved Non Local accent correlates ($d = 1.69$), $t = 4.10$. No other effects were observed. As mentioned above, half of the items contained plausible object interpretations of the subject remnant and half contained implausible ones. However, when Plausibility was added as a fixed effect to the model above, it did not add or remove any significant effects.[4] In all, the lack of a robust effect for plausibility is compatible with the idea that comprehenders' ratings were guided by pitch accent placement, semantic parallelism, and a general preference for local correlates, instead of a failed attempt to integrate the remnant as an object. Nonetheless, the issue is not fully settled, and may require additional testing using online methods.

As noted, Local and Non Local contrasts usually differed not only by animacy, but also by definiteness: sentential subjects were proper names (*Kayla*) in 65% of the items and definite descriptions denoting individual roles (*the patient*) or professions (*the math teacher*) in the remaining 35% of the items. As noted by a reviewer, accommodating an appropriate scale for proper names might be pragmatically more demanding or less constrained than for descriptions. For example in (22), finding a scale for the Non Local contrast condition, *Ryan couldn't lift the jug, let alone Cindy*, requires accommodating some relation between *Ryan*

[4]In particular, there was no main effect of Plausibility, $t = 0.47$, no interaction between Plausibility and Locality, $t = -1.56$, no interaction between Plausibility and Accent Placement, $t = 0.03$, and no interaction between all three factors, $t = 0.32$. The weak numerical trend towards an interaction between Plausibility and Locality could suggest a greater penalty for violating Locality with Implausible subject remnants (d = 1.32) than with Plausible subject remnants (d = 1.04), suggesting that pragmatically forcing the subject reading further induced even lower acceptability ratings than before. However, the potentially low power of this non-orthogonal manipulation obliges such an interpretation to remain speculative. Crucially, the planned effects of Locality, Accent Placement, and their interaction were fully significant in this model, as in the original model limited to planned contrasts.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and *Cindy* along an appropriate dimension such as strength. Finding a scale for the object contrast condition, *Ryan couldn't lift the jug, let alone the barrel*, requires the comparison of *the jug* and *the barrel*, which might be facilitated by their conventional real-world properties, e.g., weight or size. In other words, the availability of real-world knowledge might have eased the computation of the necessary scalar contrast independently of the location of the correlate.

Although we have attempted to control for this issue in the next experiment, we conducted an additional post-hoc analysis on the items, in which the non-orthogonal contrast of Subject type (*Proper name, Definite*) was added as a predictor to the linear mixed effects regression model above. The scores and resulting model are displayed in Tables 5 and 6, respectively.

Model comparison revealed that adding Subject type resulted in a marginally improved model fit, $\chi^2(4) = 9.02$, $p = 0.06$. Although the predicted results are qualitatively identical to the previous model, the more complex model revealed a significant three-way interaction between Locality, Accent Placement, and Subject type. While the penalty for violating Locality was the same within the Local accent condition for Proper names ($d = 1.68$) and Definite description ($d = 1.70$) subjects, the penalty for violating Locality with a Non Local accent was twice the size for Proper names ($d = 0.81$) compared to Definite descriptions ($d = 0.41$), $t = 2.30$. No other effects were observed.

It is important to note that the predicted interaction between Remnant type and Accent placement emerges in both types of items, with either Proper names or Definite descriptions as the subject; the Proper name simply produced a larger penalty for violating Locality. This pattern coheres with our intuition that the likely pairing between the correlate and the remnant can be affected by many factors, including the semantic parallelism between the correlate and the remnant, and the availability of a scalar contrast between the two terms (as manipulated in Harris & Carlson, 2016).

To address this concern directly, a *post-hoc* completion study manipulating the type of sentential subject (Proper name, Definite, Definite with adjective) was conducted over the Internet using Ibex Farm (http://spellout.net/ibexfarm/). Participants were instructed to complete written sentence fragments truncated after the *let alone* coordinator (e.g., *{Patrick / the slacker / the smartest slacker} didn't read the article, let alone …*) with the first natural-sounding completion that came to mind. Sentence fragments were based on the Experiment 1 items, but modified so that the object was always definite and so that subjects varied according to the three conditions. Materials were presented with 56 unrelated experimental items and 5 highly constrained catch items. Items were counterbalanced in 3 equally distributed lists and presented in individually randomized order.

The data consists of 432 responses from 24 native speakers of English who provided sensible completions of catch items, and was annotated for the syntactic category of the remnant (either VPs or NPs), and, in the case of NP remnants, for Locality. There was an overall 65% bias towards VP completions, consistent with previous studies (Harris, 2016; Harris & Carlson, 2016). A local contrast was vastly preferred in conditions with both proper name (98%) and definite description (100%) sentential subjects. When the sentential

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

subject contained an adjective (*smartest*), 65% of completions contrasted with the most local noun, significantly fewer than the other conditions, $\chi^2(2) = 8.81$, $p < 0.05$, which is compatible with Harris & Carlson's (2016) study that used similar adjectives to create non-local contrasts. As mentioned, we suspect that correlate selection is modulated by a great many factors, including the accessibility of a contextually salient scale, locality, and pitch accent location. However, there were no clear differences between responses to proper names and definite descriptions as subject correlates in our materials.

3.5 Discussion

Using a contrastive L+H* accent to mark subject or object nouns as focused, the experiment manipulated the compatibility between accent location and the parallelism of the remnant and correlate (varying the animacy of the NPs). As predicted by a theory subject to enduring focus, structures that violated Locality were rated as less natural than those that did not – i.e., remnants that contrasted with the matrix object were preferred to those that contrasted with the non-local matrix subject. The penalty for non-local correlates was reduced, though not eliminated, when the matrix subject was the only contrastively accented constituent in the matrix clause. The results offer initial support for the principles outlined above, in that the default location for focus continues to influence the interpretation of focus-sensitive constructions, even when focus-marking prosody is present. The pitch accents did affect ratings, though, so it is also clear that listeners were aware of their presence. More generally, these results point to the multiple factors that must be resolved when resolving focus-sensitive phenomena.

4. Experiment 2

The previous experiment investigated the preference for local correlates by using simple, canonical subject-verb-object structures. The study therefore does not resolve how general the Locality bias is, nor whether it could be expressed as a bias against subject correlates (or a bias for object correlates). Subject contrasts are present within corpora of *let alone* and the related *much less* coordinator, though they are not as common as objects or other elements late in sentences. In the auditory NPR corpus discussed above, 6% of the NP remnants contrasted with subjects; in a British National Corpus search for *let alone* constructions, 17% of NPs were subject contrasts; in a COCA corpus of *much less* constructions and a COCA corpus of *let alone* constructions, 15% of NPs were subject contrasts (Harris & Carlson, 2016). It is also true that subject contrasts were more likely than other NP contrasts to have the remnant placed inside the sentence instead of in clause-final position, so about 5% of the NP examples in the corpora had the exact structure tested in Experiment 1. This percentage is not large, but large enough that we believe the constructions are completely grammatical.

The primary purpose of the following experiment is to explore the generality of the Locality bias by comparing remnants that resolve to a contrast with a non-local, but highly salient, head of a relative clause (NP1) compared to remnants that contrast with more local nouns at the ends of the relative clauses (NP2), as in (24). Only subject-extracted relative clauses were used so as to avoid well-known processing difficulties associated with object-extracted

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

relative clauses (e.g., Gennari & MacDonald, 2008; Gordon, Hendrick, & Johnson, 2001; Grodner & Gibson, 2005; King & Just, 1991; Staub, 2010).

(24) $[_{CP}\ [_{NP1}$ NP head$]_1\ [_C$ who / that $[_{IP}\ t_1 \ldots$ RC verb $[_{NP2}$ embedded noun$]]]$

The second aim of the study was to determine whether Locality effects manifest solely in the case of syntactic ambiguity. To that end, we varied whether the non-local, relative clause head correlate was syntactically disambiguated in the remnant to determine if disambiguated remnants engender a similar cost in naturalness ratings.

4.1 Participants

Fifty-six subjects were recruited from AMT to complete an auditory rating and questionnaire study on the Qualtrics platform. Subjects were compensated $4 for completing the experiment, which took less than 30 minutes to complete on average. Eight subjects were removed for counterbalancing purposes, leaving a total set of 48 subjects evenly distributed into six lists, counterbalanced across conditions.

4.2 Materials

Twenty-four sextets like (25) were recorded; items crossed *Accent location* (a. Non Local accent vs. b. Local accent) and *Remnant type* (i. Local, ii. Non Local, iii. Non Local RC). As with Experiment 1, sentences with the Local and Non local levels of Remnant type were disambiguated by the animacy of the noun in the remnant. In the Non Local RC condition, the remnant began with the same noun as the Non Local condition (*a book*) but followed it with a relative clause ending in VP ellipsis. The ellipsis syntactically disambiguated the remnant towards the Non Local correlate, as both consist of a subject relative clause headed by an inanimate noun (*an article that exposed the governor* vs. *a book that did*).

(25) Although he was a highly praised journalist, John didn't write

a. an ARTICLE that exposed the governor, let alone …

b. an article that exposed the GOVERNOR, let alone …

Remnant type

i. the president

ii. a book

iii. a book that did

All sentences were recorded by the second author and analyzed for adherence to certain prosodic specifications, i.e. presence of L+H* contrastive accents on the earlier (Non Local) or later (Local) nouns and the remnants, and a L−L% or L−H% prosodic boundary before *let alone*. Continuation rises followed the initial adverbial phrases or clauses (*Although*… in (25)), and L−L% boundaries before the remnant were as common in the corpus study as L −H%s. The longer remnants (*a book that did*) had contrastive accent on the head noun (*book*) and all remnants ended in L−H% continuation rises. Any sentences that did not have the intended contour were re-recorded. Acoustic measurements to substantiate the listed features are in Tables 1 and 2. Specifically, words with contrastive (L+H*) pitch accents

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

were significantly higher and longer on average than those words in the other conditions, and all had the characteristic narrow peak shape of this type of accent.

The same *post-hoc* norming study presented in Experiment 1 also tested whether the intended prominence contour on the experimental sentences from Experiment 2 was audible to participants. The audio files for conditions (25a) and (25b) of each item were truncated before the main clause and in the pause before *let alone*, so that the remnant and contrastive construction would not influence perceptions of the main clause (e.g., *John didn't write an article that exposed the governor*.). Participants indicated the location of prominence by answering questions like (26).

(26) Which phrase did the speaker emphasize?

a. just an article

b. just the governor

c. both an article and the governor

d. other (please enter the phrase)

The options included sole prominence on the head noun of the relative clause (26a) or on the final noun in the relative clause (26b), dual prominence on both (26c), and other (26d), with a text box for participants to indicate another phrase. Subjects located prominence largely as expected. For the head NP accent sentences (25a), 79% of responses identified the prominence solely on the relative clause head, whereas for the final NP sentences (25b), 77% of responses chose sole prominence on the final noun. Subjects chose dual prominence in about 20% of responses, and the difference between conditions was not significant in a binomial test, $p = 0.74$. The same five subjects were primarily responsible for the majority of dual prominence responses as in the norming task to Experiment 1. These results suggest that for most listeners, the clauses hosting *let alone* ellipsis had clear pitch accents with the highest prominence on the intended NPs.

All of the possible correlates here were positioned within the predicate of the matrix clause, i.e., the relative clause NP (*an article that exposed the governor*). If the Locality bias in Experiment 1 simply reflects a preference for object correlates or a dispreference for subject correlates, then the Local and Non Local conditions in this experiment should not differ. If instead Locality reflects the default focus structure of a sentence, then we expect lower ratings for the Non Local conditions, in which the remnant is paired with a higher NP (*an article*).

4.3 Methods

Experimental items were presented in one of 12 counterbalanced lists in a single pseudo-randomized order, with no consecutive items of the same type or condition. The experimental items were interspersed with 84 items from unrelated experiments for a total of 108 stimuli. Each participant heard only one version of each experimental item and heard equal numbers of items in each condition over the experiment. The same Likert scale as in Experiment 1 was used to collect naturalness ratings.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

4.4 Results

Responses were collected in Qualtrics and analyzed in R; means and standard errors are presented in Table 7, and *z*-score transformed scores showing the deviation from the mean is presented in Figure 3. The data initially were subjected to a linear mixed effects regression model with random slopes and intercepts, as in the previous experiment. However, as this model failed to converge, we used a simplified model with random intercepts only. With respect to the deviation contrast coding, the Local NP (25.i) level was selected as the statistical baseline for the Remnant type factor, and the Local accent level was treated as the baseline for the Accent location factor (25b), as in the previous experiment. We thus test whether the effects of Non Local accent conditions, Non Local correlate conditions, or their interaction are greater than what would be expected from normal deviation from the grand mean.

As predicted, the Non Local NP remnant (*the president*) was rated as less acceptable than the Local NP baseline (*a book*), $t = -2.50$, which did not differ in acceptability from the Non Local RC remnant (*a book that did*), $t = 1.21$ (see Table 6). Non Local accent (pitch accent on the relative clause head) was rated less acceptable overall than Local accent, $t = -2.17$. We attribute this effect to an experimental bias toward the Non Local conditions: the Non Local noun provides the correct correlate in two of the three remnant conditions, thereby producing higher ratings overall for the accent pattern. In addition, Accent location affected naturalness ratings differently for the Non Local remnant types. The ratings penalty for Local accent locations was reversed when paired with a Local remnant (d = −0.55, $t = -3.88$), but accent location did not differentially affect Non Local RC remnants, (d = 0.30, $t = -1.40$), suggesting that the effect of Locality is limited to syntactically ambiguous forms.

The overall pattern is perhaps best understood as a penalty for a mismatch between pitch accent and the likely remnant-correlate pairing. Indeed, in a model that transformed the conditions to reflect the relative match or mismatch between the pitch accent and the remnant, there were only two main effects: one showing a naturalness penalty for remnants whose correlates are not marked with matching, i.e., contrastive, accent, and another showing a naturalness penalty for syntactically ambiguous remnants with Non Local correlates. In this model, the Match factor reflected cases in which the location of the pitch accent matched semantically with the remnant (Match: 25a.ii, 25a.iii, 25b.i) vs. Mismatch: 25a.i, 25b.ii, 25b.iii). Remnant type was specified with the levels Local NP (25.i), Non Local NP (25.ii), and Non Local RC (25.iii). The statistical baseline for model was the Match-Local condition, i.e., *John didn't write an article that exposed the GOVERNOR, let alone the PRESIDENT* in (25.b.i). Responses were given the same random effects and contrast coding as before, resulting in a decrease in naturalness ratings for Mismatching ($\hat{\beta} = -0.20$, $SE = 0.04$, $t = -5.70$), and Non Local ($\hat{\beta} = -0.13$, $SE = 0.05$, $t = -2.50$) conditions. No other fixed effects, besides the Intercept ($\hat{\beta} = 5.52$, $SE = 0.14$, $t = 39.00$) were significant, including the Non Local RC condition, which confirms the interpretation that Non Local remnants are not costly when syntactically unambiguous.

In support of this interpretation, *post-hoc* pairwise t-tests with Bonferroni correction revealed a penalty for the condition with Local accent and Non Local correlate (22b.ii) compared to conditions with matching accents (52b.i and 25a.ii, 25a.iii) in by-subject

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

analyses, *p's* < 0.05. The penalty for mismatching accent also appeared as fully significant in by-items analysis for conditions with Non Local correlates (25a.ii, 52a.iii), *p*'s < 0.05, and as a non-significant trend for the condition with a Local correlate (25b.i), $p = 0.097$. In concert with the auditory corpus study and the previous experiment, listeners appear to expect the correlate and remnant to match in their accent status, and find deviations from that expectation less natural.

4.5 Discussion

As in Experiment 1, we find effects both of overt accent placement within the matrix clause and of the locality of the correlate. Focus-sensitive coordination constructions, like other contrastive ellipsis constructions, usually appear with contrastive accents on the intended correlate and remnant. These cues are clearly highly beneficial to listeners in finding the appropriate correlate for the remnant. Nevertheless, overt prosodic cues apparently do not void the general tendency for local correlates. As the potential correlates were both in the predicate in this study, and neither had a subject role, the Locality bias cannot be reduced to an aversion for subject correlates or to a preference for objects. Instead, the results support the idea that the default focus position, i.e. the most deeply embedded constituent, is entertained as a possible focus position even when it is not explicitly marked as such. Furthermore, the penalty for Non Local correlates was fully mitigated when the remnant was syntactically disambiguated towards the non-local correlate, as in the Non Local RC condition.

As observed by a reviewer, the relative clause constructions from Experiment 2 may be relevant in addressing whether the Locality Bias should be stated in term of (a) linear distance / order or (b) structural proximity. If we assume that *let alone* conjoins vPs in cases like (24) (Hulsey 2008; Harris, 2016), then the remnant is *linearly* more local to the embedded noun (*the governor*), but *structurally* more local to the RC head (*an article*). While our findings suggest that the linearly closest correlate (*the governor*) is preferred in cases of syntactic ambiguity, we would need to test subject-extracted and object-extracted relative clauses, e.g., *"John didn't write an article$_1$ that the governor read t_1…"* to make a more direct comparison. If linear order is the main factor in Locality, then the preferences we observed in Experiment 2 should reverse for object-extract relative clauses, so that remnants related to the RC head (*an article*) would be preferred, assuming that the head is interpreted in its thematic object position. We leave this question to future research.

Finally, the prominence-marking norming studies of Experiments 1 and 2 were conducted to address the concern that the Local noun (object noun or embedded noun) received pitch accent in the Non Local accent conditions. The results clearly showed that subjects did not perceive pitch accent exclusively on the Local noun when the Non Local noun received pitch accent. In 20–30% of cases, both the Local and Non Local noun were reported as prominent. Crucially, the location of pitch accent in the audio file did not significantly affect the dual-prominence response rates. Further, these responses were provided by a minority of subjects, who consistently selected this option. These subjects may have perceived the materials in a unique way, selected the response to avoid having to choose between the options, or simply not have understood the instructions.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

It is important to note that the norming studies do not constitute a test of enduring focus or Locality; the task simply asked where the subject perceived prominence, an acoustic property, but did not ask them to assign focus, an interpretive property, to constituents in the sentence. We did not necessarily expect a bias towards perceiving prominence on the local noun, a finding which lends some weak evidence against a perceptual illusion interpretation of enduring focus, discussed below.

5. General Discussion

In a small auditory corpus study, we find that the likely correlate to the remnant of *let alone* ellipsis is almost always marked with a pitch accent, often a contrastive one. General patterns observed in the corpus, including the Locality Bias, were compatible with those reported in Harris & Carlson (2016). In two auditory rating studies, we find focus-sensitive coordination structures to be sensitive to focus in two different ways: first, comprehenders respond to the overt marking of correlates and remnants with contrastive pitch accents, and, second, they show sensitivity to the default position of focus in a persistent Locality bias. Furthermore, Locality persists when the potential noun phrase correlates are all situated within a relative clause in the main predicate, a finding that we explain by appealing to default positions for focus.

However, Locality is overturned by syntactically unambiguous remnant structures. We could imagine two distinct models of the retrieval of a correlate in focus-sensitive coordination structures: the processor either does not entertain multiple possible correlates for unambiguous remnant structures, or utilizes syntactic parallelism to generate hypotheses about likely correlate-remnant pairs. The former possibility could be formulated in terms of a filter of permissible correlates, in which syntactically incompatible correlates are simply ruled out by virtue of their form. The latter would be expected under a constraint-based approach in which all constituents would be considered as potential candidates, no matter how briefly, before the processor assesses cues to dispose of poor matches. At present, our results cannot arbitrate between these models, which might benefit from studies using methods that collect response data in real time, so that the activation (and decay) of correlates from memory could be mapped throughout the interpretation process.

An alternative explanation of the Locality bias that may deserve further exploration is that the preference for local correlates is due to a bias to retrieve the structurally or linearly closest possibility from memory. This option might be understood in terms of an independently plausible strategy for minimizing demands on working memory, in which accessing the most recent antecedent would ease the retrieval of a dependency, e.g., in Dependency Locality Theory (Gibson, 1998) or in cue-based parsing models (Lewis & Vashishth, 2005). For instance, in current cue-based models of retrieval, the features of a target and its competitors are activated in memory in parallel, and compared against the retrieval cues provided by the probe by an associative cue-matching mechanism. Although this family of models typically aims to explain interference effects in long distance grammatical dependencies, such as reflexives and long distance subject gaps (see Lewis & Vashishth, 2005; Lewis, Vasishth, & Van Dyke, 2006, for review), cue-based parsing might well be extended to retrieving a correlate for a remnant, along the lines of Harris (2015) for

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

sluicing ellipsis. Models of this type would predict that the increased similarity between a non-correlate distractor and the target would generate interference during retrieval, potentially resulting in the occasional misretrieval of a distractor as the correlate. However, cue-based models are only beginning to consider the effect of discourse status and focus on retrieval (Jäger, Engelmann, & Vashishth, 2017), and currently a wealth of options remains open. For example, if pitch accents impact the accessibility of a correlate by increasing its salience in memory or in the discourse representation, correlates in syntactically local (or more recent) positions would be predicted to have an independent advantage on this theory, matching the results observed here.

A reviewer proposed that the infrequency of non-local correlates in the *let alone* construction in corpora could explain the preference for local correlates, a concern which would be further compounded by the fact that focus occurs most frequently in object position in English and many other languages. While we have no doubt that frequency (of the construction or of focus placement in general) interacts with the ease of interpretation and dependency resolution, we have sought to explain how the Locality Bias might have emerged from general independent preferences on focus interpretation. We have argued that Locality can be explained, at least in part, by the remnants' sensitivity to focus position in the antecedent clause, an independently supported property of *let alone* ellipsis and other forms of clausal ellipsis. In support, we found that overt pitch accent location, a strong indicator of focus, modulated the acceptability of non-local correlate-remnant pairs. However, we also found that pitch accent alone is not enough to overturn the Locality Bias, presumably because it does not unambiguously signal the location of focus, as argued for elsewhere (Carlson et al., 2009; Clifton & Frazier, 2016).

This is not to say that the frequency of the configuration is irrelevant to the interpretation or acceptability of the structure, or even to the formation of information structure defaults for a particular language in the first place (Herring, 1990). Yet, by referencing plausible defaults that govern focus interpretation, we can account for not only the patterns observed in corpora but also performance in controlled experiments. The account therefore goes beyond a correlation between performance and exposure, by generating predictions that extend to less commonly observed forms. In general, one may use correlations between corpus data and performance to predict one from the other without committing to causal relationships (Brysbaert, Mandera, & Keuleers, 2017); instead, we have pursued a hypothesis that attributes a single underlying source to each. At the end of the day, these two accounts may simply be too closely intertwined to distinguish empirically, or else may capture different aspects of the data.

Some support for the default focus explanation of Locality over others comes from studies in Carlson et al. (2009) on sluicing ellipsis. In one study, presenting additional lexical material with an adjunct role between the correlate and the closest remnant did not affect the Locality bias. If locality were a purely structural notion based on memory, then the object preference should have waned when that object was separated from the remnant by an additional phrase. Another study showed that *it*-cleft structures (27) served to focus matrix clause arguments and increase their use as correlates to the same extent as pitch accent placement, even though the focused constituent in an it-cleft (e.g., *Lisa*) is placed quite early in the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

sentence. These effects would not be predicted under a retrieval account that correlates temporal distance with decay, though it remains to be seen what models of sentence interpretation with content-addressable retrieval systems (e.g., McElree, 2000, 2001, among many others), which have been recently extended to account for ellipsis (Harris, 2015; Martin & McElree, 2008, 2009, 2011), would predict for our cases.

(27) It was Lisa who Patty praised at the ceremony, but I don't know who else.

Previous studies propose that sluicing ellipsis is subject to a Locality bias driven by focus structure rather than recency alone. To the extent that different types of ellipsis behave similarly in processing, then, we might naturally extend this argument to focus-sensitive coordination structures. Nevertheless, we acknowledge that the proper tests to distinguish a focus-based explanation from a recency account have not been conducted, and that the discourse status afforded by focus marking may interact with memory representations by making their contents more salient for retrieval or richer in the encoding process.

The central theoretical contribution of this paper has been to explicate and defend the principle of enduring focus in auditory sentence processing, in which default locations for focus continue to influence listeners' interpretation of focus structure, despite the alternate locations for pitch accents. We now consider two potential explanations for enduring focus, both of which seem viable given current evidence.

The first possibility is that enduring focus arises through a perceptual illusion: hearers' expectations about where a pitch accent typically appears overrides the input, so that they perceive pitch accents in locations where they are not present acoustically. Many studies have shown that expectations about the signal can indeed produce perceptual illusions, e.g., the classic case of phoneme restoration, in which an excised phoneme is restored when it has been replaced by an extraneous noise, such as high amplitude noise or a cough (Samuel, 1981; Warren, 1970; Warren & Warren, 1970). Recent research has shown that factors influencing whether a phoneme (or which phoneme) is restored may include syntactic (Stoyneshka, Fodor, and Fernández, 2010) and pragmatic (Mack et al., 2012) constraints, in addition to low-level factors, such as lexicality, phonetic similarity, and word frequency (Samuel, 1996 for review). In a study that directly manipulated prosodic information, Bishop (2012) found that an object NP accented with a relatively non-prominent H* was perceived as more prominent (and the verb was perceived as less prominent) when a *wh*-question focusing the object preceded it. This directly illustrates the way that perceptions of accent can be influenced by the overall information structure. Although certainly not conclusive, the results of the two norming studies cast doubt on the 'illusion' theory. If the preference for more local correlates resulted from the misperception of prominence, we would also expect to have found that default focus positions would be regularly rated as prominent, regardless of the interpretation of focus. Instead, subjects tended to select the pitch marked constituent as prominent, showing a high sensitivity to the actual prosodic input.

The second possibility is that perceivers ignore or discard acoustic evidence in favor of a conventional mapping between syntax and information structure, potentially as a by-product of default placement of prosodic prominence. As noted earlier, pre-nuclear pitch accent placement is affected by many factors, including the phonological requirement that each

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

phonological phrase contain a pitch accent, or a pressure towards alternating stressed and unstressed words (e.g., Ladd, 2008; Selkirk, 1995). Thus, a pitch accent may not always mean that an element is in focus. A listener may not therefore be certain whether acoustic prominence marks the location of focus within an utterance. Following Truckenbrodt (1995), Büring (2012, 2016) proposed that default focus had a persistent role in the production of sentences, a proposal he called *prosodic inertia*:

(28) **Prosodic inertia:** Principles that determine default prosody continue to impact overtly focused structures (Büring, 2012, 2016; Truckenbrodt, 1995).

In some respects, our enduring focus principle might follow as a corollary of prosodic inertia: comprehenders might interpolate focus in positions that they expect to be marked as such in default prosody, regardless of the acoustic input, in favor of a more conventional information structure.

Indeed, prominence cannot always transparently be read directly off the input, as the way that pitch accent types convey information status may be opaque, or at least imprecise, for some language users (see Dahan, 2015, for a review). Recent experimental evidence suggests that speakers and listeners do not always distinguish between pitch accent types in terms of a conventional mapping to information status, particularly the pitch accent associated with new information (H*) and the pitch accent associated with contrastive information (L+H*) (Breen et al., 2010; Watson, Tanenhaus, & Gunlogson, 2008). Relatedly, there is disagreement as to whether those pitch accents are actually distinct or merely different points along a continuum (Bartels & Kingston, 1994; Ladd, 2008; Pierrehumbert & Hirschberg, 1990). More broadly, different speakers use different pitch accents in speech, and accent sentences different amounts, and listeners have to adapt to this wide range of variation (e.g., Grabe 2002, 2004). As such, there may be good reason to default to conventional mappings between sentence positions and discourse status.

In sum, focus assignment may not simply be an all or nothing affair. The processor appears to entertain other possible focus mappings than those that are signaled overtly. We speculate that such a loose mapping between form and discourse representation could be considered an advantage, rather than a failure, in that (i) sentence elements receive prominence for many reasons, not all of which indicate focus status, and (ii) relying on information structure defaults safeguards against potential misalignments generated by error in the signal, making for a more robust discourse processing system (see Gibson, Bergen, & Piantadosi, 2013, on the noisy channel theory). This is highly sensible for languages like English which use a multitude of phonetic correlates (primarily pitch, but also loudness, intensity, length, spectral tilt, etc.) to signal prominence (e.g., Gussenhoven, 2004, for a review, and Breen et al., 2010, for experimental evidence).

In general, an appealing null hypothesis regarding the relation between explicit and implicit indicators of focus is that foci are read off transparently from overt indicators of pitch accent (Bock & Mazella, 1983). A comprehension system that is subject to enduring focus provides an intriguing counterpoint to this null hypothesis. The results of both experiments strongly support the claim that auditory processing system is subject to enduring focus. Listeners evidently do not use the overt position of accents as their sole guide to determining the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

location of focus – not even for our constructions, which apparently mandate contrastive accents. Thus, the final argument of a clause, in the default position of focus, is always considered as potentially focused and thus remains a potential correlate despite the presence of contrastive accents elsewhere.

Many additional questions remain as to why defaults persist, and precisely how their influence should be accounted for in models of real-time language processing. Although we may not yet be in a position to completely address these admittedly broad questions, we have argued that such a comprehension system with some degree of enduring focus may actually be robust: by prioritizing more regular locations for focus, the comprehension system can find stable interpretations despite potential errors in the signal. A strategy that favors conventionalized defaults finds intuitive support in the fact that speakers may not always be entirely sensitive to the needs of their audience, i.e., speakers may sometimes engage in only *limited* audience design (cf., Clark & Murphy, 1982), task demands and contextual support permitted.

Although there is good evidence that speakers are closely attuned to the speech of their interlocutors, by monitoring phonetic variation (e.g., Kraljic, Samuel, & Brennan, 2008; Magnuson & Nusbaum, 2007), speaker identity (Nygaard & Pisoni, 1998), and even speaker class (Van Berkum et al., 2008), speakers do not always adjust their utterances to ease interpretation (e.g., Brown & Dell, 1987; Fukumura & van Gompel, 2012). Rational hearers may likewise be wary of committing to messages that deviate from the norm, unless such a message is contextually supported.

The persistence of such a strategy could be understood as in Harris & Potts' (2008) broadly game-theoretic sense: speakers tend not to deviate from established defaults, except when the intended message is sure to be recovered. And if speakers were to deviate from the norm in unpredictable contexts, they would be compelled to sufficiently heighten or modulate the production signal when such deviations occur, in order to make their intended message recoverable. Consequently, deviant formulations of the message might increase the possibility of misconstrual, and such productions may well be prohibitively risky. This is the essence of the conventional logic behind information structure: deviations from canonical or default forms are highly marked in both speech and interpretation. Our findings suggest that such defaults are indeed powerful when locating a correlate for a remnant in ellipsis structures, in that pitch accent placement does not uniquely determine focus marking that is necessary for interpreting focus-sensitive ellipsis.

Acknowledgments

Thanks to Joseph Tyler, Benjamin Lee, Dallas Cox, and Blake Clark for assistance with Experiments 1–2; Katherine Griffitts, Joe Castle, and Torianne Crouch for assistance with the norming experiments; and several anonymous reviewers for helpful comments. We also extend thanks for partial financial support to the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under grant number R15HD072713, and an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number 5P20GM103436-13. The research described here is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or any other institution.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

References

Bartels C, Kingston J. Salient pitch cues in the perception of contrastive focus. In: Bosch P, van der Sandt R, editorsFocus and natural language processing. Vol. 3. Heidelberg: IBM Deutschland; 1994. 1–10.

Bates DM. lmer, p-values and all that. 2006. Internet post retrieved from URL https://stat.ethz.ch/pipermail/r-help/2006-May/094765.html

Beaver DI, Clark BZ. Sense and sensitivity: How focus determines meaning. Wiley-Blackwell; 2009.

Beckman M, Elam G. Guidelines for ToBI transcription, vers. 2.0. Ms. and accompanying speech materials. Ohio State University; 1997.

Bishop J. Information structural expectations in the perception of prosodic prominence. In: Elordieta G, Prieto P, editorsProsody and meaning. Berlin, Germany: Walter de Gruyter; 2012. 239–270.

Bock JK, Mazzella JR. Intonational marking of given and new information: Some consequences for comprehension. Memory & Cognition. 1983; 11:64–76. [PubMed: 6855561]

Breen M, Fedorenko E, Wagner M, Gibson E. Acoustic correlates of information structure. Language and Cognitive Processes. 2010; 25:1044–1098.

Brown P, Dell G. Adapting production to comprehension: The explicit mention of instruments. Cognitive Psychology. 1987; 19:441–472.

Brysbaert M, Mandera P, Keuleers E. Corpus Linguistics. In: De Groot AMB, Hagoort P, editorsResearch Methods in Psycholinguistics. Vol. Ch. 10. London: Wiley; 2017.

Büring D. Focus and Intonation. In: Russell G, Fara DG, editorsRoutledge Companion to the Philosophy of Language. London: Routledge; 2012. 103–115.

Büring D. Intonation and Meaning. Oxford: Oxford University Press; 2016.

Cappelle B, Dugas E, Tobin V. An afterthought on *let alone*. Journal of Pragmatics. 2015; 80:70–85.

Carlson K. The effects of parallelism and prosody in the processing of gapping structures. Language and Speech. 2001; 44:1–26. [PubMed: 11430186]

Carlson K. Parallelism and Prosody in the Processing of Ellipsis Sentences. Outstanding Dissertations in Linguistics. Routledge; 2002.

Carlson K. The role of *only* in contrasts in and out of context. Discourse Processes. 2013; 50:1–27.

Carlson K, Dickey MW, Frazier L, Clifton C Jr. Information structure expectations in sentence comprehension. The Quarterly Journal of Experimental Psychology. 2009; 62:114–139. [PubMed: 18609404]

Chafe WL. Meaning and the structure of language. Chicago, IL: University of Chicago Press; 1970.

Chafe WL. Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View in Subject and Topic. In: Li CN, editorSubject and topic. New York: Academic Press; 1976. 25–55.

Chomsky N. The Minimalist Program. Cambridge, MA: MIT Press; 1995.

Chomsky N, Halle M. The sound pattern of English. Cambridge, MA: MIT Press; 1968.

Chung S, Ladusaw WA, McCloskey J. Sluicing and logical form. Natural Language Semantics. 1995; 3:239–282.

Cinque G. A null theory of phrase and compound stress. Linguistic Inquiry. 1993; 24:239–297.

Clark HH, Haviland SE. Comprehension and the given-new contract. Discourse production and comprehension. Discourse processes: Advances in research and theory. 1977; 1:1–40.

Clark HH, Murphy GL. Audience design in meaning and reference. Advances in Psychology. 1982; 9:287–299.

Clifton C Jr, Frazier L. Focus in corrective exchanges: Effects of pitch accent and syntactic form. Language and Speech. 2016; 59:544–561. [PubMed: 28008801]

Dahan D. Prosody and language comprehension. Wiley Interdisciplinary Reviews: Cognitive Science. 2015; 6:441–452. [PubMed: 26267554]

Davies M. The Corpus of Contemporary American English: 520 million words, 1990-present. 2008. URL http://corpus.byu.edu/coca/

Dickey MW, Bunger AC. Comprehension of elided structure: Evidence from sluicing. Language and Cognitive Processes. 2011; 26:63–78.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Fillmore CJ, Kay P, O'Connor MC. Regularity and idiomaticity in grammatical constructions: The case of let alone. Language. 1988:501–538.

Firbas J. Non-thematic subjects in contemporary English. Travaux linguistiques de Prague. 1966; 2:239–256.

Frazier L. Processing ellipsis: A processing solution to the undergeneration problem. In: Chang CB, Haynie HJ, editorsProceedings of the 26th West Coast Conference on Formal Linguistics. Somerville, MA: Cascadilla Press; 2008. 21–32.

Frazier L, Clifton C Jr. Comprehension of sluiced sentences. Language and Cognitive Processes. 1998; 13:499–520.

Frazier M, Potter D, Yoshida M. Pseudo Noun Phrase Coordination. In: Arnett N, Bennett R, editorsProceedings of the 30th West Coast Conference on Formal Linguistics. Somerville, MA: Cascadilla Proceedings Project; 2012. 42–152.

Frazier L, Taft L, Roeper T, Clifton C, Ehrlich K. Parallel structure: A source of facilitation in sentence comprehension. Memory & Cognition. 1984; 12:421–430. [PubMed: 6521643]

Fukumura K, van Gompel RP. Producing pronouns and definite noun phrases: do speakers use the addressee's discourse model? Cognitive Science. 2012; 36:1289–1311. [PubMed: 22671693]

Gennari SP, MacDonald MC. Semantic indeterminacy in object relative clauses. Journal of Memory and Language. 2008; 58:161–187. [PubMed: 19724662]

Gibson E. Linguistic complexity: Locality of syntactic dependencies. Cognition. 1998; 68:1–76. [PubMed: 9775516]

Gibson E, Bergen L, Piantadosi S. The rational integration of noisy evidence and prior semantic expectations in sentence interpretation. Proceedings of the National Academy of Science. 2013; 110(20):8051–8056. [10.1073/pnas.1216438110]

Gordon PC, Hendrick R, Johnson M. Memory interference during language processing. Journal of Experimental Psychology: Learning, Memory and Cognition. 2001; 27:1411–1423.

Grabe E. Variation adds to prosodic typology. In: Bel B, Marlin I, editorsProceedings of the Speech Prosody 2002 Conference; 11–13 April 2002; Aix-en-Provence. Laboratoire Parole et Langage; 2002. 127–132.

Grabe E. Intonational variation in urban dialects of English spoken in the British Isles. In: Gilles P, Peters J, editorsRegional Variation in Intonation. Linguistische Arbeiten; Tuebingen, Niemeyer: 2004. 9–31.

Grodner DJ, Gibson EAF. Consequences of the serial nature of linguistic input for sentential complexity. Cognitive Science. 2005; 29:261–291. [PubMed: 21702774]

Gussenhoven C. The phonology of tone and intonation. Cambridge, UK: Cambridge University Press; 2004.

Halle M, Vergnaud JR. Stress and the cycle. Linguistic Inquiry. 1987; 18:45–84.

Harris Jesse A. Structure modulates similarity-based interference in sluicing: An eye tracking study. Frontiers in Psychology. 2015; 6:1839. [PubMed: 26733893]

Harris JA. Processing let alone coordination in silent reading. Lingua. 2016; 169:70–94.

Harris JA, Carlson K. Keep it local (and final): Remnant preferences in "let alone" ellipsis. The Quarterly Journal of Experimental Psychology. 2016; 69:1278–130. [PubMed: 26085004]

Harris JA, Potts C. Perspective-shifting with appositives and expressives. Linguistics & Philosophy. 2008; 32:523–55.

Herring SC. Proceedings of the 16th Annual Berkeley Linguistics Society. Berkeley, CA: Berkeley Linguistics Society; 1990. Information structure as a consequence of word order type; 163–174.

Horn LR. Papers from the Fifth Regional Meeting, Chicago Linguistic Society. Chicago, Ill: Chicago Linguistic Society; 1969. A presuppositional theory of 'only'and 'even'; 98–107.

Horn LR. Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In: Schiffrin D, editorMeaning, Form, and Use in Context. Washington: Georgetown University Press; 1984. 11–42.

Hulsey S. Unpublished PhD thesis. Cambridge, MA: MIT; 2008. Focus sensitive coordination.

Jackendoff RS. Semantic interpretation in generative grammar. Cambridge, MA: MIT Press; 1972.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Jäger LA, Engelmann F, Vasishth S. Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. Journal of Memory and Language. 2017; 94:316–339.

King J, Just MA. Individual differences in syntactic processing: The role of working memory. Journal of Memory and Language. 1991; 30:580–602.

Koopman H, Sportiche D. Subjects. Ms., UCLA; Los Angeles, California: 1988.

Kraljic T, Samuel AG, Brennan SE. First impressions and last resorts: How listeners adjust to speaker variability. Psychological Science. 2008; 19:332–338. [PubMed: 18399885]

Kratzer A. Severing the external argument from its verb. In: Rooryck J, Zaring L, editorsPhrase Structure and the Lexicon. The Netherlands: Springer; 1996. 109–137.

Krifka M. Basic notions of information structure. Acta Linguistica Hungarica. 2008; 55:243–276.

Kuroda S-Y. Whether we agree or not: A comparative syntax of English and Japanese. Lingvisticae Investigationes. 1988; 12:1–47.

Ladd DR. Intonational Phonology. 2. Cambridge, UK: Cambridge University Press; 2008.

Ladd DR, Morton R. The perception of intonational emphasis: continuous or categorical? Journal of Phonetics. 1997; 25:313–342.

Ladd DR, Schepman A. "Sagging transitions" between high pitch accents in English: experimental evidence. Journal of Phonetics. 2003; 31:81–112.

Lambrecht K. Information structure and sentence form: A theory of topic, focus, and the mental representations of discourse referents. Cambridge, UK: Cambridge University Press; 1994.

Lewis RL, Vasishth S. An activation-based model of sentence processing as skilled memory retrieval. Cognitive Science. 2005; 29:375–419. [PubMed: 21702779]

Lewis RL, Vasishth S, Van Dyke JA. Computational principles of working memory in sentence comprehension. Trends in Cognitive Science. 2006; 10:447–454.

Mack JE, Clifton C, Frazier L, Taylor PV. (Not) Hearing optional subjects: The effects of pragmatic usage preferences. Journal of Memory and Language. 2012; 67:211–223. [PubMed: 22754087]

Magnuson JS, Nusbaum HC. Acoustic differences, listener expectations, and the perceptual accommodation of talker variability. Journal of Experimental Psychology: Human Perception and Performance. 2007; 33:391–409. [PubMed: 17469975]

Marantz A. On the Nature of Grammatical relations. Cambridge, MA: MIT Press; 1984.

Martin AE, McElree B. A content-addressable pointer mechanism underlies comprehension of verb-phrase ellipsis. Journal of Memory and Language. 2008; 58:879–906.

Martin AE, McElree B. Memory operations that support language comprehension: evidence from verb-phrase ellipsis. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2009; 35:1231–1239.

Martin AE, McElree B. Direct-access retrieval during sentence comprehension: evidence from sluicing. Journal of Memory and Language. 2011; 64:327–343. [PubMed: 21580797]

McElree B. Sentence comprehension is mediated by content-addressable memory structures. Journal of Psycholinguistic Research. 2000; 29:111–123. [PubMed: 10709178]

McElree B. Working memory and focal attention. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2001; 27:817–835.

Merchant J. The syntax of silence: Sluicing, islands, and the theory of ellipsis. Oxford: Oxford University Press; 2001.

Merchant J. Fragments and ellipsis. Linguistics and Philosophy. 2005; 27:661–738.

Nygaard LC, Pisoni DB. Talker-specific learning in speech perception. Attention, Perception, & Psychophysics. 1998; 60:355–376.

Phillips C, Parker D. The psycholinguistics of ellipsis. Lingua. 2014; 151:78–95.

Pierrehumbert J, Hirschberg J. The meaning of intonational contours in the interpretation of discourse. In: Cohen P, Morgan J, Pollack M, editorsIntentions in Communication. Cambridge MA: MIT Press; 1990. 271–311.

Poirier J, Wolfinger K, Spellman L, Shapiro LP. The real-time processing of sluiced sentences. Journal of Psycholinguistic Research. 2010; 39:411–427. [PubMed: 20229060]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Prince EF. The ZPG letter: Subjects, definiteness, and information status. In: Thompson S, Mann W, editorsDiscourse Description: Diverse Analyses of a Fund-Raising Text. The Netherlands: John Benjamins; 1992. 295–325.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2015. R version 3.2.3. URL http://www.R-project.org/

Sailor C, Thoms G. On the non-existence of non-constituent coordination and non-constituent ellipsis. In: Santana-LaBarge RE, editorThe Proceedings of the 31st West Coast Conference on Formal Linguistics. Somerville, MA: Cascadilla Proceedings Project; 2014. 361–370.

Samuel AG. Phonemic restoration: Insights from a new methodology. Journal of Experimental Psychology: General. 1981; 110:474–494. [PubMed: 6459403]

Samuel AG. Phoneme Restoration. Language and Cognitive Processes. 1996; 11:647–653.

Selkirk EO. Phonology and syntax. Cambridge, MA: MIT Press; 1984.

Selkirk EO. Sentence prosody: Intonation, stress and phrasing. In: Goldsmith J, editorHandbook of phonological theory. Oxford, UK: Basil Blackwell; 1995. 550–569.

Siegel MEA. Gapping and interpretation. Linguistic Inquiry. 1984; 15:523–530.

Staub A. Eye movements and processing difficulty in object relative clauses. Cognition. 2010; 116:71–86. [PubMed: 20427040]

Stolterfoht B, Friederici AD, Alter K, Steube A. Processing focus structure and implicit prosody during reading: Differential ERP effects. Cognition. 2007; 104:565–590. [PubMed: 16989798]

Stoyneshka I, Fodor JD, Fernández EM. Phoneme restoration methods for investigating prosodic influences on syntactic processing. Language and Cognitive Processes. 2010; 25:1265–1293.

Toosarvandani M. Unpublished PhD thesis. Berkeley, CA: University of California, Berkeley; 2010. Association with foci.

Truckenbrodt H. PhD Thesis. Massachusetts Institute of Technology; Cambridge, MA: 1995. Phonological phrases – their relation to syntax, focus, and prominence.

Van Berkum J, Van Den Brink D, Tesink CM, Kos M, Hagoort P. The neural integration of speaker and message. Journal of Cognitive Neuroscience. 2008; 20:580–591. [PubMed: 18052777]

Warren RM. Perceptual restoration of missing speech sounds. Science. 1970; 167:392–393. [PubMed: 5409744]

Warren RM, Warren RP. Auditory illusions and confusions. Scientific American. 1970; 223:30–36. [PubMed: 5480550]

Watson DG, Tanenhaus MK, Gunlogson CA. Interpreting pitch accents in online comprehension: H* vs. L+ H. Cognitive Science. 2008; 32:1232–1244. [PubMed: 21585451]

Ward G, Birner B. Information structure and non-canonical syntax. In: Horn L, Ward G, editorsThe handbook of pragmatics. Oxford, UK: Oxford University Press; 2004. 153–174.

Weir A. PhD Thesis. University of Massachusetts; Amherst: 2014. Fragments and clausal ellipsis.

Appendix A

Items from Experiment 1. Only item 1 shows the complete paradigm, which can be reconstructed from the remnants given (inanimate / animate) and alternating pitch accent. Pitch accented elements are printed in CAPS in item 1. In items 11–20, the animate remnants are plausible as the object of the verb.

1. PATRICK didn’t read the article, let alone JASON.

PATRICK didn’t read the article, let alone the BOOK.

Patrick didn’t read the ARTICLE, let alone JASON.

Patrick didn’t read the ARTICLE, let alone the BOOK.

2. Danielle didn't pass the quiz, let alone the final / Kayla.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

3. The patient didn't eat dinner, let alone dessert / her family.
4. The team can't run a mile, let alone a marathon / the coach.
5. Alexis wouldn’t buy that shirt, let alone that outfit / Jessica.
6. David didn’t wear a sweater, let alone a coat / Marcus.
7. The adults wouldn’t try the appetizer, let alone the main dish / the kids.
8. Mike couldn’t build a raft, let alone a boat / Chuck.
9. The trainer couldn't tame the poodle, let alone the Doberman / the owner.
10. Emily couldn't write a story, let alone a novel / Monica.
11. Jonah wouldn't send a postcard, let alone a letter / Daniel.
12. Jenny can't sketch a cat, let alone a horse / Sarah.
13. Ryan couldn't lift the jug, let alone the barrel / Cindy.
14. Nora wouldn't drive a car, let alone a motorcycle / Vicky.
15. The hikers didn't reach the trail, let alone the campsite / the campers.
16. Erica wouldn’t bring the map, let alone the directions / Rachel.
17. Jason couldn’t reach the door, let alone the lock / Anthony.
18. The salesman wouldn’t recommend the product, let alone the accessories / the manager.
19. Scott didn't enjoy the preview, let alone the movie / Robbie.
20. The math teacher couldn't explain the equation, let alone the theorem / the English teacher.

Appendix B

Items from Experiment 2. The full paradigm is shown only in item 1, where the locations of pitch accent are marked in CAPS. In the remaining items, pitch accent locations are underlined, and the alternative remnants are provided in the order of contrasting with the local correlate / the non local relative clause head / the non local relative clause head disambiguated syntactically.

1. Although he was a highly praised journalist, John didn’t write …
    - **a.** an article that exposed THE GOVERNOR, let alone the president.
    - **b.** an article that exposed THE GOVERNOR, let alone a book.
    - **c.** an ARTICLE that exposed the governor, let alone the president.
    - **d.** an ARTICLE that exposed the governor, let alone a book.
    - **e.** an article that exposed THE GOVERNOR, let alone a book that did.
    - **f.** an ARTICLE that exposed the governor, let alone a book that did.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

2. Startled by the accusation, Marianne replied that she didn't own a cat that had fleas, let alone rabies / a dog / a dog that did.

3. Because he had a bad back, Pete couldn't move a chair that was made of wood, let alone metal / a bench / a bench that was.

4. If it were up to her, Darlene wouldn't have picked a TV show that glorified violence, let alone torture / a documentary / a documentary that did.

5. A true purist, Nigel would never drink tea that contained milk, let alone lemon / coffee / coffee that did.

6. As she has terrible food allergies, Carrie wouldn't have ordered a side dish that contained nuts, let alone shellfish / a main course / a main course that did.

7. Since he only listens to heavy metal, Max probably wouldn't have downloaded a song that sounds like country music, let alone bluegrass / an album / an album that did.

8. Because he had such a sensitive stomach, Bill couldn't eat a salad that had onions, let alone jalapenos / a burrito / a burrito that did.

9. A staunch pacifist, Mica wouldn't touch a pellet gun that was loaded with blanks, let alone live bullets / a rifle / a rifle that was.

10. Given her extreme dislike of noisy music, Shelly wouldn't attend a reception that featured a string quartet, let alone a brass band / a concert / a concert that did.

11. As he was a lazy student, Milton hadn't ever taken a class that required writing chapter summaries, let alone critical papers / a graduate seminar / a graduate seminar that did.

12. Before treatment for her extreme shyness, Lydia would never have gone to a potluck that was hosted by friends, let alone strangers / a dinner party / a dinner party that did.

13. Because he was modest, Hubert didn't want to buy a painting that included nudity, let alone sex / a sculpture / a sculpture that did.

14. As she was a meticulous proofreader, Laura wouldn't turn in a paper that had typos, let alone factual errors / a dissertation / a dissertation that did.

15. With his technical knowledge, Mike wouldn't keep using a laptop that had viruses, let alone malware / a desktop / a desktop that did.

16. After the seminar on cultural sensitivity, Marla would never circulate an e-mail that contained an off-color joke, let alone an insult / a memo / a memo that did.

17. Since the CEO was a frugal traveller, he didn't request a room that included a fridge, let alone a microwave / a suite / a suite that did.

18. As he was concerned about food safety, Ian would never serve a meal that contained a roach, let alone a fly / a starter / a starter that did.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

19. In a market with inflated real estate prices, Sarah refused to visit condos that came with hardwood floors, let alone a Jacuzzi / houses / houses that did.

20. As a single dad, Matthew had trouble finding a babysitter who could handle the kids, let alone the dogs / a girlfriend / a girlfriend who could.

21. After she moved to the desert, Jenny couldn't find a fruit tree that could survive extreme heat, let alone dusty soil / a bush / a bush that did.

22. At the cast party, Bill didn't notice the actor who was wearing a mask, let alone a gorilla suit / the big star / the big star who was.

23. Despite searching for days, Maureen couldn't find a briefcase that had a comfortable handle, let alone enough space / a computer case / a computer case that did.

24. With her weight issues, Sally could never find a shirt that flattered her arms, let alone her tummy / a dress / a dress that did.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

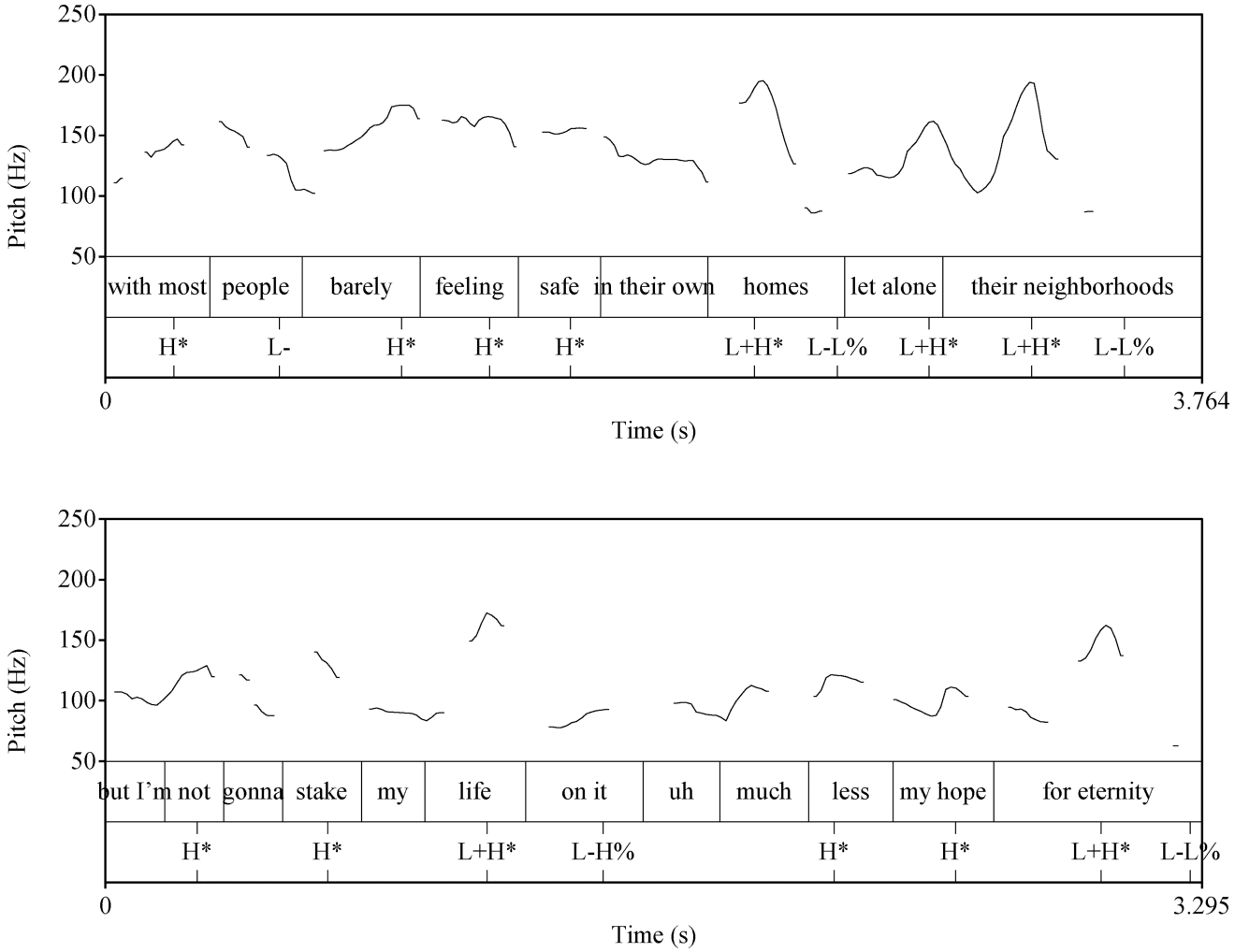

Figure 1.
Pitch track of examples with a local correlate (a) and a non-local correlate (b).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

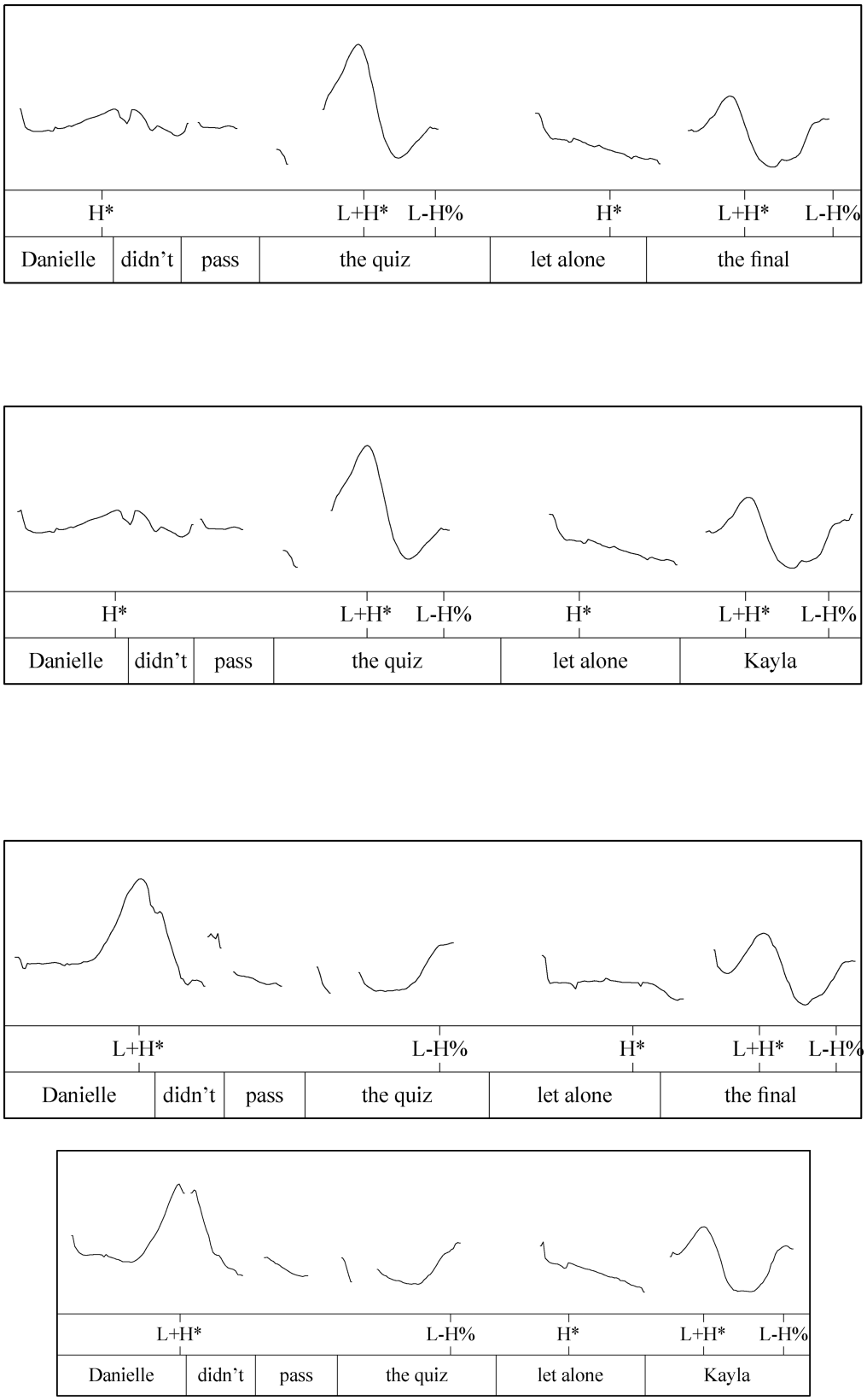

Figure 2.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Pitch excursions of (21), labeled with the Tones and Breaks Indices annotation system of Beckman & Elam (1997). The first two panels show contrastive pitch accent on the object, with either a matching object (a) or mismatching subject (b) remnant. The final two panels show contrastive pitch accent on the subject, with either a mismatching subject (c) or matching object (d) remnant.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

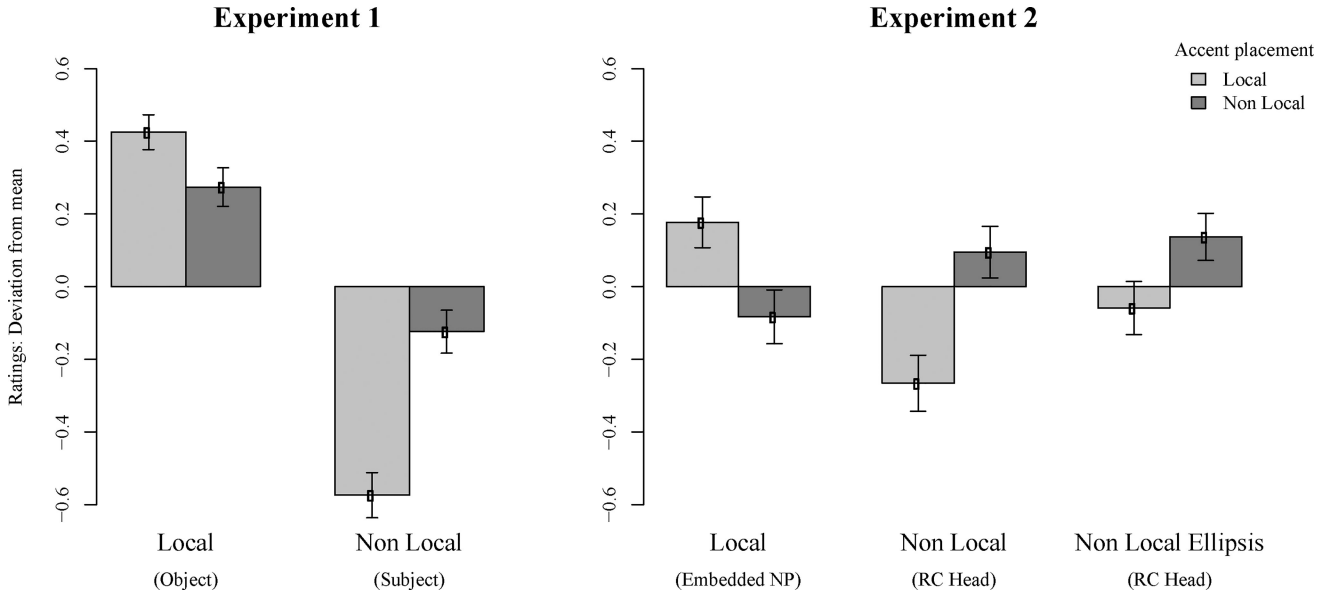

Figure 3.

Ratings for each condition expressed as deviations from the grand mean for Experiments 1 and 2.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Average fundamental frequency measurements in Hz for critical words and boundaries for Experiments 1–2 (standard deviations under 20Hz for peaks, 10 Hz for boundaries). The third column contains the average values for L− and L% or H% boundary tones that appeared at the end of the host clause before the coordinator.

condition	Subject or NP1 peak	Object or NP2 peak	Boundary tone values (L−, L/H%)	Remnant peak
Experiment 1				
a	249	318	150, 212	249
b	241	306	153, 209	243
c	347	164	154, 212	260
d	337	163	157, 211	245
Experiment 2				
a .i	284	148	141, 149	216
a .ii	284	148	141, 153	218
a .iii	282	147	141, 152	224
b .i	206	273	140, 157	214
b .ii	192	267	140, 155	214
b .iii	190	266	139, 158	222

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Average duration measurements in milliseconds (ms) for critical words for Experiments 1–2 (standard deviations under 200 ms for words, under 60 ms for pauses).

condition	Subject or NP1	Object or NP2	Pause	Remnant NP
Experiment 1				
a	395	703	141	636
b	381	722	137	714
c	458	593	116	630
d	452	586	123	715
Experiment 2				
a .i	528	603	138	
a .ii	527	602	152	
a .iii	515	599	159	
b .i	462	694	126	
b .ii	447	692	128	
b .iii	453	698	140	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Experiment 1. Raw means/z-score transformations. Standard errors are in parentheses.

	Local correlate	**Non-local correlate**	**Mean**	**Non-local penalty**
Local accent	5.84/0.43 (0.08)	4.15/−0.12 (0.10)	4.99 (0.08)	1.69
Non-local accent	5.58/0.27 (0.09)	4.91/−0.57 (0.10)	5.25 (0.07)	0.67
Mean	5.71 (0.06)	4.53 (0.07)		

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Experiment 1: Fixed effects of the linear mixed effects regression model.

Parameter	Estimate	Standard Error	t-value
(Intercept)	5.119	0.131	39.04*
Remnant type	−0.589	0.088	−6.73*
Accent location	0.126	0.051	2.48*
Remnant type × accent location	0.252	0.062	4.10*

Parameters with t values above |2| were considered significant, and are marked with an asterisk.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 5

Experiment 1: Raw means/z-scores transformations including non-orthogonal contrast of subject type. Standard errors are in parentheses.

Remnant type				
Subject	**Accent placement**	**Local**	**Non-local**	**Non-local penalty**
Proper name	Local accent	5.88/0.45 (0.10)	4.20/−0.54 (0.12)	1.68
	Non-local accent	5.58/0.28 (0.11)	4.77/−0.21 (0.12)	0.81
	Remnant type mean	5.73/0.36	4.49/−0/37	
Definite	Local accent	5.75/0.37 (0.13)	4.05/−0.63 (0.19)	1.70
	Non-local accent	5.58/0.27 (0.16)	5.17/−0.21 (0.17)	0.41
	Remnant type mean	5.67/0.32	4.61/−0.30	

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 6

Experiment 1: Fixed effects of the linear mixed effects regression model with the nonorthogonal contrast of subject type as an additional predictor.

	Estimate	Standard Error	t-value
(Intercept)	5.128	0.133	38.54*
Remnant type	−0.571	0.089	−6.41*
Accent placement	0.138	0.054	2.57*
Subject type	0.027	0.057	0.47
Remnant type × accent placement	0.285	0.059	4.81*
Remnant type × subject type	0.059	0.050	1.19
Accent placement × subject type	0.040	0.050	0.80
Remnant type × accent placement × subject type	0.111	0.047	2.30*

Parameters with t-values above |2| were considered significant, and are marked with an asterisk.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 7

Experiment 2. Raw means/z-score transformations for each condition. Standard errors are presented in parentheses.

Remnant type				
Accent location	**Local NP**	**Non-local NP**	**Non-local relative clause**	**Mean**
Non-local	5.40/−0.08 (0.10)	5.67/0.10 (0.10)	5.77/0.14 (0.10)	5.60 (0.06)
Local	5.79/0.18 (0.10)	5.12/−0.27 (0.10)	5.43/−0.06 (0.10)	5.45 (0.06)
Mean	5.59 (0.08)	5.39 (0.08)	5.58 (0.07)	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 8

Experiment 2: Fixed effects of the linear mixed effects regression model.

	Estimate	Standard Error	t-value
(intercept)	5.520	0.142	39.00*
Local accent	−0.077	0.036	−2.17*
Non-local NP	−0.127	0.051	−2.50*
Non-local relative clause (RC)	0.061	0.051	1.21
Local accent × non-local NP	−0.196	0.051	−3.88*
Local accent × non-local RC	−0.071	0.051	−1.40

Parameters with t values above |2| were considered significant, and marked with an asterisk.