Intonation and Focus Layers

Jonni M. Kanerva
Indiana University

Leslie A. Gabriele
Indiana University

Recommended Citation

Kanerva, Jonni M. and Gabriele, Leslie A. (1995) "Intonation and Focus Layers," *North East Linguistics Society*: Vol. 25 , Article 23.
Available at: https://scholarworks.umass.edu/nels/vol25/iss1/23

Intonation and Focus Layers

Jonni M. Kanerva and Leslie A. Gabriele

Indiana University

1.   Introduction

Once one looks beyond artificially isolated sentences, the study of focus becomes unavoidable. Focus, as a marker of the information structure of a sentence, is a key conceptual tool for modelling how an individual sentence contributes to the flow and growth of information in discourse. Strong and varied effects on other components of the grammar provide empirical motivation for focus, including phonological phrasing, sentential accent placement, word order, and even the selection of arguments for semantic operators. Focus is thus rapidly gaining recognition as a pivotal factor in grammar (e.g., Jackendoff 1972, Gussenhoven 1984; Rooth 1985, 1992; Kanerva 1990, 1991; Krifka 1991; Partee 1991; Steedman 1991; von Stechow 1991).

1.1.   One-layer Model

Until recently, most work on the phonology and semantics of focus assumed a one-layer model of the information structure of a sentence. Example (1) (from Rooth 1985:2) illustrates the typical representation of focus in sentence grammar—as a diacritic adornment of syntactic structure:

(1)   I only introduced [$_F$ BILL] to Sue.

The phonological and semantic components then interpret this structure independently of

335

336   JONNI M. KANERVA AND LESLIE A. GABRIELE

one another. In the phonology, the most prominent stress of the sentence ("nuclear stress", typically marked by capitalization) is realized on or within the focused constituent (e.g., Jackendoff 1972:237). In the semantics, the operator *only* "associates" with the focused element.

Although this simple example successfully correlates focus placement, phonology, and semantics, it is easy to construct more complicated examples in which the three-way connection appears to break down (e.g., Partee 1991:179; Rooth 1992:109).

1.2.   Two-layer Models

Richer data has given rise to elaborations on the overly simple one-layer structure. In one approach, Krifka (1991) adopts a superordinate topic-comment division. A sentence is first divided into a topic and comment; each of these in turn can contain a single-layer focus-ground division. Example (2) illustrates both of these layers of information structure:

(2)   [$_T$ Bill's [$_F$ YOUNGEST] sister] [$_C$ [$_F$ kissed JOHN.]]

In another approach, Vallduvi's (1992) top layer is a focus-ground division, and the ground can be further divided into a link (similar to topic) and tail:

(3)   a. {focus, {$_{ground}$ link, tail}}

         b. [$_G$ [$_L$ Bill's YOUNGEST sister]] [$_F$ kissed JOHN.]

Both of these approaches share a crucial assumption: the information structure of a (simplex) sentence is not recursive. Indeed, Vallduvi and Zacharski (1994:684) claim explicitly that "speaking of recursive focus-ground structures within a simplex sentence has little conceptual motivation from the informational or communicative perspective."[1]

We will argue just the opposite—that information structure is recursive—and we provide evidence in the form of layered focus-ground divisions. We show moreover that this recursion is motivated both conceptually, by discourse structure, and empirically, by relationships between multiple pitch accents.

2.   Information in Discourse and in Sentences

The structure of discourse is organized primarily by questions, answers, and hierarchical relationships among these (see esp. Carlson 1983, 1984). Questions represent simultaneously a current state of information and a goal for specific further

---

[1] Recursive focus-ground structures are used in grammatical accounts of the semantics of association with focus (e.g., Krifka 1991). However, this logical recursion serves only to provide appropriate arguments to semantic operators (e.g, *only* and *even*). These accounts still assume a nonrecursive informational/illocutionary structure.

information. Since we are concerned here with information structure and not speech acts, we specifically want to separate the informational/semantic aspect of a question from the illocutionary act of using a question. (Ginzburg ms. contrasts the terms "question" and "query" to make this same distinction.)

Recursion in discourse structure arises naturally in the context of questions and answers. The process of answering a question may spawn extensive question-answer subdialogues (consider, e.g., a court trial as an extended process of answering a question of guilt). The dynamic, incremental growth of information in an extended discourse would be unmanageable without the ability to nest and structure information to arbitrary depths as one proceeds.

It is usually assumed that utterances (or sentences) are the atoms from which the information structure in discourse is built up. We propose, in contrast, that individual sentences have a complex informational structure that mirrors the question-answer structure in a discourse.[2] In this way, the recursion of information structure within a sentence is nothing more than a reflection of the recursion that is possible in a full discourse.

3. Pitch Accents and Relative Prominence

Empirical motivation for recursive focus-ground structures comes from the presence and relative prominence of pitch accents. At the heart of its intonational component, English has a rich system of pitch accents—tonal targets that are located with respect to prominent syllables. (For competing proposals regarding the precise inventory of accent types in English, see, e.g., Bolinger 1958, 1986; Pierrehumbert 1980; Gussenhoven 1983; Pierrehumbert and Hirschberg 1990). A typical reading of the sentence in (4), for example, would have a high-toned pitch accent (H* in Pierrehumbert's 1980 system) associated with the primary stress syllable *lieve*:

(4)   I believe them.
              |
              H*

Alternative renditions of (4) can be achieved by a variety of means: a different accent type, a different placement of the accent, inclusion of additional accents, and different relative prominence for a given accent. This paper considers the latter three factors.[3]

To a first approximation, the relative prominence of pitch accents is reflected in the relative heights at which the pitch accents are produced. Pierrehumbert, for instance, builds this simplifying assumption into her explicit model for English intonation

---

[2] This is a small step away from fully identifying information structure with discourse structure. This view is appealing, but arguing for or against it would be beyond the scope of the current paper.

[3] The first factor, accent type, can certainly interact with focus judgments, but it introduces complexities beyond the scope of the present study.

(1980:73): "a prominence relation between two H tones is expressed as a ratio between their phonetic values." The phonetic value at issue here is fundamental frequency (F0)—the frequency of vocal fold vibration in voiced sounds.

Two further considerations complicate the relation between F0 values and pitch accent prominence. First, F0 values tend to decline over the span of a sentence. Pierrehumbert and Beckman (1988:57) suggest the cover term "downtrend" to cover a range of of factors that contribute to declining F0 values. These include downstep/downdrift (tonally triggered lowering, also called catathesis), declination (phonetic, time-triggered lowering), and final lowering (i.e., lowering at the very ends of sentences; see Liberman and Pierrehumbert 1984). Because of such prevalent lowering trends, a pitch accent of a given prominence can have a markedly lower F0 towards the end of a sentence than near the beginning. Liberman and Pierrehumbert provide explicit numerical modelling of these effects.

Second, certain pitch accent types have the intrinsic effect of lowering any accents that follow them.[4] In Pierrehumbert's pitch accent system (1980), any bitonal pitch accent will trigger downstep of the following pitch accents, as in the following example (1980:92):[5]

(5)   I really believe Ebenezer was a dealer in magnesium
        | | | | |
        H*   H+L*   H+L*   H+L*   H+L* L L%

The transcribed reading for (5) is one which sounds (and looks, on pitch tracks) like a descending staircase—each accented syllable after the first jumps down noticeably from what precedes it. In our data below, we avoid this extra complexity by restricting ourselves to simple high-toned pitch accents.

4.   Focus Layers and Relative Prominence

The core observation that leads us to posit focus layers—recursive focus-ground structures—is the following:

(6)   Differences in the relative prominence of accents can indicate
        alternative orders for packing/unpacking information.

A first line of support for this position is provided by Liberman and

---

[4] Depending on one's other assumptions, this lowering may be nothing more than a reflex of the more general pattern of catathesis. Pierrehumbert (1980) and Ladd (1983) provide contrasting models of how pitch accent types interact with lowering.

[5] The L and L% at the end of (5) are not pitch accent tones, but rather a phrasal low tone and intonational boundary low tone.

Pierrehumbert's (1984; henceforward LP) detailed phonetic study of relative prominence and F0 scaling in English. The accented target words in (7)c, *Anna* and *Manny*, were tested in two contrasting contexts, (7)a and (7)b:

(7)   a. What about Manny? Who came with him?
      b. What about Anna? Who did she come with?
      c. ANNA came with MANNY.

After adjusting for the systematic decline in F0, LP found consistently greater prominence on the accent corresponding to "the main answer". A sample pair of pitch tracks is shown below (1984:168):
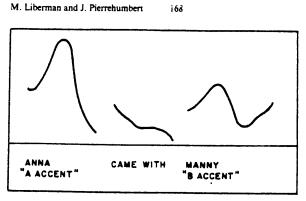
M. Liberman and J. Pierrehumbert   168

Figure 9
An F0 contour for *Anna came with Manny*, produced as a response to *What about Manny? Who came with him?*
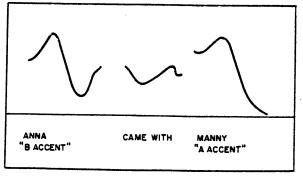

Figure 10
An F0 contour for *Anna came with Manny*, produced as a response to *What about Anna? Who did she come with?*

The analysis in terms of focus layers is just a short step from LP's original example. The nested information structure corresponds to a question asked within the informational scope of another question. In (8), we represent this structure in more discourse-like (dialogue-like) form to emphasize this connection.[6] (Recall that these information-structure questions lack illocutionary force.)

---

[6] In this and following examples, the highest-level question will be labeled as Q, and each level of nesting will prefix a "sub".

340   JONNI M. KANERVA AND LESLIE A. GABRIELE

(8)   a. Q: (Who came with who?)
       b. SubQ: Who came with him [Manny]?
       c.   A: ANNA came with MANNY.

Sentence (8)c answers both Q and SubQ, which corresponds to an information structure with two focus layers. At the outer layer (A as an answer to Q) both *Anna* and *Manny* are focused. At the inner layer (A as an answer to SubQ), only *Anna* is focused.[7]

What remains now is to relate the phonological structuring of accent prominence to the information structure in the focus layers. We propose (9) as an initial generalization:

(9)   a. Within a focus layer, focus is more prominent than nonfocus.
       b. Prominence accumulates across focus layers.

This statement covers two kinds of relative prominence: (i) one accent can be more prominent than another, and (ii) an accented constituent is more prominent than an unaccented one.

Multi-layer examples abound in real discourse. The utterance in (10), for instance, was captured from living dialogue (10/8/94, 4:15 p.m.) by one of the authors. As indicated by the numbers 1 and 2, *you* was accented and most prominent, and *first* had a less prominent accent.

(10)   In fact, I can drop you off first.
                              1   2

One test of a sentence's information structure is that it should mesh with the larger discourse structure within which it occurs. The setting for (10) was that the speaker was planning a trip to take a child to a friend's house and go with his spouse to the grocery store. In this context, the speaker was concerned with (i.e., his information goal was) how to arrange the driving; the information structure in (11) is fully appropriate for this goal of planning the driving route:[8]

(11)   a. Q: Who can I drop off in what order?
        b. SubQ: Who can I drop off first?
        c.   A: I can drop you off first.

---

[7] Note that many researchers (e.g., Steedman 1991) would treat the secondary pitch accent on *Manny* as marking topic. With focus layers, examples like this (and others below) give no reason for positing an extra theoretical entity such as topic. We leave for further research the larger issue of whether topic effects can all be derived by the structural resources of focus layers.

[8] For concreteness, (13)a and (13)b are shown here with the modal *can* already in place. We leave for further research the general issue of where such modal elements are introduced in these layered structures.

Again, the relative prominence of the two accents is determined from the focus layers ((11)c answers SubQ and partially answers Q). *You* is focused in two layers, and *first* is focused in just one layer.

The focus layers approach also sheds light on the subtle contrast between (10) above and a plausible counterpart in which the relative prominence of the accents reversed:

(12)   I can drop you off first.
                        2   1

This accent pattern requires a slightly different information structure:

(13)   a. Q: Who can I drop off in what order?
        b. SubQ: When can I drop you off?
        c.   A: I can drop you off first.

The difference is thus isolated in SubQ: this information goal is either where the driver (*I*) should drive, or when the passenger (*you*) should get dropped off. In this way, the focus layers analysis gives linguistic substance to the intuitions that (10) reflects the driver's perspective, whereas (12) would show more concern for the passenger.

5.   Focus Layers and the Presence/Absence of Pitch Accents

Information structure with focus layers can also solve puzzles in the placement, presence, or absence of pitch accents. Gussenhoven (1984:47) notes a contrast between positive and negative answers to a simple yes/no question:

(14)   a. Is he the GARDENER?
        b. YES, he's the GARDENER.
             versus
        c. NO, he ISn't the gardener.

If polarity is the new part of the response, as the accent on the verb in (14)c suggests, why does (14)b have accent on an NP that is repeated from the immediately preceding question?

Leaving the *yes* aside, we take the accent on *gardener* as a direct clue to (14)b's information structure:

(15)   a. Who/what is he?
        b. He's the gardener.

In other words, (14)b has just a single focus layer, which is formally unrelated to the overt question (14)a. (15) nevertheless fits in the ongoing discourse structure, in the following way. The overt question (14)a is already answered by the *yes*; *he's the gardener* addresses instead a superordinate information goal. The information structures

from the individual pieces can be felicitously unified in a fuller discourse structure, as shown in (16):

(16)   a. Q: (Who/what is he?)   ← not spoken
        b. SubQ: Is he the gardener?
        c. SubA: Yes,
        d. A: he's the gardener.

For the negative answer, (14)c, a similar unification is not available:

(17)   a. Q: (Who/what is he?)
        b. SubQ: Is he the gardener?
        c. SubA: No,
        d. A: ##he isn't the gardener.

The *no* already suffices to answer the yes/no question, but the negative follow-up is not informationally sufficient to answer a superordinate wh-question. It can only be a reiteration and strengthening of the *no*. Hence, an accent on *gardener* would sound discoursally incoherent.

One final example for this section, adapted from Carlson (1983:201), illustrates the striking difference a second accent and second focus layer can make.[9] The setting is a court of law, and the defendant must answer carefully in order to limit possible negative consequences:

(18)   a. What did you hit the victim with?
        b. Safer answer: I hit him with a bicycle chain.
                                                    1
        c. Dangerous answer: I hit him with a bicycle chain.
                                              2   1

The single accent in (18)b corresponds to a single-layer information structure, in which (18)b is an answer to (18)a and nothing else. In (18)c, on the other hand, the secondary accent indicates a dangerous larger information goal:

(19)   a. Q: Who did you hit with what? <dangerous question>
        b. SubQ: What did you hit the victim with?
        c.   A: I hit him with a bicycle chain.

---

[9] Carlson makes the same point but uses syntactic fronting instead of a secondary accent: *him I hit with a bicycle chain*. Indeed, many of Carlson's syntactic examples provide strong evidence for multiple focus layers.

6.   More Than Two Layers

The nonrecursive approaches to information structure, with their limited additional structure, can with some stretching handle the two-layer cases above. The embedded focus, for instance, might be handled as topic. But a recursive approach differs crucially in being able to handle structures reaching to any arbitrary depth. We present below candidates for a three-layer analysis.[10]

Example (20) below is modified from Vallduvi and Zacharski (1994:695).

(20)   a. Background:
            Manny and Larry are separated from their girlfriends,
            who are abroad doing linguistic fieldwork.
            One of the girlfriends is reachable by phone, the other isn't.
        b. Q: What do Manny and Larry do when they're lonely?
        c. A: Manny calls his girlfriend,
                but Larry can only dream about his.
                3   1   2

The information structure for the second clause in (20)c spells out a specific hierarchical order in which the information in the answer is assembled:

(21)   a. Q: Who does what with whom (when lonely)?
        b. SubQ: What does Larry do with whom (when lonely)?
        c. SubSubQ: What does Larry do with his girlfriend (when lonely)?
        d.   A: Larry can only dream about his.

The nested questions in this information structure represent a doublefold narrowing of the informational goals. The scope first narrows to Larry's activities, and then to Larry's interaction with his girlfriend. As for fitting (21) into the larger discourse, this analysis makes the claim that (20)c is presented as a comparison that starts by opposing Manny and Larry and proceeds to specify that difference as a function of how each interacts with his girlfriend.

One final example:

(22)   We have to refute their points,
            but they only have to call ours into question.
            3   2   1

As above, the three-layer information structure shows a successive narrowing of information goals.

---

[10] We believe examples with even more layers than this are possible in principle, as a matter of linguistic competence, but are vanishingly rare in actual performance. Moreover, even if they occurred, they would be hard to detect and hard to distinguish semantically.

344   JONNI M. KANERVA AND LESLIE A. GABRIELE

(23)   a. Q: Who has to do what?
        b. SubQ: What do they have to do?
        c. SubSubQ: What do they have to do with our points?
        d.   A: They only have to call ours into question.

This information structure is a natural fit for a larger discourse such as the following:

(24)   a. Q: (Is it fair?)
               (What does each side have to do?)
        b. SubQ1: (What do we have to do?)
        c.   A1: We have to refute their points

                but

        d. SubQ2: (What do they have to do?)
        e. SubSubQ: (How do they have to deal with our points?)
        f.   A: they only have to call ours into question.

Since the elaborated nonrecursive approaches can provide only two informationally prominent units (e.g., Vallduvi's focus and link or Krifka's focus-of-topic and focus-of-comment), they seem in principle unable to explain three-layer examples such as those just presented. With recursion in focus layers, however, we can handle them straightforwardly.

7.   Conclusion

    In this paper we have analyzed examples of one specific way, embedded question/answer structures, in which the information structure of a sentence mirrors that of the discourse. By explicitly representing the questions (information goals) as part of the sentence's recursive information structure, we obtain conceptual and empirical benefits. Subtle yet strong discourse intuitions can be connected with explicit elements of information structure. The relative prominence among multiple accents in a sentence can be predicted from context, or conversely can be used to predict discourse felicity. It is clear that richer phonological data will be crucial for understanding focus and information—at the very least, secondary accents and their relative prominence need to be faithfully and consistently recorded.

    This work represents also the beginning stage toward a larger goal: a unified account of the interactions of phonology, semantics, and information structure at the sentence level, together with a constrained mapping between sentential information structures and the fuller dynamic discourse. We believe that focus layers, figuratively and literally, point the way to this deeper understanding.

345

References

Bolinger, D. (1958) "A Theory of Pitch Accent in English," *Word* 14, 109-149.

Bolinger, D. (1986) *Intonation and its Parts: Melody in Spoken English,* Stanford University Press, Stanford.

Carlson, L. (1983) *Dialogue Games: An Approach to Discourse Analysis,* D. Reidel, Dordrecht.

Carlson, L. (1984) "Focus and Dialogue Games: A Game-Theoretical Approach to the Interpretation of Intonational Focusing," in L. Vaina and J. Hintikka (Eds.), *Cognitive Constraints on Communication,* D. Reidel, Dordrecht, 295-333.

Ginzburg, J. (n.d.) "Resolving Questions," ms. University of Edinburgh.

Gussenhoven, C. (1983) *A Semantic Analysis of the Nuclear Tones of English,* Indiana University Linguistics Club, Bloomington.

Gussenhoven, C. (1984) *On the Grammar and Semantics of Sentence Accents,* Foris, Dordrecht.

Jackendoff, R. (1972) *Semantic Interpretation in Generative Grammar,* MIT Press, Cambridge.

Kanerva, J. M. (1990) *Focus and Phrasing in Chichewa Phonology,* Garland Press, New York.

Kanerva, J. M. (1991) "Focus Contexts," in S. Chung and J. Hankamer (Eds.), *A Festschrift for William F. Shipley, Syntax Research Center,* Santa Cruz, 123-146.

Krifka, M. (1991) "A Compositional Semantics for Multiple Focus Constructions," *SALT* 1, 127-158.

Ladd, D. R. (1983) "Phonological Features of Intonational Peaks," *Language* 59, 721-759.

Liberman, M. and J. B. Pierrehumbert (1984) "Intonational Invariance under Changes in Pitch Range and Length," in M. Aronoff and R. T. Oehrle (Eds.), *Language Sound Structure: Studies in Phonology Presented to Morris Halle,* MIT Press, Cambridge, 157-233.

Partee, B. H. (1991) "Topic, Focus and Quantification," *SALT* 1, 159-187.

Pierrehumbert, J. B. (1980) *The Phonology and Phonetics of English Intonation,* Doctoral dissertation, Massachusetts Institute of Technology, Cambridge. [Distributed 1987 by the Indiana University Linguistics Club, Bloomington.]

Pierrehumbert, J. B. and M. E. Beckman (1988) *Japanese Tone Structure,* MIT Press, Cambridge.

Pierrehumbert, J. B. and J. Hirschberg (1990) "The Meaning of Intonation Contours in the Interpretation of Discourse," in P. Cohen, J. Morgan, and M. Pollack (Eds.), *Plans and Intentions in Communication,* MIT Press, Cambridge, 271-312.

Rooth, M. E. (1985) *Association with Focus,* Doctoral dissertation, University of Massachusetts, Amherst.

Rooth, M. E. (1992) A Theory of Focus Interpretation, *Natural Language Semantics* 1, 75-116.

Steedman, M. (1991) Structure and Intonation, *Language* 67, 260-296.

Vallduvi, E. (1992) *The Informational Component,* Garland Press, New York.

346   JONNI M. KANERVA AND LESLIE A. GABRIELE

Vallduvi, E. and R. Zacharski (1994) "Accenting Phenomena, Association with Focus, and the Recursiveness of Focus-ground," in P. Dekker and M. Stokhof (eds.), *Proceedings of the Ninth Amsterdam Colloquium, Institute for Logic, Language and Computation,* Amsterdam, 683-702.

von Stechow, A. (1991) "Current Issues in the Theory of Focus," in A. von Stechow and D. Wunderlich (Eds.), *Semantik/Semantics: An International Handbook of Contemporary Research,* de Gruyter, Berlin, 804-825.

Department of Linguistics
Indiana University
Bloomington, IN 47405

jkanerva@indiana.edu
lgabriel@indiana.edu