

THE EMPHATIC JUNCTURE: A NOVEL USE OF THE IP BOUNDARY IN ENGLISH

Bethany Sturman

University of California, Los Angeles
bsturman@ucla.edu

ABSTRACT

This paper presents evidence for a new non-syntactic use of the IP boundary in American English, the EMPHATIC JUNCTURE (EJ), annotated on the Breaks tier as *4e*. The EJ has two functions. The first is to highlight the material following the juncture as prominent. The second is to signal perspective shift. The EJ's status as an IP-type boundary is clear from significant final lengthening of the preceding word and the presence of a notable pause (sometimes >600 ms). The boundary tone sequence used to mark an EJ is a plateau (H-L%, !H-L%). It is intentionally planned by the speaker rather than a disfluency, as evidenced by the continuity of the pitch track across the juncture as well as downstep across the EJ.

Keywords: prosody, prosody-semantics interface, intonation, juncture, annotation

1. INTRODUCTION

In American English, the Intonational Phrase (IP) is the largest prosodic phrase. The right edge of an IP is marked with a boundary tone (e.g. H%, L%), final lengthening, and a large juncture after the final word of the IP that may include a pause [10, 1, 2]. In many cases, IP boundaries align with the edges of syntactic constituents ([13, 14], among others). In some cases, however, speakers insert additional IP boundaries for information structural reasons, such as adding a pause before and/or after narrow focus or adding a pause after a contrastive topic item (see [3] for an overview of both.)

In this paper, I present evidence for a new non-syntactic use of the IP boundary, the EMPHATIC JUNCTURE (EJ), annotated on the Breaks tier in MAE_ToBI [2] as *4e*. The EJ is found in various constructions such as transparent free relatives (e.g. *in what some folks call a % silver tsunami*, shown in Fig. 7), partial quotation (e.g. *Larry challenged an % "alarming rule" % at the board meeting.*, see Fig. 2), and in various speech styles, such as sermon speech or news speech.

The data shown in this paper come from three

sources. The first is a paired production-perception experiment investigating the prosody of partial quotation and its perspective shifting properties (henceforth PQ Exp.) The pair of participants were instructed to imagine they were playing a game where they scored by interpreting the sentences in the same way. The speaker was instructed to read the words on the screen after reading the item silently. This paradigm was successful in eliciting emotionally engaged speech. The second source was National Public Radio recordings (henceforth NPR), and the third source was the recording of a Baptist sermon.

2. THE PHONETIC MARKERS OF THE EMPHATIC JUNCTURE

The phonetic markers of the EJ are very similar to the canonical IP boundary but consist of a specific combination of IP boundary cues. The EJ induces final lengthening on the word preceding the juncture, and the EJ must be realized with a notable pause. The boundary tone sequence used to mark an EJ is a plateau (H-L%, !H-L% in MAE_ToBI [2]).

2.1. Final lengthening

The EJ's status as an IP-type boundary is clear from significant final lengthening of the word directly preceding the juncture. Consider, for example, the word *an* in Figs. 1 and 2, uttered by the same female speaker as part of the partial quotation experiment. Because these tokens were produced as part of an experiment with a counterbalanced design, there are no tokens uttered by the same speaker to act as a precise minimal pair by varying only the presence of an EJ due to the usage of partial quotation. These two utterances form a near minimal pair, however, given that the word of interest occurs in the same syntactic position and with very similar surrounding phonetic segments on either side of the potential juncture location ([d] _ V).

In Fig. 1, *an* is produced IP-medially, with a duration of 160 ms. In Fig. 2, *an* is produced directly before an EJ, with a duration of 439 ms. In both figures, the realization of *an* is highlighted with a box.

Figure 1: *Ron observed an eerie enigma last night.* Duration of *an* = 160 ms (data from PQ Exp.)

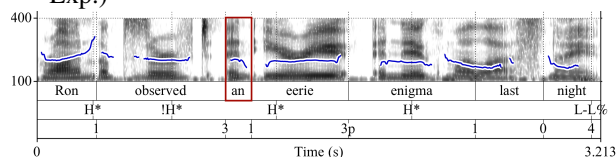
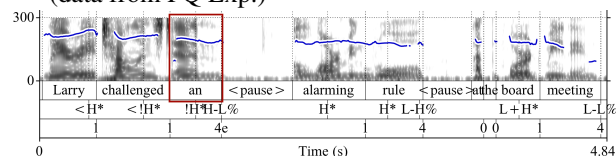


Figure 2: *Larry challenged an “alarming rule” at the board meeting.* Duration of *an* = 439 ms (data from PQ Exp.)



2.2. Obligatory pause

The EJ’s phonetic status as an IP boundary is also supported by the presence of a notable pause. Although a pause is optional as part of a canonical IP boundary, the pause is obligatory in the case of the EJ. These pauses can be rather lengthy, sometimes >600 ms. The pause in Fig. 2 between *an* and *alarming*, for example, is 606 ms. However, this pause is not due to a slowdown in phonological planning or some other disfluency, as would typically be marked on the MAE_ToBI Breaks tier with the 2 or 3*p* label. In fact, it is intentionally placed by the speaker to highlight the material following the EJ. Some phonetic evidence for this intentionality includes the continuity of the pitch track across the juncture, as can be clearly seen in Fig. 2. The intentionality from a semantic perspective will be explored further in 3.1.

2.3. Plateau boundary tone

While a canonical IP boundary can be marked with a tonal rise, fall, or plateau, the boundary tone sequence used to mark an EJ must be a plateau (H-L% or !H-L%). The plateau indicates that there is a continuation of the utterance to come after the juncture.

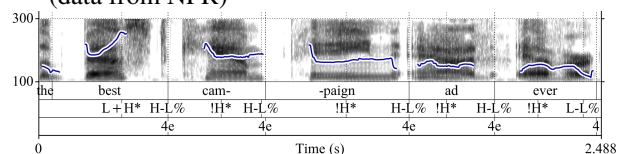
3. THE FUNCTION OF THE EMPHATIC JUNCTURE

The EJ has two primary semantic functions. The first is to highlight the material following the juncture as prominent (3.1), a strategy particularly common in certain speech styles (3.1.1). The second is to mark the following content as perspective shifted (3.2).

3.1. The EJ highlights prominence

One of the primary functions of the EJ is to highlight the material following the juncture as prominent. An example of this function is shown in Fig. 3. In this utterance, the speaker places an EJ after every syllable besides *the*, even within a single word (*campaign*), and each of these syllables is pitch-accented. These extra junctures and pitch accents emphasize the importance of the marked material. This juncture is similar to the break index 2 described in the ToBI manual, found in the example “Iraqi” ([1], ch. 2.10), but the degree of emphasis is even higher here.

Figure 3: *The best % cam- % paign % ad % ever.* (data from NPR)



3.1.1. Emphatic marking in various speech styles

The highlighting function of the EJ is particularly common in certain performative speech styles, such as news speech (see the second juncture in Fig. 4) or sermon speech (Fig. 5.) The speaker can use the EJ to signal to listeners that the material following the juncture is particularly important or noteworthy. In these styles, the EJ can also be used as a rhetorical tool, inserting pauses in unexpected locations to keep the audience engaged, similar to how the speaker can vary pitch range and speech rate for stylistic reasons. There need not be any marked construction to justify EJ insertion.

Figure 4: *In what NASA is calling % a room % with a view.* (data from NPR)

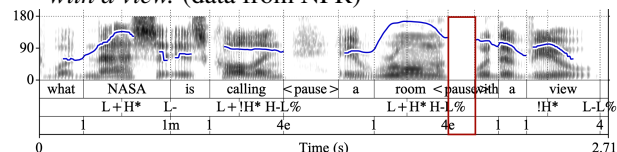
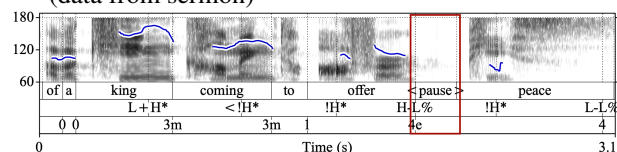


Figure 5: *Of a king coming to offer % peace.* (data from sermon)



3.2. The EJ marks perspective shift

Speakers are generally committed to the truth of their utterances by default. In certain cases, how-

ever, speakers may utter content they might not fully endorse. This is the semantic phenomenon known as perspective shift [6, 15]. Perspective shift is a pragmatically risky strategy since speakers generally wish to avoid content they do not endorse being mistakenly attributed to their beliefs [9, 6]. As such, perspective shift typically occurs in the presence of various syntactic or semantic constructions that provide a salient individual to whom the shifted material can be attributed. I will examine the role of the EJ in two such constructions, transparent free relatives and partial quotation, neither of which have been described prosodically. I will then discuss why the EJ is a consistent prosodic feature of perspective shifting constructions.

3.2.1. Transparent free relatives

Transparent free relatives (TFRs; e.g. *Allen poured what (is called/he calls) a % beergarita* [17]) provide a syntactic means to introduce an explicit attribution of an expression (*beergarita*, known generally as the pivot [5]) with a verb of saying (*calls*) or belief and an optional source for the attribution (*he*). Speakers typically mark the left edge of the pivot with an EJ, shown in Figs. 6 and 7.

Figure 6: *Allen poured what he calls a % beergarita at the party on Friday.* (data from PQ Exp.)

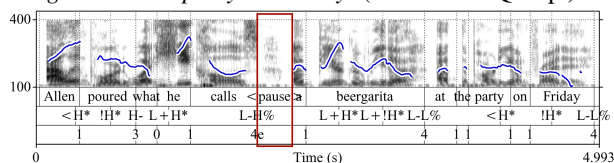
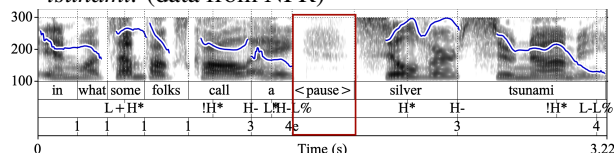


Figure 7: *In what some folks call a % silver tsunami.* (data from NPR)



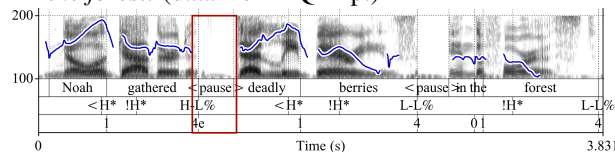
When using a TFR to perspective shift, the material the speaker wishes to attribute to another source is contained within the pivot. Note that the EJ occurs directly before the material the speaker wishes to perspective shift.

3.2.2. Partial quotation

Partial quotation (e.g. *Noah gathered "deadly berries" in the forest.*) has been proposed to facilitate perspective shift [11], following previous analyses for full clausal quotation [8]. Unlike TFRs, however, partial quotation lacks an explicit mechanism to encode the non-speaker

source for the quoted material. Speakers mark the onset of partial quotation (what would correspond to the orthographic quotation marks) using an EJ. In addition to the partial quotation example shown in Fig. 2 above, the example in Fig. 8 demonstrates the use of the EJ to mark partial quotation.

Figure 8: *Noah gathered % deadly berries % in the forest.* (data from PQ Exp.)



The right edge of the partial quotation is also being marked with a notable pause, though the boundary tone is not a plateau but a fall (L-L%). This juncture is serving to audibly delineate the right edge of the perspective-shifted material. The right edge juncture is optional, but if it occurs it is always paired with an earlier EJ.

3.2.3. Why does the EJ mark perspective shift?

As shown in 3.2.1 and 3.2.2, speakers regularly employ the EJ to mark the start of the perspective shifted material. The EJ is a signal the speaker can utilize (along with facial expressions, paralinguistic gestures, voice quality, etc.) to indicate that the marked material should be attributed to a source other than the speaker. Inserting a large pause is one of the best tools a speaker has to differentiate perspective-shifted material in the midst of an utterance that also contains content the speaker does want to endorse. The plateau preceding the pause indicates an impending continuation (i.e. that there is more material following the pause that is related to the pre-pausal material.) By combining both the plateau and the pause, the speaker is cuing they are shifting perspective in the midst of the utterance.

4. DIFFERENCES BETWEEN THE EJ AND THE CANONICAL IP BOUNDARY

4.1. Canonical IP boundary corresponds to syntax

In many cases, IP boundaries align with the edges of syntactic constituents ([13, 14], a.o.). The default placement of IP boundaries serves to mark the syntactic grouping of words, with larger prosodic boundaries more likely to occur at the edges of larger syntactic constituents. Speakers can use this convention to disambiguate certain ambiguous structures, such as the attachment location of an adjunct PP. For example, consider the sentence in (1):

- (1) The artist sketched the man with the pen.

With no medial IP boundaries, (1) is ambiguous as to whether *the pen* should be interpreted as a modifier, describing the particular man the artist sketched as the one who had the pen, or as an instrument, specifying what writing instrument the artist used. These two meanings can be disambiguated through the insertion of IP boundaries [12, 4, 16, 7]. Inserting an IP boundary after *sketched* (shown in (2)) disambiguates to the modifier interpretation, whereas inserting an IP boundary after *man* disambiguates to the instrument interpretation, shown in (3).

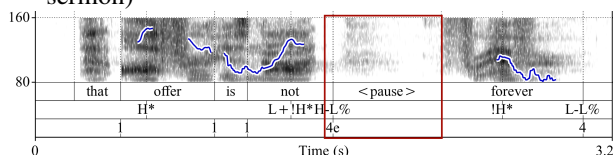
- (2) The artist sketched % the man with the pen.
 (3) The artist sketched the man % with the pen.

Thus, IP boundaries are typically informative to the underlying syntactic structure and constituency.

4.2. The EJ does not mark syntactic structure

Although canonical IP boundaries typically mark the edges of large syntactic constituents, this is not a function of the EJ. In fact, an EJ can intervene between even the most local constituency relationships. EJs can occur between a determiner and its noun, such as between *an* and *alarming* in Fig. 2 in 2.1 above, or between negation (*not*) and an adverb (*forever*), as in Fig. 9.

Figure 9: *That offer is not % forever.* (data from sermon)



Given the typical role IP boundaries play in marking syntactic structure, the placement of EJs in these positions is quite surprising. Although the phonetic realization of the EJ gives it the appearance of an IP boundary, clearly the EJ is not functioning to mark the syntactic grouping of words. This becomes even more apparent when we consider the domain of downstep in relation to the EJ.

4.3. Downstep across the EJ

In English, downstepping refers to a high target being realized lower relative to a previous high target than can be explained by declination (e.g. !H* in MAE_ToBI), resulting in a lowering of the upper bound of the pitch range. Subsequent high targets will thus be realized no higher than the pitch of the !H* target unless the pitch range is reset at

an Intermediate Phrase break. Downstepping often occurs when the information is predictable or backgrounded [10].

In the MAE_ToBI framework, the Intermediate Phrase is defined as the domain of downstep, meaning that downstepping can only occur between two high targets that are within the same intermediate phrase. The fact that the Intermediate Phrase is defined as the domain of downstep is also why pitch range reset may optionally occur at an Intermediate Phrase boundary.

Perhaps surprisingly, however, downstep is sometimes licensed across the EJ, specifically the highlighting EJ. In Fig. 9, for example, the high targets, associated with *not* and *forever*, on either side of the EJ are in a clear downstep relationship. When the pause is removed in this utterance, it sounds like a textbook case of downstep from *offer* (H*) to *not* (L+!H*) and again from *not* to *forever* (!H*).

An even more striking example is shown in Fig. 3 above. Given that EJs are being placed word internally in this utterance, this example shows that the highlighting use of the EJ does not correspond to an information structure boundary. Since this EJ does not signal an information structure boundary, downstepping is licensed across the juncture.

Downstepping is not licensed, however, across the perspective shifting EJ. Instead, pitch range is allowed to reset after the juncture. When we consider that the perspective shifting EJ is placed in order to mark the start of perspective shifted material, this divergence makes sense. Perspective shift signals a major information structure boundary, and pragmatically, the perspective shifted material should be neither predictable nor backgrounded.

5. DISCUSSION

The emphatic juncture looks very similar to an IP boundary phonetically, being realized with final lengthening and a pause. Unlike an IP boundary, however, the pause is obligatory, and the EJ does not correspond to syntactic grouping relations. Rather, the EJ has two functions: 1) highlighting the material following the juncture as prominent or 2) marking the onset of perspective shift.

Although the two types of EJ are phonetically identical, they differ in how they interact with the prosodic structure of the utterance regarding downstep. Future research could illuminate more distinctions between these two types of EJ and perhaps even further uses of the EJ, as well as other types of boundaries with non-syntactic functions.

6. REFERENCES

- [1] Beckman, M., Ayers-Elam, G. 1997. Guidelines for ToBI labelling (version 3.0).
- [2] Beckman, M., Hirschberg, J., Shattuck-Hufnagel, S. 2005. The original ToBI system and the evolution of the ToBI framework. In: Jun, S., (ed), *Prosodic Typology: The Phonology of Intonation and Phrasing, Volume 1*. Oxford University Press 9–54.
- [3] Büring, D. 2003. On d-trees, beans, and b-accents. *Linguistics and philosophy* 26, 511–545.
- [4] Fodor, J. D. 1998. Learning to parse? *Journal of psycholinguistic research* 27, 285–319.
- [5] Grosu, A. 2003. ‘transparent’ free relatives as a special instance of ‘standard’ free relatives. *From NP to DP* 1, 139–178.
- [6] Harris, J., Potts, C. 2009. Perspective-shifting with appositives and expressives. *Linguistics and Philosophy* 32, 523–552.
- [7] Jun, S. A. 2003. Prosodic phrasing and attachment preferences. *Journal of psycholinguistic research* 32, 219–249.
- [8] Kaplan, D., others, 1989. *Themes from Kaplan*. Oxford University Press on Demand.
- [9] Lasersohn, P. 2007. Expressives, perspective and presupposition. *Theoretical Linguistics* 33(2), 223–230.
- [10] Pierrehumbert, J. 1980. *The Phonology and Phonetics of English Intonation*. PhD Thesis MIT.
- [11] Potts, C. 2007. The dimensions of quotation. *Direct compositionality* 405–431.
- [12] Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., Fong, C. 1991. The use of prosody in syntactic disambiguation. *the Journal of the Acoustical Society of America* 90(6), 2956–2970.
- [13] Selkirk, E. 1986. On derived domains in sentence phonology. *Phonology* 3, 371–405.
- [14] Selkirk, E. 2011. The syntax-phonology interface. In: Goldsmith, J., Riggle, J., Yu, A., (eds), *The Handbook of Phonological Theory*. 2nd edition.
- [15] Smith, C. S. 2003. *Modes of discourse: The local structure of texts* volume 103. Cambridge University Press.
- [16] Snedeker, J., Trueswell, J. 2003. Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language* 48, 103–130.
- [17] Wilder, C. 1998. Transparent free relatives. *ZAS papers in linguistics* 10, 191–199.