

Phonetic and Phonological Correlates of Broad, Narrow and Contrastive Focus in English

Dmitry Sityaev and Jill House

University College London, UK

E-mail: dima@phonetics.ucl.ac.uk, jill@phonetics.ucl.ac.uk

ABSTRACT

This paper investigates how different types of focus in English are realised prosodically. Two production experiments were conducted to explore the phonetic and phonological correlates of broad, narrow and contrastive focus on NP constituents. The findings suggest that there are small but significant differences between the three types of focus. Words under narrow and contrastive focus tend to have longer duration. In sentences with two pitch accents, the second pitch peak undergoes a smaller degree of downstep when associated with contrastive focus. There was no clear evidence of association between a focus type and a particular pitch accent.

1. INTRODUCTION

The notion of focus has been explored in different realms of linguistics – in syntax, semantics, and phonology. This has led to sometimes competing definitions of focus. The term has a long association with the phonological phenomenon of pitch accent [1], [2], [3]. The approach preferred here is to define the notions of “focus” and “accent” independently, and to investigate the relationship between phonological structure (partially determined by pitch accent placement) and focus structure (determined by syntax/semantics). This relationship becomes clearer when we consider the sub-categories of “broad”, “narrow” and “contrastive” focus in more detail. Consider (1):

(1) She broke her LEG

where capitalization indicates that the word is accented.

(1) can be an answer to any of the questions below:

(2a) What happened?

(2b) What did she break?

(2c) Did she break her neck?

Depending on the question, different focus structures are associated with (1), namely

(3a) [_F She broke her LEG]

(3b) She broke [_F her LEG]

(3c) She broke her [_{CF} LEG]

(3a) is considered to have a *broad* focus reading; in other words the whole utterance is presented as new information. (3b) has a *narrow* focus on the word “leg”, i.e. only part of the utterance contains new information. (3c), in some accounts [4] a special case of narrow focus,

can be described as having a *contrastive* focus on the word “leg”: the speaker uses contrastive focus to introduce a contrasting element (“leg”) into the discourse with which he intends to override or correct an element (“neck”) already present in the hearer’s informational context.

It has been claimed that English sentences like (1) are potentially ambiguous between a broad focus and a narrow focus interpretation, since in a non-emphatic reading there are no reliable phonetic cues to distinguish the two [5]. This would imply either that there was no difference in phonological structure, or choice of pitch accent, or that any underlying difference was neutralised in the phonetic realisation. On the other hand, it has been claimed that the prosodic correlates of contrastive focus may be different from those of broad/narrow focus. For instance, in American English, contrastive focus is said to be marked with a L+H* pitch accent which has a different phonetic and phonological identity from the monotonal H* pitch accent used to mark new information in cases of broad or narrow focus [6]. Another study [7] failed to find support for the claim made in [6]; however, in a series of perceptual experiments the authors showed that for speakers of American English, peak height was by far the most prominent cue to contrastivity. Pitch height is also used in Scottish English to signal contrastivity: accents associated with contrastive focus, as opposed to those simply signalling new information, were observed to be extra-high, or “boosted” [8]. [1] and [9] maintain that there are no reliable phonetic cues to contrastivity.

Two production experiments were conducted to explore the prosodic correlates of broad, narrow and contrastive focus on NP constituents in British English. Experiment 1 looked at sentences where the same NP was in focus under all three conditions, and therefore accented. A monosyllable was used each time as the target word. The position of the accented word was systematically varied within the sentences. In sentences such as “My HEAD aches” there was just one pitch accent predicted. Our primary concern was to investigate properties of the accented word itself. Following the findings of [7], [8] and [10], our initial hypothesis was that contrastive focus would be associated with pitch accents with higher F0 peaks and late peak alignment.

Experiment 2 investigated sentences containing two NPs which could be predicted to carry pitch accents (e.g. “MELANIE will lean on MERRYLIN”), elicited in contexts where the second accented word would be in broad, narrow or contrastive focus. Attention was mainly directed towards the relationship between the two peaks, to see whether English showed patterns similar to those reported for Catalan [4], whereby when contrastive focus fell on the second peak downstep was blocked. No strong phonetic differences were predicted in the realisation of broad vs. narrow focus.

2. EXPERIMENT 1

2.1 Method

2.1.1 Material and subjects

The data consisted of recorded sentences containing a pitch accent realised on a monosyllabic NP. The sentences were constructed so that the position of the accented monosyllable was varied: it was placed sentence-initially (“MEL has just phoned”), sentence-medially (“I want to have my LEG examined”) and sentence-finally, (“I broke my NECK”). All were declaratives for which it was reasonable to expect a falling intonation, e.g. H* pitch accent and L% boundary tone. Each sentence was produced in broad, narrow and contrastive focus contexts. Participants in the experiment were 4 male and 2 female native speakers of Southern British English aged between 20 and 29, with no speech or hearing impediments. For each focus condition speakers recorded the same 12 sentences twice. A total of 72 sentences was obtained from each speaker.

2.1.2 Equipment and procedure

The data were recorded in an anechoic chamber, using a B&K sound level meter of the type 2231, fitted with a 4165 microphone. Speakers were given oral instructions, and seated in front of the microphone and a computer screen. During the recording session, the subject was presented with an auditory stimulus (a pre-recorded context question) via headphones, while the sentence to be used as the answer appeared as a visual prompt on the monitor. The speaker was instructed to read the answer aloud in a manner which corresponded to the question asked. The recordings made on DAT were further digitised at a sampling rate of 16 kHz (16-bit resolution) and segmented using a CoolEdit program. Each file was then transferred onto a Sun Sparcstation and analysed using XWAVES+. For each file, the waveform, fundamental frequency contour and a broad spectrogram were obtained.

The start and end of the target word were estimated using auditory and visual criteria. In each case values for peak F0 and word duration were obtained, together with a measurement of peak onset time relative to vowel onset.

2.2 Results

Results of one-way ANOVAs revealed that focus condition had no significant effect on the height of the F0 peak in the target word, or on peak alignment. Analysis of word durations, however, showed that focus condition has some effect, though weak. For some speakers, the accented words produced under both narrow and contrastive focus were significantly longer in duration than the same words in the broad focus condition (Fig. 1). This was significant for speaker AW sentence-initially ($F_{2,14}=5.026$, $p<0.05$) and sentence-finally ($F_{2,14}=8.985$, $p<0.01$) and speaker SF sentence-medially ($F_{2,14}=4.859$, $p<0.05$). For speaker DK, the word duration under contrastive focus was longer than under both broad and narrow sentence-medially ($F_{2,14}=7.499$, $p<0.01$)

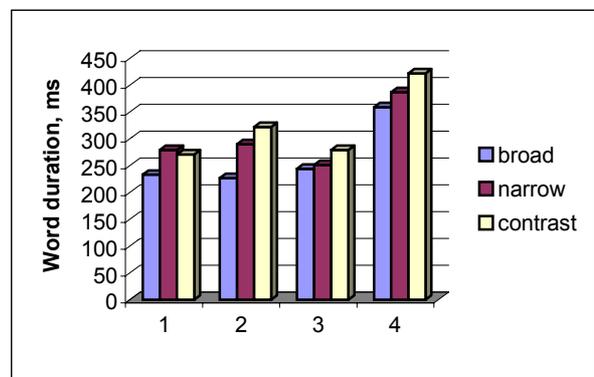


Figure 1. Mean key-word duration in various sentence positions under three focus condition. 1 = speaker AW (sentence-initially), 2 = SF (sentence-medially), 3 = DK (sentence-medially), 4 = AW (sentence-finally).

Data from the two female speakers (AW and ZH) was inspected further for differences in phonological structure, pitch accent choice and boundary tone. When the target word was placed in initial position, all sentences were realised with just the one pitch accent on that word. When the word was placed sentence-medially or finally, with few exceptions there was an opportunity for additional pitch accents to be produced on accentable words occurring earlier in the sentence. This option was taken up for the great majority of broad focus examples. The situation for narrow and contrastive contexts was less straightforward: sometimes a pitch accent would be clearly present at the potential site, sometimes it would be suppressed altogether, or flattened in contour shape. There was no clear pattern of distribution for either focus condition. Pitch accents on the target word all fell into the H* or L+H* category across focus conditions; clear-cut examples of L+H* were in the minority, best identified when the low leading tone could be associated with deep valleys between pitch accents. On a few occasions the two speakers produced a fall-rise pattern ((L+)H*L-H%), always in the contrastive context.

3. EXPERIMENT 2

3.1 Method

3.1.1 Material and subjects

The data consisted of recorded sentences containing two NPs predicted to carry pitch accents, e.g. “MELANIE will lean on MERRYLIN”. Focus contexts were varied for the second accented word. The NPs included both monosyllabic and polysyllabic words.

Four subjects, 3 female and 1 male, were all native speakers of Southern British English aged between 23 and 28, with no speech or hearing impediments. Each sentence in each context was recorded three times. In total there were 1080 sentences recorded and analysed (270 from each speaker).

3.1.2 Equipment and procedure

Equipment and procedure of this experiment were as described in 2.1.2. Analysis focussed on measuring the peak F0 values (H1 and H2) for the two accented words.

3.2 Results

The relationship (downstep/upstep) between H1 and H2 was considered as a function of focus condition. The results revealed that in the broad focus context, downstep was significant for the three female speakers (MF: $t=20.110$, $df=50$, $p<0.001$; HB: $t=7.640$, $df=52$, $p<0.001$; FS: $t=4.812$, $df=43$, $p<0.001$) and upstep was significant for the male speaker (LW: $t=-5.989$, $df=47$, $p<0.001$).

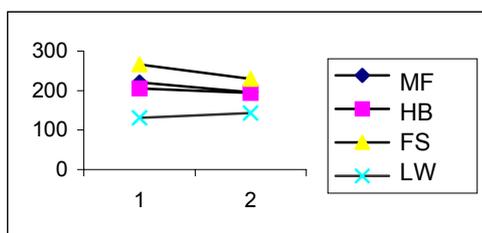


Figure 2. F0 values for H1 and H2 in broad focus conditions for each speaker.

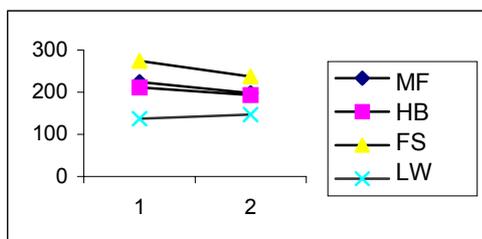


Figure 3. F0 values for H1 and H2 in narrow focus conditions for each speaker.

Narrow focus placed on the last word in the sentence showed a similar effect to broad focus: downstep was significant for the three female speakers (MF: $t=20.934$, $df=48$, $p<0.001$; HB: $t=9.858$, $df=32$, $p<0.001$; FS: $t=6.421$, $df=51$, $p<0.001$) and upstep was significant for the male speaker (LW: $t=-4.207$, $df=49$, $p<0.001$).

When contrastive focus was placed on the last word in the sentence, the relationship between the two pitch accents was characterised by downstep for the two female speakers (MF: $t=14.363$, $df=49$, $p<0.001$; HB: $t=6.124$, $df=30$, $p<0.001$) and upstep for the other two subjects – significant for male (LW: $t=-8.608$, $df=43$, $p<0.001$) but insignificant for female (FS: $t=-0.762$, $df=47$, $p>0.05$).

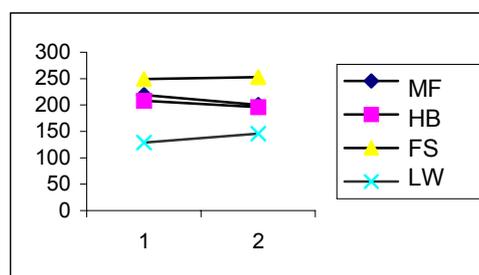


Figure 4. F0 values for H1 and H2 in contrastive focus conditions for each speaker.

The next step was to investigate any difference in scaling of the two peaks in the data under broad, narrow and contrastive focus conditions. The ratios between the two peaks H2 and H1 were obtained for each sentence, and a series of ANOVA tests were performed.

Results revealed that for two female speakers the difference in H2/H1 ratio was significantly different between contrastive and broad (MF: $p<0.01$; FS: $p<0.001$) and also between contrastive and narrow (MF: $p<0.001$; FS: $p<0.001$). For the third female speaker, difference was only significant between narrow and contrastive focus (HB: $p<0.05$). There were no significant differences in the H2/H1 ratio for the male speaker (LW).

Speaker	broad	narrow	contrastive
MF	0.89	0.88	0.91
HB	0.94	0.91	0.94
FS	0.88	0.87	1.03
LW	1.11	1.09	1.14

Table 1. H2/H1 ratios for broad, narrow and contrastive focus for each speaker.

4. DISCUSSION

The results obtained in Experiment 1 regarding peak height and peak alignment support the view expressed in [1] and [9] that there are no robust phonetic cues in the accented word itself for disambiguating between broad, narrow and contrastive focus. The hypothesis that contrastive focus would be associated with “boosted” F0 peaks was not supported. Differences in peak onset time could be attributed to sentence position, consistent with [11], but not to focus condition. There was no evidence in the English data for an association between contrastive focus and “delayed” F0 peaks, as reported in [10] for German, where it is claimed that late peaks are associated with “emphasis on a new fact” or “contrast to what should exist or exists in the speaker’s or hearer’s idea”.

However, for some speakers, the duration of the accented word increased under narrow or contrastive focus. A similar increase in duration under contrastive focus has been reported in [12]. Increased duration of this sort could be analysed as phrase-final lengthening, reflecting the presence of a phrase accent belonging to an embedded intermediate phrase and thus a difference in underlying phonological structure. Observed differences in the number and nature of pre-final pitch accents in the different contexts also suggest that further investigation of structure might be fruitful. Perceptual evidence is needed to test the robustness of durational and other effects.

Experiment 2 showed that when sentences with two pitch-accented NPs were spoken in conditions of broad or narrow focus, there was usually some downstep between the two peaks. However, placing a contrastive focus on the last word in a sentence led to some subjects produce a downstep on H2 and to some subjects produce an upstep on H2. Similar results were reported in [4] for Catalan: they concluded that broad focus is associated with downstep, while contrastive focus in sentence-final position tends to involve a non-downstepped peak. Further analysis of the scaling of the two pitch peaks in Experiment 2 revealed that even when downstep was observed, it was present to a significantly lesser degree in the contrastive context than when there was broad or narrow focus in sentence-final position. This provides some additional evidence that there is a small difference between contrastive and non-contrastive focus realisations.

In general, when compared across the two production experiments, the results show that there are more differences between contrastive focus and non-contrastive focus (broad or narrow). These differences were mostly evident when comparing sentences containing two pitch accents with narrow/contrastive focus in sentence-final position.

5. CONCLUSION

The findings of the two production experiments suggest that there are small but significant differences which speakers can make between the three types of focus. More differences were found between contrastive vs non-contrastive (narrow/broad) focus. The extent to which these cues are used by listeners is not yet known. Tests are in progress to test the perceptual robustness of the prosodic cues identified

REFERENCES

- [1] M.A.K Halliday, *Intonation and Grammar in British English*. The Hague: Mouton, 1967.
- [2] S. Prevost, *A semantics of contrast and information structure for specifying intonation in spoken language generation*, PhD dissertation, University of Pennsylvania, 1995.
- [3] M. Steedman, “Structure and Intonation”, *Language*, 67, pp. 260-296, 1991.
- [4] E. Estebas and J. Maidment, Central Catalan declaratives: the relation between focus and downstep. In *Proceedings of the ICPHS 1999*, San Francisco, pp. 1861-1864, 1999.
- [5] R. Ladd, *Intonational phonology*. Cambridge University Press, 1996.
- [6] J. Pierrehumbert and J. Hirschberg, “The meaning of intonational contours in the interpretation of discourse,” in *Intentions in communication*, Cohen, Morgan and Pollack, Eds., pp. 342-265. Cambridge, MA: MIT Press, 1990.
- [7] C. Bartels and J. Kingston, “Salient pitch cues in the perception of contrastive focus,” in *Focus and Natural Language Processing: Proceedings of the Journal of Semantics Conference on Focus*, IBM Working Papers. 1994.
- [8] G. Brown, K. Currie and J. Kenworthy, *Questions of Intonation*, London: Croom Helm, 1980.
- [9] D. Bolinger, *Intonation and its Parts*, London: Edward Arnold, 1986.
- [10] K. Kohler, “Prosody in speech synthesis: the interplay between basic research and TTS application,” *Journal of Phonetics*, vol. 19, pp. 121-138, 1991.
- [11] K. Silverman & J. Pierrehumbert, “The timing of pre-nuclear high accents in English”, in J. Kingston & M. Beckman (eds), *Papers in Laboratory Phonology 1*, CUP, pp 72-106, 1990.
- [12] W. Cooper, S. Eady and P. Mueller, “Acoustical aspects of contrastive stress in question-answer contexts,” *Journal of Acoustical Society of America*. Vol. 77(6), pp. 2142-2156, 1985.