

Single vs. double focus in English statements and yes/no questions

Fang Liu

Department of Psychology, Goldsmiths, University of London, UK

f.liu@gold.ac.uk

Abstract

This study investigates the acoustic realization of single vs. double focus in statements and yes/no questions in General American English. Four speakers produced four sets of utterances of different lengths with alternating focus and sentence type conditions. Results indicate that double focus increases max F_0 and duration of the focused word to a similar degree as single focus in both statements and yes/no questions. Furthermore, post-focus pitch range suppression occurs in both single- and double-focused statements. However, in yes/no questions, post-focus pitch range compression and raising occur only after single focus and focus 2 of double focus. In contrast, F_0 after focus 1 of double focus falls gradually until the stressed syllable in focus 2 in yes/no questions. These findings suggest that double focus interacts with sentence type in a more complicated way than single focus in shaping F_0 contours of English sentences.

Keywords: single/double focus, sentence type, English

1. Introduction

The occurrence of multiple foci in a single sentence is very common in daily conversation, as in the following example.

- Anne will read Lee the new mail. (Neutral focus)
- **WHO** will read **WHOM** the new mail? (Double focus)
- **ANNE** will read **LEE** the new mail. (Double focus)

Much research has been done on the acoustic realization of single focus in statements and questions in various languages [1, 2, 3] among others, and there have also been studies investigating the prosodic characteristics of double focus in statements in different languages [4, 5, 6]. However, to date there has been no research on how double focus is encoded in questions in the world's languages. Therefore, the first aim of the current study was to examine how double focus differs from single/neutral focus in affecting F_0 contours of yes/no questions in General American English.

According to [4], the effect of focus is *perhaps* more pronounced in shorter sentences than in longer ones in English. In order to verify/falsify this claim, the current study employed sentences of different lengths in investigating double focus in English statements and yes/no questions.

Finally, for English statements, it has been found that maximum F_0 's and durations of the focused words are comparable between single- and double-focused conditions [4]. However, post-focus F_0 is significantly lowered after single focus only, but not after focus 1 of double focus. Therefore, the current study explored whether similar patterns hold for single focus versus focus 1 of double focus in English yes/no questions.

2. Methods

2.1. Subjects

Four native speakers of General American English, three female and one male, participated in the study. They aged between 22 and 32, and self-reported no speech or hearing disorders.

2.2. Materials

Test materials included sentences of four different lengths, in which the initial, medial, and final words may be focused.

- **Short (7 syllables):** *Anne* will read *Lee* the new *mail*
- **Medium (12 syllables):** *Nina* is selling *Lily* a yellow *lemon*
- **Long (17 syllables):** *Elaine* might be introducing *Lamar* to her best girlfriend *Arlene*
- **Extra long (24 syllables):** *Amelia* has been accommodating *Ramona* with a lot of delicious *vanilla*

Each sentence was prompted by different preceding statements/questions to contain the following possible sentence type and focus conditions.

- **Sentence type (2):** statement vs. yes/no question
- **Focus (7):** neutral, initial, medial, final, initial + medial, initial + final, and medial + final

Each combination of the conditions was produced five times by every subject, resulting in a total of 1120 sentences (4 lengths \times 2 sentence types \times 7 focus conditions \times 4 subjects \times 5 repetitions).

2.3. Procedure

Recordings were done in a soundproof booth in the Department of Speech, Hearing and Phonetic Sciences at University College London. Through a JavaScript program, both the prompt and target sentences were displayed one at a time in random order in a web browser. The words to be focused were highlighted using square brackets. Subjects read aloud both the prompt and target sentences into a headset microphone that was about 2 inches away from the left side of their lips. Sampled at 44.1 kHz and with 16-bit amplitude resolution, the produced utterances were directly digitized onto a hard disk using Praat [7].

2.4. Data analysis

Using a Praat script [8], time-normalized F_0 contours of all the syllables in the utterances were extracted, and maximum and mean F_0 's (st) and durations (ms) of the syllables were also calculated. Repeated measures ANOVAs were conducted using R [9] to examine the effects of focus (neutral, initial, medial, final, initial + medial, initial + final, and medial + final) and sentence type (statement vs. yes/no question) on max F_0 , mean F_0 , and duration of the three key words (initial, medial, and final) in the sentences. In these ANOVA models, F_0 and duration measurements were the dependent variables, subject the random factor, and focus, sentence type, and keyword position the within-subject factors. For post-hoc analyses, Tukey's Honest Significant Difference method (the Studentized range statistic) was used for multiple comparisons between the means of multiple levels of the factors. Due to space limit, only a subset of graphs and statistics will be presented here.

3. Results

3.1. Graphic analysis

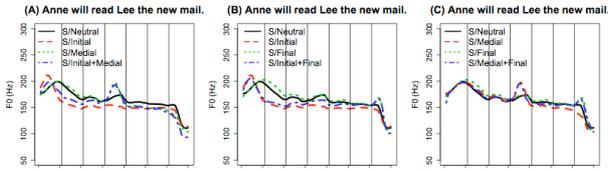


Figure 1: Neutral vs. single vs. double focus in English statements. Note: S/Neutral: a statement with neutral focus; S/Initial+Medial: a statement with double focus on the initial and medial words. The vertical lines are syllable boundaries.

Figure 1 displays time-normalized F_0 contours of neutral-/single-/double-focused English statements, each averaged across 20 repetitions by 4 subjects. As can be seen, compared to neutral focus, both single and double focus create a tri-zone F_0 modification in English statements: pre-focus pitch range neutralization, on-focus pitch range expansion, and post-focus pitch range suppression (compression and lowering). While the focused words under single or double focus seem to have similar maximum F_0 values, the post-focus pitch range suppression caused by single focus appears to have a greater magnitude and a larger scope than that by focus 1 of double focus.

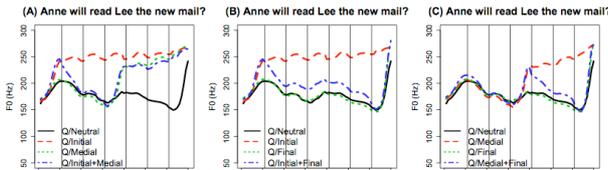


Figure 2: Neutral vs. single vs. double focus in English yes/no questions. Note: Q/Neutral: a question with neutral focus.

Figure 2 shows averaged time-normalized F_0 contours of neutral-/single-/double-focused English yes/no questions. As can be seen, post-focus pitch range is compressed and raised to a high plateau by both single focus and focus 2 of double focus. In contrast, the F_0 trajectory after focus 1 of double focus falls gradually till focus 2. However, the focused words have similar max F_0 values across single and double focus conditions.

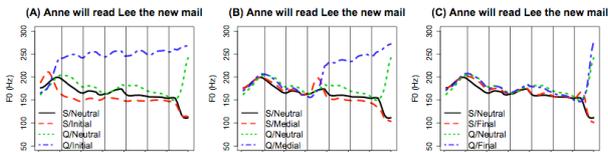


Figure 3: Neutral vs. single focus in English statements vs. yes/no questions.

Figure 3 demonstrates how English statements and yes/no questions differ under neutral vs. single focus. As can be seen from Fig. 3C, neutral-/final-focused statements and yes/no questions mainly differ in F_0 of the final word: falling in the former and rising in the latter. However, when under initial (Fig. 3A) or medial (Fig. 3B) focus, F_0 contours of these two sentence types differ not only in the focused word (falling in statements and rising in questions), but also in post-focus pitch range: compressed and lowered in statements, but compressed and raised in yes/no questions. Also, as found in [2] and will not be discussed in detail here, there is a pitch target shift for the stressed syllable of the content word between statements ([high] or [fall]) and yes/no questions ([rise]) in English.

Consequently, the F_0 difference between these two sentence types becomes salient starting from the stressed syllable of the first content word, whether or not it is focused (Fig. 3).

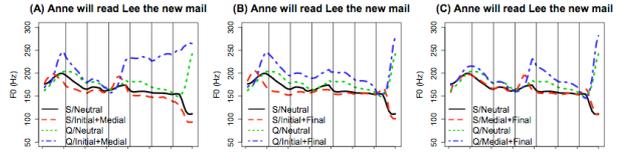


Figure 4: Neutral vs. double focus in English statements vs. yes/no questions.

Figure 4 illustrates how double focus affects the F_0 contours of statements and yes/no questions in English, with the neutral focus condition as reference. As can be seen from Fig. 4A, post-focus pitch range is compressed and lowered after both foci in double-focused statements. However, in double-focused questions, post-focus pitch range is compressed and raised after focus 2, but falls gradually after focus 1.

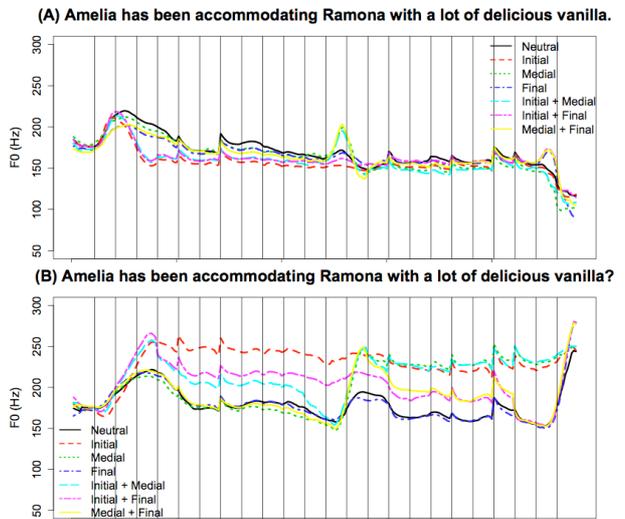
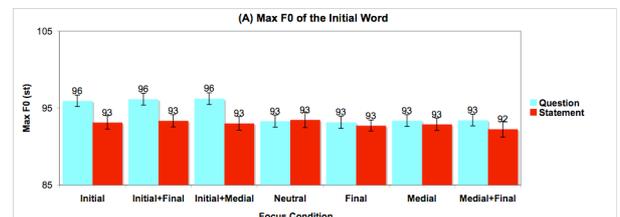


Figure 5: Neutral vs. single vs. double focus in the extra-long statements and questions.

Figure 5 displays averaged time-normalized F_0 contours of the extra-long statements and questions under different focus conditions. The overall patterns of the focus and sentence type effects on F_0 trajectories are similar to those in the short sentences discussed above, suggesting that communicative functions are encoded similarly regardless of sentence length. Furthermore, the difference in max F_0 between focus 1 and focus 2 of the double-focus conditions (initial + medial, medial + final, and initial + final) was not significantly different across different sentence lengths in either statements or questions (t -tests: all $p > 0.05$). Therefore, the following statistical analyses were conducted based on all the sentences, with the effect of sentence length not considered.

3.2. Statistical analysis

3.2.1. Max F_0



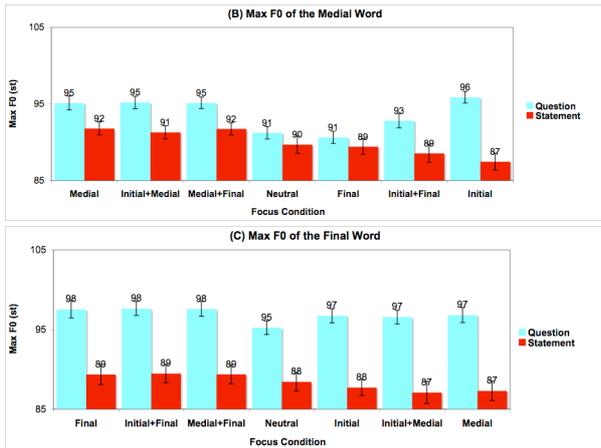


Figure 6: Max F_0 's of the initial, medial, and final words in statements and questions under different focus conditions.

As shown in Figure 6 and revealed in the repeated measures ANOVA, max F_0 's of the three key words vary significantly with its position ($F(2,6) = 15.99, p = 0.0039$), sentence type ($F(1,3) = 10.75, p = 0.0465$), and focus ($F(6,18) = 6.68, p = 0.0008$). That is, the initial key word (mean = 93.7 st) is in general higher in max F_0 than the final key word (mean = 92.6 st), which is in turn higher in max F_0 than the medial key word (mean = 91.9 st). Furthermore, the three key words generally have higher max F_0 's in questions (mean = 95.0 st) than in statements (mean = 90.5 st), and they also have different max F_0 's under different focus conditions. The two-way and three-way interactions between the three factors are all statistically significant (sentence type \times focus: $F(6,18) = 12.23, p < 0.0001$; sentence type \times key-word position: $F(2,6) = 19.85, p = 0.0023$; focus \times key-word position: $F(12,36) = 12.08, p < 0.0001$; sentence type \times focus \times key-word position: $F(12,36) = 3.07, p = 0.0045$). The following post-hoc analyses reached statistical significance ($p < 0.05$).

Firstly, the initial word has significantly greater max F_0 in questions than in statements under initial, initial + final, and initial + medial focus conditions (Fig. 6A). This is due to the pitch target shift of the stressed syllable in the initial word from [high] or [fall] in statements to [rise] in questions, as mentioned earlier. For the question-initial word, its max F_0 is significantly greater when it is focused (under initial, initial + medial, and initial + final focus conditions) than when it is pre-focus (under medial, final, and neutral focus conditions), demonstrating both on-focus pitch range expansion and pre-focus pitch range neutralization. Furthermore, the question-initial word under initial + final and initial + medial focus conditions also has significantly higher max F_0 than when it is under the medial + final focus condition, again demonstrating both on-focus pitch range expansion and pre-focus pitch range neutralization.

Secondly, for the medial word (Fig. 6B), its max F_0 is significantly greater in questions than in statements when it is post-focus (under initial and initial + final focus condition), or on-focus (under medial, initial + medial, and medial + final focus conditions). This is either due to post-focus pitch range difference between statements (compressed and lowered) and questions (compressed and raised) after the initial focus, or the pitch target shift of the stressed syllable in the medial word between statements ([high] or [fall]) and questions ([rise]). For the statement-medial word, its max F_0 is significantly greater when it is focused (under medial, initial + medial, and medial + final focus conditions) than when it is post-focus (under initial and initial + final focus conditions), demonstrating both

on-focus pitch range expansion and post-focus pitch range suppression in statements. The question-medial word has significantly higher max F_0 when it is post-focus (under initial focus) or on-focus (under medial, initial + medial, and medial + final focus conditions) than when it is pre-focus (under final and neutral focus conditions), again due to the tri-zone pitch range modification of focus in questions. Interestingly, the question-medial word also has significantly higher max F_0 when it is under initial focus than when it is under the initial + final focus, suggesting that there is post-focus pitch range suppression after focus 1 of double focus in questions.

Thirdly, for the final word (Fig. 6C), its max F_0 is significantly greater in questions than in statements under all focus conditions. This is either due to post-focus pitch range difference between statements (compressed and lowered) and questions (compressed and raised) after the initial/medial focus, or the pitch target shift of the stressed syllable in the final word between statements ([high] or [fall]) and questions ([rise]).

Finally, as shown in Fig. 6A-C, each of the three key words does not have significantly different max F_0 when it is single-focused than when it is double-focused in either statements or questions. That is, the on-focus pitch range expansion is similar in magnitude across single and double focus conditions.

3.2.2. Duration

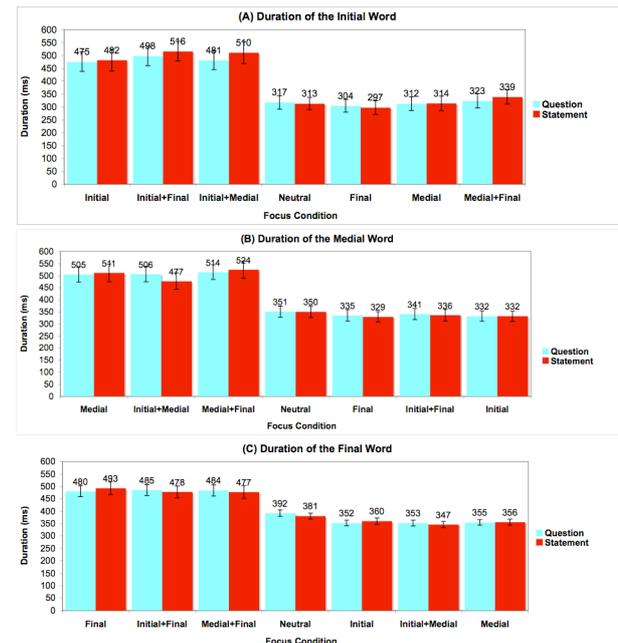


Figure 7: Durations of the initial, medial, and final words in statements and questions under different focus conditions.

As shown in Figure 7 and revealed in the repeated measures ANOVA, durations of the three key words vary significantly with focus ($F(6,18) = 13.68, p < 0.0001$), but not with its position ($F(2,6) = 2.37, p = 0.17$) or sentence type ($F(1,3) = 0.02, p = 0.90$). That is, the three key words have different durations under different focus conditions, but their durations are similar across key-word positions and sentence types. In addition, the two-way interaction between focus and key-word position is also statistically significant ($F(12,36) = 14.60, p < 0.0001$), which will be explained by the following post-hoc analyses (only comparisons that reached statistical significance, $p < 0.05$, are shown below).

Firstly, as shown in Fig. 7A, both statement- and question-initial word has significantly longer duration when it is focused (under initial, initial + final, and initial + medial focus conditions) than when it is not focused (under medial, medial + final, final, and neutral focus conditions).

Secondly, as shown in Fig. 7B, both statement- and question-medial word has significantly longer duration when it is focused (under medial, initial + medial, and medial + final focus conditions) than when it is not focused (under initial, initial + final, final, and neutral focus conditions).

Thirdly, as shown in Fig. 7C, the statement-final word has significantly longer duration when it is focused (under final, initial + final, and medial + final focus conditions) than when it is under the initial + medial focus condition. Furthermore, the statement-final word under final focus also has significantly longer duration when it is under initial/medial focus. For the question-final word, its duration is significantly longer when it is focused (under final, initial + final, and medial + final focus conditions) than when it is under the initial or initial + medial focus condition. Furthermore, the question-final word also has significantly longer duration when it is under double focus (initial + final and medial + final) than when it is under medial focus. As also found in [4], these patterns seem to suggest that earlier focus tends to shorten the duration of the final word.

Finally, as shown in Fig. 7A-C, single- and double-focused key words do not differ in duration in either statements or questions. Namely, single and double focus increase the duration of the focused word in a similar amount.

3.2.3. Single vs. double focus

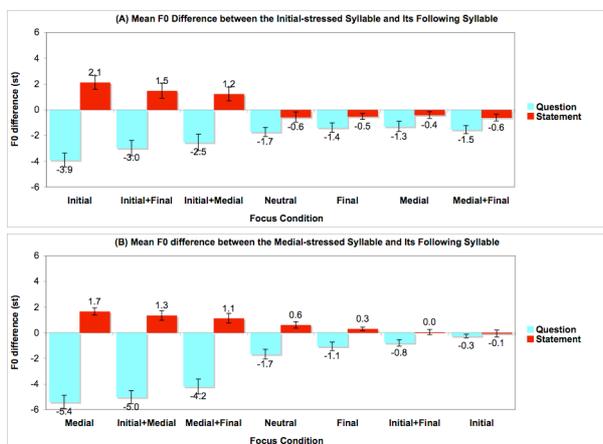


Figure 8: Mean F_0 differences between the stressed syllable in the initial (A) and medial (B) words and its following syllable in statements and questions under different focus conditions.

The difference in mean F_0 (in st) between the initial-/medial-stressed syllable and its following syllable was calculated as an indicator of the magnitude of post-focus F_0 drop/increase in statements and questions under single vs. double focus (see the first three pairs of bars in Fig. 8A-B). Post-hoc analyses in the repeated measures ANOVAs of mean F_0 difference on focus and sentence type indicated that this difference is numerically, but not significantly, bigger (in absolute value in the case of questions) in single-focused conditions than in double-focused conditions.

4. Discussion and conclusions

In summary, this study suggests that in English statements, double focus has similar effects as single focus on max F_0 and duration of the key words. Specifically, max F_0 's and

durations of the focused words are increased to the same degree by single and double focus. Furthermore, as demonstrated by the immediate mean F_0 drop between the stressed syllable in initial/medial focus and its following syllable (Section 3.2.3), the magnitude of post-focus pitch range suppression after focus 1 of double focus is not significantly different from (although is slightly smaller than) that after single focus.

In yes/no questions, max F_0 's and durations of the focused words are also increased to the same degree by single and double focus. Furthermore, focus 2 of double focus shows similar effects as single focus in raising and compressing the pitch of post-focus words. However, as demonstrated by the significantly lower max F_0 for the medial word when it is under initial + final focus than when it is under initial focus (Section 3.2.1), there is post-focus suppression after focus 1 of double focus in English questions. Namely, F_0 falls gradually from the stressed syllable in focus 1 until the stressed syllable in focus 2 in double-focused English questions.

Finally, although further analysis is warranted, in the current study max F_0 differences between the two foci in double-focused sentences do not vary with sentence length, suggesting that the effect of double focus is similar across sentences of different lengths.

In summary, the above findings indicate that single and double focus are realized similarly in English statements, but differ in terms of post-focus pitch modification for focus 1 of double focus in yes/no questions. This is likely due to the fact that focus and sentence type modify F_0 of the between-focus words in opposite directions in double-focused yes/no questions.

5. Acknowledgements

This work was supported in part by the Economic and Social Research Council [Grant number PTA-026-27-2480].

6. References

- [1] Eady, S. J. and Cooper, W. E., "Speech intonation and focus location in matched statements and questions", *Journal of the Acoustical Society of America*, 80:402-416, 1986.
- [2] Liu, F., "Intonation systems of Mandarin and English: A functional approach", Ph.D. Dissertation, University of Chicago, 2009.
- [3] Pell, M. D., "Influence of emotion and focus on prosody in matched statements and questions", *Journal of the Acoustical Society of America*, 109:1668-1680, 2001.
- [4] Eady, S. J., Cooper, W. E., Klouda, G. V., Mueller, P. R. and Lotts, D. W., "Acoustical characteristics of sentential focus: Narrow vs. broad and single vs. dual focus environments", *Language and Speech*, 29(3):233-251, 1986.
- [5] Rump, H. H. and Collier, R., "Focus conditions and the prominence of pitch-accented syllables", *Language and Speech*, 39(1):1-17, 1996.
- [6] Wang, B., "The prosodic realization of dual focus: Comparing Mandarin and German", *The 8th Phonetic Conference of China and the International Symposium on Phonetic Frontiers*. Beijing, 2008.
- [7] Boersma, P. and Weenink, D., "Praat: doing phonetics by computer (Version 5.0.08) (Computer program)", Retrieved February 11, 2008, from <http://www.praat.org/>.
- [8] Xu, Y., "_TimeNormalizeF0.praat", available from <http://www.phon.ucl.ac.uk/home/yi/tools.html>, 2005-2009.
- [9] R Development Core Team, "R: A language and environment for statistical computing", R Foundation for Statistical Computing, <http://www.R-project.org>, 2009.