Focus and penultimate vowel lengthening in Zulu

Danfeng Wu[1], *Yadav Gowda*[1]

[1]Massachusetts Institute of Technology

`dfwu@mit.edu, ysg@mit.edu`

Abstract

Many Bantu languages exhibit fixed placement of focus at the Immediately-After-the-Verb (IAV) position, which has been argued to be related to this position's prosodic prominence. Elements in this position appear at the right edge of a prosodic phrase, and are subject to penultimate vowel lengthening, which we take to be a form of phrasal stress which occurs at the right edge of every prosodic phrase. Previous literature has claimed that in Zulu, focus cannot be the most prominent element in a sentence. We present evidence from a production study in Zulu showing the contrary, i.e. the degree of penultimate vowel lengthening at the IAV/vP-final position is greater than at any other prosodic phrase edge, lending phonetic support to the claim that this position is prosodically prominent in a sentence. We further show that the vP-final position is prominent regardless of whether or not it is focused, which implies that Zulu has a fixed position that realizes sentential prominence.

Index Terms: focus, prosodic phrasing, penultimate lengthening, duration, production, Zulu, Bantu

1. Introduction

Zulu is a Bantu language (S42) spoken by over 10 million speakers in the southern part of the African continent, primarily in South Africa. Like in many other Bantu languages, focused elements must occupy the Immediately-After-the-Verb (IAV) position, and must be the only non-verbal elements in vP. Unfocused elements that originate in vP must move out of vP to ensure satisfaction of this requirement.

It is puzzling why focused elements must occupy the IAV position. Cheng and Downing [1], [2] have suggested that the IAV position is prosodically prominent, and since focus must occupy a prominent position, it must be in the IAV. If prosodic prominence can be realized acoustically, we should be able to find phonetic evidence for this claim. There is a further question of what the domain is within which the IAV/focused material must be prominent.

One way to realize prominence in Zulu is through phrasal stress. It has been assumed that phrasal stress in Zulu can be marked by penultimate vowel lengthening (PVL) [3], [4], [5]. PVL refers to lengthening of the penultimate vowel of the last word in a prosodic phrase. A higher degree of PVL should correlate with stronger stress and, therefore, higher prominence.

The IAV position is both immediately after the verb, and final in vP. We provide evidence from a production study showing that elements in vP-final position carry the greatest degree of PVL in a sentence. Since PVL is a cue to phrasal stress, this suggests that the vP-final position carries the strongest stress in a sentence. Therefore, the requirement for focus to occupy the vP-final position amounts to alignment of focus to prosodically the strongest position in a sentence.

2. Focus in Zulu

Focus occupies the IAV/vP-final position in Zulu. Take the ditransitive construction as an example. The canonical word order of a ditransitive construction under broad focus is S-V-IO-DO, with the indirect object (IO) *isitha* 'enemy' and the direct object (DO) *ubuthi* 'poison' both in vP, and the subject (S) *umongameli* 'president' outside vP:

(1) [In answer to 'What happened?']
Um-ongame:li) u-thumelele isi-tha ubu:-thi]$_{vP}$)
1-president s1-send.to 7-enemy 14-poison
ngeSo:nto).
on.Sunday
The president sent poison to the enemy on Sunday.

Focused elements must occupy the IAV/vP-final position, and must be the only non-verbal elements in vP. If the IO is focused (2), then the DO must evacuate the vP, and vice versa (3). Evacuation is indicated by a verbal object marker which agrees in class with the moved object [6]. For example, when the DO is focused in (3), the IO moves rightward out of the vP, creating the S-V-DO-IO order. The verb also agrees with the dislocated IO overtly, as is shown by the object marker that matches the class of the IO.

(2) [In answer to 'Who did the president send poison to on Sunday?']
*Um-ongame:li) u-bu-thumelele **isi:-tha**$_F$)]$_{vP}$ ubu:-thi)*
1-president s1-o14-send.to 7-enemy 14-poison
ngeSo:nto).
on.Sunday
The president sent poison to **the enemy** on Sunday.

(3) [In answer to 'What did the president send to the enemy on Sunday?']
*Um-ongame:li) u-si-thumelele **ubu:-thi**$_F$)]$_{vP}$ isi:-tha)*
1-president s1-o7-send.to 14-poison 7-enemy
ngeSo:nto).
on.Sunday
The president sent **poison** to the enemy on Sunday.

Cheng and Downing [1] derive this pattern with a high-ranked constraint, **C$_{PVL}$**: *foci must have PVL*, which derives a prosodic phrase-final position. We propose a revised version of **C$_{PVL}$** and provide support for it.

As seen in (1)-(3), PVL not only occurs in the IAV/vP-final position, but in other positions too, e.g., the pre-verbal subject. If foci are only required to have PVL, they should be allowed to occur in any position with PVL. This is not the case, however. C$_{PVL}$ predicts that the pre-verbal subject should be able to stay in-situ under narrow focus, as this is a position with PVL. Contrary to this prediction, the focused subject must be clefted instead [1], [4], [7]:

(4) [In answer to 'Who is carrying the basket?']
(ù-Nhlâ:nhla) (ó-thwél' ú-bhasikí:di).
COP.1-Nhlanhla REL.1-carry 1-basket
It is Nhlanhla who is carrying the basket.   *(from [7])*

The clefted subject immediately follows the copula, and is phrased with the copula prosodically [4]. Assuming that the copula is verbal, this suggests that the focused subject must occupy the IAV/vP-final position, just as the focused object does (2)-(3). Thus, although many phrase-final positions have PVL, the IAV/vP-final position must be special in some way.

Suppose that the vP-final position is more prominent than other phrase-final positions, and assume that relative degree of PVL reflects relative prosodic prominence [5], then it should have more PVL than any other phrase-final position. In this paper we compare the length of the penultimate vowel of a word when it is focused/in the vP-final position vs. when it is unfocused but in a position with PVL, and find that the degree of PVL is greater in the vP-final position than in any other position in a sentence.

These findings suggest that $\mathbf{C_{PVL}}$ must be constrained further (*contra* Downing [5]) to $\mathbf{C'_{PVL}}$: foci must undergo the greatest degree of PVL *in a sentence*. In other words, not only must foci have PVL and phrasal prominence, but they must have the greatest PVL in a sentence and therefore sentential prominence.

After showing that under narrow focus the vP-final position is the most prominent in a sentence, we investigated the origin of this prominence. One can imagine that focused elements are lengthened just due to focus, i.e. focused elements have a higher degree of PVL than non-focused elements. Thus, if we find that the vP-final position has the greatest degree of PVL in narrow focus sentences, it is possible that this is not an effect of this position, but an effect of narrow focus. Thus, we also want to know whether the vP-final position receives the same degree of PVL when it is narrowly-focused as compared to when it is not narrowly-focused and just happens to be vP-final (e.g. in cases of broad focus). We found no significant difference in the degree of PVL between these two conditions, which suggests that it is not focus that makes a word prominent, but rather, sentential prominence invariantly falls on a fixed position, vP-finally.

3. Methods

3.1. Materials

The materials that test the hypothesis of sentential prominence of the vP-final position consist of three parts. Each part compares the penult vowel duration in the vP-final position with that in an unfocused position with PVL. There are three unfocused positions that we compared the vP-final position with: post-vP sentence-medial, post-vP sentence-final, and pre-verbal.

Part 1 compares the vP-final position with the unfocused post-vP sentence-medial position. It involves 60 target sentences (20 sets x 3 focus conditions). There are three possible focus conditions for each target sentence, namely broad focus, IO focus and DO focus. Focus was elicited with a leading wh-question, as in (1), (2) and (3), and the speaker was to say the leading question and the target sentence (both in Zulu) in a pair. By putting the same word under focus in the vP-final position (e.g. *isitha* 'enemy' in (2) and *ubuthi* 'poison' in (3)) and in the unfocused post-vP sentence-medial position that also undergoes PVL (*isitha* 'enemy' in (3) and *ubuthi* 'poison' in (2)), we can compare the duration of the penultimate vowel in these two positions. We will discuss the use of broad-focus sentences in section 6.

Part 2 has 40 target sentences (20 sets x 2 focus conditions). The two focus conditions are IO focus and DO focus. These tar-

get sentences are identical to those with IO focus and DO focus in the first part of the experiment, except that they do not have a sentence-final adjunct, as in (5) and (6). Focus was elicited with a leading wh-question in the same way. We compare the duration of the penultimate vowel of a word in the vP-final position (e.g. *isitha* 'enemy' in (5) and *ubuthi* 'poison' in (6)) with that of the same word in the unfocused post-vP sentence-final position (*isitha* 'enemy' in (6) and *ubuthi* 'poison' in (5)).

(5) [In answer to 'Who did the president send poison to?']
Um-ongame:li) u-bu-thumelele ***isi:-tha$_F$**]$_{vP}$) ubu:-thi).*
1-president   S1-O14-send.to   7-enemy   14-poison
The president sent poison to **the enemy**.

(6) [In answer to 'What did the president send to the enemy?']
Um-ongame:li) u-si-thumelele ***ubu:-thi$_F$**]$_{vP}$) isi:-tha).*
1-president   S1-O7-send.to   14-poison   7-enemy
The president sent **poison** to the enemy.

Part 3 involves 40 target sentences (20 sets x 2 focus conditions). The focus conditions are locative focus and object focus, as in (7) and (8). Notice that the locative phrase can appear preverbally when unfocused, and that furthermore, in this position, it undergoes PVL. Thus, we compare the duration of the penultimate vowel of a word in the vP-final position (*eTheki* 'Turkey' in (7)) with that of the same word in the unfocused pre-verbal position (*eTheki* 'Turkey' in (8)).

(7) [In answer to 'Where did Sipho drink coffee?']
u-Si:pho) u-li-phuze ***e-The:ki$_F$**]$_{vP}$) i-kho:fi).*
1-Sipho   S1-O7-drink   LOC-Turkey   7-coffee
Sipho drank coffee in **Turkey**.

(8) [In answer to 'What did Sipho drink in Turkey?']
u-Si:pho) e-The:ki) u-phuze ***i-kho:fi$_F$**]$_{vP}$).*
1-Sipho   LOC-Turkey   S1-drink   7-coffee
Sipho drank **coffee** in Turkey.

In addition, we collected another 10 target sentences (5 sets x 2 focus conditions), whose use will be discussed in section 6.

Out of 150 target sentences in total, 5 sentences from Part 1 and 2 sentences from Part 2 had to be discarded due to unforeseen syntactic issues with the stimuli, as well as recording issues.

3.2. Data collection

We conducted a production experiment in Boston with a single native speaker of Zulu from Durban, South Africa. The participant is male, in his 50s, speaks Zulu regularly with his family, and visits South Africa frequently. He was remunerated a small sum for his time, and granted his written consent to being tested.

The data was recorded with a head-mounted microphone and a Zoom H4n digital audio recorder at the sampling frequency of 44.1kHz in a quiet, non-reverberant room. The stimuli were presented in Zulu in slides. All the stimuli were pooled together and presented in random order. The participant was directed to speak as naturally as possible and repeat in case of coughing, stuttering or laughter. The experiment lasted 2 hours, with appropriate breaks given intermittently.

The raw sound data were first chunked into individual utterances, and then the length of relevant vowels and words were labeled and measured manually in Praat [8].

3.3. Data analysis

We fitted three linear mixed effects models, corresponding to three parts of the data. The first two models (for Parts 1-2) had the duration of the penultimate vowel of a word as the depen-

dent variable in each model, and focus (whether the sentence focused the IO or the DO) and argument (whether the length value comes from the IO or the DO) as fixed effects. In the third model (for Part 3), the dependent variable was the penultimate vowel length of the locative phrase, so there was only one fixed effect, i.e. focus (whether the sentence focused the locative argument or the object). All the models included random intercepts and slopes for the target words. For linear mixed models, we used the *lme4* [9] and *lmerTest* [10] packages in R [11]. Figures were generated with *ggplot2* [12], *ggsignif* [13], and *tikzDevice* [14].

4. Results

We found from the first model that the penultimate vowel in the vP-final position is on average 62 ms longer than the penultimate vowel in the post-vP sentence-medial position ($p < 0.001$).
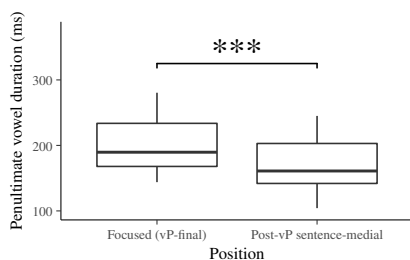

Figure 1: *Penultimate vowel duration vP-finally vs. in the post-vP sentence-medial position.*

Given this data, there is the possibility that the penultimate vowel is longer under focus due to an overall lengthening effect on the containing word. To account for this possibility we also measured the length of entire containing word. We ran another regression with the *ratio* of the length of the penultimate vowel to the length of the whole word as the dependent variable, and with the same fixed and random factors as used previously. We found that this ratio is greater by 2% in the vP-final position than in the post-vP sentence-medial position ($p < 0.01$). This indicates that there is significantly more PVL in the vP-final position than in the post-vP sentence-medial position, and that this PVL effect is independent of the possible lengthening effects on the whole word.

We found from the second model that the penultimate vowel in the vP-final position is on average 54 ms longer than the penultimate vowel in the post-vP sentence-final position ($p < 0.001$).
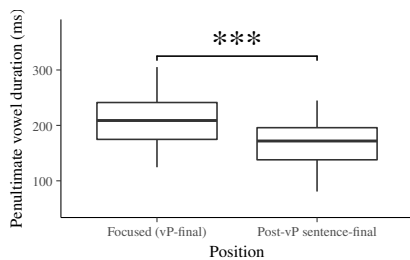

Figure 2: *Penultimate vowel duration vP-finally vs. in the post-vP sentence-final position.*

Results from the third model showed that the penultimate vowel in the vP-final position is on average 46 ms longer than the penultimate vowel in the pre-vP position ($p < 0.001$).
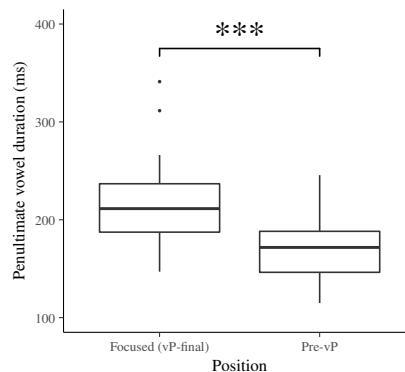

Figure 3: *Penultimate vowel duration vP-finally vs. in the pre-vP position.*

These results together indicate that the vP-final position has a higher degree of PVL than other positions which undergo PVL, supporting the claim that it undergoes the highest degree of PVL in a sentence, and therefore is the most prosodically prominent.

5. Interim discussion

If we take PVL to be a cue to phrasal stress, the experimental findings in §4 indicate that focused elements in the vP-final position carry the strongest stress in a sentence. Two possible theories about the origin of this stress are: (1) the vP-final position always has sentential stress, no matter whether it is focused or not; or (2) focused elements independently receive stress, regardless of their position. In the following section, we test these two theories and argue that (1), that the vP-final position always has sentential stress regardless of focus, accounts for the data best.

6. Further analyses

To test the hypothesis that sentential stress falls on a fixed position, vP-finally, regardless of focus, we compared the degree of PVL in the vP-final position under narrow focus with the vP-final position that is not narrowly-focused (under broad focus). If they do not differ, then we cannot reject this hypothesis, and prominence occurs in a fixed position.

6.1. Data

The materials used in this part of the analysis were elicited in the same session as the materials before. They consist of two parts, using ditransitive and monotransitive constructions respectively.

The first part uses ditransitive constructions to compare the vP-final position that is narrowly focused with the position under broad focus that happens to be vP-final. There are 40 target sentences to compare totally (20 sets x 2 focus conditions). The two focus conditions are broad focus, as in (1) and DO focus, as in (3). We compared the penultimate vowel duration of *ubuthi* 'poison' in (1) and (3). Though it is vP-final in both cases, it is not narrowly focused and just happens to be vP-final in (1), whereas it is narrowly focused in (3).

The second part makes the same comparison in monotransitive constructions. There are 10 target sentences (5 sets x 2 focus conditions). The focus conditions are broad focus and object focus, as in (9) and (10). For example, we compared the penultimate vowel duration of *ixoxo* 'frog' in (9) and (10).

(9)   [In answer to 'What happened?']
       U-kho:zi) lu-bulale i-xo:xo]$_{vP}$).
       10-eagle s10-kill 5-frog
       The eagle killed the frog.

(10)   [In answer to 'What did the eagle kill?']
       *U-kho:zi) lu-bulale **i-xo:xo**$_F$]$_{vP}$).*
       10-eagle s10-kill 5-frog
       The eagle killed **the frog**.

6.2. Data analysis

We fitted two linear mixed effects models for the two parts of the data respectively. They had the length of the penultimate vowel of the DO and the object as their respective dependent variables, and focus as a fixed effect (broad focus vs. narrow focus). Both models included random intercepts and slopes by target word.

6.3. Results

Results from the first model showed that there is not a significant difference in length of the penultimate vowel of the DO in DO focus vs. broad focus.
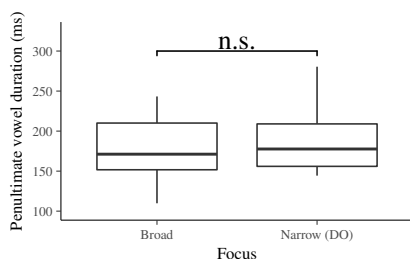

Figure 4: *Penultimate vowel length of DO under broad focus and narrow focus on DO.*

We saw similar results for the second model. The degree of penultimate vowel lengthening applied to the vP-final position position is not significantly different between broad focus and narrow focus cases.
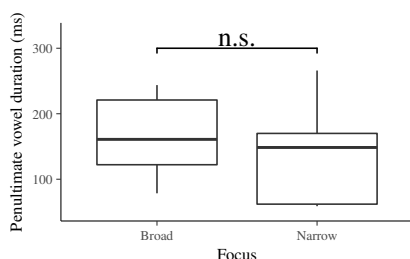

Figure 5: *Penultimate vowel length of the object in broad focus and narrow focus.*

Additionally, in order to account for the possibility that the entire word lengthens rather than just the penultimate vowel, we compared the ratio of the penultimate vowel of the word in the vP-final position to the antepenultimate vowel, and again, did not find a significant difference.

These results suggest that the same degree of PVL is applied to the vP-final position regardless of whether or not the element in this position is under narrow focus, and that therefore there is no specific effect of focus on PVL. Rather the effect on PVL that we see is applied uniformly to elements occupying the vP-final position.

7. Discussion

We have shown not only that the vP-final position is the most prominent in a sentence, i.e. it is more prominent than any other position that also undergoes PVL, but also that the vP-final position is prominent regardless of whether or not it carries narrow focus. This answers the question of why focused elements must be vP-final: focused elements must occupy the most prosodically prominent position in a sentence.

Our empirical findings not only shed light on focus in Bantu, but also informs a broader cross-linguistic view on focus. Languages differ in how they realize focus. In languages where the location of pitch accent is flexible (e.g. English), pitch accent placement can be manipulated in order to ensure that focus has the strongest pitch accent in a phrase. Other languages that do not have this flexibility use other ways to mark focus. One famous example is Hungarian, which has a dedicated pre-verbal position for focus, to which contrastively-focused elements and elements with answer focus move overtly. Szendrői [15], [16], [17] has argued that this movement is driven by prosody. Elements in this pre-verbal position receive sentential stress, so focused elements must move to this position in order to receive this stress.

These facts appear strikingly similar to Zulu, as noted by Cheng & Downing [2]. Parallel to Hungarian focus, focused elements in Zulu must receive sentential stress, which falls on the vP-final position. Zulu also has prosody-driven movement, but instead of moving the focused element, in Zulu, non-focused elements move out of the way, ensuring that the focused elements receive sentential stress.

8. Conclusion and future work

This paper has provided evidence from a production study for the claim that the vP-final position is the most prosodically prominent in a sentence. Focused elements, which must be prosodically prominent, need to occupy this position. We hope to conduct future research to determine whether the other defining aspect of the IAV position, its placement after the verb, is also prosodically driven. We also hope to verify these results by conducting the same study with additional speakers. Finally, we plan to conduct a perception study to determine whether speakers are sensitive to the differences we found to be statistically significant in this study.

9. Acknowledgements

10. References

[1] L. Cheng and L. J. Downing, "Where's the topic in Zulu?" *The Linguistic Review*, vol. 26, no. 2-3, pp. 207–238, 2009.

[2] ——, "Against FocusP: arguments from Zulu," in *Contrasts and positions in information structure*, I. Kučerová and A. Neeleman, Eds. Cambridge University Press, 2012, pp. 247–266.

[3] J. S. M. Khumalo, "An autosegmental account of Zulu phonology," Ph.D. dissertation, University of the Witwatersrand, Johannesburg, 1987.

[4] L. Cheng and L. J. Downing, "The prosody and syntax of Zulu relative clauses," *SOAS Papers in Linguistics*, vol. 15, pp. 51–63, 2007.

[5] L. J. Downing, *Focus and Prominence in Chichewa, Chitumbuka and Durban Zulu*. Universitätsbibliothek Johann Christian Senckenberg, 2008.

[6] L. C. Buell, "Issues in Zulu verbal morphosyntax," Ph.D. dissertation, UCLA, 2005.

[7] L. Cheng and L. J. Downing, "Clefts in Durban Zulu," *Cleft structures*, pp. 141–164, 2013.

[8] P. Boersma and D. Weenink, "Praat: Doing phonetics by Computer," 2019.

[9] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[10] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest package: Tests in linear mixed effects models," *Journal of Statistical Software*, vol. 82, no. 13, pp. 1–26, 2017.

[11] R Core Team, *R: A Language and Environment for Statistical Computing*, Vienna, Austria, 2019.

[12] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.

[13] C. Ahlmann-Eltze, *ggsignif: Significance Brackets for 'ggplot2'*, 2019, r package version 0.6.0. [Online]. Available: https://CRAN.R-project.org/package=ggsignif

[14] C. Sharpsteen and C. Bracken, *tikzDevice: R Graphics Output in LaTeX Format*, 2019, r package version 0.12.3. [Online]. Available: https://CRAN.R-project.org/package=tikzDevice

[15] K. Szendrői, "Focus and the syntax-phonology interface," Ph.D. dissertation, University College, London, 2001.

[16] ——, "A stress-based approach to the syntax of Hungarian focus," *The Linguistic Review*, vol. 20, pp. 37–78, 2003.

[17] ——, "Focus movement (with special reference to Hungarian)," *The Blackwell companion to syntax*, pp. 272–337, 2006.