A Downstep Analysis of Swahili Prosody: a Case of the Exclusive Particle TU

Shigeto Kamano

International Christian University

1 Introduction

This paper introduces an analysis of Swahili prosody, focusing on the exclusive particle TU. Swahili is widely spoken as a lingua franca in Eastern Africa, but the prosody of Swahili remains understudied. Previous studies report that TU means 'only' or 'just,' appears at the end of a sentence, and is pronounced with a falling intonation pattern. However, the latest study with recordings of a Swahili native speaker suggests that TU is produced with a high pitch, and the semantic domain of TU corresponds to the pitch compression range. In the current paper, in contrast to pitch compression, I propose that downsteps better characterize the prosodic domain of TU.

First, background information on TU is introduced (Section 2). Section 3 discusses Kamano et al. (2022), which suggest the pitch compression of TU. Methodological details of recordings for the current study are shown in Section 4. After demonstrating the results in Section 5, I propose a downstep analysis as an alternative analysis of Swahili prosody in Section 6. This paper concludes with discussions of the analyses and suggestions for future research.

2 TU in Swahili

This section introduces the syntactic/prosodic characteristics of the exclusive particle TU. TU means 'only, exactly, simply, merely' and originates from the adverb *túpú* 'only, in vain' in Proto-Bantu (Meeussen 1967). There are two types of hosts that TU is attached to: a preceding verbal host in a VP and a phrasal-final nominal host in a DP. 1 and 2 represent these two types of hosts, respectively.

(1) Example of TU attached to VP
Wa-tu wa-me-lela tu.
cl2-person cl2-PRF-sleep TU
'The people are just sleeping.' (Ashton 1944)

(2) Example of TU attached to DP
N-a-taka ma-wili tu.
I-ABS-want cl6-two TU
'I want only two (eggs).' (Ashton 1944)

Kamano et al. (2022) argue that the difference in the hosts of TU marks different semantic domains of TU. Using Swahili corpora, Kamano et al. demonstrate how the position of TU affects the meaning of a sentence. Their arguments are (i) TU means 'just do it (just V O)' if attached to VP, and (ii) TU means 'do it only (V O only)' if attached to DP. 3 and 4 are a minimal pair that are contrastive in the position of TU.

(3) Example of 'just V O' – V *tu* O
Nipe tu chakula hicho - pumba na kibuzi.
give TU food this
'Just give me this food - some chaff and a small goat.' (Kamano et al. 2022)

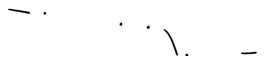
* This proceeding is based on a multi-author presentation at the 7th Asian Junior Linguists Conference. I am deeply indebted to Yuko Abe, Kumiko Miyazaki, and Seunghun J. Lee, who gave me insightful advice and extensive support as co-authors of the presentation. I would also like to show my gratitude to organizers, chairpersons, and editors of the AJL 7. This work is supported by KAKENHI (C) 20K00578 to Seunghun J. Lee (PI) and JICUF grant [Spring 2022-2023 | 77] to Shigeto Kamano.

(4) Example of 'V O only' – V O *tu*
Nipe chakula hicho tu - pumba na kibuzi.
give food this TU
'Give me this food only - some chaff and a small goat.' (Kamano et al. 2022)

When TU is located immediately after a verb, the domain of TU is VP in 3. On the other hand, TU modifies a noun and is in the sentence-final position in 4. When going back to 1 and 2, TU appears in the sentence-final position in both cases. Since an intransitive verb is used in 1, the position of TU exhibits no apparent difference among these examples. Kamano et al. (2022), however, make use of a transitive verb to show the domain difference with regard to the position of TU.

Previous studies other than Kamano et al. (2022) have not mentioned the prosodic prominence of TU. Ashton (1944) states that the sentence-final TU has a falling intonation pattern. Dots and lines in 5 represent the pitch transition of 1.

(5) Pitch transition of 1 (Ashton 1944: 75)



Watu wamelela tu.

Ashton suggests that TU is produced with no prosodic prominence, as shown in 4. This description of the pitch of TU is in accord with Krifka (1998), who reports that TU cannot be stressed. However, native speakers of Swahili anecdotally describe the intonation as different, which is also the case with the instructions given in Swahili classes. Also, the descriptions in Ashton (1944) or Krifka (1998) include only the sentence-final TU but not the sentence-medial TU, the existence of which is reported by the latest corpus study (Kamano et al. 2022).

Unlike those previous studies, Kamano et al. (2022) suggest that TU is a prosodically prominent particle. Recording one Swahili speaker, Kamano et al. demonstrate that TU is produced with a high pitch regardless of the sentence length: even when the sentence is composed of multiple intonational phrases, the sentence-final TU is produced with a high pitch (see Figure 1).

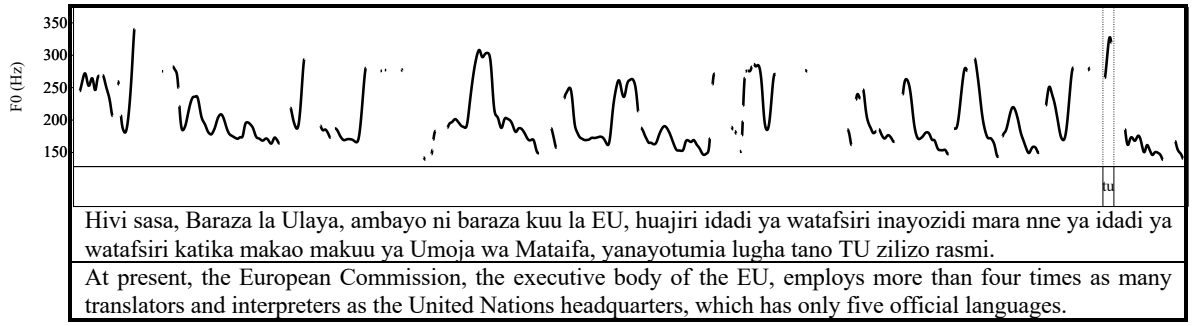


Figure 1: Example of the prosodically prominent TU (Kamano et al. 2022)

In addition to the sentence-final TU, Kamano et al. (2022) report that sentence-medial TU also has prosodic prominence. This current paper follows the observation of TU in Kamano et al. but later points out an alternative view for intonation patterns of the phrases preceding/following TU.

3 Pitch compression analysis

On top of the prosodic prominence of TU, it is reported that TU compresses the pitch of sentences around the particle. According to Kamano et al. (2022), pitch compression exists in phrases that precede or follow TU and corresponds with the semantic domain of TU. When TU appears immediately after a verb, *i.e.*, is attached to the preceding VP, what is compressed is DP after TU (see Figure 2).

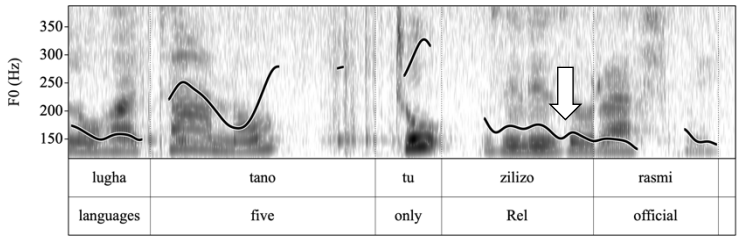


Figure 2: Pitch compression of the phrase *zilizo rasmi* 'that are official' following TU (Kamano et al. 2022)

On the other hand, TU causes pitch compression of DP before TU when TU is located in the sentence-final position (see Figure 3). 6 summarizes their arguments.

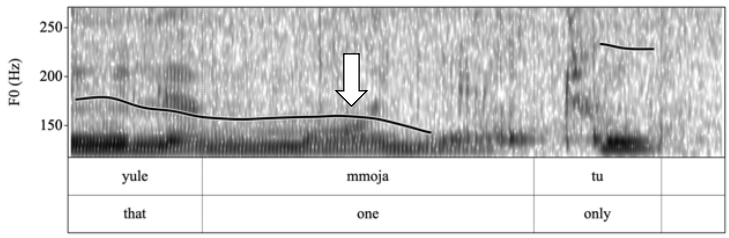


Figure 3: Pitch compression of the phrase *yule mmoja* 'the one' preceding TU (Kamano et al. 2022)

(6) Pitch compression caused by TU (Kamano et al. 2022)
   a. Sentence-medial TU (V *tu* O)
      - Meaning: 'just V O'
      - Semantic domain: VP ($_{VP}$[V *tu* O])
      - Pitch compression: DP after TU
   b. Sentence-final TU (V O *tu*)
      - Meaning: 'V O only'
      - Semantic domain: DP (V $_{DP}$[O] *tu*)
      - Pitch compression: DP before TU

As summarized in 6 above, recordings of one native speaker demonstrate that pitch compression is caused by TU, and the semantic domain of TU corresponds to the range of pitch compression.

The pitch compression analysis of Swahili is reasonable in that it is compatible with post-focal compression in Japanese. In Japanese, focused elements, such as Wh-elements, narrow the pitch range in the post-focal area (Ishihara 2015). Since in Kamano et al. (2022), the domain of TU is marked not only by the syntactic position of TU but also by the prosodic pattern, their analysis corresponds to the post-focal pitch compression in Japanese[1].

However, more evidence is required to support the pitch compression analysis of Swahili. First, as Kamano et al. (2022) only record one native speaker, their analysis does not mention the possibilities of inter-speakers variations. Also, they do not compare the pitch excursion of TU sentences with the pitch excursion of non-TU sentences. Without a comparison of these sentences, it is not clear that the pitch compression of the phrases adjacent to TU really stems from the domain of TU. Therefore, the current study aims to confirm what is a factor of pitch compression by recording more speakers and using contrastive sentences for stimuli, which are introduced in the following sections.

4 Methods

4.1 *Stimuli and speakers* We consider three types of sentences for stimuli: No-TU, Non-final-TU, and Final-TU. The base structure of all stimuli sentences is S + V + O. No-TU-conditional sentences exclude the TU particle (see 7a). The Non-final-TU sentences follow the S + V + *tu* + O structure (see 7b), and the Final-TU sentences

[1] The latest discussion of the post-focal compression in Japanese appears in Kawahara et al. (2022).

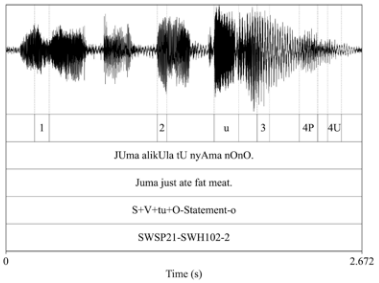
are composed of S + V + O + *tu* (see 7c).

(7) Examples of the stimuli
a. No-TU condition with final [o] vowel
Juma a-li-kula nyama nono.
Juma 3SG-PST-eat meat fat
'Juma ate fat meat.'
b. Non-final-TU condition with final [o] vowel
Juma alikula nyama nono tu.
'Juma just ate fat meat.'
c. Final-TU condition with final [o] vowel
Juma alikula nyama nono tu.
'Juma ate only fat meat.'

Each condition has five types of object DP that consist of a noun and an adjectival modifier, which differ from one another in the vowel quality of the sentence-final syllable. Overall, 15 items are created: 3 TU-positional conditions x 5 object DP types.

Six native speakers of Zanzibar Swahili were asked to read the 15 items twice. The age of participants was 23 to 49 at the time of recording. All recordings were conducted in a quiet room in Zanzibar, Tanzania. Audacity on a MacBook computer set was used at a 44.1 kHz sampling rate and 16-bit depth.

4.2 *Annotation and analysis* Praat (Boersma & Weenink 2018) is used to process and annotate the recordings. The vowel of the target TU and vowels of the penultimate syllable of each word are annotated as in Figure 4.



Juma alikula tu nyama nono.
'Juma just ate fat meet'
*The annotated vowels are capitalized in the figure.

Figure 4: Example of the annotated sentences

For instance, the [u] vowel in the penultimate syllable of the second word *alikula* 'he/she ate' is marked with 2. Only penultimate syllables are annotated because Swahili speakers put an accent on the penultimate syllables, and the accents are realized as a high pitch[2]. The maximum F0 from each annotated interval is extracted using the Praat algorithm. Then R (R Core Team 2020) is used to process these acoustic data.

5 Results

In our recordings, the prosodic prominence of TU is observed in most of the tokens (109/117)[3]. The pitch-lowering pattern, however, turns out that it occurs in No-TU sentences and varies according to the position of TU. I will compare intonation patterns between No-TU sentences and Non-final-TU sentences in Subsection 5.1 and between No-TU sentences and Final-TU sentences in Subsection 5.2. The results call for alternatives to the pitch compression analysis in Kamano et al. (2022).

5.1 *No-TU vs. Non-final-TU* First, let us compare Non-final-TU sentences with No-TU sentences. The pitch-lowering pattern of Non-Final-TU sentences is observed but on par with No-TU sentences. Figure 5 demonstrates high-pitch transitions in one of the recordings of SWH104[4].

[2] The fourth word is further annotated into 4P that stands for the penultimate syllable, and 4U for the ultimate (final) syllable in order to compare the prosody between declaratives and interrogatives. This is not the focus of the current paper.

[3] All the data of TU without the pitch prominence belong to one speaker.

[4] SWH*XXX* stands for speaker IDs (from 101 to 106).

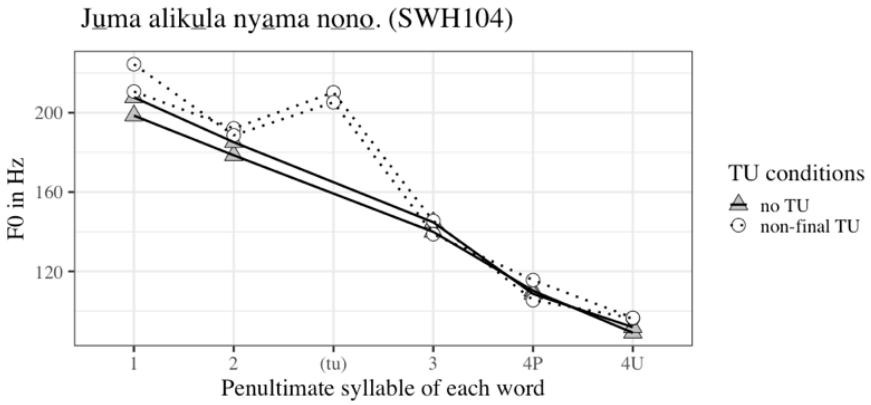


Figure 5: High-pitch transition of SWH104 (No-TU vs. Non-final-TU)
*4P = penultimate syllable of the fourth word, 4U = ultimate syllable of the fourth word

In Figure 5, numbers on the x-axis represent the penultimate syllables of each word, and the maximum pitch is plotted on the y-axis. Solid lines and shaded triangles stand for the high-pitch transitions of No-TU sentences, and dotted lines and white dots for Non-Final TU sentences. When focusing on Non-final-TU sentences, we can observe the prosodic prominence of TU and the pitch compression of the fourth word. However, as shown in Figure 5, the pitch-lowering patterns are also true of No-TU sentences.

5.2 *No-TU vs. Final-TU* In this subsection, No-TU sentences are compared with Final-TU sentences. The intonation pattern before the final TU varies more greatly among speakers. For example, some speakers produce the preceding word of the final TU with a higher pitch (see Figure 6). In the following plots, solid lines and shaded triangles stand for the high-pitch transitions of No-TU sentences, and dotted lines and white dots for Final TU sentences.

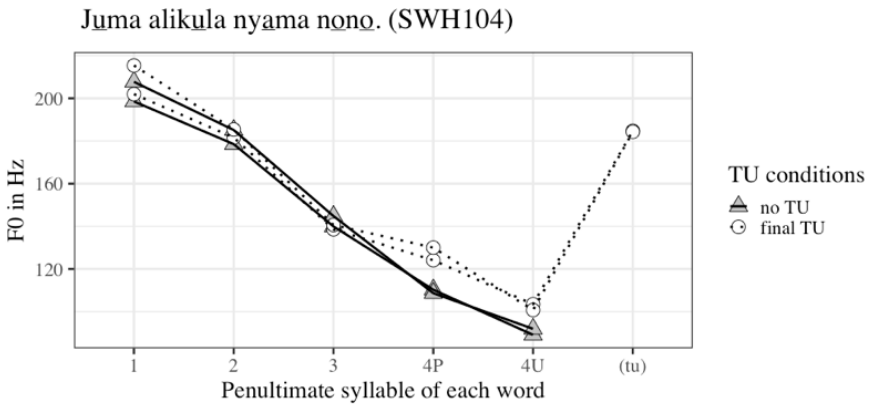


Figure 6: High-pitch transition of SWH104 (No-TU vs. Final-TU)
*4P = penultimate syllable of the fourth word, 4U = ultimate syllable of the fourth word

In the fourth word, the white dots for Final-TU sentences are plotted above the shaded triangles for No-TU sentences. Three speakers exhibit this raised pitch in the preceding word of the final TU. On the other hand, as shown in Figure 7, other speakers do not show any significant difference in prosody between No-TU sentences and Non-final-TU sentences.

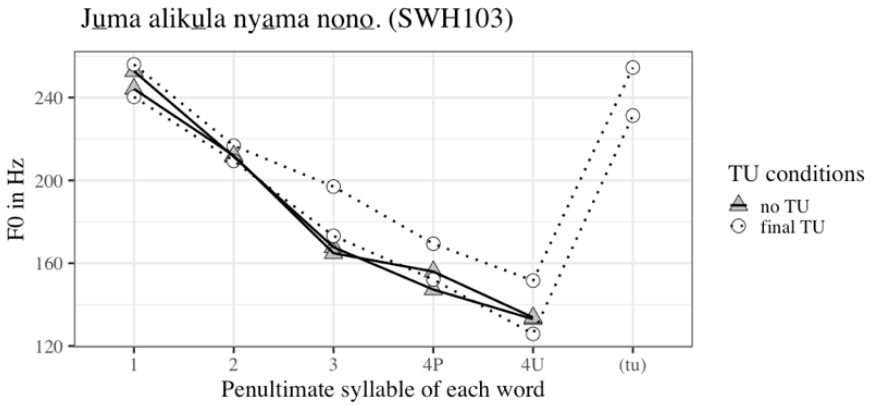


Figure 7: High-pitch transition of SWH103 (No TU vs. Final TU)
*4P = penultimate syllable of the fourth word, 4U = ultimate syllable of the fourth word

For example, in the utterance of SWH103, the high pitch of the fourth word in No-TU sentences (dots) and Final-TU sentences (triangles) mark similar values. Three speakers use this intonation pattern.

6 Downstep analysis

In this section, I propose an alternative analysis of the intonation pattern in Swahili. The pitch-lowering pattern of Non-final-TU sentences is true of No-TU sentences. In No-TU sentences, the intonation pattern has nothing to do with the pitch compression of TU since they do not contain TU. Therefore, it is better to say that the high pitch of penultimate syllables is influenced by downsteps that are independent of TU; downsteps are neither generated nor interrupted by the existence of TU. Also, when comparing Final-TU sentences with No-TU sentences, the inter-speaker variation is observed in the pitch of the preceding word. This variation suggests that the degree of downstep is not fixed nor obligatory. In the case that speakers produce the fourth word even with a higher pitch in Final-TU sentences, the semantic domain of TU is no longer reflected in the prosodic domain, which opposes the pitch compression suggested in Kamano et al. (2022).

One thing to note is that downstep here is defined as "the lowering of a high tone in certain specifiable circumstances" (Connell 2001). It would be different from downstep in Bantu tonal languages because it is said to be influenced by a low tone or floating low tone (Van de Velde et al., 2019), but Swahili is an accentual language. The pitch-lowering pattern in Swahili might be just a phonetic effect called declination. As Figure 8 shows, the range of a high pitch in each penultimate syllable certainly decreases from the beginning to the end of a sentence.

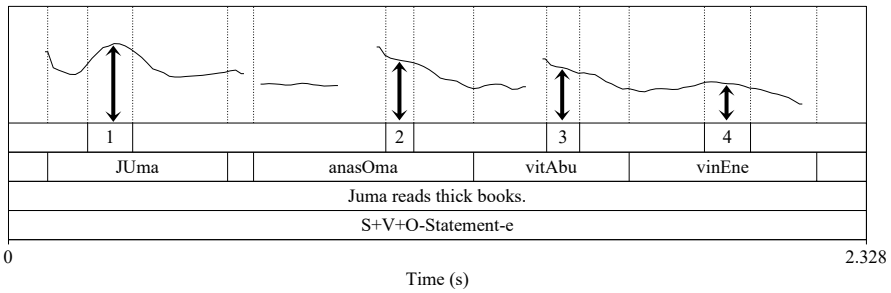


Figure 8: Example of pitch lowering patterns in No-TU sentences (SWH101)

Since all the word is accented in stimuli sentences, it is unclear whether the cause of the pitch lowering is an interaction between high tones or a phonetic condition for creating high tones: downstep applies to the former case and declination to the latter case.

7 Conclusion

This paper has discussed the prosodic patterns in Swahili and proposed an alternative way to analyze them. The prosodically prominent particle TU appears in (i) the sentence-final position to modify a DP or (ii)

immediately after a verb to modify a VP. A previous study reports that TU compresses the pitch of its adjacent phrases, which correspond to the semantic domain of TU. The prosodic prominence of TU, which Kamano et al. (2022) report, is observed in the current study. However, this study reveals that the pitch-lowering pattern appears in sentences without TU and exhibits inter-speaker variations in sentences with a sentence-medial TU. These results suggest that pitch compression in Kamano et al. (2022) can be subsumed under a downstep analysis. For further discussion, it is necessary to focus on the phonetic interaction between downstep and the prosodic prominence TU, and the phonological analysis of prosodic structures TU should attach to.

8 Appendix

Here are our stimuli for the recordings. Six participants were asked to read each sentence twice. The appendix only includes IDs that are concerned in this paper.

ID	Syntax	Swahili	English
SWSP001	S + V + O	Juma alimfuata watoto wanane.	Juma followed eight children.
SWSP002	S + V + O + *tu*	Juma alimfuata watoto wanane tu.	Juma followed only eight children.
SWSP003	S + V + *tu* + O	Juma alimfuata tu watoto wanane.	Juma just followed eight children.
SWSP007	S + V + O	Juma anasoma vitabu vinene.	Juma reads thick books.
SWSP008	S + V + O + *tu*	Juma anasoma vitabu vinene tu.	Juma reads only thick books.
SWSP009	S + V + *tu* + O	Juma anasoma tu vitabu vinene.	Juma just reads thick books.
SWSP013	S + V + O	Juma anapenda shepu ya kifuniko.	Juma loves the shape of the cap.
SWSP014	S + V + O + *tu*	Juma anapenda shepu ya kifuniko tu.	Juma loves only the shape of the cap.
SWSP015	S + V + *tu* + O	Juma anapenda tu shepu ya kifuniko.	Juma just loves the shape of the cap.
SWSP019	S + V + O	Juma alikula nyama nono.	Juma ate fat meat.
SWSP020	S + V + O + *tu*	Juma alikula nyama nono tu.	Juma ate only fat meat.
SWSP021	S + V + *tu* + O	Juma alikula tu nyama nono.	Juma just ate fat meat.
SWSP025	S + V + O	Juma alitafuta kiwango cha nuru.	Juma checked the level of light.
SWSP026	S + V + O + *tu*	Juma alitafuta kiwango cha nuru tu.	Juma checked only the level of light.
SWSP027	S + V + *tu* + O	Juma alitafuta tu kiwango cha nuru.	Juma just checked the level of light.

References

Ashton, Ethel O. 1944. *Swahili Grammar (including intonation)*. Longmans.

Boersma, P. & Weenink, D. 2018. *Praat: doing phonetics by computer*. Version 6.0.29. Retrieved May 29, 2020. Available from www.praat.org.

Ishihara, Shinichiro. 2015. Syntax–phonology interface. *Handbook of Japanese phonetics and phonology*, 569–618. De Gruyter Mouton.

Kamano, Shigeto, Yuko Abe, Kumiko Miyazaki & Seunghun J. Lee. 2022. Prosodically prominent clitic: the exclusive particle tu in Swahili. Paper presented at the Prosody and Grammar Festa 6, Online.

Kawahara, Shigeto, Jason A. Shaw & Shinichiro Ishihara. 2022. Assessing the prosodic licensing of wh-in-situ in Japanese: A computational-experimental approach. *Natural Language & Linguistic Theory* 40(1), 103–122. doi:10.1007/s11049-021-09504-3. https://link.springer.com/10.1007/s11049-021-09504-3 (February 13, 2023).

Krifka, Manfred. 1998. Additive Particles under Stress. *SALT VIII*, 111–129.

Meeussen, Achille Emile. 1967. Bantu grammatical reconstructions. *Africana linguistica* 3(1). Persée-Portail des revues scientifiques en SHS, 79–121.

R Core Team. 2020. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.