Serial verb constructions and their event representations in Akan

Laura Domine

June 2019

Abstract

Serial verb constructions (SVC) are often assumed to describe a single event in the speaker's mind. The event representations of SVC have been investigated using various methods in psycholinguistics. We focus on a video segmentation task and attempt to replicate the experiment conducted by Cole (2016). It originally involved Lao speakers but we chose to work on the Akan language, one of the historically first language where SVC were diagnosed. The original results showing that there was less segmentation in regions described with an SVC as opposed to a coordinate construction (CC) could not be reproduced and we discuss several potential reasons for this.

1 Serial verbs

According to Aikhenvald (2018) serial verb constructions (SVC) appear in more than two thirds of the world languages. However, because they do not appear in most Indo-European languages, they only started to be recognized by linguists as a separate linguistic category in the second half of the nineteenth century. A German missionary Christaller (1875) who wrote one of the first grammar of Twi (Asante and Fante dialects) described it this way: *"many verbal notions that are expressed with a simple verb in English and other European and Asiatic languages, are expressed by syntactical combinations of verbs."* The expression "serial verb" appeared first in a grammar of Fante Akan by Balmer and Grant in 1929.

A possible definition of serial verbs is outlined by Aikhenvald (2018) and we only remind here some of the fundamentals properties usually assigned to SVC: a serial verb need to involve two or more verbs, each of which could be the sole verb of a clause. There should not be any marker of coordination, subordination or any other sort of dependency. It should behave as a single predicate: for example TAM marking will be shared among the verbs of the SVC. The SVC has its own transitivity which may depend on the components verbs and the type of SVC. In addition they share at least one core argument such as syntactic subject, object or obliques. The last property evoked by Aikhenvald (2018) relates directly to the topic of this paper: *"The serial verb construction is conceived as describing a single event."* Notably the definition of "event" is not straightforward: *"The event described by a serial verb construction may be simple or complex. It may cover a set of interconnected activities, or states."* Some of these properties have noted exceptions. The definitions of SVC have been multiple and are still an object of debate.

2 Event representations of SVC

Usually the single eventhood of SVC is examined using intuition, translation or cultural restrictions on SVC as pointed out by Defina (2016). However a few attempts at investigating specifically the event representations of SVC can be outlined.

The first is the pioneering work of Givón (1991) who uses prosody on three serializing languages of Papua New Guinea. The assumption is that pauses in speech correspond to event boundaries. It was found that pauses within SVC are no more likely to

occur than within a single lexical item. In addition they were less frequent than pauses between clauses. According to the authors this points to a single event representation of SVC.

On the other hand, Baker et al. (2010) studied crosslinguistic correlations of semantic structures and meanings of SVC. They assume that individual lexical verb can be identified with single events. They find that the meaning of SVC expression was closer to that of complex clauses than single verbs, coming to the opposite conclusion than that of Givón (1991).

Finally Defina (2016) studied co-speech gestures in Avatime, a Kwa language from Ghana. The experiment showed that SVC tended to be accompanied with single gestures, compare to complex clauses with would go along with multiple gestures, thus supporting the hypothesis of single event representation of SVC.

Video segmentation tasks in psycholinguistics were pioneered by works such as that of Zacks et al. (2001). Here we try to replicate the experiment conducted recently by Cole (2016). Lao speakers were asked to divide videos into events. The assumption is that grammatical constructions such as serial verbs could influence the video segmentation behavior of speakers. In this experiment, it was found that if they were shown a SVC description of the video beforehand, they divided the video into fewer events than if they were shown a coordinated construction. This evidence supports the claims of Givón (1991) and Defina (2016) regarding the single eventhood of SVC.

Following Cole (2016) I decided to investigate the eventhood of SVC using video segmentation behaviors on the Akan language, which is a Kwa language from the Niger-Congo family. There are several dialects in Akan, mainly Asante Twi, Fante Twi and Akwapim Twi. Most of the participants spoke the Asante Twi dialect which was selected for the experiment material. Just like Cole (2016), we assumed a single event representations of SVC. As a consequence, our behavioral prediction was that participants primed with an SVC description of a video (as opposed to CC) would divide it into fewer events.

3 Experiment

3.1 Methods

Participants were recruited from a 3rd and a 4th year class at Kwame Nkrumah University Of Science and Technology (Kumasi, Ghana), thanks to Samuel Nkansah (Twi language instructor at Stanford) and Juliet Oppong-Asare (professor in Ghana). A total of 58 native Akan speakers were recruited.

The experiment materials were made of 5 videos kindly shared by Douglas Cole in order to replicate Cole (2016) as closely as possible. Table 1 describes the content of each video and table 2 reports the SVC and CC primes that were used for each of them. Table 3 reports the filler questions that accompanied each video. Table 4 reminds the definition of regions of interest (ROI) in each of the video. All of them are unmodified or only translated into Akan from the experiment materials of Cole (2016).

The experiment took place remotely. The link to the online hosted experiment was shared with the participants. They were compensated for their participation as well as for the cost of taking part into the experiment with Airtime (sent through the Xoom service).

The experiment materials, anonymized data and analysis scripts can be found on Github. The experiment preregistration is available on OSF.

It should be noted that this procedure differs in a few important ways from the original one:

- Considerably less participants were recruited (Cole (2016) had 105 subjects).
- The experiment was run remotely (as opposed to in person).
- I didn't enforce a threshold (minimum number of segmentation marks made by the subject) in the practice video.
- Although the videos are the same as in the original experiment, they are culturally less adapted to the participants, which might have added some confusion.

Regarding the last point, this clearly stands out as a weakness of this reproduction. For example Aikhenvald (2018) highlights that *"close and culturally appropriate association of events (and not just 'single eventhood') is crucial for combinability of verbs within a sequential symmetrical serial verb constructions in Alamblak"* as well as in other languages such as White Hmong. Future work could include making new videos whose content would be more appropriate to the cultural context in Ghana.

The experiment of Cole (2016) had a "between-subjects" design. Participants were assigned randomly to three different groups:

- Control group (does not see any prime).
- Experimental group 1 (videos 1-3-5 with SVC prime, videos 2-4 with CC prime)
- Experimental group 2 (videos 1-3-5 with CC prime, videos 2-4 with SVC prime)

The whole steps to go through the experiment were the following:

1. Legal notice and participant registration
2. Random assignment to one of the three groups
3. Practice video (90s)
4. For each of the videos:
   (a) Play the video and ask participants to segment it using either the space bar (if they were on a computer) or tapping on the video (if they were on a phone). They are not allowed to replay it.
   (b) Filler: two yes/no questions that were related to the videos and made sure the participants paid some attention to the videos.
5. For each of the videos;
   (a) Play ROI of the video. They are allowed to replay it.
   (b) Ask for the participant's preference between SVC and CC descriptions of the ROI.
6. Collect some demographic informations and ask about the global experience of the participants during the experiment.

3.2 Results

In all the figures shown in this section, the error bars are statistical.

Exclusion criteria When counting participants we write $X + Y + Z$ as a shortcut for X participants in the control group, Y in the experimental 1 group and Z in the experimental 2 group. Among the 58 (19+16+23) total participants, there were 29 (11+5+13) participants left after applying the first exclusion criterion and 21 (7+3+11) after the second criterion. The exclusion criteria were the following:

- Never segmented any video (29 participants)
- Never segmented at least once in a ROI (8 participants)

Overall clicking behavior Figure 1 shows the distribution of the count of participants clicks in a video. The wide distribution shows the need to control for clicking behavior when we analyze the data.

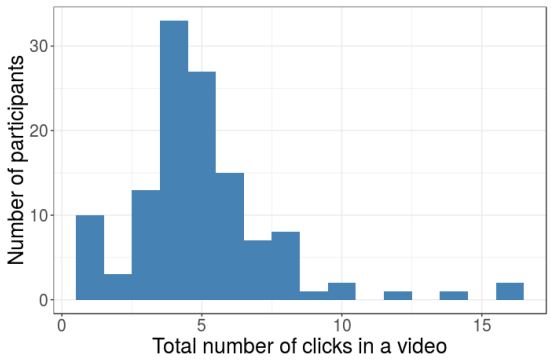

Figure 1: Distribution of a participant's total number of clicks in a video.

Figure 2 shows how many clicks were inside a ROI as opposed to outside. It also distinguishes between the three different groups (control, experimental 1 and 2).

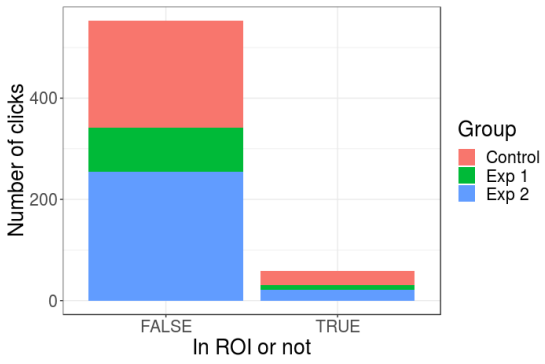

Figure 2: Total number of clicks in and outside of the ROIs, depending on the group (control or experimental 1, 2).

SVC vs CC One can compare the count of clicks inside a ROI per participant and per video for SVC and CC primes (see Figure 3). Due to poor statistics and overall low pressing behavior most of them lie at 1 click in the ROI.

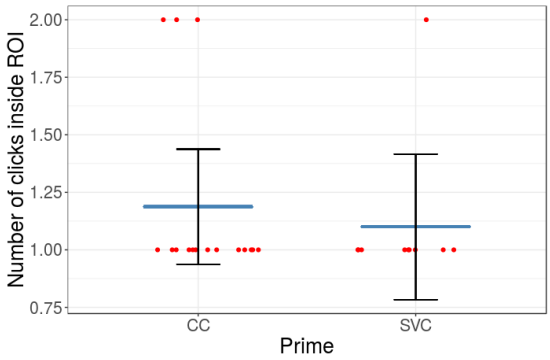

Figure 3: Selecting only the experimental groups 1 and 2, we look at the number of clicks inside the ROI per participant, per video. Exclusion criteria were applied, and result in the low statistics. A horizontal jitter was applied to the points in order to disambiguate them.

The mean for SVC primes ($m = 1.1$) is slightly lower than for CC ($m = 1.1875$), however the error bars do not allow us to conclude. Since we are dealing with count data, a Poisson regression is appropriate to analyze the relationship between the clicks inside a ROI and the type of priming. However the small number of participants left after the exclusion criterions makes it difficult for the model to converge and we ran into singularity issues when trying to include random effects into account for the subjects and video variability.

Influence of priming Showing a prime, regardless of its SVC or CC content, might affect the overall segmentation behavior. Figure 4 shows that at first sight the number of clicks in a video for participants in the control group seem to be higher compared to the experimental 1 and 2 groups. However accounting for the participants and video items variability as random effects in a Poisson regression of the click counts as a function of the priming shows that this is not a significant effect ($\beta = -0.2062$, $SE = 0.1355$, $z = -1.522$, $p < 0.128$).

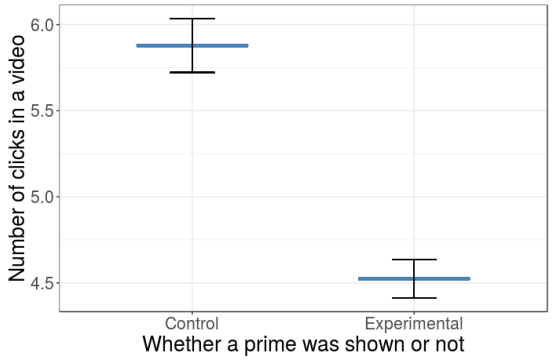

Figure 4: Whether the priming affected overall clicking behavior. Does not control for participants and video items variability yet.

Preferences 87% of participants preferences (across both videos and participants) are SVC, which confirms that SVC were legitimate descriptions of the material. An unexpected observation was that the time spent by participants on the page asking to decide between SVC and CC descriptions of the ROI video seem to be significantly correlated to the SVC/CC final choice (see Figure 5). Participants

seem to make their choice faster when they decide upon a SVC description. A mixed effects logistic regression model accounting for the participants and video items variability as random effects is able to predict the duration of the choice from the preference ($\beta = -1.6536$, $SE = 0.4384$, $z = -3.772$, $p < 0.00016$).

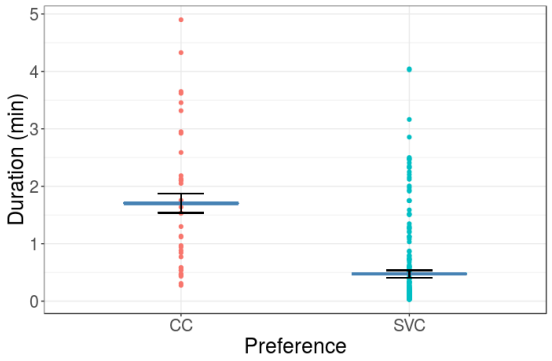

Figure 5: Duration of the choice between SVC and CC preference

4 Conclusion

In order to determine whether SVC are represented as a single event, we attempted to replicate the video segmentation experiment of Cole (2016) on a different language, namely the Akan language. We made every effort to follow it as closely as possible but some caveats were mentioned in section 3.1. The main drawbacks of this replication are the potentially inappropriate cultural context of the videos, and the remote operation of the experiment which has two noticeable consequences: first it makes it difficult to control directly for the participants' understanding of the instructions, and second it introduces uncertainty mainly because of internet speed in Ghana.

The results are so far inconclusive. As of the current data that was collected to date, there was no significant difference in segmentation behavior between SVC and CC primes. This contradicts the results of Cole (2016). We found that the overall clicking behavior does not seem to be impacted by the priming, in agreement this time with the original experiment. Finally the preferences of participants were in majority SVC descriptions of the videos, as expected. The time spent on the preference question seem to be bigger when the participants chose an SVC description rather than CC, which remains to be explained. It is dubious whether this can be directly related to the event representation properties of SVC.

To improve the quality of these results, first of all more statistics would be needed. Solving the previously mentioned caveats would also certainly help in improving the data quality. However this replication attempt might be considered as a promising pilot experiment.

5 Acknowledgements

This work could not have been possible without the generous help of Samuel Nkansah, Twi instructor at Stanford University, both to translate the experiment materials into Twi and to find partners in Ghana. Dr. Juliet Oppong-Asare, professor at the Kwame Nkrumah University Of Science and Technology in Kumasi, Ghana, was a great and crucial help in recruiting the participants of the experiment. I want to thank both of them here for their time and commitment.

References

Aikhenvald, A. Y. (2018). *Serial verbs*. Oxford Studies in Typology and.

Baker, B., Harvey, M., et al. (2010). Complex predicate formation. *Complex predicates: Cross-linguistic perspectives on event structure*, pages 13–47.

Christaller, J. G. (1875). *A grammar of the Asante and Fante language called Tshi Chwee, Twi based on the Akuapem dialect with reference to the other (Akan and Fante) dialects*. Basel evang. missionary society.

Cole, D. J. (2016). Lao serial verb constructions and their event representations.

Defina, R. (2016). Do serial verb constructions describe single events?: A study of co-speech gestures in avatime. *Language*, 92(4):890–910.

Givón, T. (1991). Serial verbs and the mental reality of 'event': Grammatical vs. cognitive packaging. *Approaches to grammaticalization*, 1:81–127.

Zacks, J. M., Tversky, B., and Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of experimental psychology: General*, 130(1):29.

Video id	Target sequence	Description	Snapshot
0	Cook rice eat	A man washes vegetables at a sink and cuts them up. Then he mashes peppers in a spicy sauce. Next he rinses the sticky rice at the sink and puts the cooker on the stove. When the rice is done cooking he moves it to a smaller basket and sits at the table to eat with his family. Finally, he washes dishes and sweeps the floor.	
1	buy motorbike ride	A man hails a tuk-tuk (similar to taxi) and rides to a bank. He withdraws some money from a teller and rides the tuk-tuk to a motorbike shop. After talking with the salesman, he hands over the money and drives away. He then arrives at his relatives' home to show them his new motorbike.	
2	wash clothes hang up	A woman gathers some clothes and puts them in a basket. Next she fills a large tub with water from a hose. After that she washes the clothes by hand, wrings them out and hangs them on a line. Lastly, she irons the clothes and puts them away in drawers.	
3	grill fish sell	A man stands in a river throwing a fishing net. Then he cleans the fish, adds salt, and stuffs it with herbs. Next he puts the fish on a grill, while a customer walks up. The customer hands him money, the man puts the fish in a banana leaf and hands it to the customer. Finally, he mends his fishing nets.	
4	borrow hat wear	A farmer plows his field with a small walk behind tractor. He wipes his forehead and rests in a small hut. A friend walks up wearing a rice hat, and the man points to the hat, indicating he wants to borrow it. The friend gives him the hat, he puts it on, and continues plowing the field. The man finishes plowing and walks up the steps of his house.	

Table 1: List of videos descriptions, from Cole (2016).

Video id	Verbs involved	SVC description	CC description
0	Cook / Eat	Papa no noa ɛmoo di.	Papa no noa ɛmoo na wadi.
1	Buy / ride	Papa no tɔɔ sakri no twiieɛ.	Papa no tɔ ɔsakri no ɛna ɔtwiieɛ.
2	Wash / hangup	Maame no sii nneɛma no hataeɛ.	Maame no sii nneɛma no na ɔhataeɛ.
3	Grill/sell	Papa no toto nsuomu nam tɔn.	Papa no toto nsuomu nam a ɔtɔn.
4	Borrow/wear	Papa no kɔserɛ ɛkyɛ hyɛ.	Papa no kɔserɛ ɛkyɛ na wahyɛ.

Table 2: Primes used in the experiment

Video id	Target sequence	Question 1	Question 2
0	cook rice eat	What color was the rice cooking basket? 1. black 2. white	How many people ate together? 1. two 2. three
1	buy motorbike ride	What did the man ride to the bank? 1. taxi 2. tuk-tuk	How many people worked at the bank? 1. one 2. three
2	wash clothes hang up	What color was the washing bucket? 1. black 2. white	How many items did the woman wash? 1. one 2. three
3	grill fish sell	What was the man wearing? 1. rice hat 2. baseball cap	How many fish did the man sell? 1. one 2. two
4	borrow rice hat wear	What did the man plow the field with? 1. tractor 2. buffalo	What was the weather in the video? 1. hot 2. cold

Table 3: Filler questions, from Cole (2016).

Video id	Target sequence	ROI (ms)	Total (ms)
0	cook rice eat	23133-27133 and 38221-42221	8000
1	buy motorbike ride	35334-43053	7719
2	wash clothes hang up	25429-33378	7949
3	grill fish sell	33219-39351	6132
4	borrow rice hat wear	20346-26464	6118

Table 4: Regions of interest (ROI) in each video, as in Cole (2016).