Open Review

( ) I would not like to sign my review report

(x) I would like to sign my review report

English language and style

( ) Extensive editing of English language and style required

(x) Moderate English changes required

( ) English language and style are fine/minor spell check required

( ) I don't feel qualified to judge about the English language and style

	Yes	Can be improved	Must be improved	Not applicable
Does the introduction provide sufficient background and include all relevant references?	(x)	( )	( )	( )
Are all the cited references relevant to the research?	(x)	( )	( )	( )
Is the research design appropriate?	( )	(x)	( )	( )
Are the methods adequately described?	(x)	( )	( )	( )
Are the results clearly presented?	(x)	( )	( )	( )
Are the conclusions supported by the results?	( )	(x)	( )	( )

Comments and Suggestions for Authors

In the abstract, there is an error in the syntax of the following sentence:
Many measures of these quantities, either rely on a statistical notion of information, are difficult to compute, or can only be applied to strings.
the alternative "either ... or" is complicated by the insertion of a superfluous colon : "are difficult to compute"

Reply: Thank the reviewer very much for this comment. We intended to mean that “are either difficult to compute, rely on the statistical notion of information, or can only be applied to strings”. We have revised this sentence accordingly. We also revised some grammar errors and typos here and there.

The more I think about this article, the more I think that it grasps too much. The sections 4.3. and 4.4. dedicated to the complexity of life seem superfluous. I would suggest renouncing altogether the

parts that delve into the complexity of life and to recenter the whole structure of the article to computational linguistics.

Reply: We highly appreciate this comment and suggestion from the reviewer.

Surely Sections 4.3 and 4.4 are not full-length discussions, but our main point there is to raise the question on origins of life and "why life is ordered" and mention the potential connections with ladderpath. We hope to convey the idea that ladderpath can provide a different angle when investigating this type of questions (this intention was also stated in the main text there). Besides, these two sections are also intended to link together our previous works on chemical space and origins of life (Yu Liu et al, Exploring and mapping chemical space with molecular assembly trees. *Science Advances* 2021, 7, eabj2465), which relates to this ladderpath approach.

On the other hand, although most of the examples we provided to illustrate our approach are strings or linguistic examples (note that 2D patterns were also used), the only reason is that strings are the simplest cases, and we think using the simplest examples to explain a new approach step by step would be a good choice. The main point of our manuscript is to introduce this new approach, rather than applying it to a particualr field or entity (as it can be applied to many types of entities), although computational linguistics would be a super relevant and interesting field.

After seriously considering the reviewer's suggestion, we do think it's worthwhile to mention the conncetions and various potential applications, rather than purely foucsing on computational linguistics. Indeed, much more future work is needed (actually we are moving forward based on this work) to answer particular questions. Please also allow us to quote Reviewer #1's comments on this regard: "I find this paper very interesting, and I find the final discussion about the origin of life based on the remark that repetition makes easier the invention of life very attractive." We hope this could help ease the reviewer's concern about "grasps too much".

Thanks again for your time.