Speakers must coordinate the processing of ideas, words, and movements over time, but they have flexibility in how they do so. They can start utterances with a single phonological word prepared or after preparing and buffering multiple words [1]. Two experiments demonstrate that speakers can use a correlate of word length to estimate the amount of time they will have available for preparing later words during speech. They can thereby minimize word preparation & buffering prior to speaking while maintaining a high level of fluency.

When speakers describe scenes, they tend to gaze at referents during the second before uttering their names. These gazes reflect the time needed to select names and retrieve phonological forms [2].

Speakers take almost a second to prepare an object’s name, but articulating it takes less time. So, in naming two objects, speakers must either pause between names or buffer the first name while preparing the second name.

Speakers can prepare long names as quickly as short names [3], but longer names take more time to say, of course. Longer names provide speakers with more time to prepare the following words. If speakers use word length information in timing their speech, they should prepare second names before saying long names compared to short names. Therefore, speakers may start saying long first names earlier than short ones.

Alternatively, if long names take more time to prepare than short ones, speakers should start saying long first names later than short ones.

Finally, if speakers entirely prepare both names before speaking, they should start saying long and short first names after similar latencies (assuming the names take similar amounts of time to prepare).

32 pairs of highly codable objects with mono- and multi-syllabic names were matched for word frequency and initial phonemes. Second objects were counterbalanced.

First object name   | List 1  | List 2
---|---|---
Short   | “scarf” |  
Long    | “skeleton” |