The Role of Numerical Grammar in Efficient Communication

Research in cognitive science consistently demonstrates evidence for communicative efficiency pressures in language evolution. In information theory, signal efficiency is measured in units of information (usually bits) transferred per unit of signal length. Information theorists argue that modern languages evolved to favor shorter, compressed messages that optimize for maximal information transfer (Gibson et al, 2019; Piantadosi et al, 2011).

Different languages have adapted distinct mechanisms for achieving this goal. Many european languages, for example, utilize gender casing and marking. Nouns, both animate and inanimate, are attributed to feminine, masculine, or neutral gender. While gender assignment seems somewhat arbitrary (*"water"* is feminine in Russian but masculine in Spanish), information theorists hypothesise that gender attribution facilitates more efficient signal encoding and decoding during communication. Consider a simple example phrase *"where is the book*?", contrasted with the spanish translation *"dónde* (where) *está* (is) *el* (the, masculine) *libro* (book)". The use of a masculine determiner (*el* vs. *la*) in Spanish provides clues to the listener about the nature of the noun that may follow. Put this way, the phrase *"dónde está el"* in Spanish successfully transmits more information than its Enlgish complement "where is the".

To support this theoretical paradigm, researchers have relied on eye tracking experiments to demonstrate the effect that gender marking, prenominal color adjectives, and contextual cues have on the speed of word prediction tasks in children and adults (Dahan et al, 2000; Lew-Williams & Fernald, 2007; Fernald et al, 2010; Sedivy et al, 1999). Generally, participants are instructed to listen to an audio recording of a sentence and identify the subject of the sentence out of several options as soon as they become aware of it. The results of these studies are in-line with the communicative efficiency hypothesis. French listeners, for example, were found to fixate on the correct target images quicker when exposed with gender marking clues in the form of gender determinants.

While modern English lacks mechanisms such as gender marking and case conjugation, recent work by Dye et al (2018) hypothesized that descriptive prenominal adjectives in english might serve a similar communication-efficiency goal. In the example they give, the prefix *"I would like a nice cold"* immediately primes the listener to expect nouns like *"beer"* or *"coke"* to follow (in contrast to nouns like *"car"* or *"blanket"*). Dye et al (2018) present theoretical grounding to back up their claims and carry out analysis of how inclusion of such adjectives decreases the overall entropy (or surprisal) of a sentence. Thus, they argue that similar to gender casing, adjectives can limit the scope of possible sentence continuations, decreasing cognitive load and increasing efficiency.

In this study, I propose to explore the role that numerical grammar plays in communication efficiency. English language conveniently differentiates between singular and plural objects via numerical indicators "is" and "are". In the simplest form, these indicators are used to differentiate between singular and multiple occurences of an item: "where *is* my pen" vs. "where *are* my books". More interestingly, some words assume plurality even when behaving as singulars: "where *are* my glasses", "where *are* the scissors". Following the format of previous studies, I aim to observe the effect that the presence of numerical indicators (*is*, *are*) has on the speed of word prediction tasks. Observed difference in fixation latency across the two groups will serve as empirical evidence in support of Dye et al's communication-efficiency theory for English speakers.

Method

Procedure. A fixed set of common objects (~10) will be selected to serve as the target set. Objects can be animate or inanimate, as long as they are easily represented and recognizable in image form. Participants will be tasked with listening to a sentences of the form

"There [is/are] [a/the] [object]"

and asked to select the *object* mentioned in the sentence out of two images presented on the screen. The images will appear on the screen at the onset of the audio signal. Participants will be instructed to select the correct image as quickly as possible and rewarded for correct selections and quick response time. Rewards will be monetary, with the exact amount determined based on the experimental budget.

Ideally, if we have access to an eye tracking experimental set up at Berkeley, we would follow the protocol described in previous work and monitor the participants' gaze as they listen to the audio prompt. The measurement will capture the time from the start of audio input until fixation on the image that is ultimately selected. We hypothesise that this latency period will be shorter for the treatment group, where subjects are able to use numerical cues to select the correct target before hearing the target word pronounced.

The difference in treatment and control groups will manifest in the image options that the participants are presented with for target selection. Both groups will be instructed to select one out of two images. In the control group, the plurality of objects in both images will coincide with the numerical indicator in the input audio. Concretely, the audio *"there is an elephant"* will be paired with (1) an image of <u>one</u> elephant and (2) and image of <u>one</u> car (not an elephant). In contrast, for treatment group, the second image would contain multiple cars, or another plural entity. Each participant will engage in an identical number of trials and the set of target words and paired true/distractor objects and images will be shared by the two groups. The order of trials will be shuffled for each participant. Assuming a large enough participant sample size, this will remove any potential perception biases caused by the order in which the target objects are presented.

In the absence of an eye tracking setup, we can modify the experimental design to measure the time it takes for participants to select (click) the target image instead. This measurement will not be as accurate as time to fixation, but could still yield interesting results. In this approach, we can consider making the recognition task more challenging by pairing objects that share the first phoneme (e.g. "*ch*air", and "*ch*ocolate") for each trial. This should force the participants to listen to the sentence in completion in the absence of further numerical cues (in treatment group), extending the window of time that we have to observe the potential treatment effect.

Participants. Native Enligh speakers will be recruited to participate in the experiment. UC Berkeley students are an ideal target group since participants would be of approximately the same age and ability. English comprehension will be tested in a preliminary word-image matching task where subjects will be instructed to listen to a sentence and select the object mentioned in the sentence on the screen. This will also serve as a computer proficiency, vision, and hearing test. All objects selected for the main experiment will be used in the preliminary test to ensure that participants are able to discriminate among the targets. Subjects who score low on a word-to-image matching task will not advance to the next stage of the experiment.

Controlling for age is important primarily to enforce a minimum age requirement. Young children whose language comprehension is still developing should not participate in the experiment. Since we intend to measure reaction time will be the measure of the treatment effect, it is also important to ensure that attention span is roughly equal across the subjects. We make the assumption that, on average, college students have roughly similar reaction times to natural speech.

We restrict the subject pool to native speakers to eliminate any effect of native language semantics that could influence one's perception of numerical concepts. As discussed earlier, different languages have adapted different methods for communicative efficiency. It is possible that in some languages numerical grammar does not exist (like gender does not exist in English).

In the event that we choose to administer the modified version of the test (without eye tracking), we can recruit a larger participant pool and administer the test on MTurk. In this case, the preliminary comprehension test will be of utmost importance to filter out non-native speakers.

Participants will be randomly assigned into treatment and control groups. I don't anticipate the need for blocking and clustering, unless we have a very uneven sex or age-group representation in the participant pool. Though I am not aware of specific differences in cognitive processes of male and female that could affect the study, its best to play it safe and control for an even split between the treatment and control groups.

The results of this experiment will provide further insight into how modern English language evolved under cognitive pressure and whether or not numerical grammar is a mechanism that was adapted in favor of communicative efficiency. It's interesting to note that the Old Engligh precursor to modern English contained gender marking similar to some other European languages discussed earlier (Dye et al, 2018). It is curious to explore how and why the language evolved away from the seemingly efficient gender marking system, and whether or not communicative efficiency was traded off for something else or simply deemed redundant.

References

Dahan, D., Swingley, D., Tanenhaus, M. K., & Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in french. *Journal of Memory and Language*, 42(4), 465–480.

Dye, M., Milin, P., Futrell, R., & Ramscar, M. (2018). Alternative solutions to a language design problem: The role of adjectives and gender marking in efficient communication. *Topics in Cognitive Science* 10(1): 209-224

Fernald, A., Thorpe, K., & Marchman, V. A. (2010). Blue car, red car: Developing efficiency in online interpretation of adjective-noun phrases. Cognitive Psychology, 60(3), 190–217.

Gibson, E., Futrell, R., Piandadosi, S., Dautriche, I., Mahowald, K., Bergen L., Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences*. Vol. 23, No. 5

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, 18(3), 193–198

Piantadosi, S.T., Tily, H., & Gibson, E. (2011) Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, 108(9), 3526-9

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, 71(2), 109–147