**Introduction**

Our overall aim is to provide clues on how linguistic processes related to mechanisms which integrate the process of perception of visual scenes and production of utterance. As an initial effort to do this, we report on two sets of eye-tracking experiments, while proposing mechanisms for interpretation which synthesizes existing theories in the literature.

The figure above, which is produced from analysis software EyeParser, contrasts eye movements and utterances gathered from two subjects who were exposed to the same scene but exhibited very different eye movements and utterance patterns: macro-to-micro pattern and micro-shift pattern.

### Two Dichotomous Views

**Structural Guidance**

Perceptual saliency does not play a significant role in utterance order as holistic apprehension of scenes precedes formulation of sentences.

**Perceptual Guidance**

Perceptual saliency influences the order of utterance of production as apprehension and formulation are critically integrated and occur in parallel.

Our claim is that a single mechanism with a change of policy, rather than a radical change of the whole mechanism, can support both of the two seemingly opposing scenarios.

### References


### Acknowledgement

This research was supported in part by the National Science Foundation under Grant No. 0624674 (M.A. Arbib, Principal Investigator), and in part by a Research Grant from the Okawa Foundation.

**Experiment 1**

A pilot experiment was designed and conducted in order to test hypotheses on how semantic representation is built from the perceived visual information and how it influences the produced utterances.

**Participants**

8 native English speakers with normal vision.

**Visual Stimuli**

Photographs of complex, natural live-action scenes.

**Procedures**

Online Description: to keep describing aspects of the scene as quickly as possible while seeing it for 15 sec.

Offline Description: to describe the scene after the scene disappears 10 sec later.

**Observations**

1. **macro-to-micro pattern** was found (Structural Guidance).

2. **micro-shift pattern** was found (Perceptual Guidance).

**Conclusion**

Normal BG/Direct perspective case induces an eye-gaze and utterance pattern that is more similar to the macro-to-micro pattern. The online task seems to influence the eye-gaze and utterance pattern for some case, but it is not significant. The general conclusion is that the seemingly opposing two patterns may be the result of different scene and task settings.

**Experiment 2**

Our claim is that there is a single mechanism that can support both of seemingly opposing gaze and utterance patterns. This indicates that we can control the resulted pattern by changing parameters involved with policy selection, such as threshold, scene arrangements, or task requirements, etc., as similarly addressed by Kuchinsky (2009).

**Participants**

15 native English speakers with normal vision.

**Visual Stimuli**

Photographs of complex, natural live-action scenes with background manipulation (Normal-BG for macro-to-micro, and blurred-BG for micro-shift)

**Procedures**

Similar to Experiment 1, we used two task types, which we assume influence subject's threshold value.

**References**