An Experimental Study of Indonesian Voice

Andrew J. Gallagher
University of Oregon

1.0 Introduction

The experiment described in this paper is an attempt to resolve two complex issues of discourse analysis within a general theoretical framework of functional grammar. One task of analysts of discourse is to determine how mental constructs are coded or mapped onto grammatical form in order to communicate some set of propositions. This has led researchers to engage in attempts to determine the theme or topic of a discourse, two notions notoriously difficult to define. At the discourse level, it is tempting to determine the "aboutness" of the discourse, and it seems that everyone from the Prague School onwards has succumbed to this temptation, with very limited success. In fact, it has proved impossible to define these terms to everyone's satisfaction.

A similar problem has plagued the study of Indonesian discourse. Indonesian can be described as having two voices, active and passive marked by the verbal prefixes meN- and di- respectively. The di- prefix is only a part of a split person passive system, in which di- is used for 3rd person passive constructions while 1st and 2nd person constructions rely on word order and unmarked verb roots to indicate the passive voice (Moeliono 1988; Siu 1976; Wolff 1986a, 1986b). This paradigm is not without controversy however, since some Indonesianists ally Indonesian with other Austronesian languages using a focus rather than voice system (Dardjowidjojo 1974; Poedjosoedarmo 1985), or at least sufficiently unlike the Indo-European paradigm to be labelled as other than "passive" (Oglobin 1983; Verhaar 1976). Some have considered Indonesian to have a backgrounding and foregrounding system (Hopper 1979) or to exhibit syntactic or discourse ergativity. Many of these claims have gone unproven due to the lack of case marking on the NPs which

leads to inconclusive analyses, a matter summarized by Cumming and Wouk (1987). Other summaries of discourse-related issues can be found in works of Kaswanti Purwo (1986, 1988). These controversies notwithstanding, for the purposes of this paper, the terms active and passive will be used for simplicity.

Thus, in discourse analysis of Indonesian we are left with the task of determining the function of the meN- and di- prefixes. If they are some type of markers of voice alternations we still must explain the reasons for these alternations in real-world language use. One must consider the grammatical relations, the semantic roles and the pragmatic motivations involved. An alternate view of language use in general is to consider that the tired notions of theme or topic are, in fact, heuristic labels used to describe what is happening in the cognitive processing system of the speaker. Tomlin (1994) after years of experimental studies has taken a couple of steps back from this approach and uses a finer grain model of language production. Tomlin (to appear) argues convincingly that the notions of theme or topic are superfluous.

> In this model there is no need for a linguistic category of theme or topic. Instead the grammar merely looks at the event representation directly and maps the current attentional focus onto subject. Since attentional focus is needed anyway, the overall grammar can be kept simple, and the concept of theme or topic is rendered superfluous to a theory of language or language production.

This was one of the motivations in designing the present experiment. Instead of struggling to describe and define a topic on a global discourse level, we have attempted to describe what is happening at the level of cognition. Analyzing the notion of voice in linguistics is one way of studying cognitive processes. Voice, as defined by Crystal (1985:329) is "a category used in the grammatical description of sentence or clause structure, primarily with reference to

verbs, to express the way sentences may alter the relationship between subject and object to the verb, without changing the meaning of the sentence." In common usage, in an active voice construction such as John ate the apple, the grammatical or syntactic subject John is also the actor or agent, while in a passive voice construction such as The apple was eaten by John, the syntactic subject apple is also the patient, i.e. that which is acted upon by the agent John. It is fairly easy to describe the movement of agent and patient to different positions in a clause or sentence, but it is another matter to explain in a real-world text or conversation, why one voice construction was chosen over the other.

2.0 Early Experiments

Studies have shown a link between voice and discourse and there have been attempts to determine the factors in the selection of voice, primarily by researchers in the tradition of psychology and cognition experiments. Johnson-Laird, as early as 1968, performed such experiments and found that "the **situation** of a communication (its socio-physical setting and linguistic context) probably exerts a decisive influence upon the form it takes and the way it is understood" (Johnson-Laird 1968:8). He also noted that active and passive voice constructions differ in their emphasis on the syntactic subject, determined by the word order. i.e. whether the agent or patient is in subject position. Turner and Rommetveit (1968:548) in an early study of active and passive sentences, used pictures to prime subjects to elicit active or passive voice, found that "the most salient semantic element, whether the actor of [sic] the acted-upon element, tended to be become the subject of the sentence." Similarly, Tomlin (1994:528) reports that Flores d'Arcais (1975) in a study of Italian, used word cues to prime agent and patient, and found that "priming the agent led to actives 77% of the time; priming the patient led to passives 67% of the time."

For the present experiment I used a computer animated film designed by Tomlin (1994) involving colored fish

engaging in simple events. The fish were given one of five colors (black, white, red, green and yellow) but drawn in exactly the same shape and style. In the original pilot experiment, the fish participated solely in eating events, i.e. only two fish were visible on the computer screen at one time, and one fish inevitably ate the other. Just prior to the eating event, an arrow would appear on the screen pointing to one of the two fish. The subjects were instructed to keep their eye on the fish with the arrow, which in effect would allow a manipulation of their focal attention on a particular fish. The phrase attention detection is used to indicate that which is given attention at particular moments of utterance production. The subjects were also instructed to describe the event while it **happens**, in what is referred to as on-line description, in order to prevent the subject from having time to think and possibly reformulate the utterance with a newly attended referent. This also neatly reduces the effects of memory storage and retrieval on the language production (Cowan 1988, 1993).

Following previous voice studies, Tomlin (1994) investigated the production of active and passive voice constructions. Since the action of eating requires a transitive verb in English, Tomlin predicted that if the agent were primed, i.e. the fish that would **eat** is given the arrow, the subject would use active voice to describe the event. On the other hand, if the patient were primed, i.e. the fish that would be **eaten** is given the arrow, the subject would use the passive voice to describe the event. In the original study, Tomlin performed the experiment with speakers of eight different languages. The general finding was that languages grammatically code the focally attended referent in different ways. In English, Burmese and Indonesian, the focally attended referent is coded as the syntactic subject, i.e. the priming of agent resulted the active voice, and the priming of the patient resulted in passive voice, both more often than chance. However, in Mandarin, Polish, Russian, Bulgarian and Akan, there seems to be no grammatical coding of a focally attended referent. In some cases, agentivity, for instance, seems to override any other coding strategy. See