THE COMPLEXITY OF SIMPLE SENTENCES

Maria Black

Abstract

The performance of experimental subjects on a sentence-picture matching task shows that passive sentences can be as easy to process as active sentences. Passive sentences are only more difficult when particular types of verbs are involved.

1 Introduction

There is a common assumption in both the psycholinguistic and neurolinguistic literature that passive sentences are more difficult to process than their active counterparts. Much of the evidence, however, comes from studies which have used 'off-line' tasks where subjects are asked to judge or memorize sentences, or match sentences to pictures. Since in these tasks performance is usually measured in terms of accuracy (error rates), what we may be measuring is the outcome of 'on-line' comprehension processes and not necessarily 'on-line' processes per se. In a sentence-picture matching task, we are not only measuring the end product of linguistic processing, but also the result of non-linguistic processes since picture recognition, and its interaction with language, is also involved. Therefore, experimental findings based on these tasks do not necessarily show that 'on-line' processing of passives is more complex. As Carrithers (1989) points out, 'the results from these experiments suggest that there is something about the semantic representation of passives, the end product of on-line comprehension processes, that complicates subsequent judgements and memory tasks....' (Carrithers 1989: 76).

In this paper, I will discuss some results from a sentence-picture matching task.[1] These results show that there might indeed be something about the semantic/ conceptual properties of passives that makes subjects' performance less accurate. But the factors involved do not apply equally to all passives. Some passive sentences are just as easy as their active counterparts or other simple sentences.

2 What is involved in the task

What does a subject have to do in order to select the appropriate picture in this type of task? If a sentence is presented first, the subject has to form a representation of the input sentence and must hold that representation long enough to guide and constrain picture selection. Once the pictures are revealed, each picture has to be interpreted, and a representation of the depicted event or

[1] This experiment was carried out in collaboration with Sally Byng and Lyndsey Nickels, as part of a larger study funded by a MRC Grant to Dr. S. Byng.

state must be constructed. The representation of the pictures then has to be compared against that of the sentence and a decision about 'the best fit' must be made.

In comprehending simple active sentences with two noun phrases and a verb, a listener has to go through at least the following processing 'stages':[2]

i) The incoming string has to be processed acoustically and phonologically so that the input string can be segmented into its component elements - phonological words and phrases. As each phonological word is identified, lexical access and recognition take place.

ii) The initial process of segmentation and recognition serves as the basis for the construction of a syntactic representation which specifies the linear order of elements and the major structural relations between them. This syntactic representation serves as the input to a set of 'mapping' functions or procedures which link particular structural positions, such as the Subject/ Object positions or the External/Internal argument positions (Williams 1981), with particular thematic roles.

iii) The particular thematic roles to be mapped onto structural positions are selected by reference to the lexical representation of the verb. If, for instance, the verb expresses an action (e.g.kiss), the thematic roles to be mapped will be Agent and Theme/Patient, while if a verb of perception or psychological state is involved (e.g.see), the relevant thematic roles will be Experiencer and Stimulus.

iv) Each thematic role has to be mapped onto a particular structural position, e.g., the Agent might be mapped onto the Subject position, and the Theme/Patient onto the Object position. The mapping might be determined by item-specific information, and/or by general procedures which apply to particular pairs or sets of thematic roles, irrespective of the particular verb or lexical item involved.

v) The semantic 'content' of the phrases must be combined and integrated to form a full semantic representation of the sentence, specifing the precise nature of the event or state expressed by the sentence, and the precise nature or identity of the participants that play certain roles in that event or state.

vi) This semantic representation serves as the basis for inferential processes, judgements of semantic or pragmatic plausibility, and the identification of particular features that might guide or aid the process of picture recognition. Although very little is known about how picture recognition after hearing a sentence might differ from picture recognition without linguistic input, it is reasonable to assume that hearing a linguistic description of an event or state affects our perception of a picture of that event or state; for instance, one of the pictures in the array might be immediately disregarded and not fully processed simply because it does not contain some feature or property associated with the particular kind of action or state expressed by the sentence. Or we might concentrate on different aspects of the picture depending on whether we hear a verb that 'focuses' on the effect of an

[2] Depending on one's theoretical assumptions, the processing of passive sentences could involve more complex processing at the syntactic level and/or in terms of the mapping of thematic roles. However, I will not deal with these issues as they are not relevant to the findings reported in this paper.

action (e.g.fill) as opposed to a verb that 'focuses' on the action itself (e.g.pour).

This is obviously a very simplified sketch of the main 'stages' involved in performing this task, and it begs a number of important theoretical questions about the precise nature of the representations and procedures involved, and the temporal relations between different levels of processing. But it should at least make clear that performance could be affected by factors at any one of these levels. Selection of the wrong picture could mean that the sentence has been incorrectly interpreted and the picture to match that incorrect representation has been selected; or that the sentence has been correctly interpreted but a picture that does not match it has been selected because something has happened in the translation from language to picture, or in the process of picture recognition itself. Not all errors are necessarily errors of sentence comprehension.[3]

Most psycholinguistic and neurolinguistic studies which have used this task have focused on 'early' stages of processing (e.g. stages (ii) and (iv)) and we still know very little about stage (iii) or 'later' levels of processing. In the next section, I will discuss some experimental results which suggest that factors at 'later' levels affect subjects' performance.

3 Normal performance

3.1. Subjects. 21 subjects were tested on a sentence-picture matching task. All subjects were native British English speakers, whose ages ranged from 15 to 80. Three of the subjects were under 20, three were in their 50s, nine in their 60s, five in their 70s, and one was eighty. Nineteen of the subjects were female and two were male. None of the subjects had any formal education beyond the age of 16.[4]

3.2 Procedure and materials. Six sets of 10 sentences were constructed:

i) Active actional: the verbs in this set are all actional and assign the thematic role of Agent to the Subject and Theme/Patient to the Object (splash, scold, expel, photograph, protect, grab, help, find, wash, weigh).

ii) Active non-actional: the verbs in this set express perceptual or psychological states and assign the thematic roles of (a) Experiencer to the Subject and Stimulus to the Object (dread, admire, notice, hear, see); (b) Stimulus to the Subject and Experiencer to the Object (surprise, astonish, delight, shock,

[3] Black, Nickels, and Byng (in press) show that aphasics' errors on this task can be due to quite different underlying deficits.

[4] These subjects were tested to provide control data for a group of aphasic patients; therefore, their age range and educational level was determined by that of the aphasic subjects they were matched with. However, there were no significant age-related differences in their performance. Some features of the task were also required by its use with aphasic patients, e.g. the use of highly distinctive characters that were not semantically related or visually confusable.

impress).[5]

iii) Passive actional: these sentences were derived from those in set (i) by passivizing the verb and maintaining the order of the NPs (e.g. The judge weighs the pilot => The judge was weighed by the pilot); or by passivizing the verb and changing the order of the NPs (e.g. The pilot was weighed by the judge).

iv) Passive non-actional: these sentences were derived from those in (ii) in the way outlined above.

v) Adjectives: these sentences contained adjectives expressing psychological states or dispositions (shy, fond, sad, sorry, friendly, nasty, rude, polite, cruel, pleasant).

vi) Locative: these sentences contained spatial prepositions (in, on, under (twice), inside, over, after, above, below, behind).

The verbs in (i) and (ii) were matched for frequency, using 'total verb frequency' ratings from Francis and Kucera (1982). The adjectives in (v) were frequency matched with the verbs as far as possible, but no frequency match was possible for the prepositions in (vi) because of the limited range of items.

Except for the locative sentences, the Subject and Object NPs were both human. The characters were chosen so that they were equally likely, or unlikely, to perform an action or be in a certain state, and were clearly visually distinct. The pairings of characters were: clown/astronaut; boxer/cowboy; judge/pilot; cook/dancer; vicar/sailor; nun/queen; swimmer/workman.

The order of presentation of the sentences was randomized both within and between categories, and this order was maintained for each subject.

For each sentence, 3 pictures appeared on a page, in a vertical array (see Figure 1). These pictures consisted of a target picture, a reverse role picture and a lexical distractor picture. Each picture portrayed the same characters; the target and reverse role pictures involved the characters in the same action/state, but the roles of the characters were reversed (e.g. Target: the nun splashes the queen; Reverse Role: the queen splashes the nun). However, the action/state was portrayed in different ways in each picture to prevent the possible use of cues from picture similarity (e.g. the nun splashes the queen by jumping into a puddle, but the queen splashes the nun with water from a cup). The lexical distractor pictures involve the same characters but in a different action/state (e.g. the queen touches the nun). The active and passive forms of a verb used the same picture sets. The left/right order of the characters in the pictures was varied; the vertical order of the target, reverse role and lexical distractor pictures was also varied.

The test was administered in two equivalent versions. Version B was derived from Version A by keeping the pictures constant but changing the order of NPs in half of the sentences (e.g. Version A: the judge weighs the pilot; Version B: the pilot weighs the judge), with appropriate adjustments to maintain all the design features outlined above.

Each subject heard a sentence or read one on a card. S/he was then shown the array of three pictures. Subjects were instructed to point to the picture that 'went best' with the sentence. Eleven subjects were tested on the spoken

[5] These verbs, and some of the adjectives in (v), can also have an actional interpretation. However, the possibility of such an interpretation does not seem to have significantly affected subjects' performance (see Section 3.3).

version of the test, and ten on the written version.

3.3 **Results.** As there was no significant difference between subjects' performance on the auditory and written version of the task (Mann-Whitney, $z = .321$, $p < .1$), the data from the two versions were collapsed and analysed together. There were significantly more correct responses than error responses and significantly more reversal errors than lexical errors (Wilcoxon, $z = 3.8$, $p < .001$). Figure 2 shows the proportion of correct responses for each task condition.

For actional verbs, there was no significant difference between active and passive sentences (Mann-Whitney, $p > 0.05$; Wilcoxon, $p > 0.05$). But the pattern of errors could indicate that some passives might be more difficult depending on the verb involved, although the data is too limited to draw any firm conclusions in this respect. Subjects made a total of 6 errors on the active version of splash, photograph, and protect and no errors on the passive, while for scold, expel, grab, and help a total of 5 errors were made on the passive and no errors on the active. No errors were made on either active or passive versions of wash, while there was one error on both versions of weigh and find.

For non-actional verbs, there was a significant active/passive difference only for the subset of the non-actional verbs which, in the passive, assign the thematic role of Stimulus to the Subject noun phrase, and Experiencer to the noun phrase in the by-phrase (e.g. see, admire) (Mann-Whitney, $p < 0.01$; Wilcoxon, $p < 0.01$). There was no significant active/passive difference for the other subset of non-actional verbs, nor was there any significant difference between the two non-actional subsets in the active version.

Subjects' accuracy was affected by the semantic or conceptual properties of the verb in the sentence. There were significantly more errors on non-actional sentences than on actional sentences, and this was true of both active and passive sentences (Mann-Whitney, $p < 0.05$ and 0.001; Wilcoxon, $p < 0.05$ and 0.001 respectively). There were also significantly more errors on the adjective sentences, which also express a perceptual or psychological state, than on the actional active sentences (Mann-Whitney, $p < 0.05$; Wilcoxon, $p < 0.05$), while there was no significant difference between the adjective sentences and the non-actional active sentences. These differences, however, cannot be reduced to a difference between actions and states because error rates on the locative sentences, which express a state, were not significantly different from those on the actional active sentences, while significantly less errors were made on the locative sentences than on the non-actional active sentences (Mann-Whitney, $p < 0.01$; Wilcoxon, $p < 0.05$). So, there is something specific to perceptual and psychological predicates and/or their pictorial representation that affects accuracy on this task.

These results show that subjects are not always more accurate on active sentences: error rates vary depending on the verb, or type of predicate, involved.

It is unlikely that these results could be described in syntactic terms. For instance, there are no systematic syntactic differences between the active actional sentences and active non-actional sentences we have used. Even if we take into account some recent syntactic analyses of psychological verbs (see Belletti and Rizzi 1988), we could not give an adequate account of all the relevant results. Similarly, there appear to be no syntactic properties that are common to the verbs on which subjects tended to make more errors , nor any syntactic

explanation of why a subset of non-actional verbs should be particularly difficult in the passive. Furthermore, if the actional/non-actional difference was explained in terms of factors at 'early' levels of processing, one would be interpreting the high error rates in the non-actional conditions as errors of 'on-line' sentence comprehension. One would effectively be claiming that about a third of English speaker-hearers do not reliably comprehend sentences with these verbs - a somewhat implausible conclusion.

On the other hand, the results can be quite naturally described in semantic and/or conceptual terms, which strongly suggests that semantic/conceptual factors play a role in this task, during 'off-line' processing. Visual/conceptual factors might also be involved: the difference between actional verbs and verbs and adjectives expressing perceptual and psychological states might be partly due to a difference in 'picturability' as psychological states are undoubtedly more difficult to depict successfully. But 'picturability' cannot be the whole explanation because similar differences between actional and non-actional verbs have been found in experiments with children where pictures were not involved and in children's spontaneous speech (see Sudhalter and Brain 1985 ; Pinker 1990). And 'picturability' alone cannot explain the verb-specific effects in relation to the passive.

What these results highlight is that measuring accuracy of performance on this task might tell us more about 'late' stages of sentence interpretation than about parsing and 'on-line' comprehension (see Tyler 1987; Carrithers 1989). To clarify this point, I will now discuss the verb-specific effects in more detail.

4 Verb-specific effects

Many theorists have claimed that passive sentences differ in meaning from their active counterparts (see Pinker 1990, and references therein), and it is generally agreed that passives have additional pragmatic properties. According to Pinker (1990), the 'core meaning' of the passive is roughly equivalent to 'X is in the state/circumstance defined by Y having acted on it.' The ease with which a verb can be passivized, or the ease with which a particular passive sentence can be interpreted, will therefore depend on how well the meaning of a verb lends itself to such an interpretation. Some verbs will not fit into this interpretation at all and cannot be passivized (e.g. *$5 is costed by the book), while others may undergo passivization but the resulting passive may be more or less acceptable or easy to interpret depending on the semantic and pragmatic properties of the noun phrases involved, and/or on the pragmatic context (e.g. The thief was seen by the police vs ? The flower was seen by John). Pinker distinguishes two main aspects which have to be taken into account in assessing the compatibility of verb meanings with the interpretation of the passive - what he calls 'the agency effect' and 'the predication effect.'

What is meant by 'agency effect' is that '....the passive forces an interpretation whereby the existence of an agent-like argument or party responsible for the circumstance predicated of the Theme is asserted. .... in the passive, the by-object should be more agentive (or at least not less agentive) than the surface subject' (Pinker 1990: 91, 93). It should be stressed that 'agentive' or 'agent-like' does not mean that the by-object must have the thematic role Agent, since by-objects with a variety of thematic roles can appear in the passive. The crucial notion here is that of relatively greater responsibility

or 'causal weight'.

Several researchers have provided empirical evidence that verbs expressing interactions between entities have a sort of inherent causality in the sense that one entity is seen as more causally responsible or weighted than the other (Brown and Fish 1983; Caramazza, Grober, Garvey, and Yates 1977; Au 1986; Corrigan 1988). Although there is little theoretical consensus on the source of these causal attributions, the empirical findings are quite strong and reasonably consistent. It is also clear that these relative causal ratings are not entirely determined by the thematic roles involved. For instance, it is not the case that Agents are always attributed more causal weight than Themes/Patients, e.g. with verbs like thank, congratulate, praise, punish, reproach, and scold, Themes/Patients are attributed more causal weight.[6] Consistent causal ratings are also assigned even if there is no Agent at all. In studies of psychological verbs involving a Stimulus and an Experiencer, it was consistently found that experimental subjects attributed significantly higher causal ratings to the Stimulus, regardless of whether the Stimulus was the subject or object of a sentence and of the animacy of the noun phrases involved (Brown and Fish 1983; Au 1986; Corrigan 1988). These studies provide us with causal ratings for 7 out of the 10 non-actional verbs we used, though it is reasonable to assume that similar causal ratings would apply to the other 3 verbs since the same results have been obtained with a variety of perceptual and psychological verbs.

Given these causal ratings, 'the agency effect' can be used to explain why subjects are less accurate on non-actional passive verbs where the Stimulus is in subject position (e.g. see, admire). If phrases assigned the thematic role of Stimulus were regularly interpreted as referring to an entity with greater responsibility or causal weight, then verbs that assign Stimulus to the subject position and Experiencer to the by-object would go against 'the agency effect' because the by-object would be less 'agentive' than the surface subject.

What I am suggesting is that, in the process of understanding a sentence, certain causal inferences are drawn, for instance, about the relative causality of the participants in the situation expressed by the sentence. Some of these inferences will be based on the meaning of the verb and the relative causality attributed to each of its arguments which are differentiated by their particular thematic role, e.g. verbs of perception or psychological state assign greater causal weighting to the Stimulus argument as opposed to the Experiencer argument. Other causal inferences will be based on the interpretation of the sentence as a whole, e.g. the meaning of a passive sentence forces or licenses the inference that the subject is less causally weighted. If different causal inferences are mutually compatible or reinforce each other, then a causal 'schema' or conceptual model of the situation expressed by the sentence can perhaps be constructed more quickly and with less difficulty. Such a 'schema' can then be used to guide and facilitate the recognition and selection of the appropriate pictures. However, if the different causal inferences conflict, it might

[6] Many of the verbs with causal attributions to the Theme/Patient seem to fall within the category that Parker (1990) labels 'cause-focus' verbs, i.e. verbs that encode information 'that the event was caused by some antecedent agent-patient relation'. So perhaps the 'focus' of the verb also enters into these causal attributions.

take longer or be more difficult to construct a coherent conceptual model as one or other set of inferences might have to be disregarded or some other way of resolving the conflict might have to be found. In a sentence-picture matching task, this might mean that experimental subjects are less sure about which picture they should select or they might be less accurate in their selection.

Similar or related factors might affect subjects' responses on active and passive sentences with actional verbs. For instance, an active sentence with a verb expressing a simple action that does not result in a change of state for the object (e.g. kiss) might be easier to interpret and match to a picture than one with a verb expressing an action that results in a change of state for the object (e.g. splash). On the other hand, verbs like kiss might not be easier in the passive. If the Subject of the passive is interpreted as being in a state or circumstance which is brought about by another entity, then verbs that express an action with a specific effect on the Object would fit in best with the meaning of the passive, and indeed many natural languages restrict passivization of actional verbs to these (Keenan 1985). In English, verbs that do not involve a change of state such as verbs of contact (e.g. touch, kiss) and verbs of motion and contact (e.g. kick, hit, slap, grab) can be passivized, but their interpretation and acceptability might involve additional inferences and pragmatic assumptions (e.g. compare Mary was kissed by John vs. Mary was kissed by Nelson Mandela). English speaking children seem not to produce spontaneously passive sentences with these verbs. Pinker (1990) claims that 'every one of these passives involves a verb with a specific effect. Passives of common actional verbs, like slap, touch, kiss, and feel, are entirely absent' (Pinker 1990: 315).

If this analysis is correct, it would imply that the relative difficulty of active and passive sentences with actional verbs would also depend on the interaction or compatibility of the semantic/conceptual properties of the verb and the interpretation of the sentence as a whole. It is, therefore, interesting to note that the most commonly quoted studies which have found a difference between actives and passives in sentence-picture matching tasks have predominantly used verbs with no specific effect on the Object. Gough (1965, 1966) used only motion and contact verbs (kick, hit). Slobin (1966) found an active/passive difference only for reversible sentences where, judging from the examples given, most of the verbs did not involve a specific effect. He found no active/passive difference for non-reversible sentences and concluded that '....non-reversibility eliminates the problem of passivity, with kernels and passives roughly equal; in fact, the passive form generally tended to be somewhat easier' (Slobin 1966: 226). However, the verbs in the non-reversible condition appear to be predominantly verbs that do imply a change of state for the Object, and it might be the type of verb rather than non-reversibility that 'eliminates the problem of passivity.' The fact that we failed to find any active/passive difference for actional verbs might well be due to the semantic or conceptual properties of the verbs we used.

Given the results discussed in this paper, it seems that there is no overall difference between active and passive sentences. Or, if there is such a difference, this type of task is not sensitive to it. There are, however, specific differences between some passive sentences and their active counterparts, which might be due to the interaction of a number of semantic/conceptual factors. Depending on the task, even the processing of simple sentences can be a complex matter.

References

Au, T.K.F. (1986) 'A verb is worth a thousand words: the causes and consequences of interpersonal events implicit in language.' *Journal of Memory and Language* 25, 104-22.

Belletti, A. and Rizzi, L. (1988) 'Psych-verbs and θ-theory.' *Natural Language and Linguistic Theory* 6, 291-352.

Black, M., Nickels, L. and Byng, S. (in press) 'Patterns of sentence processing deficits: processing simple sentences can be a complex matter.' *Journal of Neurolinguistics*, Special Issue on Aphasia.

Brown, R. and Fish, D. (1983) 'The psychological causality implicit in language.' *Cognition* 14, 237-73.

Caramazza, A., Grober, E., Garvey, K. and Yates, J. (1977) 'Comprehension of anaphoric pronouns.' *Journal of Verbal Learning and Verbal Behavior* 16, 601-9.

Carrithers, C. (1989) 'Syntactic complexity does not necessarily make sentences harder to understand.' *Journal of Psycholinguistic Research* 18 (1), 75-87.

Corrigan, R. (1988) 'Who dun it ? The Influence of actor-patient animacy and type of verb in the making of causal attributions.' *Journal of Memory and Language* 27, 447-65.

Francis, W.N. and Kucera, H. (1982) *Computational analysis of present-day American English.* Providence, R.I : Brown University Press.

Gough, P.B. (1965) 'The verification of sentences.' *Journal of Verbal Learning and Verbal Behavior* 4, 107-11.

Gough, P.B. (1966) 'The verification of sentences: the effects of delay of evidence and sentence length.' *Journal of Verbal Learning and Verbal Behavior* 5, 492-96.

Keenan, E.O. (1985) 'Passive in the world's languages.' In T. Shopen (ed.) *Language typology and syntactic description, I : clause structure.* New York: Cambridge University Press.

Pinker, S. (1990) *Learnability and cognition: the acquisition of argument structure.* Cambridge, Mass: The MIT Press.

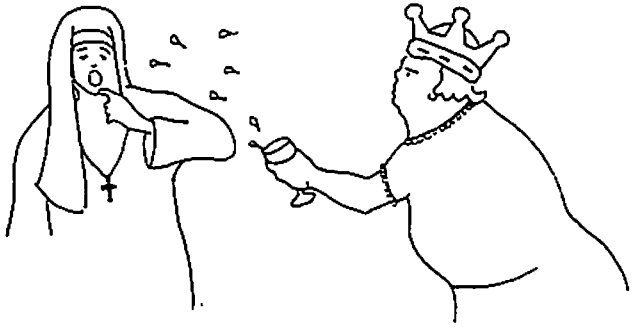
Slobin, D.A. (1966) 'Grammatical transformations and sentence comprehension in childhood and adulthood.' *Journal of Verbal Learning and Verbal Behavior* 5, 219-27.

Sudhalter, V. and Braine, M.D.S. (1985) 'How does comprehension of passives develop ? A comparison of actional and experiential verbs.' *Journal of Child Language* 12, 455-70.

Tyler, L.K. (1987) 'Spoken language comprehension in aphasia: a real-time processing perspective.' In M. Coltheart, G. Sartori, and R. Job (eds.) *The cognitive neuropsychology of language.* London: Lawrence Erlbaum Associates.

Williams, E.S. (1981) 'Argument structure and morphology.' *The Linguistic Review* 1, 81-114.

Figure 1



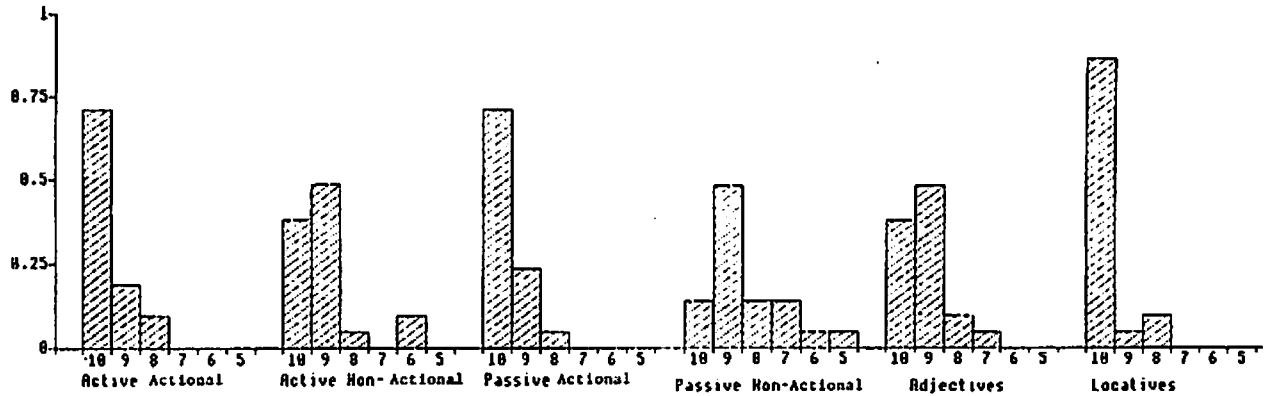


Figure 2. The vertical axis shows the proportion of subjects, and the horizontal axis the number of correct responses.