The role of empathy in sentence production: A functional analysis of aphasic and normal elicited narratives in Japanese and English

Lise Menn, University of Colorado; Akio Kamio, Dokkyo University; Makoto Hayashi, University of Colorado; Ikuyo Fujita and Sumiko Sasanuma, International University for Health and Welfare, Ohtawara; Larry Boles, University of Hawaii*

Introduction

Linguists still tend, for the most part, to be uncertain as to the theoretical relevance of work with abnormal speakers (Menn & Obler 1982, Menn 1989). In this paper, we argue as follows: in the attempt to account for the error patterns of aphasic patients in our picture-description experiments, we have had some success through invoking a concept of empathy. We have been strongly influenced by Kuno's work (Kuno and Kaburaki 1977, Kuno 1978, Kuno 1987); however, we have treated the concept of empathy as a psychological primitive - that is, as an attitude or state of mind towards an entity; specifically, an attitude of 'identification with' a participant in an event. That attitude may or may not have an overt linguistic manifestation; in addition, a speaker may have no particular empathic reaction to a given circumstance. Therefore, an utterance may have no empathic focus.

This psychological/cognitive notion of 'empathy' has been helpful in describing the response patterns of both aphasic and normal control subjects. We can therefore argue that it is an empirically testable and psychologically real notion. However, attempts to analyze elicited narratives and manipulate 'empathy' experimentally have led us to realize how complex a notion it is. So in the conclusion of this paper, we discuss some problems that have arisen and our proposed research directions. In our more ambitious moments, we conceive of this proposed line of research as part of a more general program of experimental functional syntax.

I. Background

A. 'Empathy', speaker, and hearer in a cognitive model of sentence production.

Current psycholinguistic models of language production (e.g. Levelt 1989) explicitly acknowledge that the selection of a particular syntactic/lexical form for the expression of a proposition begins with pre-linguistic cognitive processes involving the speaker's evaluation of the addressee and the setting of the discourse. These processes include discourse tracking, which requires maintaining and updating a representation of the knowledge that is believed to be possessed by or available to one's conversation partners. Such a representation is needed in order to know what the topic is, and to determine when to use a pronoun or one of many possible full noun phrases as referring expressions. Selection among forms of expression is also affected by sociolinguistic considerations of politeness, setting, group structure, etc.; evaluation of these must also be a cognitive process.

This group of discourse factors is *hearer-oriented*; that is, they concern how the speaker takes the hearer's needs into account. But the normal speaker is not a perfect rhetorician; natural discourse does not always afford smooth referent tracking and perfect coherence. Even if the speaker maintains a perfect model of the hearer's mental state - which in itself seems unlikely - he or she does not always utilize that model properly.

Models of pre-linguistic choices for language processing therefore also need to include *speaker-oriented* factors: that is, factors which reflect the speaker's impulses and which may interfere with the speaker's evaluation of the hearer and the setting, at least in rhetorically naive, unedited spoken language. Our recent research has focused on some of these factors which seem to be matters of 'speaker's impulse', at least in rhetorically-naive, unedited spoken language. Empathy (in the psychological sense of identification with a person or creature involved in the situation under discussion) is one such factor; two others that are found in our data appear to be an entity's motion or other exhibited power to cause an event, and its novelty/unexpectedness. In ordinary, rhetorically-unskilled speakers, the integration of these factors can be clumsy, leading to revisions and afterthoughts. For example, in English and Japanese, speakers often use a

pronoun or other deictic word, and then supply the referent as a repair. Japanese conversation also contains OVS and other non-verb-final sentences in which the speaker gives new information, and then adds the old referent if he or she thinks the hearer might have difficulty in retrieving it (Ono & Suzuki 1992).

These psychological factors which affect the speaker's impulse - empathy, motion/causal potential, and novelty - may be subsumed under 'salience', and all three were manipulated to some extent as variables in the most important experimental work in this area to date, Sridhar (1989). Sridhar's study of the relation of psychological salience to linguistic encoding in ten languages was based on order of mention in speakers' descriptions of objects and events in single filmed episodes. The objects varied in position, motion, and humanness,[1] and the languages varied in canonical word order (SOV, SVO). While the data showed many language-related complexities, the study showed (p. 223) that "Entities rendered salient by virtue of their intrinsic meaningfulness (e.g. humanness), or perceptual focus, tend to be expressed sentence-initially, at or near the beginning of the sentence in SVO languages, leading to the use of object-fronted topicalized sentences."

The empirically noted 'impulse factors' of empathy, power to cause events, and novelty are, then, elements of the situation which attract the speaker's attention; in our data for Japanese as well as English, they tend to result in early explicit mention of the attention-getting entity. An evaluative predicate may also be mentioned early in an emotive utterance: Ono & Suzuki (1992: 439) claim that Japanese predicates are sometimes uttered before topics (without breaks in intonation or any other indications of being afterthoughts/repairs) to express "the speaker's urgency in expressing his/her inner feeling" in natural Japanese conversations:

I.1) *...yaa da na kono kokonatsu*
disgusting COP PTCL this coconut
'Tastes awful this coconut drink.'

They therefore argue (p. 441) for "...the importance of affect as one of the motivations of word order variability in Japanese."

The linguistic data are, necessarily, complicated by the fact that the speaker's impulse factors interact with the hearer-oriented factors already mentioned. The speaker should make sure that the hearer can integrate new information into his/her existing mental representation of the events being discussed; if re-evocation of old information is needed for this, the sufficiently skilled speaker postpones the mention of new information, as is well known (Chafe 1976).

Work with aphasic patients who have reduced syntactic abilities shows that they mention new information (Bates & Wulfeck 1989) even when they can do little else; so do very young children (Greenfield and Smith 1976). This has typically been discussed in terms of attunement to the hearer's needs (if any attention has been paid at all to the speaker-hearer distinction), but it is equally well accounted for in terms of the effect of novelty on salience to the speaker. The work of Karmiloff-Smith (1979, 1986) on young children's pronoun use certainly supports a conservative estimate of children's ability to anticipate hearers' information needs and, thence, our distinction between speaker-based and hearer-based functional considerations.

Consideration of experimental and natural discourse in speakers with limited skills makes it painfully evident that there must be a process of cognitive integration of speaker impulse with consideration for the hearer, which then leads to or interacts with syntactic selection. The final syntactic choice is additionally influenced by the relative availability of the various syntactic frames which might be suitable for conveying the information in question, probably by structures recently used (Bock & Loebell 1990), and of course by other language-specific factors. We envision this whole process of integration as taking place in a parallel-processing constraint-satisfaction model (Rumelhart, McClelland and PDP group, 1988), but this claim is independent of those above.

B. Empirical work on aphasia

In recent years, a group of aphasia researchers led by Elinor Saffran and Myrna Schwartz (Saffran, Schwartz, & Marin 1980; Linebarger, Schwartz, and Saffran 1983) obtained several sets of findings which were difficult to reconcile both with each other and with ordinary clinical experience. First, they found that so-called 'agrammatic' aphasic patients - that is, patients

whose speech is slow, usually poorly articulated, and linguistically characterized by simplified syntax, a high percentage of functor omissions, and use of morphologically unmarked forms - seemed to make word-order errors in describing the action of one animate figure on another (boy chases dog), the relative configuration of two inanimates (suitcase is behind chair), and, most especially, the action of an inanimate on an animate (ball hits boy). In contrast, there were few such errors in description of the action of an animate on an inanimate (boy hits ball). These errors suggested that the patients were unable to deal with the argument structure of a sentence in syntactic terms, and that their successes in the Animate-Action-Inanimate sentence type were due to a strategy of choosing the animate referent as subject. This strategy could only work above chance level in the 'boy hits ball' case, and would give below-chance results in the case of 'ball hits boy'.

However, these results were difficult to reconcile both with the absence of reported word-order errors in the free conversations of patients of this type, and with the second set of wholly unexpected findings of this same research group (Linebarger, Schwarz & Saffran 1983, since replicated and extended by Wulfeck 1988), that such patients had near-normal abilities in making many types of grammaticality judgements.

Study of Japanese speakers' difficulty in describing a particular cartoon sequence in the Standard Language Test of Aphasia used in Japan (Fig. 1 - “Hat”) led us to consider the following possible explanation of the Saffran and Schwartz group's results on 'boy hit ball' vs. 'ball hit boy': Agrammatic aphasic speakers have limited syntactic production resources, by definition.

INSERT FIG.1 ABOUT HERE

Fig. 1: Hat

Perhaps they attempt to start sentences by mentioning the empathic focus first, but can basically only produce active-voice sentences and/or agent-first sentences. This fits in with experimental and narrative data showing that aphasics of all types tend to produce sentences with canonical word order and simple verb forms (Bates & Wulfeck 1989). Suppose, in particular, that the stimulus picture is 'ball hits boy', and that no animate entity other than the undergoer 'boy' is illustrated. In that case, the empathic focus is presumably the person affected (undergoer). The aphasic speaker, like the normal speaker, might therefore begin by mentioning the undergoer, but then, being unable either to passivize or to topicalize, would have only two choices: to abort the sentence or to continue with an erroneous 'reversed' production. In the opposite case of describing 'boy hits ball', however, there would be no word-order errors, as the narrator, after beginning with the empathic focus 'boy', could go on in the canonical word order, using active voice.

As mentioned, Saffran et al. (1980) also reported reversal errors in agrammatic aphasics' attempts to describe Animate-Animate and Inanimate-Inanimate interactions/configurations ('the cat chased the dog'; 'the suitcase is behind the chair'). When such pictures are presented in isolation, we assume that there is probably no empathic focus; the reported reversal errors could result from the patient having no empathic or other cognitive guidance as to which entity to start with. In a cognitive model of sentence production, this would cause the arousal of competing, equally-appropriate sentence frames which could blend and cause errors.

The fact that reversal errors have not been reported from agrammatic narratives (Menn & Obler 1990: 1386) would be easier to explain under the empathic-focus-first hypothesis than under the hypothesis that agrammatic aphasics have lost the ability to compute thematic relations. Under the empathic-focus-first hypothesis, empathy or other salience factors would be automatically and appropriately determined in narrative or free conversation about real things that matter. That is, there would always be 'salience-based guidance' as to which referent to mention first, so reversal in Animate-Animate or Inanimate-Inanimate situations would not occur. Complete, fully reversed sentences involving Inanimate acting on Animate, according to

this hypothesis, would be vanishingly rare because self-monitoring would be adequate to insure that sentences containing errors of that magnitude would be aborted. Aborted sentences mentioning the undergoer first might indeed occur, but would be inconspicuous among other incomplete noun-phrase-only utterances.

C. The linguistic expression of empathy.

We follow Kuno in our analysis of the linguistic manifestations of the speaker's empathy or viewpoint. However, as we need the term 'empathy' for an affective state, we will use 'viewpoint' to describe the cognitive orientation decodable from the form of an utterance. In general, the speaker shares the viewpoint of the grammatical subject of the sentence rather than that of an object or oblique argument, according to Kuno's surface structure empathy hierarchy (Kuno 1987:207). Thus, we interpret Kuno as implying, at least for English, that the speaker typically encodes the entity which attracts his/her empathy as the subject of a clause rather than as some other argument. This is how the active/passive contrast can give explicit indications of viewpoint: contrast, again, 'The ball hit the boy' (neutral or following the trajectory of the ball) with 'The boy got hit by the ball' (viewpoint, and presumably empathy, shared with the boy). The cooperative hearer, in decoding, then has his/her empathy attracted by the sentence subject, other things being equal.

However, they are not always equal, as lexical choices also give empathic information. For example, deictic expressions (e.g. demonstratives like 'here'/'there', 'this/that'; a few motion verbs such as 'come'/'go', transfer verbs like 'bring'/'take', 'send'/'receive') give us lexical indications of the speaker's viewpoint independent of syntax or context. Consider the hypothetical contrast: 'George took them to Mary'/'George brought them to Mary'. In the first sentence, the speaker shares George's viewpoint or is neutral; in the second, the speaker clearly shares Mary's point of view. Explicit indication of the state of mind of a referent is also generally taken as indication that the speaker shares the viewpoint of that referent. However, there is no generally agreed-upon exhaustive listing of indicators of shared viewpoint.

II. The experimental study

A. Hypothesis

The hypothesis of the present study is that aphasics and normals are alike in preferring to begin sentences by mentioning or referring to the empathic focus first, and further, that the agency and animacy effects found by Saffran et al. are consequences of this preference, rather than resulting from a default 'animacy strategy' that is used because verb-argument structures can no longer be computed or semantically interpreted.

The 'empathic focus first' proposal says nothing directly about agency or animacy, but the empathic focus of an utterance is most often a person, and very often a person doing something. To test the claim that empathy is really the key factor, rather than animacy or agency, speakers must be induced to talk about events in which a plausible empathic focus is not an agent. Furthermore, either the agent and the undergoer must both be animate, or the agent/cause of the action must be inanimate.

Stimulus Materials. Together with colleague Barbara Fox, we commissioned four 3 or 4-frame narrative cartoon strips in addition to the Japanese 'Hat' strip, resulting in a total of nineteen frames (Figs. 2a-d) which showed the four possible basic patterns of interactions between animates and inanimates (inanimate subjects acting on animates, two animates, two inanimates), plus several agentive and non-agentive intransitive actions. The stories also varied other factors, such as how many humans were involved, whether there was a single clearly-defined protagonist, how 'volitional' the protagonist was, and whether the protagonist (the presumed empathic focus of the story) or another person was the agent in Animate-Animate (A-A) interactions.

INSERT FIG.2 ABOUT HERE

2a: Ball 2b: Race 2c: Boat 2d: Apple

In four of the stimulus stories, there is a clear protagonist, i.e. a person whom the story is 'about'. In 'Hat' (Fig. 1), 'Boat' (Fig. 2c), and 'Apple' (Fig. 2d), there is in fact only one person, who is visible in all or almost all of the frames. In 'Race' (Fig. 2b), a girl is the only person in the

first frame, and she is visually foregrounded in the other three frames. In 'Ball' (Fig. 2a), however, the story changes focus; it starts out with the boy, but then follows the ball in through the window and looks at the man's reactions. At the end, as the man looks out the broken window, one is returned to thinking about the boy, but he is invisible. 'Protagonist' is therefore not clearly defined by the cartoon; this is reflected in the stories which our subjects tell, as will become apparent.

Elicitation Method. English and Japanese aphasic patients and controls were presented with these pictures, and asked to tell the stories; they were prompted and aided as little as possible. Their narratives were tape recorded, transcribed, and analyzed for error patterns and sentence forms chosen or attempted. Sample English and Japanese narratives are given in Appendix A.

Subjects. The aphasic subjects were unselected for diagnostic type, in order to see whether the type of aphasia had an effect either on the syntax attempted or on the errors made. Data have been analyzed to date from nine English-speaking patients capable of giving narratives (preliminary characterization: two moderate Broca's, one moderate mixed non-fluent, one anomic, and three mild fluent aphasics) and nine Japanese patients (preliminary characterization: moderate Broca's, mixed non-fluent and anomic), plus ten English-speaking controls and four Japanese-speaking controls. The ten English-speaking controls were healthy normals from ages 50 to 80; the four Japanese controls were healthy middle-aged adults. Not all the Japanese subjects narrated all the stories; the number of responding subjects for each story varies from 3 to 7 aphasics, and from 2 to 4 controls.

Analysis

Event analysis. There are, of course, many ways of describing a real or pictured event. The event analysis given in Table 1 reflects our subjects' typical responses: neither normals nor aphasics, for example, describe Frame 1 of Apple (Fig. 2d) by saying "The boy is standing on tiptoe under an apple tree with his arm held up in the air"; instead, almost without exception, they interpret his posture in the context of the whole story. For example, normals said, "The boy

is trying to get an apple but he can't reach it"; a moderate agrammatic aphasic said "The boy like to have apple - but apple is too high." Similarly, no subject describes Frame 3 of Ball (Fig. 2a) by saying "The man's mouth and eyes are wide open and his eyebrows are raised"; instead, they interpret these features, imputing an emotional state - here, surprise or annoyance. (Perhaps the actual level of description that is most frequently used, e.g. "The boy is trying to get an apple but he can't reach it", should be considered a coherent but context- dependent, intersubjective 'Basic Event Level', comparable to the Basic Object Level of Rosch's prototype theory.)

INSERT TABLE 1: ANALYSIS OF EVENTS IN ELICITATION CARTOON FRAMES - ABOUT HERE

The propositions were divided into 'foreground' propositions, which were analyzed fully (columns 1-6), and 'other' propositions (column 7). These 'other' propositions were those that were only given as additions to one or more 'foreground' propositions, and seemed to be ancillary to the main story line. (There is of course room for disagreement with this subjective classification; it could be modified by using a naive reader judgement procedure if a more objective method were needed.) Note here that cartoon narratives divide stories into discrete 'events', separated by gaps for which intervening information must be inferred. On the one hand, therefore, the narrator of a cartoon strip bears a greater cognitive burden than the narrator of a comparable filmed story in that he or she must interpret conventional cues for motion and imagine the trajectories of objects that 'move'. On the other hand, the narrator of a cartoon strip is partially relieved of another cognitive burden: he or she is more clearly directed than the narrator of a film as to which events to focus on. The explicitly pictured events are automatically 'foreground' while most of the states (pictured or implicit) and the implicit events are 'background'.

The column headers of Table 1 give the semantic relations of the people and objects mentioned or referred to as indicated by the verbs used and by the word order (for English) or the

case particles (for Japanese). The semantic descriptors in the column headings (agent/cause, 1st entity affected, 2nd entity affected) were chosen to be independent of the particular syntactic form that a speaker might have used. The second through fourth columns show which person(s) and/or objects in the frame were mentioned or referred to. This includes entities which were referred to by pronouns or zero-anaphora.

In the 'modality/time' column, all entries are assumed to depict ongoing action (usually encoded by the progressive aspect), unless 'instant' is indicated in the column. 'Present' means the exact moment depicted, 'past' is slightly before that moment. Modality is marked as 'attempt' in two frames, Hat 3 (proposition b, man tries to get hat) and Apple 1 (boy reaches for/tries to get apple); all other events that we have considered to be 'foreground' are usually interpreted as actually realized, though in a few cases they occur 'between frames' and must be inferred (e.g. Hat 3/4 - the moment where the hat hits the water is not shown).

Some arbitrariness is unavoidable in all of these decisions: after all, a schematization must be nearly as complex as the [language + context] being schematized in order to capture all the differences between choices of words. In addition, many of the relations between verb and argument - e.g. between 'win' and 'race' - are not well described in the usual semantic frameworks. The present analysis is therefore intended only as a convenience for comparison of subjects' responses.

Event encoding analysis. The event analysis of the pictures was then used as a framework for analysis of the narrative contents. Every clause or phrase produced in response to each frame was coded for the propositions which were used. Zero-anaphora is of course very common in all the Japanese narratives, as in the following sample (for abbreviations used, see Notes to Appendix A).

II. 1. Hamaguri-san:

Tsue	*o*	*tsukat-te*	*joozu-ni hippari+age-ta.*
cane	PART:OBJ	use:PRED-CONJ	skilfully:ADV pull up:PRED-PERF

using cane - skilfully - □ - pulled [it] up)

Zero anaphora is also found, ungrammatically, in the narratives of more severe English-speaking aphasics (Mr. Zebra, Hat 4: "cane...then pull out water"), and, grammatically, in conjoined sentences of English-speaking normals (Mr. Badger, Race 3/4: "She runs the race and wins first prize").

The number of descriptors that the subjects gave for each frame reflects the complexity of the event pictured, as well as its centrality to the narrative: Frame 1 of Ball (Fig. 2a), for example, was always described with the boy as subject of an active verb (either the concrete verb 'kick', the more general verb 'play with' or the more interpretive 'play soccer'), but Frame 3 of Ball was described by following the ball's trajectory, what happens to the lamp, the man's reactions, or various combination of these.

Propositions were further coded for whether they were simple descriptions ('The man stands in front of the window') or involved inferences such as mental state ('The man looks out to see who did it'), and whether they included projections of the probable consequences of the pictured events ('The lamp breaks').

Viewpoint analysis. From the subjects, verbs, and verb voices used by our narrators, the various viewpoints they expressed were determined, to the extent possible. Viewpoint analysis is a matter involving a great deal of interpretation, and we regard our attempt in this area as only exploratory. The reasoning used is based on the discussion given earlier: empathy, as an affective mental state of the speaker, may or may not be encoded in a particular utterance; moreover, as Kuno indicates, an utterance may be made from a neutral point of view. Therefore, there may be no overt indication of viewpoint, either because the speaker fails to mark his or her viewpoint, or because he or she takes a neutral point of view. It follows that we cannot always determine the speaker's attitude from looking at what he or she has said.

In analyzing the viewpoints expressed by our four subject groups, we have therefore divided utterances into two classes: those where there is some overt indication of the viewpoint of the speaker (marked empathic focus), and those in which the speaker might be neutral or might be taking the viewpoint of the grammatical subject of the sentence (unmarked empathic

focus). We have taken the following pragmatic, semantic and syntactic phenomena as putative markers of empathic focus: explicit attribution of mental state or mental contents (the boy is happy; the man is scolding), imputation of effort or intention (the boy is trying; the man is looking to see who did it), evaluation of good or bad fortune (unluckily; the man had an accident), judgement of quality of performance (carefully), topicalization, passivization, use of deictic verbs (the ball comes in), and direct discourse (Hey! Sorry!). Aphasic speakers produced a few semantically ill-formed active-voice sentences in which the undergoer is the subject ('He hits on the head - it - the apple'); these we considered as indicating an empathic focus which the speaker was unable to encode grammatically.

We have not taken order of mention as a marker of viewpoint unless topicalization or passivization has occurred. This is partly because 'order of mention' is difficult to evaluate in Japanese. A protagonist is often mentioned explicitly only once in a story by the normals. And as the examiner often provides the first mention of the protagonist for the aphasic subjects, they may not provide any explicit references to the protagonist at all.[2]

If there was no marking of empathy such as those just listed, subject-initial sentences and sentences with zero-anaphor subject were considered to be unmarked for empathic focus. The viewpoint was taken as indeterminate between 'neutral' and the viewpoint of the entity encoded as the main clause subject, whether explicit ('the ball breaks the window) or implicit in the verb ('. . . and falls'). This procedure was applied uniformly to all speakers. When no subject could be determined because only a fragmentary noun phrase or sound effect ("bump!") was produced, no viewpoint was attributed.

III. Results.

Characterization of aphasic narratives

The individual aphasic subjects in general transmitted less information than the normals. Specific error types found in both languages were omissions/fragmentary encodings (e.g. omission of verbs), failed attempts to integrate multiple propositions into a clause, and word-

choice errors (including perseverations from one story into the next). Examples of these error types are as follows.

English

III. 1. Fragmentary encoding/omission

Mr. Horse, fluent, Hat 3 He tried.

III. 2. Failed attempts to integrate multiple propositions

syntactically fragmentary:

a) Mr. Moose, mod. non-fluent, Hat 1 Walking and a cane and a hat

b) Mr. Zebra, non-fluent, Hat 3 Then the one -'round - water. spin the - 'round - water

c) Hat 4 Cane - then pull out water

syntactically smooth:

d) Mr. Wolf, severe anomic, Hat 4 He's getting his cane to put it on, he's getting it out that way.

III. 3. Word-choice error

a) Mr.Wallaby, mod. fluent, Hat 2 Wind comes and pulls his hat away

b) Hat 3 He pulls his scarf into the water

Japanese

III. 4. Fragmentary encoding/omission

a) Ms. Tampopo, mixed, Race 4 *Yuushoo.*

victory

b) Mr. Kamoshika, mixed non-fluent, Ball 4 *Soshitara* "*kora.*"

then:CONJ "Hey!"

c)	Ball 1	*Otoko*	*no*	*ko*	*ga*	*sakkaa*	*o . . . .*
		man	PART:POSS	child	PART:NOM	soccer	PART:OBJ
d)	Ball 2	*Mado*	*o,*	*mado*	*o*	*kashan.*	*[to waru]*
		window	PART:OBJ	window	PART:OBJ	sound of crash	

III. 5. Failures to integrate multiple propositions into a clause

Mr. Kamoshika mixed non-fluent Ball 4

Gashan	*to*	*kodomo ga*	*. . .*
sound of crash	PART:QUOT	child	PART:NOM

III. 6. Perseverations and other word-choice errors

Ms. Tampopo, mixed non-fluent

a) Perseveration Ball 1 *kodomo ga, booru ga, booru o tor-u, tor-u.*

child PART:NOM ball PART:NOM ball PART:OBJ take:PRED IMPF

(The Ball series was presented immediately after the Apple series; the perseveration is the use of the verb *toru* 'get, take', appropriate to the 'Apple' story.)

b) Word-choice error

Ball 3 *Otoko no hito ga booru o hum-u.*

man PART:POSS person PART:NOM ball PART:OBJ step on:PRED-IMPF

Additional error types possible only in Japanese.

Japanese patients used ellipses (grammatically permissible omissions, Kuno 1978) which did not seem to be pragmatically appropriate in the discourse context. In 9 out of 25 sessions with the patients, the main characters in the pictures, which are typically expressed as the subjects of the first utterances, were not expressed even when they were not provided by the examiner. In none of the sessions with the controls, on the other hand, did such failure to mention the main characters occur. (Note: because examiner and patient are looking at the pictured stimuli together, the 'shared information'/'new information' contrast is blurred, and determination of which ellipses are permissible is correspondingly difficult.)

Frequent revisions of particle choices by Japanese patients indicate much difficulty in clause planning. In many instances, the verb was not produced, making it impossible to be sure whether the case particles had been chosen correctly. Occasionally, as in one of our previous studies (Morishima et al. 1992), no verb exists that would have been compatible with the case particles chosen.

III. 7. Particle Revisions

Mr. Oshidori, Wernicke, Apple 3

hooki o sono ringo no ki ni, ringo no

broom PART:OBJ that:DEM apple PART:POSS tree PART:GL apple PART:POSS

ki o hooki de, u-, hooki de, hooki o,

tree PART:OBJ broom PART:INSTR uh, broom PART:INSTR broom PART:OBJ

hooki de butsuke-te

broom PART:INSTR knock:PRED-CONJ

Viewpoint results. We found a high degree of similarity of viewpoints across all four subject groups. The same was true of the use (or attempted use) of marked-viewpoint forms. (Recall that our putative markers of empathic focus were: explicit attribution of mental state or mental contents, imputation of effort or intention, evaluation of good or bad fortune, judgement of quality of performance, topicalization, passivization, use of deictic verbs, and use of direct discourse.)

Combining the English and Japanese results, Figure 3 shows that aphasics' choice of marked vs. unmarked forms for a given frame was highly correlated with the markedness choice made by normals. If we consider the use of unmarked forms only (open squares and circles), the correlation is .82; if we consider the use of marked forms only (black squares and circles), the correlation is .80. These results are of course two sides of the same coin: if subjects use more marked forms, they use fewer unmarked ones. (The correlations differ and the percent of marked and of unmarked forms used for each frame does not always sum to 100% because not all subjects gave a response for each frame.)

Figures 4a-8a show the same data, but graphed in such a way as to permit comparison of markedness choices with the corresponding cartoon frames. Note especially the shift to marked forms in Boat 3, Race 4, and Ball 4.

However, the particular syntactic and lexical devices used to mark the viewpoints differed. For example, quite a number of the more impaired patients marked their viewpoint by the use of direct discourse:

III. 8 Ms. Tampopo, mixed non-fluent, Ball 4

Ntoo	*nto*	*ne*	"*gomen ne.*"
well:INTERJ	well:INTERJ	PART:SFL	"sorry PART:SFL"

A few of the deictically marked forms are interesting to examine further. In Apple Frame 4, three of 4 normals used ochite kuru 'comes falling', rather than just ochiru 'falls'. For one of these, atama ni 'on (his) head' was the first phrase, but two of them began the statement with ringo ga 'apple (subj)'. This could indicate competing perspectives; but as this word order seems

completely natural to native speakers of Japanese, we suggest instead that ochite kuru indeed indicates taking the boy's viewpoint. The constituent order reflects the additional factors of cause/effect iconicity, picture scanning order, and the fact that the boy's desire for the apple, rather than just the boy, may be regarded as the topic of the story.

The fourth control used the perspective-neutral ringo ga atama no ue ni ochimashita 'the apple falls onto (his) head'. As for the patients, 6/7 mentioned ringo 'apple' first (usually marked correctly with ga/wa, but once with o). Four of the 6 produced the verb; two of the 4 also used ochite kuru 'comes falling', and the other two who expressed the verb used forms of ochiru 'fall'.

Encoding choice results

The choice of which propositions to encode was very similar over members of all four subject groups.[3] Overall, the aphasics and the control subjects were also very similar in the way they chose to assign syntactic roles (subject, object) to the participants in the stories (main character, subsidiary character). The choice of which referents to code as topics - either by topicalization or by use as subject - was again similar; we have graphed this set of choices for normals versus aphasics, combining the results from English and Japanese in Figures 4b-8b. Note the switches back and forth among 'man', 'hat', and 'wind' in Fig. 4b; between the girl and the other person in Fig. 5b Frame 2 (but not Frame 4); among the boy, the ball, and the man in Fig. 6b; and so on: again, the frame-by-frame similarity between the aphasics and the normals in their encoding choices is overwhelming.

Much of the linguistic discussion of orientation and empathy has centered around the choice of active vs. passive voice. In the English data, a few patients and normals used get-passives in Race 2, and three patients used variants of 'is thrown out' in Boat 3. However, in the Japanese data, neither patients nor normals made much use of passive voice (one for patients, five for normals) or OSV order (one instance for each group). Both groups often began utterances with locative expressions.

We next examined the effect of animacy and other factors on topic choice, as shown in Table 2.

INSERT TABLE 2. ANIMACY VS. MARKEDNESS, ALL SUBJECTS - ABOUT HERE

Several factors seemed to control topic choice across the four subject groups, as can be determined from comparison of Table 2 with the cartoon pictures. Unsurprisingly, Animates (columns 1 and 3) were chosen as subjects/topics most of the time; 403 animates vs. 164 inanimates. When an Inanimate (columns 2 and 4) was chosen as subject/topic of a particular scene, it was most likely to be a cause (wind, ball, boat, impact of boat, apple) or a freely moving object (ball, lamp, hat, boat, apple). However, a few items coded as subjects/topics are taken from existentials, which introduce them because they are about to play important roles in the story (water, broom). Several others are inanimate objects heavily affected by the action, namely, the broken window and the falling lamp. These factors are summarized in Table 3.

INSERT TABLE 3. FACTORS IN CHOICE OF INANIMATE SUBJECT/TOPIC - ABOUT HERE

One might expect that inanimate subjects/topics would never be marked as empathic foci. However, our definition of 'marked-viewpoint forms' allows for this as a logical possibility: inanimates might be topicalized, used with deictic verbs, or used as the subjects of passive-voice verbs. They might also be used as subjects of metaphorical expressions in which the inanimate object is treated as having a mental state.

Comparison of the subjects/topics which were marked for empathic focus (according to our definition) with the subjects/topics which are unmarked and might be neutral gives a striking result. While 34% of the subjects/topics of the unmarked clauses are inanimate, only 6% of the subjects/topics of the marked clauses are inanimate: 3.5 references to the ball and 1 to the apple, which are freely moving causes; .5 to the hat, a moving undergoer; and 1 to the broken window

glass. (There were two ambigous referring expressions, which were counted as .5 for each possible referent.) Four of the six were marked by deixis, 1.5 were passivized, and .5 was marked by the judgmental phrase ayamatte "accidentally"; no metaphors were used.

Again, aphasics and normals were very much alike in the choices of what to encode as subject/topic; the comparison between the groups is given in Table 4.

INSERT TABLE 4. APHASIC AND NORMAL CHOICES OF SUBJECT/TOPIC ABOUT HERE

Animacy and clause structure.

There were no completed subject-object reversal errors in Japanese or English in these five story contexts, in spite of the fact that an 'animacy strategy' would have worked in only half of the frames, namely those in which an animate being acts on (or attempts to act on) an inanimate object: Ball Frame 1 'boy kicks ball'; Race Frame 1 'girl ties shoe'; Apple Frame 1 'boy reaches for apple'; Frame 2 'boy picks up broom', Frame 3 'boy hits branch with broom'; Hat Frame 4 'man retrieves hat'). There were no apparent problems of subject choice in the Animate-Animate frames (Race 2, 3, 4), even though two of these were complicated by being 'dative' events (i.e., events in which possession of an object is transferred from person to another). With the protagonist recipient, normals and aphasics in both languages successfully used active-voice 'receive/accept' verbs (Race 4: "she gets a prize"). In a few instances, both control and aphasic subjects used the secondary character as the sentence subject (e.g. Race 2: "Somebody gave her a number").

The Inanimate-Inanimate frames posed some problems for some of the Japanese aphasics who tried to begin with the object affected (the window in Ball Frame 2, the hat in Hat Frame 2), but English speakers had fewer problems with them. We think that this is because the syntax and morphology of English, in these particular instances, made it possible for English speakers to deal with these frames in the active voice. In the case of the window, most of our patients, like

the normals, followed the ball's trajectory through the window or took the boy as subject (Mr. Hyrax "an' he puts it through the plate window"; in the case of the hat, several of them used the intransitive construction "His hat blows off."

Although there were no full reversal errors, some attempts at Inanimate-Animate frames indicated problems. A clear example comes from one of the few aphasic sentences which had incompatible subject and verb, e.g. Mr. Wallaby, "He hits on the head" for the final frame of the Apple strip, Fig. 2a. These are of particular interest, as they appear to result from blends of two sentence structures. In this example, the patient, starts with the boy as subject, but continues with a predicate that follows the trajectory of the apple. Since this error could be seen more prosaically as substitution of active 'hits' for get-passive 'gets hit', it is worth pointing out that Mr. Wallaby was a rather mildly impaired fluent aphasic who had few functor problems: for example, he began his narrative by saying "Here he is trying to reach the apple, if it's an apple."

In some cases, Japanese normals used rather different constructions from Japanese aphasics. For example, in Race Frame 2, intended to show the girl getting a number put on her back, 3/4 normals said zekken o tsukete iru 'she has a number'. In this phrasing, the action of placing the number is not explicit; the result of the action is presumably more important than the action in the context of preparation for the race. The remaining normal used ....zekken o tsukete morau, roughly 'she gets a number', which has the function of a passive beneficial to undergoer. Patients either had some difficulty in interpreting the frame (e.g. nambaa o miru 'looks at a number'; undoo saseru 'makes her do exercise') or in finding words to encode the event (e.g. zekken o dasu roughly 'shows a number' [lit: makes a number appear]). Only two of 7 patients used either of the constructions used by the normals.

IV. Summary of findings

Summing up all these observations obtained from our analysis of these data, we may conclude as follows. First, the striking similarity of viewpoints across all four subject groups shown in Figure 3 suggests that aphasics, both in English and in Japanese,

take fundamentally the same point of view as do normals in both languages. This means that in producing an utterance in response to a series of cartoon pictures, aphasic speakers' capacity to perceive empathy is within normal limits.

Second, the interaction of animacy with markedness and subject/topic choice suggests an encoding hierarchy. Humans are strongly preferred as subject/topic (all animates in these stories are human), and are occasionally chosen even when they are not visible in the frame being described (e.g. 'He hits the dock,' Boat 2). As indicated in Table 3, when an inanimate is chosen as subject/topic, it is most likely to be an entity with causative power, in motion, not under current human control (wind, the ball in flight, the apple falling). Being in free motion and an object of concern (the hat rolling, the lamp) also encouraged choice of an inanimate as subject/topic; so did being a severely affected object of value, the broken window.

Explicit syntactic or lexical indications of empathic focus are overwhelmingly reserved (94%) for animates. The hat, the ball, the apple, and the broken window were the only inanimates that were ever elevated to the status of empathic foci, and this happened very rarely.

V. Discussion.

A pervasive but often unacknowledged problem in functional syntax is the following: functional syntax attempts to develop explanations for syntactic choices and syntactic interpretations in terms of concepts like 'topic', 'focus', 'about', 'information' etc. These notions appear to refer to non-linguistic concepts, and the explanations thus appear to make claims about the relation between these concepts and specific linguistic choices. Therefore, logically, testing these explanations requires that we determine the non-linguistic state of affairs independently of the linguistic facts. Only then can we test whether the claimed correlation holds between the non-linguistic state and its proposed linguistic expression.

However, when we analyze narratives apart from an analysis of the real-word events being encoded, we have no independent information as to what happened. Even when we know what the stories are, as in this and other experimental studies, we have no independent evidence

as to the speaker's state of mind (or what state of mind the speaker wishes to create in the hearer). One way to deal with this is to redefine these notions operationally in linguistic terms, as many functionalists have done (e.g. Givón 1984), but then their explanatory power, being entirely intra-linguistic, is diminished. An alternative is to treat them as psychological primitives, and then attempt to manipulate them experimentally, in narratives of varying degrees of naturalness and control as in our work, in the studies by Karmiloff-Smith, by Sridhar and by Bates et al. cited above, and in the ongoing cross-linguistic "Frog Story" studies of children's narratives by Slobin and associates.

Consider 'empathy' in its ordinary sense of a feeling of identification with a person, implying the shared viewpoint. What stimulus factors might one try to manipulate as experimental variables in order to increase a narrator's empathy with a cartoon character? Everyday experience as consumers of comic strips and cartoon films, and now video games, suggests that contributory factors include animacy, volitionality, consciousness, power, perceived attractiveness, and perceived similarity to the narrator on any of a large number of social dimensions. The visible manifestations of the first four factors include motion, effect on other entities, and portrayed emotional reactions.

It seems likely that 'perceived similarity to narrator's self' is the rubric under which all of the above can be summarized. This is supported by informal observation: different people react quite differently to the various characters in a complex story, and the tendency to share the viewpoint of the character most like oneself is quite evident. However, this 'perceived similarity to self' cannot be equated with objectively measured similarity to the narrator; our love for powerful heroes suggests that some idealized self is the one from which distance is measured.

In addition, biological 'releasers' - that is, physical properties that produce automatic reactions in the observer, such as large eyes with large pupils - also play a role in elicting empathy, and they are consciously manipulated by popular artists. (Disney characters have increased eye-size over the years; and Schodt (1983: 91) notes "... by far the most striking visual aspect of [Japanese] girls' comics is the orblike eyes of the characters...Both males and females

in the comics are given this treatment, although the eyes of the males are somewhat smaller, as are those of cold, evil people."

A systematic research program, then, would use narration of sets of animated cartoons or cartoon strips as stimuli, and systematically manipulate variables like similarity to narrator and eye size, looking for corresponding changes in linguistic responses to examine subconscious empathy. Perhaps other methods (questionnaires - 'Who was your favorite character?') could be used to examine conscious empathy.

A final meta-consideration on the problem of defining (or refining) 'empathy' as a technical term: Whenever a term in general usage is adopted as a term of art, its boundaries (even if they remain fuzzy) are sharpened and shifted; when a term defined in one area of science makes connections with notions in a related area, similar shifts are inevitable. This paper attempts to take the study of empathy in language beyond the linguistic basis given to it by Kuno, by providing it with a new foundation in cognitive psychology.

As we do so, we face choices. If empathy in the non-technical sense has to do with a sense of likeness-to-self, how far should it extend as a technical term? In particular, should a technical notion of empathy extend to inanimate objects? In everyday behavior, there is evidence that it should: some speakers at moments of stress certainly go beyond 'Oh, the poor dog!' to 'Oh, my poor car!'; they respond with 'Ouch!' to the sound of a collision, and cheer the 'efforts' of a tow-truck or a construction crane. (Note our factors of damage, motion, and causal efficacy in these imagined but plausible examples.)

The approach taken in the present paper, as indicated in section I.C above, is to treat the use of certain marked syntactic and lexical devices as linguistic evidence that a particular referent is the empathic focus of the sentence. The devices included passives, get-passives, expressions of judgement and mental state, and deictic expressions. However, as we stated, there is no generally agreed-upon list. More to the present point,

if we are trying to put the notion of empathy on a psychological basis, any such list must be tentative. Whether a particular linguistic expression in fact reliably indicates a particular attitude of the speaker (or the attitude that the speaker wishes to be perceived as holding) is an empirical matter. Furthermore, it is a matter that is difficult to test psychologically, although that is what we wish to do by manipulating variables like volitionality, animacy, attractiveness, and value in our drawings. Obtaining experimental support for linguistic intuitions is a very slow and clumsy enterprise, requiring continual revision and refinement of the notions involved. But this is the typical course of scientific development, and it is how linguistics will become a real part of cognitive science.

In review: our work with aphasic patients drew attention to the potential utility of an 'empathy' concept, and then to problems in employing it as a purely linguistic construct. The fact that the initial subject population is language-disordered has thus become almost incidental. But it is precisely the work with an exceptional population that has made plain the need for deeper analysis of a theoretical notion. We hope that, in the near future, our experimental analysis of abnormal language will make a substantial contribution to the study of functional syntax in general, and to the cognitive and linguistic notions of empathy in particular.

*We gratefully acknowledge the assistance of Audrey L. Holland and her research group at the Center for Neurogenic Language Disorders, University of Arizona, in affording us access to their patient population, and of Yasunori Morishima, who translated and carried out a preliminary analysis of the Japanese narratives. This work was partly supported by NIH grant PHS-DC00730-02, the Japan Society for the Promotion of Science, and the Japan-U.S. Educational Commission (Fulbright Commission). Data were collected while Dr. Fujita was at the National Institute for Rehabilitation of the Disabled in Tokorozawa, Dr. Sasanuma was at the Tokyo Metropolitan Institute of Gerontology, and Dr. Boles was at the University of Arizona.

Notes

1. Viewers of animated films impute animacy and volition on the basis of little more than motion cues: Russell Tomlin at Oregon has been working with computer programs in which two outlined fish move towards one another from opposite sides of a computer screen, one opens its mouth, the other disappears into the mouth, and the first closes its mouth again.

2. It is even harder to analyze order of mention in natural Japanese conversations than in these narratives, because of the extent of the ellipses: conversations studied by author M.H. have more zero anaphora than the picture narratives of our normal subjects.

3. A language-specific exception to this is that English speakers have the option of expressing within one clause the trajectory of an entity which is encoded as the object of a sentence ['the boy kicks the ball through the window']; this is not natural in Japanese.

References

Aphasia Research Group (Nirayama conference) (no date). Hyoo-jun shitsugoshoo kensa (Standard Aphasia Test).

Bates, E., and B. Wulfeck. 1989b. Crosslinguistic studies of aphasia. In B. MacWhinney and E. Bates (eds.), *The Cross-Linguistic Study of Sentence Processing*. Cambridge: Cambridge University Press.

Bock, K. and H. Loebell. 1991. Framing sentences. *Cognition* 35:1-39.

Chafe, W. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C.N. Li (ed.), *Subject and Topic*, 25-55. New York: Academic Press.

Greenfield, P. M., and J. H. Smith. 1976. *The Structure of Communication in Early Language*. New York: Academic Press.

Givón, T. 1984. The pragmatics of referentiality. In D. Schiffrin (ed.), *GURT'84: Meaning, Form, and Use in Context - Linguistic Applications*. Georgetown University.

Karmiloff-Smith, A. 1979. *A Functional Approach to Child Language*. Cambridge: Cambridge University Press.

Karmiloff-Smith, A. 1986. Some fundamental aspects of language development after age five. In P. Fletcher and M. Garman (eds.), *Language Acquisition* (2nd ed.), 455-474. Cambridge: Cambridge University Press.

Kuno, S. 1978. *Denwa no Bumpoo*. Tokyo: Taishukan.

Kuno, S. 1987. *Functional Syntax: Anaphora, Discourse, and Empathy*. Chicago: University of Chicago Press.

Kuno, S., and Kaburaki, E. 1977. Empathy and syntax. *Linguistic Inquiry* 8:627-672.

Levelt, W. 1989 *Speaking*. Cambridge, MA: MIT Press.

Linebarger, M. C., M. F. Schwartz, and E. Saffran. 1983. Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition* 13:361-392.

Menn, L. 1989. Some people who don't talk right: Universal and particular in child language, aphasia, and language obsolescence. In N. Dorian (ed.), *Investigating obsolescence: Studies in language contraction and death.*, pp. 335-345. Cambridge: University Press.

Menn, L., and L.K. Obler. 1982. Exceptional language data as linguistic evidence: An introduction. In L. K. Obler & L. Menn (eds.), *Exceptional Language and Linguistics*, pp. 3-14. New York: Academic Press.

Menn, L., and L.K. Obler. 1990. Chapter 20, Conclusion: Cross-language data and theories of agrammatism. In L. Menn and L.K. Obler. (eds.), *Agrammatic Aphasia, Vol. II*, pp. 1369-1389. Amsterdam: John Benjamins.

Morishima, Y.; L. Menn; I. Fujita; A. Kamio; and S. Sasanuma. 1992. Bunsanshutsu no 'on-line processing': nihongo oyobi eigo ni okeru shitsugoshoo kara no shoorei. [Online processes of sentence production: Evidence from aphasic speakers of Japanese and English.] Proceedings of 9th Annual Meeting of the Japan Society for Cognitive Science, pp.106-107.

Rumelhart, D. E.; J.L. McClelland; and the PDP group. 1988. *Parallel Distributed Processing*. Cambridge, MA: MIT Press.

Ono, T., and R. Suzuki. 1992. Word order variability in Japanese conversation: Motivations and grammaticization. *Text* 12:429-445.

Saffran, E.; M.F. Schwartz; and O.S.M. Marin. 1980. The word order problem in agrammatism: II. Production. *Brain and Language* 10:263-280.

Schodt, Frederick. 1983. *Manga Manga: The World of Japanese Comics*. Tokyo and New York: Kodansha International.

Sridhar, S.N. 1989. Cognitive structures in language production. In B. MacWhinney and E. Bates (eds.), *The Cross-Linguistic Study of Sentence Processing,* pp. 209-224. Cambridge: Cambridge University Press.

Wulfeck, B. 1988. Grammaticality judgements and sentence comprehension in agrammatic aphasia. *Journal of Speech and Hearing Research* 31:72-81.