

HOW CHILDREN ACQUIRE LANGUAGE: THE MOTOR THEORY ACCOUNT

Patricia Kuhl, in her Nature review article **37**, surveys the research carried out over many years into all aspects of the acquisition of language by children. She recognizes that 'the mystery' is not yet solved; although substantive progress has been made on some aspects of infants' speech development, notably of the phonology of the parent language. How far in fact has the extensive research program into child language taken us and how plausible and helpful so far are theories of child language acquisition? Over the last few decades research into child language acquisition has been revolutionized by the use of ingenious new techniques which allow one to investigate what in fact infants (that is children not yet able to speak) can perceive when exposed to a stream of speech sound, the discriminations they can make between different speech sounds, different speech sound sequences and different words. Infants' perception of speech develops a good way ahead of their capacity to produce speech sounds, no doubt a reflection of the longer time it takes for the motor capacity for speech to mature. . However on the central features of the mystery, the extraordinarily rapid acquisition of lexicon and complex syntactic structures, little solid progress has been made.

Problems

As Saffran, Senghas, and Trueswell⁴⁰ strikingly put it: "You must discover the internal structure of a system that contains tens of thousands of units, all generated from a small set of materials. These units, in turn, can be assembled into an infinite number of combinations. Although only a subset of those combinations is correct, the subset itself is for all practical purposes infinite. Somehow you must converge on the structure of this system to use it to communicate. And you are a very young child. This system is human language. The units are words, the materials are the small set of sounds from which they are constructed, and the combinations are the sentences into which they can be assembled. Given the complexity of this system, it seems improbable that mere children could discover its underlying structure and use it to communicate. Yet most do so with eagerness and ease, all within the first few years of life."

More specifically,

"Children learn implicitly. By 18 months of age, 75% of typically developing children understand about 150 words and can successfully produce 50 words."

1. How infants learn language with such remarkable speed remains a mystery.
2. How do infants acquire and produce the speech sounds (phonemes) of the community language?
3. How do infants find words in the stream of speech?
4. How do infants link words to perceived objects or action, that is, discover the meanings of words? "The mechanism that controls the interface between language and social cognition remains a mystery."
5. "Why do we not learn new languages as easily at 50 as at 5?"
6. "Why have computers not cracked the human linguistic code?"

(quoted extracts from Patricia Kuhl)

The Motor Theory Account of Child Language Acquisition

1. Finding the phonemes

On the motor theory [Note 1], each speech-sound is the product of an articulatory gesture [Note 2]. Articulatory gestures are the exapted products of innate motor programs which evolved in mammals for the generation of a set of specific arm movements or postures. An infant is sensitized to speech-sounds which generate neural motor programs which match the innate set of motor programs for arm postures and movements. This makes it possible for very young infants to discriminate categorically between heard speech-sounds (as demonstrated by research using HAS (high amplitude sucking) or head-turning paradigms. The Motor Theory of Speech Perception (Alvin Liberman¹⁶⁴ and colleagues) indicated how heard speech accesses motor programs required for the production of the heard speech (a cross-modal transfer). A child becomes able to produce specific speech-sounds as the motor organization of the articulatory system matures in close association with muscular and neural motor development for bodily action generally (including control of arm postures and movements). The infant's early ability to discriminate speech-sounds more extensive than those in the phonology of the parent language is narrowed down by exposure over months to the more limited range of speech-sounds found in the parent language (as described by Patricia Kuhl and other researchers).

2. Finding the Words

"The word falls, one is tempted to explain, into a **mould of my mind** long prepared for it" (Wittgenstein¹⁸¹). Similarly, Chomsky⁷ said: "language acquisition commonly proceeds on course even without any concern on the part of the human models; the precision of phonetic detail a child acquires cannot possibly be the result of training. The speed and precision of vocabulary growth has to be explained by a biological endowment for language; the child somehow has the concepts available before experience with language and basically learns **labels** for already existing concepts." (emphasis added). What can be the link between the mould and the word, between the already existing concept and the word?

What account can the motor theory give of the acquisition of words by children, not only the recognition, discrimination of the individual word, but also of the link which the child establishes between a word and the object or action to which it refers; how does the child acquire the meanings of words? First of all, the child has to be able to pick out the individual word of the parent language from the stream of speech-sound to which it is exposed. To approach this, consider: what is a word? A sequence of speech-sounds combined into a unity, separable from the flow of heard speech, formed from the phonemes which constitute the phonology of the ambient language. At the same time, a word is the product of the combination of articulatory gestures (by the speaker), of motor patterns formed from specific positions and movements of the speaker's articulatory system, from the respiratory apparatus, to the larynx, the mouth, tongue and lips. A word is a neural motor program which can be instantiated to produce a specific patterning of sound in time and which can be heard as integrated (like a music theme) by the hearer, by the child. An adult can perceive a spoken word as a familiar sound pattern but to the child the sound pattern of the word will be novel, unknown. So there must be some other process which enables a child to latch on to the novel, unknown word. This involves the neural representations of perceived objects and actions.

3. Finding the objects and actions (to which words are to be attached)

The ability of humans to recognize a nearly unlimited number of unique visual objects must be based on a robust and efficient learning mechanism that extracts complex visual features from the environment. The basic building blocks of adult spatial vision are in place, if not fully mature, during the first few months of life. Studies often reveal impressive perceptual skills in infants, despite physical limitations and a lack of world experience. Eye-related motor control improves dramatically over the first few months of life. Infants eventually show their perception of objects by reaching for them, and at four months old, they can grasp objects. By four to six months of age, they can estimate an object's distance, orientation, and size, as shown by appropriate grasps for the objects. In general, children seem to categorize objects based on shape rather than size or texture; shape is important for children. The implication is that shape generalizations can facilitate learning of categories for objects. Object segregation requires that infants both perceive (or understand) that the items in question are objects, and distinguish between these objects based upon their perceptual differences. Infants as young as 2 months of age can perceive a moving object as unified. Infants younger than 4.5 months are capable of using featural cues to discriminate between objects, or other test items. Considerably younger infants, 2 months in fact, have some notion of "objectness".

4. Attaching the words to the objects or actions

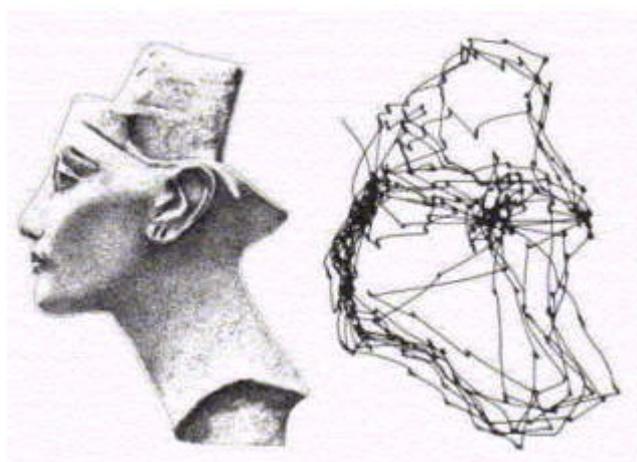
In the brain, what might be the neural 'mould' or 'concept pattern' into which the word can fall, or fit like a key into a lock? The infant's task is to find the appropriate words in the ambient language to fit what it already knows. Before the infant knows words, it will be familiar with many objects, actions, sounds, colours etc. in the external world as well as in terms of its own body, states, feelings, emotions, patterns of its own action. For the infant to acquire the appropriate word for an object, the neural representations of these known objects must be linked with the neural representations of the appropriate words.

Things which are to be attached to words are the result of many different forms of perception. This is reflected in the different categories of words found in the lexicon: nouns, verbs, adjectives, adverbs, prepositions and conjunctions, the closed set of function words (in English – in other languages inflections, affixes etc). Nouns and verbs may be concrete or abstract, refer to external perception or internal perceptions. Adjectives include colours, shapes and sizes – and there are touches, tastes, smells, sounds (besides speech sounds), haptic (touch) experiences.

To see how neural representations of words can become linked to neural representations of objects, it is necessary to take account of what research has discovered or suggested about the neural representation of objects. To start with visual objects. Vision is by far the most important source for the infant's growing knowledge of the world. It is remarkably rapid; humans can recognize an object within a fraction of a second, even if there are no clues about what kind of object it might be. Vision is the most richly represented sensory modality in the cortex. Experiments with monkeys have revealed at least thirty separate visual cortical areas, occupying about one-half of the total cortex and representation of vision in the human cortex is at least as extensive. Not surprisingly, infants concentrate on shapes; words for visual objects are first acquired;

There is a large body of research into the neural representation in the cortex of visual objects. "We are beginning to make progress in identifying the distributed cortical networks associated with semantic object representations, and the networks underlying our ability to retrieve, select and operate upon them" (Martin and Chao**133**). A common feature of all concrete objects is their physical form. The shape of an object is correlated with perceiving objects in many different recognition tasks. Neuroimaging studies have shown that regions in the cortex involved in representing and perceiving objects represent object shape rather than contours or other low-level features. Common neural representation is observed when objects have the same shape but different contours, but not when contours are identical and the perceived shape is different. (Ishai, Ungerleider, Martin and Haxby**60**; Ungerleider and Haxby**69**; Kourtzi and Kanwisher**128**; Welchman, Deubelius and Kourtzi**151**).

Visual perception of objects is a motoric process. The moulds into which, in the language development of the infant, the appropriate words of the ambient language will fall, or be fitted, are formed by the neural representations of the shapes of the concrete objects extracted from the infant's stream of visual experience by the remarkable motor processes of the eye. The eye scans the object by a rapid succession of movements (motor commands for the eye muscles produce movement of the eye up or down, from side to side and obliquely). Eye movements are composed of saccades and fixations; a saccade is a rapid movement of the eye to foveate (the fovea is the central and most sensitive part of the retina) one salient feature after another; a fixation is a pause in the movement of the foveated eye when the finer detail of the object at the point at which the movement is halted is to be obtained. The pattern of the actions of the eye is the result of a complicated neural system, heavily researched but not yet fully described or explained. The outcome of the scanning process is a network of movements and halts, more easily illustrated than described in the work of the Russian physiologist Yarbus**70** (as continued by Noton and Stark**65**). See the famous illustration (recorded by Yarbus) of movements of the eye in scanning a photograph of a bust of Nefertiti. It is from the motor record of the scanning of the object that the shape of the object is derived and finds its neural representation in the dorsal and ventral streams of the brain's cortical visual system. This is how the 'mould' or first outline of the conceptual store for an object is formed into which the neural representation of the appropriate word for the object is to 'fall' or be fitted.



NEFERTITI

5. The recognition of the appropriate word for an object

How is the appropriate word formed and found to be appropriate, in the sense of fitting into or being directly associated with the neural mould constituted by the representation of the visual object in the cortex? Consider what the situation for the infant or young child is. The child has been surrounded from birth, enveloped from birth in the normal case, in a stream of speech sound; the child has been able to distinguish and be responsive to speech sound as distinct from other sound; speech sound has often occurred at the same time as the child is placed amongst objects to which words are eventually found to refer. This process can be taking place anywhere in the world, in any language area. In each different language area a child will be exposed to a different set of words found to be appropriate for the visual objects which the child has already acquired. How can all these words in thousands of different languages, a multitude of different words for the same visual object, be appropriate? This is to be dealt with later; for the time being the question is to be considered for a single language, the ambient language of the developing infant. The question is how the particular word in the language for a particular visual object, the word which eventually goes to fit the 'mould' or neural store already developed by the infant for the object, is in some way specially appropriate and recognised as such by the infant.

The question how a word is appropriate to its meaning leads back to the way in which words emerged for particular objects in the history of any given language. How does a word emerge? At some stage in the history of a human group a new word emerged for the particular object; the new word came into general use and survived because it seemed the fittest word for the object to members of the originating group. In what sense could a word be fit for the object to which it referred? In what sense could the word be said to match the object? Let us consider the perception of any object. For many objects, on seeing them, we can, without saying a word, perform a hand and arm movement to indicate the object. For example, we can indicate a circle by forming or performing a circle with our arms. For a tree, we can, with some accuracy, even indicate the kind of tree by using our hands and arms. We can indicate other objects by pointing. to them, to our head, our foot, our ear, our eye and so on. Homely visual objects, a bowl, a cup, a plate, can be indicated by miming the particular shape. Other items can be indicated by the appropriate contour, a step, an edge, a hill. For many objects we can perform actions to indicate the objects to other persons in our group. Objects are represented by patterns of action for which we have acquired the neural representations needed to perform the actions. Once we have in our neural store (motor memory) the pattern of action representing the object, we can, by a universally available process of motor equivalence[[Note 3](#)], transfer this bodily pattern of action to the articulatory system; an externally perceivable bodily gesture becomes an articulatory gesture, producing a sequence of speech sounds, a unified word, equivalent to the action. We can do this because, as considered in the section above on phonemes, speech sounds are evolutionarily derived from bodily action, from innate programs for elementary movements and postures of the arms. This does not mean or require that at any point in the history of a human group there must have been a developed gestural language as a precursor to spoken language. An appropriate word for an object can be generated simply by imagining how an object might be physically represented or by concentrating on the visual perception of the object and transferring this imagined or visual motor pattern to form an articulatory gesture and so constitute an appropriate word.

We now have arrived at a set of speech sounds forming a word which matches the object, a word which is appropriate for the object to which it refers. How does an infant acquire the word for a visual object on hearing it? From the motor theory of speech perception, the hearing of a word by an adult is perceived in terms of the motor program required to produce the word; the word is cross-modally transformed into the articulatory motor program for producing the word. Similarly for the infant the word is cross-modally transformed (from auditory to motor). There are well-known examples of cross-modal transformation by infants, notably the transformation which must take place when an infant reproduces in its own face the facial expression of the experimenter. See the familiar illustration of a very young child doing this:



from Meltzoff and Moore 94, 95

The motor program so generated by the infant on hearing the word matches the motor program already acquired as a neural representation for the object from the motoric visual processes involved in perceiving the object. In the same way as we recognise a visual object by reacquainting ourselves with the motor shape of the object, so the infant 'recognizes' the motor structure of the word for the particular object in the ambient language. The two fit together, are associated together neurally. This is how the infant acquires the word and its meaning. The word 'falls into the mould' constituted by the neural representation of the object, or, in Chomsky's terms, the infant acquires the appropriate 'label' for the object, a label which is not random or arbitrary but designed to match the motor shape of the object to which it refers.

6. Action word acquisition by children

The next large class of words acquired by young children are action words. These are usually classified as verbs but many nouns are also in effect action words, most obviously nouns formed from verbs. There is evidence (Pulvermüller¹⁴³) that action nouns and verbs are stored together in the cortex. A principal section of action words relate to the child's own bodily activities, eating, drinking, biting, walking, talking, chewing, touching, giving, taking, pushing, pulling, climbing, grasping, and these activities as seen in the behaviour of other children and adults. The formation of words for these actions in the ambient languages in many cases seems especially transparent, particularly for words involving movements of the mouth and face. Saying the word 'spit' almost amounts in itself to the action of spitting; the same seems true also for 'bite', 'chew', 'gnaw', 'snarl', 'sniff'. The infant will have acquired the neural representations for most actions at an early stage in its development, no doubt with 'suck' and 'cry' as primordial (innate) actions. How did the words for actions in any language emerge as early items in the lexicon acquired in a group? The pattern of action generated the articulatory gesture for the action and within a small group (probably a close family group) the word was recognised as being appropriate for the action, fit to survive in the developing lexicon of the group. The motor program for the action, transferred by motor equivalence to the articulatory system, generated the articulatory gesture constituting the word referring to the action. The perception of a given action performed by others no doubt as research into mirror neurons has shown would be recognised because, very similarly to the motor theory of speech perception, there would be an equivalent motor response in the individual perceiving the action which again would be available to generate a word structurally related to the action. In different and separated groups different words for the actions could have emerged because each group might have acquired different sets of phonemes and in any case the particular word generated in a group would depend on the physical characteristics of the individual from whom the word first emerged.

7. Acquisition of words by children for things or processes perceived by other senses than vision or action organisation

An extensive range of words relates to sounds heard other than speech sound (loud, shrill, whistles, growls, bells, wind, things touched (hot, cold, rough, the felt shapes of things), things tasted (sweet, sour, bitter, salty, sugar, honey, salt, lemons), smells, things felt (pains, aches, sores). There are also the words which relate to internal feelings, emotions, wants, thirst, hunger (sad, happy, fearful, anxious, lonely). The normal child will early on have acquired neural representations for many of the things to which these words refer. The child will have ready the moulds into which appropriate words from the surrounding language are to be fitted. The appropriateness of the words in the surrounding language derives again from the first emergence of each word in the originating group, where an individual transferred the neural patterning derived from the feelings, sensations, to the articulatory system by motor transformation and so produced a word which for that individual was structurally related to the particular feeling or sensation and which was accepted by other members of the group (probably closely related with closely similar physical and brain formation, particularly for the articulatory apparatus).

8. Acquisition of closed class words

Words considered above, for things perceived by the main sensory systems, internal as well as external, belong to what are described as 'open classes': there is no limit on the number of things, percepts, action patterns, feelings, etc for which words can be added to the lexicon. By contrast function words form a closed class; in English there are about 200 - many of them of central importance in the formation of sentences and in grammatical structures generally. Other languages use devices equivalent in their grammatical effects to function words in English (inflections of nouns, verbs and adjectives, gender and agreement, affixes, suffixes and infixes and other special features). Function words appear at a late stage in an infant's language development. Whilst a satisfactory account can be given of how, in the grouped list below, deictic, spatial, time and quantity function words can be acquired by children, explaining how words in the 'unclassified' group can be taught is a problem, very much found to be so in practice by parents and teachers of autistic children. How does one explain or demonstrate to a small child 'yet' 'though' or 'if' ?

Deictic a an that the there these this those he her him his I it its me our she their them they us we which who whose you your what

Spatial about above across against along amid among anywhere around at back before behind beside between by down far fore from front further hence in inside into off on onto out outside over round thence through throughout to up upon where wherever with within without

Time after again ago already always during first never now since soon still then today towards until when while whilst

Quantity all another any both each either enough less many more most much neither one other part some somewhat two whatever whatsoever

Affixes -able -al -ary - ed -en -ete -ful -ian -ible -ic -ice -ing -ive -less -ly -nce -ness -ous 's -s -sion -thing -tion -ward

Auxiliaries am are be can cannot do get has have is make may ought seem seems "" been could did done got had made might must seemed shall should was were will would

Unclassified also although and as because but due even except for how however if indeed in order that just like moreover nevertheless no nor not notwithstanding of or otherwise perhaps quite rather same so such than though thus too unless unlike very well whether why yet

Items are allocated to groups on what seems to be their basic function though most spatial function words can also be used for time functions.. Function words and lexical words are not sharply distinct categories. Items often have dual functions; 'to' is basically a directional word but it is also an important grammatical word in the formation of infinitives; 'by' is causal as well as referring to position in space. It is generally difficult to define, that is to explain, a grammatical function word, as any dictionary will show, but nevertheless it contributes substantially to the meaning of any sentence in which it appears.

Linguistics over the years has tended to treat function words as of less theoretical significance than larger syntactic forms. In phrase structure, and the transformational-generative approach, syntax was primary and lexicon (including function words) secondary. More recently, with the shift back from syntax to lexicon of the minimalist program, grammatical function words can now be seen as main structural building blocks of the sentence. The changed emphasis makes it more important to consider how children can come to comprehend the function words of the parent language, to acquire them and to become able to use them appropriately. This is a most important stage in the child's progress towards general grammatical competence.

Many linguists have noted the progress of their children's acquisition of language. Parents' reports provide the bulk of information on the first production of words; free speech analysis and controlled experimental techniques have also been used. However, assessing how early and extensively children comprehend function words is more difficult. Research has been directed towards the general growth of the content word lexicon and less to function words. The extensive cross-language research (notably by Slobin^{41, 42} and his collaborators) has been concerned mainly with wider grammatical patterning of child language though language-specific patterns have been found in the acquisition of words in the locative set of function words. In the early stages of lexical development, when vocabularies are under 200 words, grammatical function words represent less than 5% of all words. At the primitive syntax two-word stage, some function words ('all', 'I', 'no', 'more', 'other', 'off', 'by', 'our', 'away') are found but unstressed function words ('of', 'the', 'on', 'and', 'does') as well as inflections like -ed, -ing, are omitted. Simpler forms are used before more complex ones, e.g. the plural marker -s in English (e.g. cats) before the present tense marker -s (he walks).

Between 18 months and 2 1/2 years, children move on from first words to grammar. The content word vocabulary increases rapidly. Function words are acquired more gradually, which particular words are acquired depending on their frequency, regularity, salience and usefulness for the child. When the expressive vocabulary is between 300 and 500 words, the number of grammatical function words increases in step with overall vocabulary size and with various indices of grammatical productivity. Children at this stage often may use interrogative words such as 'who', 'what' and 'where', form relative clauses, and use comparatives, negations, complements, conjunctions, and passives. In other languages, children acquire the special grammatical features as quickly: free word order, SOV and VSO orders, rich systems of case and agreement, strings of agglutinated suffixes, ergative case marking, gender markings. Onset and growth of inflectional morphology can vary markedly from one language to another, starting as early as the one-word stage in some richly inflected languages.

On the order in which individual function words are acquired a study by Caselli, Casadio and Bates³³ is relevant. Cross-linguistic similarities and differences in early lexical and grammatical development, as reported by parents, were analysed for 1001 English-speaking children and 386 Italian-speaking children between 18 and 30 months. Though the two language lists were not identical, there was a high degree of similarity in the order of acquisition. 'Mine!' (Italian 'mio') was the first item in the pronoun class in both languages, and 'more' (Italian 'ancora') was the first quantifier. Singulars tended to appear earlier than plurals in every relevant class (e.g. pronouns and auxiliaries). Marking of person for pronouns followed the same sequence in both languages; 'we', 'he' and 'she' preceded 'us', 'her' and 'him'. Question words appeared in roughly the same order ('what', 'where', 'why', 'how', 'when', 'which'), though 'where' seemed to be earlier in English while 'chi' ('who') was earlier in Italian. Connecting words in the two languages followed the same sequence ('and', 'because', 'so', 'but', 'then', 'if'), except that 'because' preceded 'and' in Italian. Prepositions and locatives showed a number of parallels: words for direction or location of a single element came first (e.g. 'down', 'up', 'on', 'out', 'here' and 'there'), followed by locatives for a simple relationship of one entity to its base ('on', 'inside', 'under', 'over'); the locatives that appeared last were those expressing a relationship between two entities or a relationship requiring assumptions about the orientation of the array relative to the speaker and listener (e.g. 'next to', 'beside', 'behind'). The article included a comparative table of function words by order of acquisition within each category (pronouns and pronominal determiners, question words, prepositions, quantifiers and articles, connecting words, auxiliary verbs). [table (slightly altered) is [here](#)].

Universal features of language emerge from universal features of the human mind as constituted by the structure and functioning of the human brain. Before humans acquired language, there were patterns of action and perception which were the precursors of the system of function words (or grammatical devices other than or in addition to function words in languages other than English). Both for the group in which a language was developing and for the infant acquiring its native language, pre-verbally there are distinct structural patterns of action and perception which would come later to be represented by specific function words (or parallel forms). These neural precursors of function words are present in other species. 'If' and 'then' for an animal are present in the situation "if you come nearer then I will attack you"

The emergence or the acquisition of a new function word is not a matter of obtaining a new action tool but of latching on to a verbal equivalent for an existing tool. The function of function words is to allow expression of what is in our minds in all its complexity. In the development of a language and in infancy we acquire a huge array of neural representations of patterns of action and perception for which content words are found. The process of our minds (of the active brain) is in the interconnection and interaction of the neural representation of what content words refer to. Function words derive from the innate modes in which concepts are interrelated and interconnected in daily life experience. Through function words and the links they make between content words the complex moment to moment cognitive processes are mirrored in articulated speech sounds; the internal is externalised.

In the same way as motor and perceptual experience generated structurally equivalent words for particular objects and actions through the expression of the associated motor programs via the articulatory system, so in the development of a language the innate concept-structuring system generated particular functional word-forms to match the different elements of the system. Though the words and word-elements differ between languages; what particular sound structures emerged in any particular language depended on 'accidental' characteristics of the particular human group in which the language was developing. In the case of function words there are and were no externally perceivable objects or actions giving rise to the words. Instead there was the living experience of the motor and perceptual activity of the individual in whom a first version of a function word was generated. Particular function words result from the transmission of neural motor patterning to the articulatory system to form motoric articulatory gestures, speech sounds integrated into specific word-patterns.

The process by which in a developing language words generated for objects and actions went to form the lexicon of a particular community is the same process as that by which children acquire their lexicon. Other members of the community recognised the appropriateness of a particular word for what it referred to. Children acquire function and other words because the motor structure of the word matches the motor structure of the neural representation of what the word refers to. The child 'recognises' the grammatical elements as fitting with its already existing neural grammars of action and perception; the function word in the stream of speech is accepted as an appropriate representation of the child's own functioning.

9. Acquisition of sentence-structure

Word-structure is specific to the language community into which the child is born. Acquiring sentence-structure means acquiring the accepted ways into which already acquired content and function words are fitted together to produce acceptable sentences in the parent language. Production of grammatical sentences comes relatively late in the child's development but comprehension comes much earlier. 12-14 month old babies can recognize when sentences are meaningful and grammatically correct and show confusion when they are presented with grammatically incorrect word order. Between the second and third years children produce 2 or 3 word utterances and more extended sentence forms come as the lexicon increases rapidly both for content and for function words. It is notable that children can respond appropriately to normal patterns of speech with no simplifications as in the earlier stages of motherese or telegraphic speech.

Between languages there are many differences in sentence-structure besides differences in words for objects, events and functions. The differences are seen in the ordering of words in the sentence, in the grouping of words into noun and verb phrases, in clause structures. In different language communities children are equally quick in comprehending and later producing acceptable sentences.

A process in the acquisition of sentence-structure seems to occur similar to that already seen in the acquisition of the phonemes and the function words of the language, a narrowing down from an initial ability to perceive a range of possible sentence-structures to arrive at the appropriate sentence-structure for the language. In the case of phonemes and function words the progression has been from an innate general basis to the specific final forms. The question is the character of the precursor non-verbal pattern from which sentence-structures are derived and which is narrowed down to match the sentence-structure experienced by the child in the surrounding stream of speech. Chomsky's proposal is that there is a specifically **linguistic** range of forms for which the child is genetically pretuned and the child selects those first experienced in the community into which it is born. This seems improbable since it would require that the infant should have in its pre-linguistic neural organisation all the complexities of sentence-structure in other languages, and that it would arrive at the appropriate language forms by discarding, for example, Chinese analytic grammar, the

Russian inflectional system, polysynthetic structures and many other forms; how the many different sentence-structures could be generated from a small number of parameters has not been satisfactorily explained. What is more plausible is that, as in the case of phonemes and function words, the precursor of sentence-structure is innate neurophysiological organisation unrelated to language. As others have suggested, sentence-structure in all languages derives from perceptual and motor organisation, the perceptuo-motor system. The motor theory of language origin and function extends beyond the lexicon to the grammatical system, and specifically to sentence-structure. The aspects of sentence-structure which have their origin in the vision or action neural systems include predicate form, subject-verb-object /subject-object-verb order, noun and verb phrase ordering, prepositional and postpositional forms, free order or serial order, topic and/or subject, clause structures, modal and conditional forms. The neural sources for these forms are to be found in what Gregory⁵⁵ named the 'grammar of vision' or what Lashley¹³⁰ described as the 'grammar of action'.

Action is necessarily serially ordered and speech, as a form of action, also has to be expressed in a serial string. Grouping of noun with adjective and determiner in the noun phrase derives from ordering in visual perception where the scanning eye moves from one salient feature to the next; at each stage, that is in the course of each fixation, the qualities associated with the particular object are acquired. In visual scanning, the object-quality order is an equally possible order as the quality-object order depending on the relative salience of quality and object, as guided by the distribution of visual attention. What the eye sees together remains together in the process of translating the visual to the motor image which precedes speech. In the grammar of action the same applies; the manner of an action (adverbial) goes with the specific category of action experienced or observed (to be expressed as the verb). In this way word orders in the sentence reflect orders implicit in innate vision and action syntaxes.

The infant is able to comprehend sentence-structure long before it can produce structured sentences; this is not surprising given that the vision and action systems already form part of its overall organisation; the neural structures on to which the parental sentence-structures are to be mapped already exist. As the infant enlarges its lexicon of content words in step with the array of function words, production of acceptable sentences in the parent languages becomes useful and easy. Different sentence-structures in different languages are alternative methods of mapping on to the same perceptuo-motor neural bases. In the same way as actions can be performed in different but equally effective ways, or different pictures can be drawn to present the features of a given visual scene, so a specific act of perception or a specific motor pattern of action can be expressed in the different ways offered by the sentence-structure of each language. The child can understand the formation of sentences by others (and so arrive at the meaning they intend to convey) by a process of matching the structure of their sentence with its own processes for forming sentences. By a reverse process to that by which we ourselves convert our unitary instantaneous meaning-pattern into a serial string of words in the sentence, we are able to integrate into one meaning the serial string of words in the sentences we hear.

10 . Community acquisition of sentence structure

A child is able to acquire a language because the language which it acquires originated in neural processes similar to those operating in the child's acquisition of the language. In the group in which a language originated, each new word was generated by a single individual; the motor articulatory structure of the word matched the motor structure of a perceived object or event. The word could be accepted by closely related members of the group as being appropriate for its meaning. Over time a substantial lexicon of content words would be built up; along with, but more slowly, a collection of function words or equivalent affixed particles. The capacity to order the string of speech would, as in the infant, be derived from the structures and functioning of the visual and action systems in the individual. The group language would be continuously refined and extended until in due course it arrived at the complexity and effectiveness of modern languages. The only essential evolutionary precursor for the development of a language would be the fact that in humans, but not in related species, a direct neural link had come into existence between the complex motor system and the articulatory structure, on the lines indicated by Jürgens¹⁶¹. There need have been no single ancestral language from which all others were derived. The capacity to develop a new and distinct language would exist in any close human group. Differences between emerging languages would reflect anatomical and neurophysiological differences between human groups, but more specifically would reflect the special character of the individual through whom each word first appeared. Different words between languages would have been the product of individual differences in perception and action, individual differences which are still very much in evidence in modern psychological research. Though the lexicons and syntaxes of languages which developed in different human groups may differ greatly from the lexicon and syntax of English, the fact that all languages are inter-translatable means, as Jakobson¹⁷ pointed out, that at a fundamental level they must be isomorphic, using superficially different ways to express the same range of meanings. In each group, the development of different lexicons and syntaxes is accounted for by differences in the perceptual and motor functioning of the individuals from whom particular content and function words or particles originated. Psychological research has demonstrated how, as a manifestation of individual differences, there can be equally valid alternative ways of scanning a visual scene, performing an action or describing an event. Osgood¹⁷³, for example, in a series of experiments with English-speaking subjects, demonstrated this; subjects observed the same experimental event and then had to report in writing what they had seen. In one case the simple event was that a black ball was rolled across a surface so that it hit a blue ball. The 26 subjects taking part produced 26 different sentences to describe what they had seen. In the case of vision, it has been found that visual

scanpaths for a given pattern vary from individual to individual. Experimental psychology and vision research have provided evidence for widespread differences between members of any single language community in many perceptual and cognitive tasks but there is evidence also for perceptual, cognitive and cultural differences between communities which also will have played a part in generating diverse languages structures. Taking the individual and cultural differences together, there will have been ample sources of variation to explain the diversity of world languages.

11. Acquiring the ability to construct the meaningful sentence: from thought, perception and action to the utterance

The child's acquisition of language is not complete when it has acquired the phonemes, the content and function words and the sentence-structures of the parent language. It still has to be able to fit together the different categories of words in a meaningful way within the acquired sentence-structures. The appropriate nouns must go with the appropriate verbs, adjectives and adverbs; the function words must be inserted at the appropriate points in the sentence; the sentence must be framed to take account of the distinct categories into which, in practice, the lexicon is divided, action and visual words, concrete and abstract, animate and inanimate, touch, smell and taste words, feeling and emotional words. In short, the child has to acquire the ability to transform the specific (momentary) neural patternings of its perceptions, thoughts and feelings into equivalent sentences. It was once said that language was the only window we had into the working of the brain. Nowadays there are also the windows into the brain provided by neuroimaging in all its forms. We are still a long way off from being able to give any satisfactory account of how the brain converts its internal functioning into the external patterning which spoken language constitutes. However, research in progress, particularly by Levelt^{162, 163}, Pulvermüller^{25, 144}, Lamb¹⁸ and others^{58, 126} working along similar lines, is beginning to suggest how the step from the internal to the external, from the simultaneous to the serial can take place.

Questions: How is the thought or perception or action translated into the appropriate, meaningful and grammatically correct sentence? What is the neural process which constructs the motor/sentence image from which the sentence is derived? What is the neural associational structure which determines the appropriate relation between noun and verb, noun and adjective, verb and adverb - the clausal structure? What evidence is there about the neural representation of objects, events and functions which may result in the so-called ' syntactisation' or the ' lemma' of individual words (Levelt's terms)

We have to consider what the answers to these questions may be for adults before thinking about children's acquisition of the meaningful sentence. The problem comes before language. It is implausible to think that language determines the structure of the thought, the perception, the event. Eventually the matter has to be dealt with by neuroscience (though psychology and philosophy may offer pointers to profitable neuroscientific research). So far, psychological research, using reaction time studies, and neurological research, using neuroimaging as well as lesion studies, have not thrown much light on the neural precursors of the meaningful sentence; attention has been concentrated on individual words. Some philosophical approaches (particularly in the Kantian tradition^[Note 4]), have relevance. Universal Grammar theory has not helped much in so far as it suggests innate specifically **linguistic** brain operations as the foundation of the sentence. The transition from thought to sentence is more probably a matter of Universal Cognition, of universally shared pre-linguistic human brain structures and processes.

At this stage no complete or other than hypothetical account can be given of how the thought, the visual perception, the planned or performed bodily action can be transduced into the sentence, the serial string of articulatory gestures forming the utterance - but how to get started? A first step may be to construct an inventory of what might be relevant, of what is known or reliably suggested about the neural representations, the neural assemblies or webs, for objects, events, perceptions, and functions. A list might include: 1. Objects, perceptions, events, are stored in the brain in ways related to the mode of origin of the objects, events, and other items.^[Note 5], Different categories of words are topographically distributed in relation to the external or internal sensory origin of the actions, of the visual, haptic or gustatory perceptions, of feelings, emotions and kinaesthetic experiences. 2. All words have a motor basis. Their structure can, by motor equivalence, be expressed in the form of bodily action, gestures formed from postures or movements of the arms. The associated gestures show a clear relation to the meaning of the words. The structure of the word is not arbitrary but derived from its meaning. 3. Research into motor control has proposed that all actions are preceded by a motor image, which in turn goes to form a serial motor program, constructed from discrete motor primitives, as the action is executed. As an action, an utterance is preceded by a motor image, which generates the string of articulatory gestures formed from discrete articulatory primitives, and so the spoken sentence. 4. Neuroimaging research has shown that the brain responds differently to grammatically correct and grammatically incorrect sentences.^(Fromkin11)

Before one considers how a thought (perception, intended action, desire etc) is converted into a sentence, an utterance, what can be said about ' what a thought (etc.) is? The thought is what is accessible and actually accessed at the point where the constantly-moving internal attentional scanning process is fixated at a particular moment of internal time. The neural process generating the continual movement of attention (conscious or not, directed inwards or outwards) is similar to, or perhaps even the same as that manifested in the continual scanning by the eye of the visual environment. Attention moves from salient point to salient, with fixation at each point for the acquisition of the content at that point. It is this that is convertible, via a motor process, into the equivalent structuring of the articulatory motor system. It is this which in turn is perceived externally as the utterance, the spoken sentence. Kant spoke about 'apperception' ,

essentially meaning ' perception of perception'- not a general state but a specific moment at which what is perceived (internally or externally since all vision is part of total perception) is perceived. Neuroscience research is struggling to give an account of the interplay of attention and perception in terms of neural localisation and networks (work of Koch, Crick and others). Research into the functioning of visual attention and visual awareness may well throw light on the inwardly directed process of attentional scanning and awareness which constitute ' conscious experience' , the moments of consciousness which are thoughts. Velmans' account of the unity of internal and external perception is particularly relevant. [Note 6]),

Research so far has not thrown much light on the relation of thought and sentence in the brain. Little is known about the neural basis of object-based attentional control¹⁴⁹ and almost nothing is known about the cerebral substrates of sentence-level production processes¹²⁵. Following Llinas' account (summarised in Note 6), thoughts represent the ongoing state of play of our cognitive systems, the products of a "continuous humming brain" where "rapidly moving electrical storms represent, internally, the fast and ever-changing external reality" with which a purely hierarchical connectivity alone is too slow and unwieldy to keep pace; thought is internalised movement. We do not prepare our thoughts or consciously preform our sentences. The thought exists as part a total interconnected network, formed from the conceptual and emotional systems and integrated with the functional and motor systems. Though the word network is necessarily less extensive, it parallels the conceptual network and is integrated with it. As a motor network, it can instantly be expressed as a motoric complex, forming the patterns of articulatory gestures which constitute spoken sentences (utterances). The forms of the motor programs from which utterances are derived are constrained by the sensory organisation of the conceptual system or perhaps not so much constrained as more easily linked with aspects of the conceptual system with similar sensory origins. For example, colours and shapes are more closely associated for concepts derived from visual input. Closeness within a sensory network also means that differences and distinctions within the network reduce the probability that incompatible elements will be linked together.

Each word associated with a concept figuring in the network carries with it associations from past experiences of uttered speech and heard speech and the role played by the word in the utterance is not determined but made more probable by these associated links. Each word may have grammatical tags from the ambient language but no mandatory and unchangeable lemmas. It is easy to construct sentences where a particular word can be realised as different parts of speech and different parts of the sentence, with different relations to prepositions and associated forms such as adjectives and adverbs. For example, the word ' love' in a sentence can be a noun or a verb; it can be adjectival or part an adverbial form; it can be subject, object, topic or predicate or indirect prepositional form. Even the structure or ordering of the sentence in which the word appears can take different forms. The elements of the thought go to form not a serial conjunction but an immediate network assembly, a neural ' image' similar to the motor image preceding an action. The thought-image is externalised as the expression of the motor image constituted by the word-motor patterns associated with the neural ' concept' patterns. within the functional framework (matched by the set of function words). asked ' What do you see?' the immediate precursor is the structured visual patterning (motoric in character). Similarly, asked ' what are you doing?' or ' What are you going to do?' , ' What are you thinking?' the utterance is the most prob expression of the momentary state of neural interconnections at that point. As vision seems to involve the integration of snapshots, so the utterance emerges as a snapshot of the ongoing flow of consciousness.

Production and perception of sentences are radically different. The sense of the produced sentence derives from the effective integration of the ongoing process of thought, carrying with it the word associations of the active thought elements. Understanding a heard sentence is a matter of integrating unfamiliar material into the ongoing process of thought, with more or less difficulty. The nature in terms of synapses, dendrites, excitatory and inhibitory neurons of the rapid changes in the brain which instantiate a momentary state of awareness, to be transduced into a new utterance, a new sentence, is a so far unresolved problem.

CONCLUSION

On the questions (from Patricia Kuhl's review) posed in the opening section:

- The motor theory offers a plausible account of the acquisition of the various aspects of language by children.
- Why computers so far have been unable to ' crack' the language problem becomes obvious. Computers can have no natural relation between words and their meanings. Computers do not have the conceptual store to which the network of words is linked. Computers do not have the innate aspects of language functioning - represented by function words. Computers do not have the direct link between speech sounds and movement patterns. Computers do not have the instantly integrated neural patterning underlying thought - they necessarily operate serially and hierarchically.
- Adults find the acquisition of a new language much more difficult than children do because they are already neurally committed to the link between the words of their first language and the elements in their conceptual store. A second language being acquired by an adult is in direct competition for neural space with the network structures established for the first language. Acquisition of a second language by young children is easier because they acquire it in the same way

as their first language (by involvement in ongoing use of the second language in everyday life) and because there is more uncommitted neural space for the construction of a parallel word/concept network. That this is so has been shown directly by experimental investigation of bilingual speakers in the course of treatment for epilepsy.

One of Wittgenstein's early remarks (in the Notebooks 1914-1916**182**) was: "Die Sprache ist ein Teil unseres Organismus und nicht weniger kompliziert als dieser" - "Language is a part of our organism and no less complicated than it". This is true in broad terms though our overall conceptual system is more extensive and more interconnected than the language system. Language is integrated into the overall brain functioning which maintains a continuous representation of all our past and present experience, bodily and mental, and operates through a scanning process which allows momentary conscious awareness to be externalised in speech. Underlying all the functioning of language is the close integration of the motor and perceptual systems, a necessary integration both in evolutionary terms and in terms of everyday life. A possible line forward for better understanding of the transduction of the thought into the utterance, is from vision research in terms of the parallel transduction of the perceived scene into an integrated understanding of the visual scene.

Apart from Wittgenstein, philosophers generally have devoted surprisingly little attention to language in the extensive discussion over the centuries of the relation of consciousness, mind and body. This despite the reality that language has been the instrument they all have used for presenting their speculations and their systems and the recognition, ever since Aristotle, that language defines humanity, is what makes humans human. Kant has no mention of language at all. Hegel has a good deal to say about phrenology in *The Phenomenology of the Spirit* but nothing substantial about language. William James wrote illuminatingly in his *Principles of Psychology* on many aspects of the functioning of mind but had only a rather superficial 2 1/2 pages on language in his chapter on reasoning.

The list of references and bibliography is long because the papers and books have been excerpted and examined for positive or negative relevance and as general background.. The scanning of the literature has necessarily extended over many distinct areas of research, from neuroscience, neuroimaging, lesion studies, motor control, vision to linguistics, language acquisition, child development, language origins, to philosophy, the consciousness debates. For a fuller account of motor theory and associated topics, and animated demonstrations of the relation between word structures and bodily movement, see the papers on Language and Evolution at <http://www.percepp.demon.co.uk>.

NOTES

1. MOTOR THEORY²

The motor theory is a theory of the origin and functioning of language. The theory is that the structures of language (phonological, lexical and syntactic) were derived from and modelled on the pre-existing complex neural systems which had evolved for the control of body movement. Motor control at the neural level requires pre-set elementary units of action which can be integrated into more extended patterns of bodily action - neural motor programs. Speech is essentially a motor activity (a stream of articulatory gestures). Language made use of the elementary pre-set units of motor action to produce equivalent phonological units (phonemic categories). The neural programs for individual words were constructed from the elementary units in the same way as motor programs for bodily action. The syntactic processes and structures of language were modelled on the motor 'syntax'. No separate theory is needed for the origin and functioning of gesture which is itself a motor activity controlled by the same cerebral motor control systems governing all bodily movement.

2. ARTICULATORY GESTURE "In articulatory phonology, the basic units of phonological contrast are gestures ... Utterances are modeled as organized patterns ... of gestures, in which gestural units may overlap in time. The phonological structures defined in this way provide a set of articulatorily based natural classes" (Browman and Goldstein**156** 155) "we show that such gestures not only can characterize the movements of the speech articulators but also can act as phonological primitives" (Browman and Goldstein**155** 313). From Haskins Speech Laboratories: "Articulatory phonology takes seriously the view that the units of speech production are actions, and therefore that they are dynamic, not static. This approach merges a phonological model based on gestural structures with an approach called task dynamics that characterizes speech gestures as coordinated patterns of goal-directed articulator movements. At the heart of both of these approaches is the notion of a gesture, which is considered in this context to be the formation of a constriction in the vocal tract by the organized activity of an articulator or set of articulators."

3. MOTOR EQUIVALENCE

That the same movement can be executed by different effectors was called 'motor equivalence' by Lashley¹⁹; Bernstein⁷² described the concept of motor equivalence to achieve specific kinematic goals. The motor image or motor idea floats free of the particular limb or particular muscle/joint assembly usually employed to execute the motor program. The same pen-stroke can be realised by an infinite number of joint rotation patterns. "I can write the letter A with my hand, with my foot, or even with my mouth; I could even make an A by walking on the beach." (Berthoz⁷³) Subjects wrote their signature with their dominant index finger and ipsilateral big toe. fMRI showed that movement parameters for this movement are stored in secondary sensorimotor cortices of the dominant hand. These areas can be accessed by the foot and are therefore functionally independent from the primary representation of the effector. [Rijntjies et al.¹⁰¹] The present results demonstrate the existence of motor equivalence in a combined upper and lower extremity task. [Marteniuk et al. 2000⁹²] Dynamic cortical activity in the human brain reveals motor equivalence. [Kelso et al.⁸⁶]

The essential idea in the detailed development of the hypothesis of phonological/semantic equivalence is that the gross muscular expression of the word/articulatory pattern can be observed and analysed in the form of gesture and that complex gestures can be broken down into gestural elements associated with particular sound-elements. Motor equivalence can operate from speech to gesture or from gesture to speech. It also seems likely that it can operate between other modalities and speech - or more precisely between motor programs for other modalities and motor programs for speech. Every articulatory program can be redirected (through motor equivalence) to produce an equivalent movement of the hand and arm. Every gesture structured by a perceived object or action can be redirected to produce an equivalent articulatory action. Motor equivalence can function between speech and perception and between perception and motor action because perception is also a motor activity. There are motor programs and motor primitives for vision. These can be transferred by motor equivalence to other muscle/joint assemblies e.g. imitating the shape of a scanned object or a perceived movement by a gesture.

Motor equivalence is demonstrated most remarkably in the relation between speech and gesture. Speech and gesture arise as interacting elements of a single system. [McNeill⁹³] It seems that the speech-accompanying movement is produced along with the speech as if the speech production process is manifested in two forms of activity simultaneously - in the vocal organs and also in bodily movement, particularly in movements of the hands and arms." (Kendon⁸⁷ 205). "For many years we have known in a general way that speech and limb movements are related" (Munhall⁹⁶ 174 reviewing Hammond: Cerebral control of speech and limb movements⁸⁴) "comparing findings on the motor organisation of speech with the organization of voluntary movements about the elbow ... We have found that the kinematic patterns for movements of the tongue dorsum were similar to those of voluntary flexion-extension movements about the elbow" (Ostry and Cooke⁹⁹ 223). "the task dynamic model we are using for speech was exactly the model used for controlling arm movements, with the articulators of the vocal tract simply substituted for those of the arm." (Browman and Goldstein¹⁵⁵ 314) "such gestures not only can characterize the movements of the speech articulators but also can act as phonological primitives" (Browman and Goldstein¹⁵⁵ 313)

Hurley¹²⁴'s useful account drew attention to the new understanding in neuroscience of the relation between perception and action; recent neurophysiological evidence at both single-cell and cell population levels suggested a shared coding for perception and action (as in the motor theory of speech perception). Decety ¹⁰⁹and his colleagues summarised usefully the numerous studies in neurophysiology which lead one to postulate a functional equivalence between producing an action, imaging it, verbalising it and observing it, an attractive model offering a parsimonious explanation of the cognitive mechanisms underlying the generation of an intentional action and the recognition of actions performed by others, Imitation based on the neonate's capacity to represent visually and proprioceptively perceived information in a form common to both modalities. Observations in six

newborns- one only 60 minutes old - suggest that the ability to use intermodal equivalences is innate (Meltzoff and Moore 1995)

4. KANT'S CATEGORIES AND FUNCTION WORDS

<p>I <i>Quantity of Judgments</i> Universal. Particular. Singular.</p>		<p>III <i>Relation</i> Categorical. Hypothetical. Disjunctive.</p>
<p>II <i>Quality</i> Affirmative. Negative. Infinite.</p>		<p>IV <i>Modality</i> Problematical. Assertory. Apodictic.</p>

TABLE OF CATEGORIES		
<p>I <i>Of Quantity</i> Unity. Plurality. Totality.</p>		<p>III <i>Of Relation</i> Of Inherence and Subsistence (<i>substantia et accidentia</i>). Of Causality and Dependence (cause and effect). Of Community (reciprocity between the active and passive).</p>
<p>II <i>Of Quality</i> Reality. Negation. Limitation.</p>		<p>IV <i>Of Modality</i> Possibility. Impossibility. Existence. Non-existence. Necessity. Contingency.</p>

[From the Critique of Pure Reason]

This Note is concerned with the relation between Kant's Categories as transcendental forms of thought and function words as relating to innate pre-linguistic patterns of brain organisation. The Tables above are from Max Müller's 176 translation; Kemp Smith's 178 translation is the same except for minor points such as 'reciprocity between agent and patient' in place of 'reciprocity between the active and the passive'. Kant's procedure was to derive the Table of Categories from the preceding table of Judgements. The contents of both are a priori, that is they are constituents of the functioning of the human mind before any application of them to externally derived aspects of the world. As Sir William Hamilton pointed out (below), Kant's categories are totally different from Aristotle's Categories (predicaments) which were concerned with a quasi-zoological categorisation of the real objects of the world. Kant is classifying the functions of the human mind as applied to thoughts, perceptions, internal mental activity of all kinds. The Categories are effectively innate constituents of the human mind [though Kant does not use the word 'innate'] in the same way as Time and Space are a priori conditions under which all human perception takes place.

Sir William Hamilton commented: "It is a serious error to imagine that, in his Categories, Aristotle proposed, like Kant, 'an analysis of the elements of human reason'. The ends proposed by the two philosophers were different, even opposed. In their several Categories, Aristotle attempted a synthesis of things in their multiplicity – a classification of objects real, but in relation to thought; Kant, an analysis of mind in its unity – a dissection of thought, pure but in relation to its objects. In reality, the whole Kantian categories must be excluded from the Aristotelic list, as determinations of thought, and not genera of real things [Sir W.J. Hamilton Essays and Discussions in Meiklejohn 177 p. 80]. The confusion and misunderstanding, resulting from Kant's borrowing the term Categories from Aristotle, is analysed by other commentators, for example, George MacDonald Ross

[<http://www.philosophy.leeds.ac.uk/GMR/hmp/modules/kant0304/notes/generalknotes.html>].

In investigating how language is possible, how a thought can be transduced into an utterance, a spoken sentence, Kant's classification of the processes of the brain/mind is relevant. Sellars 179 has pointed out that one way of considering the Kantian categories is to think of them as essentially grammatically-derived: "We are construing mental judgements as analogous to sentences. Kant's categories are grammatical classifications. The category of causality, for example, is the form ' X implies Y' ; there is no image of causality as there is an image of a house. The categories with which Kant is concerned in the Critique are the pure categories, specialized in their turn to thought about spatio-temporal objects; we cannot abstract the categories from sensations or images. Kant's categories are forms and functions of judgment. They are grammatical *summa genera*. Aristotle's list of categories is haphazard [confusing] them with generic concepts of entities in the world. A theory of such concepts must be carefully distinguished from the grammar of thought." [emphasis added]

If Kant's Categories, as Sellars suggests, are to be thought of as aspects of a necessary grammar, then this establishes a link with theorising about the biological bases of language (Chomskyan UG etc.) To say, as Sellars does, that the pure categories are specialised to thought about spatio-temporal objects is too narrow, unless 'objects' is taken in a very wide sense as the potential content of all thoughts, all operations of the mind. How did Kant arrive at these quasi-grammatical categories, or more generally these pre-linguistic modes of functioning of the mind? The transition Kant proposed was from forms of judgment (with no origin outside the mind) to forms of thought generally (the Categories). However, the transition, as Kant formulated it, was, as has often been pointed out, in many ways unsatisfactory. Explanation and justification later in the Critique is confusing, even confused. Kant, rather unhelpfully, says: "I purposely omit the definitions of the categories in this treatise. In a system of pure reason, definitions of them would be with justice demanded of me, but to give them here would only hide from our view the main aim or our investigation, at the same time raising doubts and objections, the consideration of which, without injustice to our main purpose, may be very well postponed till another opportunity." One looks, without much success, in the latter part of the Critique for the promised clarification.

In the absence of Kant's clarification, what exactly are the contents of the Categories as pure forms of thought and where do they come from? To construct a plausible account of what each category means, or refers to, one has to make use of function words from which, by hypostatisation or abstraction, terms such as 'hypothetical', 'necessity', 'contingency', 'possibility', 'disjunctive', 'plurality', 'reality', 'existence', are derived. The categories, seem to derive from the set of function words which each language contains: 'if-then', 'must-may', 'can-cannot', 'either-or', 'more-less', 'is-is not', 'why-because', 'some-all'. Rather than Kant's formulaic presentation of the Categories, it is the array of functions, labelled by function words, which constitutes the 'pure' basis of all thought and thoughts, the means by which different concepts, different elements in the mind are linked together, put in relationship with one another. Function words cannot be derived empirically from any external perception or experience and constitute the innate a priori pre-linguistic structure from which all grammars are derived, a manifestation of universal brain-processes, human neurophysiology, to which names have become attached in world languages. There remains the question, touched on in the section on the acquisition of function words by children, how these brain functions came to be labelled with specific word-forms. To say that we find the words in the ambient language which children acquire is no sufficient answer. Children acquire the words because the functions precede the words and the words are judged appropriate for the functions which children already have. But how in the original development of any language, in a given community, did the function word come to be attached to the function? How did a member of the community come to name a function with the word 'if' or 'or'? The answer must be the same as for the labelling of any internal, subjective, experience, that the neural structure or process, the patterning of the experience, was transduced in the form of a motor patterning externalised as an articulatory gesture, so producing a word-sound structurally related to the neural patterning of the experience. The function word generated by one individual could be recognised by others as referring to the particular function because the sound of the word matched the neural patterning of

the function for them as it did for the originator of the word. For example, the word ' if' (or the equivalent word in any other language) was first generated by a single individual; there is nothing external from which 'if' can be derived, so 'if' must have been derived from pre-existing neural patterning. Function words can be seen as marking real categories of human neural functioning on which the grammatical organisation of language has been constructed.

5. WORD AND CONCEPT CORTICAL TOPOLOGY

It appears from lesion evidence, from direct stimulation of the brain (in the treatment of epilepsy) and from neuroimaging research that the brain recognises traditional parts of speech as different classes or, more precisely, that words in the brain are organized according to semantic categories; the semantic knowledge base is categorical in its organization. This requires cell assemblies with distinct cortical topographies as the biological counterparts of words. Distinct neural systems were required for the retrieval of words denoting actions versus those denoting objects and for living things as compared with inanimate objects; PET studies using pictures and words also indicated different brain networks for living and non-living entities with a crucial role of the left fusiform gyrus in the processing of animate entities and of the left middle temporal gyrus for tools. Lesion studies have also shown patients with deficits for abstract as compared with concrete words and for living things as against inanimate objects. Surprisingly specific deficits have been found for words for different types of animals, fruit and vegetables; in one case a person lost the ability to name round fruits (apples, oranges, plums) when shown them (though he called them "fruit"), but had no trouble with bananas; he could not name any birds, members of the cat family, or of the horse family, or raccoons but could name elephants and giraffes. In direct electrical stimulation of the cortex there were sites where induced changes were confined to closed class words related to syntax; conjunctions, prepositions and verb endings were altered but not other words; this was found also in lesion studies: when asked to read a list of homophone-pairs where one member of the pair was a content word and the other a function word, "H.T. was able to read ' four' but not ' for' V.S. read ' sum' but not ' some' ; and J.D. read ' for' and ' some' but not ' four' or ' sum' ". refined topographical categorisation of words must be associated with an equally refined topographical categorisation of the concepts, objects or functions to which the words relate. [for supporting material see references section Neuroimaging and Lesion Studies]

6. VISION AND THOUGHT

The following is extracted and condensed from Velmans' *Understanding consciousness*¹⁸⁰:
The phenomena of which we are conscious at any given moment are the contents of consciousness. The world as perceived is part of the contents of consciousness; the external phenomenal world is constructed from what we see, hear, touch taste and smell. In a typical awake state consciousness includes beside the external physical world as perceived, various body and inner experiences, such as thoughts and images , and awareness of an extended three-dimensional body. With this expansion of consciousness to include all that we experience, the contents of consciousness seem to form a kind of ' psychological present' immediately accessible for report.

What one experiences at a given moment depends on how one directs one' **attention**. With open eyes the contents of consciousness stretch to one' s visual horizons. [Quoting William James] not only do the sense organs themselves select, in that they respond to just a portion of the energies described by physics, but also selective attention ' out of all the sensations yielded, picks out certain areas as worthy of its notice and suppresses all the rest. The mind is at every point a theater of simultaneous possibilities. Consciousness consists in the comparison of these with each other, the selection of some, and the suppression of the rest by the reinforcing and inhibiting agency of attention.' The mechanisms of selection and choice which determine what we attend to are preconscious. Under normal conditions, attentional processing results in conscious experience. Focal-attentive processing is far more sophisticated than nonattended processing. The difference between focal-attentive and nonattended processing accounts for the functional differences between

so-called "conscious processing" and "non-conscious processing". Thoughts represent the ongoing state of play of our cognitive systems, feelings represent our internal (positive and negative) reactions to and judgements about events.

What is common to the complex processes that enable one to read, think, speak, and so on is that they operate, and "become conscious", only if they are at the focus of attention. The detailed motor programs controlling the musculature in speech or in complex adjustments to a changing environment have little manifestation in awareness. When we speak, the words that we hear ourselves utter are the result of prior semantic, syntactic and phonemic planning and consequent motor control; "by the time you are conscious of this sentence, you will already have read it." Covert speech and overt speech have a similar relation to the planning processes that produce them. The processes that enable one to read, think, speak, and so on operate, and "become conscious", only if they are at the focus of attention. The relation between thoughts in the form of "covert" or "inner speech" and the cognitive processes which generate them is similar to that between the words we express and the processes which generate overt speech. "It is only when I hear what I say that I know what I think".

Attention is a central subject in vision research. If, as Velmans suggests, our visual perception of the world is to be thought of as part of our total field of consciousness, then research into visual attention is inextricably linked with investigation of thought and consciousness, and so with the processes by which speech can make manifest the thought to which attention is directed. In many respects Velmans' discussion goes well with the account of the mind, consciousness and language given by the neurologist, Rodolfo Llinas²⁰. The following condensed extracts are from *I of the Vortex: From Neurons to Self* (2002):

Motoricity is of central importance in the origin and functioning of the capacities of the human brain; thinking is internalised movement; the externalisation of any internal image can only be carried out through movement; the premotor events leading to expression of language are in every way the same as those premotor events that precede any movement that is executed for a purpose. The integration and significance of specific sensory occurrences is dependent on the internal of the brain that we generally refer to as attention (a momentary functional disposition); consciousness is a noncontinuous event. The internal functional space that is made up of neurons must represent the properties of the external world - it must somehow be homomorphic with it; the binding of sensory information into a single cognitive state is implemented through the temporal coherence of inputs from specific and non-specific thalamic nuclei at the cortical level. The system that addresses the external world is not a slumbering machine to be awoken by the entry of sensory information, but rather a continuous humming brain; rapidly moving electrical storms represent, internally, the fast and ever-changing external reality. A purely hierarchical connectivity alone is too slow and unwieldy to keep pace; categorical representation is achieved by the activity of populations of cells and not by the electrical activity of any one (face or grandmother) cell. Findings from many systems make questionable the simplistic view of a phrenological type of modular organisation that permeated neurology for many years and that seemed to be supported by the misuse of noninvasive imaging techniques, "neophrenology" though modularity, as dissipative functional structure, is not to be entirely discarded.