UNIT – I

GENERAL INTRODUCTION

1.1 PREAMBLE TO LINGUISTICS

Communication in general can be defined as the transmission of information (= transfer of a message) between a source and a receiver by means of signs. A sign is something physical, which represents something other than itself. Signs have an exponent, a meaning and a set of referents. The exponent of a sign is its physical manifestation, something which can be perceived (heard, seen, touched, etc.) by the receiver of the message; e.g. a gesture, a facial expression, a picture, a roadsign, the sounding of a horn, a word, the smell of burning, etc. The individual things, qualities, actions, states in the world to which a sign refers are the referents (= denotata) of the sign, and these together constitute the reference (= extension) of the sign. In addition to having an exponent and reference, a sign is also associated with meaning. The meaning of a sign is the concept which it evokes in its users and which can be identified with a set of semantic features; this set can be called the sign's intension. For instance, the English word girl evokes a concept that includes the features 'young, female, human'. The signs used in a communication system constitute a code. Signs can be divided into three basic kinds: they can be symbolic, iconic and symptomatic.

• When the exponent of a sign bears an arbitrary relationship to the sign's referents, the sign is a symbol. For instance, the colours used in traffic lights are symbolic: it is a matter of arbitrary convention that the red light means 'stop' and the green light means 'go', in principle it could be the other way around. The vast majority of the words of human languages are symbolic signs: their physical form (pronunciation) and their referents are arbitrarily associated, consider

e.g. English table, Russian stol, German Tisch, Spanish mesa, etc., which all refer to the same kind of thing but all sound different.

However, when there is a natural resemblance between the exponent and the referents of the sign, the sign is an icon. For instance, the stylised silhouette of a man or a woman on a public lavatory door is an iconic sign. A small minority of the words in languages, viz. onomatopoeic words, e.g. English buzz, dingdong, miaow, cuckoo, etc., are signs that are partly iconic. (But even these are partly arbitrary, i.e. partly symbolic, because their counterparts in other languages are never quite the same.)

• And finally, when the exponent of a sign is mechanically linked to its source in such a way that it is a spontaneous reflection of the state of the source, the sign is a symptom (= indicium). Blushing is the symptom of embarrassment, a trembling voice is the symptom of excitement, smoke is the symptom of fire, etc. Symptoms do not need interpretation in a language, they are interpreted simply by there being a direct association between two states of affairs. They are not chosen and sent deliberately but follow automatically from certain states of affairs. (Human beings, however, are capable of producing some symptoms deliberately. This happens in play acting or deceiving,

e.g. when somebody deliberately assumes a tired or a drunken way of walking, even when they are not tired or drunk at all.) 1.2 Linguistic communication: the use of language Human beings communicate in lots of ways, but the most effective way of human communication is linguistic communication, i.e. the use of language. The basic signs used in linguistic communication are words.2 Words (and morphemes) are predominantly symbolic signs, though a small subset of them (onomatopoeia) are partly iconic. The linguistic signs and the rules for their combinations used by a community constitute a linguistic code (a language). Linguistic communication takes place in the

following way. Speaker A, in his/her mind, selects words from the language and combines them according to the rules of the language, i.e. encodes the message. Then his/her articulatory organs – or, in the case of writing, his/her hands – realise the signs: transmit the string of words into a physical signal, a stretch of sound or writing. This signal is perceived by the ears – or, in the case of writing, by the eyes – of Speaker B, who then decodes the message, i.e. reconstructs the message in his/her mind. These steps are shown in (1).

(1) Linguistic communication

nervous system	articulatory organs or hands	sound waves or writing	auditory organs or eyes	nervous system
Speaker A		→ signal →	Speaker B	

If Speaker A and Speaker B do not share the linguistic code (do not speak the same language), linguistic communication cannot take place between them. While animal communication happens automatically and instinctively, conditioned by the situation, relying mostly on symptomatic (and iconic) signs, linguistic communication between human beings is predominantly intentional, relying mostly on symbolic signs, and can be totally independent of the situation. A comparison of human languages and animal communication systems can be made in terms of the so called design features of language (a set of features characterising human languages).3 • We have already mentioned one of these, viz. arbitrariness, i.e. the absence of a natural bond between sign-exponents and their referents (cf. 1.1). Arbitrariness is present in the words (and morphemes) of human languages (not counting the onomatopoeic ones, which are partly natural). Interestingly, a certain degree of arbitrariness is present in bee-dancing, too, which is a term for the special movements that bees perform with their wings and bodies in order to communicate to their fellow-bees about the direction and distance of a source of nectar. This can be regarded as arbitrary because there is no obvious connection between the form of the dance and the distance from the hive, but this arbitrariness is of a very limited kind, manifesting itself only in connection with localising food. • Probably the most important design feature of language is duality (= double articulation). By this we mean that every human language is organised into two layers. The first is a layer of basic sounds called phonemes, such as /Q/, /k/, /t/, which are meaningless in isolation and take on meaning only when combined in certain ways.4 The second is a layer of meaningful units (morphemes, words, sentences), which result from combining the basic sounds,

as in e.g. /kaet/ cat, /taek/ tack, /aekt/ act, or from combining meaningful units, as in e.g. un-friend-li-ness, or in The man saw the lion. There is no evidence that duality is present in any animal communication system.

• Another important design feature is patterning. This means that every language has certain permitted ways of combination, i.e. ways in which phonemes can be combined into words and words into sentences. For instance, in English we can have /kaet/, /taek/, /aekt/, but not */ktae/ or */tkae/ or */aetk/.5 Similarly, we can have He is happy., Is he happy?, and even Happy he is., but not *Is happy he? This feature, too, is missing from animal communication systems. Other design features are creativity and displacement.

• Creativity (= productivity) means that human beings can produce and understand an infinite number of new messages that they have never heard before. This feature is present in bee-dancing, because the exact location of nectar can always be different and so new messages are always possible, but this is a very limited kind of creativity, existing in connection with localising food again.

• Finally, displacement can be defined as the ability to use language in connection with things and events remote in space and time. For instance we can talk about past, future and distant events, hypothetical objects, and we can even lie.

Displacement is certainly present in bee-dancing, but only in connection with food, and it is only spatial displacement, not temporal. (Bees can inform their fellow-bees about the whereabouts of a source of nectar at the time of the communication, but they cannot pass on any information about the availability of a source of nectar in the future or in the past.) The following table sums up those design features that we have discussed, and the corresponding values in bee-dancing. (2) Design features of human language and an animal communication system

(2) Design features of human language and an animal communication system

	human language	bee-dancing
duality	+	–
patterning	+	–
creativity	+	limited
arbitrariness	+	limited
displacement	+	limited

THE DISCRETE NATURE OF LANGUAGE

Linguistic communication, i.e. the use of language, is characteristically vocal and verbal behaviour, involving the use of discrete language elements.

• It is vocal because it is crucially associated with the articulatory (vocal) organs.

• It is verbal because words play a central part in it.

• Thirdly, it involves the use of discrete language elements, which differ from one another discretely (on an either-or basis) rather than gradually (on a more-or-less basis). Words as lexical items are discrete because they differ from one another on an either-or basis. This means that two word-realisations either represent the same word or two different words. For instance, whenever English speakers utter the word pen, there will always be some physical difference between the realisations, but these variations are not only gradual and hardly noticeable but also insignificant, and so each rendering will be taken as realising the same word pen.

However, when the English words pen, pan, car, picture and camera are uttered, the words are discretely different, regardless of the fact that in physical terms pen and pan, both starting with /p/ and ending with /n/ are more similar to each other than to the others. In the relevant sense, they are all discretely different because if you change one of them for another in a particular sentence you may get a completely different sentence, with a completely different meaning. Words are composed of basic sounds called phonemes. The latter are discrete, too, because two phoneme-realisations either represent the same phoneme or two different phonemes. The phonemes of a language are those sounds that are capable of distinguishing otherwise identical words. If you replace one phoneme with another in a particular word, you may get a different word which no longer means the same.

PARALANGUAGE

In the preceding section we have seen that the use of language is characteristically vocal and verbal behaviour, involving the use of discrete elements. However, accompanying and occasionally even replacing language, we also find behaviour which is not vocal or, if vocal, not verbal and not discrete.

Variations in this kind of behaviour, used during and instead of linguistic communication, are called paralinguistic features or paralanguage. 8 Among the non-vocal features of paralanguage we have to mention gestures, bodily movements, facial expressions which we make while we are speaking or instead of speaking. These include e.g. bowing, waving, winking, raising our eyebrows, putting our finger across our lips, shaking our head, nodding, etc. The vocal paralinguistic features include various meaningful noises, such as hm, pff, the wolf whistle (a not very polite way of young men's whistling to a pretty girl), throat clearing, etc. Some of these are intentional, some are unintentional. Throat clearing, for example, can be either. When it is unintentional, it is merely symptomatic, and shows that you have a cold. But it can be used deliberately (symbolically), too, e.g. as a warning. The suprasegmental part of human communication, which we sketched in 1.3, is vocal, non-verbal behaviour, and it is partly linguistic, partly paralinguistic. The stress degrees are linguistic because they are discrete, but the general loudness-level of a certain part of an utterance is gradable, and so paralinguistic. The pitch patterns used in word-melodies and in intonation are linguistic because they are discrete, but their vertical extent (range) and the general pitch height of certain parts of utterances (key) are gradable, i.e. paralinguistic variations. Similarly, tempo, pause-length, voice-quality variations are gradable, and however important and informative they may be, they belong to paralinguistic rather than linguistic communication.

Exercises, problems, and other tasks

1. What is communication?

2. How do we define a sign and what are the three parts of a sign?

3. What is the exponent, the meaning and the reference of the English word tree?

4. What are symbols, icons and symptoms?

5. What are onomatopoeic words?

6. What are the English counterparts of the following Hungarian words: kukurikú, bimbam, tiktak. What is your conclusion?

7. When a dog opens its mouth and shows its teeth in a threat to bite, the sign it uses is partly iconic, partly symptomatic. Explain.

8. How does linguistic communication take place?

9. What do we mean by arbitrariness as a design feature of human language? Illustrate it.

10. What is duality?

11. What is patterning?

12.How do the English pronounce the letter combinations kn, ps, mb in knee, knowledge, psychology, psalm, comb, dumb, acknowledge, rhapsody, cucumber? Try to find an explanation.

13. What do we mean by creativity and by displacement as design features of human language?

14 Linguistic communication is characteristically vocal and verbal behaviour, involving the use of discrete language elements. Explain the key words vocal, verbal, discrete.

1.2 PHILOLOGY

Comparative Philology was the dominant kind of language study in the 19th century. It was scientific in several respects. However, it narrowed down the concept of language study to a study of the history and genetical relationships of languages and of the written records that were available. 24 This kind of linguistics emerged after the discovery that Sanskrit was related to Latin and Greek. The discovery was made in 1786, by a British government official working in India, Sir William Jones. Throughout the 19th century, language scholars tried to establish genetical relationships between languages. That was the time when the various language families and branches were discovered, for example the Germanic branch (of which English is a member) and a Proto-Indo-European parent language was reconstructed. In Comparative Philology the study of language was beginning to develop towards an autonomous, independent branch of study. Language began to be studied for its own sake. Besides, this kind of language study had an objective method: it was based on textual evidence, i.e. E-language facts, found in earlier written records of language, and it also tried to show language change in a systematic way, as a process determined by rules. (In the last quarter of the 19th century, a group of scholars in and around Leipzig, nicknamed the Neogrammarians, claimed that language changes were not just accidental events or optional tendencies, but "laws".) Meanwhile, the study of the contemporary state of languages went on in the non-scientific (or not sufficiently scientific) framework of Traditional Grammar.

1.3 WRITING SYSTEM AND CLASSIFICATION

Writing normally takes the form of visible marks on the surface of a relatively permanent object. The fact that writing is designed to be taken in by the eye leads to some characteristics that are essentially universal in modern writing. We discuss these properties of writing's outer form in what follows. Visual processing, whether of writing, spoons, or coins, requires a certain amount of time and attention. Perceivers must determine the number of visual objects that are present, the location of the objects relative to one another, and the identity of the objects. They must often act before they have had time to perform a full analysis of the incoming information. For example, readers may not take in all of the visual information that is available to them, perhaps not fixating on a word at all or not processing all of the letters in a word on which they do fixate. Readers may make decisions about which words are present before they have fully processed the information they have taken in. Some universal tendencies of writing systems have developed to make reading possible despite people's limitations in vision and attention, to allow them to do a reasonable job even when going quickly. One property of modern writing systems that aids visual processing is that there is a reasonable degree of contrast among the basic elements of a script. The distinction between the closed curve shape of ‹O› and the open curve shape of ‹C› is easy enough to see, but one would not expect a writing system to include several different C-like shapes that differ only by minuscule differences in the size of the opening. Nor would one expect to find a writing system in which a version of ‹C› with a narrow line and a version of ‹C› with a broader line constituted two different letters. To use a term traditional among historians of writing (Evans, 1894), the symbols of modern writing systems are linear. They require no shading, no fill, no color other than that needed to distinguish the writing from the background, and no distinctions between lighter and darker lines or wider and narrower lines. Another visual property of writing that helps people to distinguish its elements is that most scripts contain few or no elements that differ only in their left–right orientation. English has the mirror-image pair ‹p› and ‹q›, and Korean has ㅏand ㅓ, but most writing systems do not include any pairs of this sort. This is probably because people find it hard to learn to assign objects that differ only in orientation to different categories (Kolinsky et al., 2011). Yet another visual property of writing that helps both beginning and skilled readers is the visual redundancy within the elements of a script. In many cases, this redundancy can allow an element to

be successfully identified even if some of its visual characteristics are overlooked. For example, a reader of English who fails to detect the crossbar on ‹A› can still identify the letter because there is no other letter ‹Λ› in the Latin alphabet. According to one estimate, in fact, the identity of the elements of modern writing systems can be determined when, on average, half of the strokes are removed (Changizi & Shimojo, 2005). Within a writing system, the elements show a certain stylistic consistency. For example, Chinese characters, such as 圆 'round' and 球 'ball', are squarish and angular; they do not include full circles or semicircles as do a number of the letters of the Latin alphabet. The similarities among the elements of a script reflect the importance of the aesthetic qualities of writing, above and beyond the message that it conveys. The set of forms in (1a) is more pleasing to the eye than the set of forms in (1b) for example, (a) showing a set of Hebrew letters, which are stylistically similar to each other, and (b) showing Hebrew letters interspersed with Arabic letters. As we discuss later in the chapter, stylistic consistency can benefit readers, allowing them to develop a familiarity with the visual patterns that are shared across the elements of their writing system.

Writing also reflects many of the properties of the specific language that it represents. For example, it would be rare to see ‹the› at the end of an English sentence or を at the beginning of a Japanese sentence; this Japanese symbol in reality appears only immediately after the direct object of a verb. Likewise, readers of English would notice something peculiar about a word beginning with ‹ng›, ending with ‹pv›, or totally lacking a vowel letter. Such words are odd because the sounds they would represent violate the phonotactics of English, namely, the constraints on the ordering and position of phonemes in words. English words cannot begin with [ŋ], end with [pv], or lack a vowel sound. There should be no reason to come up with the spellings that represent these sounds, unless one is designing an experiment to tap people's processing of illegal sequences. Although the restriction against words like ‹ngim›, ‹lupv›, and ‹scvnkls› is motivated by the phonotactics of English, nothing would stop the beginning reader from learning about such restrictions as graphotactic irregularities—namely, those that violate the normal patterns by which letters are assembled

1.4 SYNCHRONIC AND DIACHRONIC ASPECTS OF LANGUAGE

He separated the synchronic and diachronic aspects of lsnguage study, and argued for the primacy of the former by saying that the synchronic aspect deals with language as a collection of simultaneous facts, existing as a state at a particular point of time, whereas the diachronic regards language as a succession of states, so it is the states that have to described first. • According to Saussure, linguistic signs enter into two kinds of relationship: syntagmatic and paradigmatic. The syntagmatic relationship is a linear (horizontal, chain) relationship, which exists between the signs that follow one another in a complex unit. For example, the four words in This coffee is strong. are in a syntagmatic relationship: they are placed one after the other along the syntagmatic axis, and each of the words has a particular environment or CONTEXT which consists of the other words on its left and right. The paradigmatic relationship is a vertical (choice) relationship, which exists between a sign present in a particular environment and all the other signs that could replace it while still yielding a well-formed complex unit. For instance, coffee in the above sentence is in a paradigmatic relationship with tea, student, girl, wall, light, whisky, cigar, etc

Leonard Bloomfield, and his followers, the Bloomfieldians, thought that a linguist should collect observable data, i.e. real utterances, and analyse these data, i.e. segment and classify the physical features of the utterances collected. A body of such data (a set of observed and collected utterances) is a corpus. Using a corpus for linguistic investigation is called the "corpus-based" or inductive procedure. In Chomsky's terminology this means that American structuralism was preoccupied with discovering and describing the E-language aspect of natural languages. The Bloomfieldians dealt with

phonetics, phonology, morphology, and syntax, but rejected semantics, thinking that the study of meaning would only be possible when human knowledge had become far more advanced. The only aspect of meaning that they paid attention to was whether two forms (signs or signcombinations) had the same meaning or different meanings. They used a strictly formal analysis. This was an analysis without reference to meaning, and it was based on an examination of distribution and constituency.
• The distribution of a language element (i.e. of a phoneme or morpheme or word) is the sum of all the environments in which it occurs. If two language elements always occur in different environments, i.e. they occur in mutually exclusive environments, then there is not even one environment in which one could replace the other. To put it differently: they never enter into a paradigmatic relationship with each other. In this case we say that the two language elements have totally different distributions: they are in complementary distribution. This means that where one of them can occur, the other cannot occur, and vice versa. For instance, the English phoneme /l/ has two variants, and they are in complementary distribution.1 The "clear" variant [l] (which is like Hungarian /l/) occurs before vowels, e.g. [»helIn] Helen, and the "dark" variant […] (which is pronounced with a cupped tongue, i.e. with a raising of the back part of the tongue) occurs elsewhere, i.e. before consonants, e.g. [he…p] help and in word- final position, e.g. [he…] hell. By contrast, if the distributions of two language elements are not entirely different, i.e. there is at least one common environment in which one could replace the other, the two elements are not in complementary distribution. In this case they are either in contrast or in free variation. Two language-elements are in contrast in a particular unit if replacing one by the other changes the meaning of the unit. For example, English /e/ and /i:/ in the environment /m—t/ are in contrast because /met/ does not mean the same as /miˇt/. If however replacing one language element by another in a particular unit does not change the meaning of the unit, they are in free variation in that unit. For instance, the same two phonemes /e/ and /iˇ/ are in free variation in the environment /—kˊ»nÅmIks/, because /iˇkˊ»nÅmIks/ means the same as /ekˊ»nÅmIks/.

• The other important method of formal analysis which the Bloomfieldians introduced was constituent analysis. (The Bloomfieldians themselves called it "immediate constituent analysis" or "IC analysis".) This means cutting syntactic units (or words) into their constituents, then the constituents into their constituents, and so on until we reach the individual words (or morphemes). Cutting a unit into its constituents is based on the test of substitution (replacement). For instance, the sentence My friend ran home. can be divided into two: [My friend] and [ran home] because My friend can be replaced by a simpler constituent, e.g. Peter, as in Peter ran home; and because ran home can also be replaced by a simpler constituent, e.g. slept, as in My friend slept. So we divide the sentence into [My friend] and [ran home], and then, through further applications of the substitution test, these parts can be divided into even smaller constituents. Constituent analysis can be visualised in essentially two ways, viz. by bracketings The constituents in the representations in (1) and (2) are labelled, S stands for Sentence, NP for Noun Phrase, VP for Verb Phrase, Det for Determiner, N for Noun, V for Verb, and Adv for Adverb. Trees and bracketings do not have to be labelled but the labelled ones are more informative than the unlabelled ones. For further information on syntactic representations. Constituent analysis was suitable for resolving certain ambiguities, by showing different constituent structures, e.g.: (old (men and women)) vs. ((old men) and women).

1.5 GENERATIVE LINGUISTICS

However, there were lots of ambiguities which constituent analysis could not resolve. For instance, The lamb is ready to eat. has two distinct meanings (is ambiguous), but the American structuralists could give it only one analysis: ((The lamb)(is (ready (to eat)))). Their analysis remained on the surface and could not disambiguate structures which were different in the deep. The growing dissatisfaction

with the limitations of structuralist linguistics led to the emergence of a radically new type of linguistic analysis towards the end of the 1950s. This has become known as transformational-generative linguistics, or just generative linguistics (= generative grammar), for short. This kind of analysis distinguishes two levels of syntactic analysis: a surface structure or Sstructure (which was recognised by the structuralists, too) and an underlying abstract deep structure or D-structure (which was not recognised by the structuralists). Transformational-Generative grammar is transformational because it explains surface structure as being derived from deep structure by a series of changes: transformations.

1.6 THE 'DESIGN FEATURES' OF LANGUAGE

Questions about the nature of language have preoccupied linguists for a long time. Answers have varied, and often reflect the biases and blind spots of various schools of Linguistics. Nevertheless, there is a degree of consensus. One particularly influential set of answers was provided by the linguist Charles Hockett, who refers to the things all languages have in common as the 'design features of language'. His work has been used as a basis for discussion of this topic in many later works (e.g. Yule, 2006; Widdowson, 1996; Aitchison, 1998). In the next few pages, we shall work through a selection of some of the main 'design features', with brief definitions. 'Design Feature' 1: Spontaneous use People do not need to be made to use language. Quite the opposite! It is sometimes hard to stop them doing so. Think how much time teachers and parents spend trying to get children to stop talking, and how the children wish the adults would 'stop going on so much'! And think of how noisy gatherings of adults usually are! There are two aspects to spontaneity:

1 Spontaneous acquisition:

Children do not need to be encouraged or coerced to acquire the language or languages around them. 2 Spontaneous use: Everyone – children and adults alike – uses language spontaneously, whether or not there is a practical need to do so.

'Design Feature' 2: Turn taking

Informal conversation is normally characterised by alternation between one speaker and another. In some other genres of language use one person engages in an uninterrupted monologue. Lectures and speeches are obvious examples. However, these monologues tend to be accompanied by elaborate social conventions, and to have clearly marked beginnings and endings, reinforcing the notion that frequent turn taking is the more 'normal' case. Writing in particular allows extended single-turn use of language. Nevertheless, even monologic writing is structured by the imagined presence and demands of the listener. (In this part, though I am doing all the 'talking', I am trying to involve you as much as I can with questions and exhortations!) Some have argued that turn taking is the first language skill evident in babies, on to which other skills are built (Locke, 1993, pp. 49–55). 'Conversations' are frequent between mothers and babies well before they have begun to use any words. Here, for example, the mother treats her baby's noises as though they are 'turns' in a conversation. Mother: Whatcha gonna tell me? Baby: (gurgling noise) Mother: Come on Whatcha gonna tell me? Baby: (Two gurgling noises. Squeak. Blows air through lips) Mother: That's a nice story. What else are you gonna tell me? Come on. Baby: (Gurgling and dribbling noise) Mother: Let's hear some more. This excerpt has been taken from Cook, G., (2003)

'Design Feature' 3: Redundancy

This term refers to the fact that languages are not the most economical possible communication systems. At every level, languages use more time and space than is necessary to encode information.

They could easily be cut down in size. Take for example 'I am'. The rules of English (unlike some 'pro-drop' languages) will not allow us to drop the pronoun. Someone called Sally cannot say '*am Sally' 1 as she might in Italian ('Sono Sally'). Nor can she drop the verb as in some other languages. She cannot say '*I Sally' as she might in Russian ('Ya Sally') or in Arabic ('Ana Sally'). Yet 'I' and 'am' duplicate the same information. Either one would do. This rule of English, in other words, demands a 'redundant' word. This example is from English, but there is nothing peculiarly English about this phenomenon. There is redundancy in all languages. Italian, Russian and Arabic may be more efficient when it comes to saying 'I am', but have plenty of redundancies elsewhere. Some linguists have claimed that up to 10% of any language system is redundant

These examples concern efficiency in expressing information. There is also however a social aspect of redundancy. 'Open the door' is a good deal shorter and more efficient than 'Would you be kind enough to open the door for me please'. Clearly there are other forces at work in language use other than mere efficiency.

'Design Feature' 4: Displacement

Languages enable us to talk with quite extraordinary precision about things which are not here or now. For example: I will meet you at 9.30 a.m. on the 5th April 2015 outside the café in the northwest corner of the square at the bottom of the road leading down from the Alhambra Palace in Granada in Andalusia in Spain. This utterance could enable two people to meet at exactly this place and time – even without mobile phones! There are a number of grammatical devices in language systems (which you will be studying in Part 2) to make this precision possible. This sentence, for example, uses embedded prepositional phrases, i.e. it puts one inside another: (at 9.30 a.m. (on the 5th April 2015)).

'**Design Feature' 5: Dual structure**

Language creates meaning through the combination of discrete units. These units, however, exist on two quite different levels. On the first level the units are a closed and small set called phonemes. Each phoneme is meaningless in itself, though there are cases where a word ('I' or 'a' for example) consists of a single phoneme. In English there are about forty. The actual sound which is identified as a particular phoneme will vary considerably. For example, the English phoneme /r/ is pronounced in a variety of ways, depending on region and dialect, or even, by speakers who cannot make any of these sounds, as the sound associated with the letter W. But from the linguistic point of view none of these variations alters its phonemic identity. All that matters is that it be recognised as a token of the phoneme /r/. On the second level, phonemes combine to form words, and words combine to form larger grammatical units. The word 'rat', for example, is a succession of three phonemes /r/ /æ/ /t/. This sign can in turn combine with others to form sequences with meanings of their own. This combinatory power enables languages to create a virtually infinite set of meaningful words and word combinations.

'Design Feature' 6: Discreteness

Language uses discrete signs. They are either one thing or another. A sound is perceived as one phoneme or another; there are no intermediate cases. (Though one could create a continuum of sounds between, say, /b/ and /p/, a speaker of a language with this distinction would perceive a sound along that continuum as one or the other.) Words, composed of phonemes, inherit this absolute quality of their components: a word is either 'bat' or 'pat' or another word, but there are no intermediate cases. Many acts of non-verbal communication, on the other hand, are graded. Like the dogs' barking you considered in Activity 4, they are a case of more or less, not either/or. If, for example, I smile at you while speaking, squeeze your hand, or laugh, I may increase or decrease the

breadth of my smile, the strength of my squeeze, or the loudness of my laugh, thus signifying more or less of whatever it is I mean by these actions.

'Design Feature' 7: Semanticity

The units created on the second level – whether they are words or combinations of words – have semanticity. That is to say, they mean things. They refer to something other than themselves: entities, ideas, states of affairs, feelings and so on. 'Rat' means a particular kind of rodent, and can be used to refer to one when it appears, or to talk about one which is not there. 'Design Feature'

8: Arbitrariness

The relationship between these meaningful units and their meanings, however, is arbitrary. The meanings exist by convention rather than for any logical or motivated reason. The word 'rat' means a particular type of rodent not because there is a connection between this sound sequence and the animal, but only because speakers of English know that it does, and know that other speakers of English know that it does. Only for these reasons, can we use the word successfully! (In this respect words are like bank notes.) There are a limited set of words which seem to break this principle, the so-called onomatopoeic words, whose sounds seem to mimic their meaning: 'whisper', 'tick', 'miaow', 'cuckoo'. However, many words which seem onomatopoeic to speakers of one language do not seem so to speakers of another, even when those languages belong to the same family (e.g. Russian 'ga' and English 'honk' for the noise of a goose). On the other hand, there are sometimes striking similarities between distant languages (the Japanese for 'miaow' is 'nyao').

'Design Feature' 9: Structure dependence

Phonemes combine to form words, and words combine to form larger units with meanings greater than, and independent of, the words which compose them. Perception of these composite meanings in these longer strings depends upon perception of underlying structures. For example: (The President) (is eating) (a hamburger). has the same structure – Subject Verb Object – as the much longer (The persistent Great White Shark which has followed our raft for ten days and nights) (is grabbing) (one of the goats which was tethered too close to the edge).

'Design Feature' 10: Prevarication

Language is used to convey information about states of affairs and states of mind. Yet, as we all know, these are not necessarily true. Language can just as easily be used to withhold information, or to give false information – to prevaricate – and deception is as universal as language itself. Some theorists have gone so far as to suggest that this capacity for deception is at the heart of the development of language both for the species and for individuals. Roger Brown (1973), a leading researcher in child language, pointed out that children are frequently told off for telling untruths but rarely corrected for their grammar. Nevertheless, they all grow up to speak grammatically and tell lies. Robin Dunbar (1996) has suggested that the origins of language may be less to do with its capacity to convey information, and more with its ability to form competing social networks, in which trust and deception are key factors. 'Design Feature'

11: Reflexivity

Language is used to talk about the world and our states of mind. It can also be used to talk about itself, that is, reflexively. This use of language is also referred to as metalanguage, and it takes many forms: • comment on a particular language item, e.g. a discussion of word meaning; • comment on a particular use of language, e.g. saying that Henry James used very complex sentences, or this politician

talks in clichés; • comment on a particular language, e.g. Finnish has fourteen cases, 'chaise' is the French for 'chair'; • comment on language itself, i.e. linguistics! (In this module, you are engaging in reflexivity – using language to talk about language!) 'Design Feature'

12: Productivity

Language is productive in two senses. First of all there are mathematical properties which allow it, as the nineteenth century linguist Wilhelm von Humboldt put it, 'to make infinite use of finite means' (Chomsky, 1996, p. 8). Dual structure, structure dependence, and grammatical operations such as recursion and coordination (which you will encounter again in Part 2) mean that the set of possible sentences in a language is infinite. It is always possible to insert or add on another bit. Things are said which have never been said before. This first kind of productivity, however, does not create new units or break rules for their combination. Another kind of productivity in language, however, is the capacity to create new items and new ways of combining them, to be, in other words, creative (Carter, 2004). People often depart from the rules creatively – to be poetic or humorous, or to assert identity. As so often this is most easily seen at the lexical level in the coinage of new words. Lewis Carroll's famous poem 'Jabberwocky', for example, used the invented words 'chortle' and 'galumph', which have now entered the language. Advertisements are full of such playfulness. An advert for the motoring organisation the R.A.C., for example, uses its name as a verb: 'We'll R.A.C. to it'.

1.7. PSYCHOLINGUISTICS

Psycholinguistics is the study of how the human mind produces and understands language. Although all language is at its heart creative, there are specific types of creative language that have attracted attention from language researchers due to an assumed deviation from a utilitarian mode of communication. Specifically, this entry reviews psycholinguistic research into figurative language (e.g., metaphor and irony) and humour, and summarizes current psycholinguistic investigations into how creative meaning is processed in the human mind. The entry ends by making connections between creative language and the larger field of creativity research.

Psycholinguistic training is an approach to training people in processes that they are believed to be deficit in. It was at one time a major intervention approach for students with learning disabilities. Training approaches were often based on results from a very popular test called the Illinois Test of Psycholinguistic Abilities (ITPA). This test measured integrative, receptive, and expressive linguistic abilities through presenting test subjects with information through visual and auditory channels. The assumption was made in this treatment approach, as in others, that discrete psycholinguistic abilities can be measured directly and then remediated. This assumption has been the subject of intense research scrutiny throughout the 1970s and 1980s. There were those who claimed with data that psycholinguistic training was generally not effective, and others with data who refuted these claims, stating that psycholinguistic training provided discrete benefits. The claims and counterclaims become a bit confusing. However, based on a review of the existing literature by this author, it appears that at least some of the areas measured by psycholinguistic assessments can be enhanced by psycholinguistic training. Particularly in the "expressive" areas of manual expression and verbal expression, evidence appears to confirm that these areas can be improved moderately through direct training. In the other 10 subareas measured by the ITPA, it appears that the effects of training are more modest.

Despite these moderately positive findings, a number of researchers have questioned the practical utility of gains of the magnitude reported. That is, it may be possible that gains in specific psycholinguistic variables are possible. The question is, do these gains translate into important gains

in functioning in other areas of a person's life such as reading or language use? In the absence of these types of evidence, we must question whether these psycholinguistic interventions should be the highest priority for persons with learning disabilities.

Computational psycholinguistics is a sub discipline of psycholinguistics, which is the scientific discipline that studies how people acquire a language and how they comprehend and produce this language. The increasing complexity of the models of human language processing which have been evolved in this discipline makes the development and evaluation of computer implemented versions of these models more and more important to understand these models and to derive predictions from them. Computational psycholinguistics is the branch of psycholinguistics that develops and uses computational models of language processing to evaluate existing models with respect to consistency and adequacy as well as to generate new hypotheses. Based on a characterization of the different tasks in human language processing, the article presents different computational models of these tasks. The architectural basis, the processing strategy, and the predictions made by the programs are described to show the merits of computer modelling in psycholinguistics.

1.8 SOCIOLINGUISTICS

Language is one of the most powerful emblems of social behaviour. In the normal transfer of information through language, we use language to send vital social messages about who we are, where we come from, and who we associate with. It is often shocking to realize how extensively we may judge a person's background, character, and intentions based simply upon the person's language, dialect, or, in some instances, even the choice of a single word.

Given the social role of language, it stands to reason that one strand of language study should concentrate on the role of language in society.

Sociolinguistics has become an increasingly important and popular field of study, as certain cultures around the world expand their communication base and intergroup and interpersonal relations take on escalating significance.

The basic notion underlying sociolinguistics is quite simple: Language use symbolically represents fundamental dimensions of social behaviour and human interaction. The notion is simple, but the ways in which language reflects behaviour can often be complex and subtle. Furthermore, the relationship between language and society affects a wide range of encounters--from broadly based international relations to narrowly defined interpersonal relationships.

For example, sociolinguists might investigate language attitudes among large populations on a national level, such as those exhibited in the US with respect to the English-only amendment--the legislative proposal to make English the 'official' language of the US. Similarly, we might study the status of French and English in Canada or the status of national and vernacular languages in the developing nations of the world as symbols of fundamental social relations among cultures and nationalities. In considering language as a social institution, sociolinguists often use sociological techniques involving data from questionnaires and summary statistical data, along with information from direct observation.

A slightly different concern with language and society focuses more closely on the effect of particular kinds of social situations on language structure. For example, language contact studies focus on the origin and the linguistic composition of pidgin and creole languages. These special language varieties arise when speakers from mutually unintelligible language groups need a common language for

communication. Throughout the world, there are many sociohistorical situations that have resulted in these specialized language situations--in the Caribbean, Africa, South America, Asia, and the Pacific Islands. In examining language contact situations, it is also possible to examine not only the details of a particular language but also the social and linguistic details that show how bilingual speakers use each language and switch between them.

Another approach to language and society focuses on the situations and uses of language as an activity in its own right. The study of language in its social context tells us quite a bit about how we organize our social relationships within a particular community. Addressing a person as 'Mrs.', 'Ms.', or by a first name is not really about simple vocabulary choice but about the relationship and social position of the speaker and addressee. Similarly, the use of sentence alternatives such as *Pass the salt*, *Would you mind passing the salt*, or *I think this food could use a little salt* is not a matter of simple sentence structure; the choice involves cultural values and norms of politeness, deference, and status

In approaching language as a social activity, it is possible to focus on discovering the specific patterns or social rules for conducting conversation and discourse. We may, for example, describe the rules for opening and closing a conversation, how to take conversational turns, or how to tell a story or joke.

It is also possible to examine how people manage their language in relation to their cultural backgrounds and their goals of interaction. Sociolinguists might investigate questions such as how mixed-gender conversations differ from single-gender conversations, how differential power relations manifest themselves in language forms, how caregivers let children know the ways in which language should be used, or how language change occurs and spreads to communities. To answer these questions related to language as social activity, sociolinguists often use ethnographic methods. That is, they attempt to gain an understanding of the values and viewpoints of a community in order to explain the behaviours and attitudes of its members.

Two trends have characterized the development of sociolinguistics over the past several decades. First, the rise of particular specializations within this field has coincided with the emergence of more broadly based social and political issues. Thus, the focus on themes such as language and nationalism, language and ethnicity, and language and gender has corresponded with the rise of related issues in society at large. Second, specialists who examine the role of language and society have become more and more interested in applying the results of their studies to the broadly based social, educational, and political problems that probably gave rise to their emergence as sociolinguistic themes to begin with. Sociolinguistics thus offers a unique opportunity to bring together theory, description, and application in the study of language.

1.9. NEUROLINGUISTICS

Neurolinguistics is the study of how language is represented in the brain: that is, how and where our brains store our knowledge of the language (or languages) that we speak, understand, read, and write, what happens in our brains as we acquire that knowledge, and what happens as we use it in our everyday lives. Neurolinguistics try to answer questions like these: What about our brains makes human language possible – why is our communication system so elaborate and so different from that of other animals? Does language use the same kind of neural computation as other cognitive systems, such as music or mathematics? Where in your brain is a word that you've learned? How does a word ‘come to mind’ when you need it (and why does it sometimes not come to you?)

If you know two languages, how do you switch between them and how do you keep them from interfering with each other? If you learn two languages from birth, how is your brain different from the brain of someone who speaks only one language, and why? Is the left side of your brain really ‘the

language side'? If you lose the ability to talk or to read because of a stroke or other brain injury, how well can you learn to talk again? What kinds of therapy are known to help, and what new kinds of language therapy look promising? Do people who read languages written from left to right (like English or Spanish) have language in a different place from people who read languages written from right to left (like Hebrew and Arabic)? What about if you read a language that is written using some other kind of symbols instead of an alphabet, like Chinese or Japanese? If you're dyslexic, in what way is your brain different from the brain of someone who has no trouble reading? How about if you stutter?

As you can see, neurolinguistics is deeply entwined with psycholinguistics, which is the study of the language processing steps that are required for speaking and understanding words and sentences, learning first and later languages, and also of language processing in disorders of speech, language, and reading.

How our brains work?

Our brains store information in networks of brain cells (neurons and glial cells). These neural networks are ultimately connected to the parts of the brain that control our movements (including those needed to produce speech) and our internal and external sensations (sounds, sights, touch, and those that come from our own movements). The connections within these networks may be strong or weak, and the information that a cell sends out may increase the activity of some of its neighbors and inhibit the activity of others. Each time a connection is used, it gets stronger. Densely connected neighborhoods of brain cells carry out computations that are integrated with information coming from other neighborhoods, often involving feedback loops. Many computations are carried out simultaneously (the brain is a massively parallel information processor).

Learning information or a skill happens by establishing new connections and/or changing the strengths of existing connections. These local and long-distance networks of connected brain cells show plasticity http://merzenich.positscience.com/?page_id=143 – that is, they can keep changing throughout our lives, allowing us to learn and to recover (to some extent) from brain injuries. For people with aphasia http://www.asha.org/public/speech/disorders/Aphasia.htm (language loss due to brain damage), depending on how serious the damage is, intense therapy and practice, perhaps in combination with transcranial magnetic stimulation (TMS), may bring about major improvements in language as well as in movement control; see the Aphasia section below, and the links posted there. Computer-based methods for enabling such intense language practice under the supervision of a speech-language pathologist are becoming available.

Where is language in the brain?

This question is hard to answer, because brain activity is like the activity of a huge city. A city is organized so that people who live in it can get what they need to live on, but you can't say that a complex activity, like manufacturing a product, is 'in' one place. Raw materials have to arrive at the right times, subcontractors are needed, the product must be shipped out in various directions. It's the same with our brains. We can't say that language is 'in' a particular part of the brain. It's not even true that a particular word is 'in' one place in a person's brain; the information that comes together when we understand or say a word arrives from many places, depending on what the word means. For example, when we understand or say a word like 'apple', we are likely to use information about what apples look, feel, smell, and taste like, even though we aren't aware of doing this. So listening, understanding, talking, and reading involve activities in many parts of the brain. However, some parts of the brain are more involved in language than other parts.

Most of the parts of your brain that are crucial for both spoken and written language are in the left side of the cortex of your brain (the left hemisphere), regardless of what language you read and how it is written. We know this because aphasia is almost always caused by left hemisphere injury, not by right hemisphere injury, no matter what language you speak or read, or whether you can read at all. (This is true for about 95% of right-handed people and about half of left-handed people.) A large part of the brain (the 'white matter') consists of fibers that connect different areas to one another, because using language (and thinking) requires the rapid integration of information that is stored and/or processed in many different brain regions.

Areas in the right side are essential for communicating effectively and for understanding the point of what people are saying. If you are bilingual but didn't learn both languages from birth, your right hemisphere may be somewhat more involved in your second language than it is in your first language. Our brains are somewhat plastic – that is, their organization depends on our experiences as well as on our genetic endowment. For example, many of the 'auditory' areas of the brain, which are involved with understanding spoken language in people with normal hearing, are used in (visually) understanding signed language by people who are deaf from birth or who became deaf early (and do not have cochlear implants). And blind people use the 'visual' areas of their brains in processing words written in Braille, even though Braille is read by touch

Bilingual speakers develop special skills in controlling which language to use and whether it is appropriate for them to mix their languages, depending on whom they are speaking to. These skills may be useful for other tasks as well.

Aphasia

What is aphasia like? Is losing language after brain damage the reverse of learning it? People who have difficulties speaking or understanding language because of brain damage are not like children. Using language involves many kinds of knowledge and skill. People with aphasia have different combinations of things that they can still do in an adult-like way and things that they now do clumsily or not at all. In fact, we can see different patterns of profiles of spared and impaired linguistic abilities across different people with aphasia.

Therapy can help aphasic people to improve on or regain lost skills and make the best use of remaining abilities. Adults who have had brain damage and become aphasic recover more slowly than children who have had the same kind of damage, but they continue to improve slowly over decades if they have good language stimulation and do not have additional strokes or other brain injuries.

Dyslexia and stuttering

What about dyslexia, and children who have trouble learning to talk even though they can hear normally? Why do people have reading difficulties? Research suggests that dyslexics have trouble processing the sounds of language and have difficulty relating the printed word to sounds. Genetic differences and genetically-based brain differences have been found in families with dyslexia and developmental language disorders, and research in this area is helping us understand how genes act in setting up the initial 'wiring' of all of our brains. There is solid evidence that appropriate language-based therapy is effective for children with developmental disorders of reading and language, including stuttering.

How neurolinguistic ideas have changed

Many established ideas about neurolinguistics – in particular, roles of the traditional 'language areas' (Broca's area, Wernicke's area) in the left hemisphere of the brain - have been challenged and in some

cases overturned by recent evidence. Probably the most important recent findings are 1) that extensive networks involving areas remote from the traditional language areas are deeply involved in language use, 2) that the language areas are also involved in the processing of non-language information, such as some aspects of music that the correlations of particular areas of the brain with particular language impairments are much poorer than had been thought. This new information has become available because of major improvements in our ability to see what is happening in the brain when people speak or listen, and from the accumulation and analysis of many years of detailed aphasia test data.

How neurolinguistic research has changed?

For over a hundred years, research in neurolinguistics was almost completely dependent on the study of language comprehension and production by people with aphasia. These studies of their language ability were augmented by relatively crude information about where the injury was located in the brain. Neurologists had to deduce that information, such as it was, by considering what other abilities were lost, and by autopsy information, which was not often available. A few patients who were about to undergo surgery to relieve severe epilepsy or tumors could be studied by direct brain stimulation, when it was medically needed to guide the surgeon away from areas essential for the patient's use of language.

Early-generation computerized x-ray studies (CAT scans, CT scans) and radiographic cerebral blood-flow studies (angiograms) began to augment experimental and observational studies of aphasia in the 1970s, but they gave very crude information about where the damaged part of the brain was located. These early brain-imaging techniques could only see what parts of the brain had serious damage or restricted blood flow. They could not give information about the actual activity that was taking place in the brain, so they could not follow what was happening during language processing in normal or aphasic speakers. Studies of normal speakers in that period mostly looked at which side of the brain was most involved in processing written or spoken language, because this information could be gotten from laboratory tasks involving reading or listening under difficult conditions, such as listening to different kinds of information presented to the two ears at the same time (dichotic listening).

Since the 1990s, there has been an enormous shift in the field of neurolinguistics. With modern technology, researchers can study how the brains of normal speakers process language, and how a damaged brain processes and compensates for injury. This new technology allows us to track the brain activity that is going on while people are reading, listening, and speaking, and also to get very fine spatial resolution of the location of damaged areas of the brain. Fine spatial resolution comes from magnetic resonance imaging (MRI), which gives exquisite pictures showing which brain areas are damaged; the resolution of CT scans has also improved immensely. Tracking the brain's ongoing activity can be done in several ways. For some purposes, the best method is detecting the electrical and magnetic signals that neurons send to one another by using sensors outside the skull (functional magnetic resonance imaging, fMRI; electro-enecephalography, EEG; magnetoencephalography, MEG; and event-related potentials, ERP). Another method is observing the event-related optical signal, EROS; this involves detecting rapid changes in the way that neural tissue scatters infra-red light, which can penetrate the skull and see about an inch into the brain. A third family of methods involves tracking the changes in the flow of blood to different areas in the brain by looking at oxygen concentrations (BOLD) or at changes the way in which the blood absorbs near-infrared light (near-infrared spectroscopy, NIRS). Brain activity can also be changed temporarily by transcranial magnetic stimulation (stimulation from outside the skull, TMS), so researchers can see the effects of this stimulation on how well people speak, read, and understand language. NIRS, EROS, ERP, and EEG techniques are risk-free, so they can ethically be used for research on normal speakers, as well as on

people with aphasia who would not particularly benefit by being in a research study. TMS also appears to be safe.

It is very complicated to figure out the details of how the information from different parts of the brain might combine in real time, so another kind of advance has come from the development of ways to use computers to simulate parts of what the brain might be doing during speaking or reading.

Investigations of exactly what people with aphasia and other language disorders can and cannot do also continue to contribute to our understanding of the relationships between brain and language. For example, comparing how people with aphasia perform on tests of syntax, combined with detailed imaging of their brains, has shown that there are important individual differences in the parts of the brain involved in using grammar. Also, comparing people with aphasia across languages shows that the various types of aphasia have somewhat different symptoms in different languages, depending on the kinds of opportunities for error that each language provides. For example, in languages that have different forms for masculine and feminine pronouns or masculine and feminine adjectives, people with aphasia may make gender errors in speaking, but in languages that don't have different forms for different genders, that particular problem can't show up.

1.10 COMPUTATIONAL LINGUISTICS

Computational linguistics is the scientific and engineering discipline concerned with understanding written and spoken language from a computational perspective, and building artifacts that usefully process and produce language, either in bulk or in a dialogue setting. To the extent that language is a mirror of mind, a computational understanding of language also provides insight into thinking and intelligence. And since language is our most natural and most versatile means of communication, linguistically competent computers would greatly facilitate our interaction with machines and software of all sorts, and put at our fingertips, in ways that truly meet our needs, the vast textual and other resources of the internet.

Computational linguistics (in historical perspective) delves in some detail into the essential concepts of linguistic structure and analysis (section 2), interpretation (sections 3–5), and language use (sections 6–7), as well as acquisition of knowledge for language (section 8), statistical and machine learning techniques in natural language processing (section 9), and other miscellaneous applications.

Goals of Computational Linguistics

The theoretical goals of computational linguistics include the formulation of grammatical and semantic frameworks for characterizing languages in ways enabling computationally tractable implementations of syntactic and semantic analysis; the discovery of processing techniques and learning principles that exploit both the structural and distributional (statistical) properties of language; and the development of cognitively and neuro-scientifically plausible computational models of how language processing and learning might occur in the brain.

The practical goals of the field are broad and varied. Some of the most prominent are: efficient text retrieval on some desired topic; effective machine translation (MT); question answering (QA), ranging from simple factual questions to ones requiring inference and descriptive or discursive answers (perhaps with justifications); text summarization; analysis of texts or spoken language for topic, sentiment, or other psychological attributes; dialogue agents for accomplishing particular tasks (purchases, technical trouble shooting, trip planning, schedule maintenance, medical advising, etc.);

and ultimately, creation of computational systems with human-like competency in dialogue, in acquiring language, and in gaining knowledge from text.

Methods of Computational Linguistics

The methods employed in theoretical and practical research in computational linguistics have often drawn upon theories and findings in theoretical linguistics, philosophical logic, cognitive science (especially psycholinguistics), and of course computer science. However, early work from the mid-1950s to around 1970 tended to be rather theory-neutral, the primary concern being the development of practical techniques for such applications as MT and simple QA. In MT, central issues were lexical structure and content, the characterization of "sublanguages" for particular domains (for example, weather reports), and the transduction from one language to another (for example, using rather ad hoc graph transformation grammars or transfer grammars). In QA, the concern was with characterizing the question patterns encountered in a specific domain, and the relationship of these question patterns to the forms in which answers might stored, for instance in a relational database.

By the mid-1960s a number of researchers emboldened by the increasing power and availability of general-purpose computers, and inspired by the dream of human-level artificial intelligence, were designing systems aimed at genuine language understanding and dialogue. The techniques and theoretical underpinnings employed varied greatly. An example of a program minimally dependent on linguistic or cognitive theory was Joseph Weizenbaum's ELIZA program, intended to emulate (or perhaps caricature) a Rogerian psychiatrist. ELIZA relied on matching user inputs to stored patterns (brief word sequences interspersed with numbered slots, to be filled from the input), and returned one of a set of output templates associated with the matched input pattern, instantiated with material from the input. While ELIZA and its modern chatbot descendants are often said to rely on mere trickery, it can be argued that human verbal behaviour is to some degree reflexive in the manner of ELIZA, i.e., we function in "preprogrammed" or formulaic manner in certain situations, for example, in exchanging greetings, or in responding at a noisy party to comments whose contents, apart from an occasional word, eluded us.

A very different perspective on linguistic processing was proffered in the early years by researchers who took their cue from ideas about associative processes in the brain. For example, M. Ross Quillian (1968) proposed a model of word sense disambiguation based on "spreading activation" in a network of concepts (typically corresponding to senses of nouns) interconnected through relational links (typically corresponding to senses of verbs or prepositions). Variants of this "semantic memory" model were pursued by researchers such as Rumelhart, Lindsay and Norman (1972), and remain as an active research paradigm in computational models of language and cognition. Another psychologically inspired line of work was initiated in the 1960s and pursued for over two decades by Roger Schank and his associates, but in his case the goal was full story understanding and inferential question answering. A central tenet of the work was that the representation of sentential meaning as well as world knowledge centered around a few (e.g., 11) action primitives, and inference was driven by rules associated primarily with these primitives; (a prominent exponent of a similar view was Yorick Wilks). Perhaps the most important aspect of Schank's work was the recognition that language understanding and inference were heavily dependent on a large store of background knowledge, including knowledge of numerous "scripts" (prototypical ways in which familiar kinds of complex events, such as dining at a restaurant, unfold) and plans (prototypical ways in which people attempt to accomplish their goals) (Schank & Abelson 1977).

More purely AI-inspired approaches that also emerged in the 1960s were exemplified in systems such as Sad Sam (Lindsay 1963), Sir (Raphael 1968) and Student (Bobrow 1968). These featured devices

such as pattern matching/transduction for analyzing and interpreting restricted subsets of English, knowledge in the form of relational hierarchies and attribute-value lists, and QA methods based on graph search, formal deduction protocols and numerical algebra. An influential idea that emerged slightly later was that knowledge in AI systems should be framed procedurally rather than declaratively—to know something is to be able to perform certain functions (Hewitt 1969). Two quite impressive systems that exemplified such a methodology were shrdlu (Winograd 1972) and Lunar (Woods et al. 1972), which contained sophisticated proceduralized grammars and syntax-to-semantics mapping rules, and were able to function fairly robustly in their "micro-domains" (simulated blocks on a table, and a lunar rock database, respectively). In addition, shrdlu featured significant planning abilities, enabled by the microplanner goal-chaining language (a precursor of Prolog). Difficulties that remained for all of these approaches were extending linguistic coverage and the reliability of parsing and interpretation, and most of all, moving from microdomains, or coverage of a few paragraphs of text, to more varied, broader domains. Much of the difficulty of scaling up was attributed to the "knowledge acquisition bottleneck"—the difficulty of coding or acquiring the myriad facts and rules evidently required for more general understanding. Classic collections containing several articles on the early work mentioned in the last two paragraphs are Marvin Minsky's Semantic Information Processing (1968) and Schank and Colby's Computer Models of Thought and Language (1973).

Since the 1970s, there has been a gradual trend away from purely procedural approaches to ones aimed at encoding the bulk of linguistic and world knowledge in more understandable, modular, re-usable forms, with firmer theoretical foundations. This trend was enabled by the emergence of comprehensive syntactico-semantic frameworks such as Generalized Phrase Structure Grammar (GPSG), Head-driven Phrase Structure Grammar (HPSG), Lexical-Functional Grammar (LFG), Tree-Adjoining Grammar (TAG), and Combinatory Categorial Grammar (CCG), where in each case close theoretical attention was paid both to the computational tractability of parsing, and the mapping from syntax to semantics. Among the most important developments in the latter area were Richard Montague's profound insights into the logical (especially intensional) semantics of language, and Hans Kamp's and Irene Heim's development of Discourse Representation Theory (DRT), offering a systematic, semantically formal account of anaphora in language.

A major shift in nearly all aspects of natural language processing began in the late 1980s and was virtually complete by the end of 1995: this was the shift to corpus-based, statistical approaches (signalled for instance by the appearance of two special issues on the subject by the quarterly Computational Linguistics in 1993). The new paradigm was enabled by the increasing availability and burgeoning volume of machine-readable text and speech data, and was driven forward by the growing awareness of the importance of the distributional properties of language, the development of powerful new statistically based learning techniques, and the hope that these techniques would overcome the scalability problems that had beset computational linguistics (and more broadly AI) since its beginnings.

The corpus-based approach has indeed been quite successful in producing comprehensive, moderately accurate speech recognizers, part-of-speech (POS) taggers, parsers for learned probabilistic phrase-structure grammars, and even MT and text-based QA systems and summarization systems. However, semantic processing has been restricted to rather shallow aspects, such as extraction of specific data concerning specific kinds of events from text (e.g., location, date, perpetrators, victims, etc., of terrorist bombings) or extraction of clusters of argument types, relational tuples, or paraphrase sets from text corpora. Currently, the corpus-based, statistical approaches are still dominant, but there appears to be a growing movement towards integration of

formal logical approaches to language with corpus-based statistical approaches in order to achieve deeper understanding and more intelligent behaviour in language comprehension and dialogue systems. There are also efforts to combine connectionist and neural-net approaches with symbolic and logical ones.