Squire, L.R. and Zola-Morgan, S. 1991. "The medial temporal lobe memory system". *Science* 253.
Subramaniam, S., Biderman, I., and Madigan, S. 2000. "Accurate identification but no priming and chance recognition memory for pictures in RSVP sequences". *Visual Cognition* 7.
Swinney, D. 1979. "Lexical access during sentence comprehension: (Re)consideration of context effects". *Journal of Verbal Learning and Verbal Behavior* 18: 645–59.
Treisman, A.M. and DeSchepper, B. 1996. "Object tokens, attention and visual memory". In *Attention and Performance XVI: Information Integration in Perception and Communication*, T. Inui and J. McClelland (eds). Cambridge, Mass.: MIT Press.
Treisman, A.M. and Kanwisher, N.G. 1998. "Perceiving visually presented objects: Recognition, awareness, and modularity". *Current Opinion in Neurobiology* 8: 218–26.
Tucker, D.M 1991. "Developing emotions and cortical networks". In *Developmental Neuroscience, Minnesota Symposium on Child Psychology*, Volume 24, M. Gunnar and C. Nelson (eds). Hillsdale, NJ: Lawrence Erlbaum Associates.
Ullman, S 1994. "Sequence seeking and counterstreams: A model for bidirectional information flow in the cortex". In *Large-Scale Neuronal Theories of the Brain*, C. Koch and J.L. Davis (eds). Cambridge, Mass.: MIT Press.
Ungerleider, L.A. 1985. "The cortical pathways for object recognition and spatial perception". In *Pattern Recognition Mechanisms, The Pontifical Academy of Sciences*, C. Chagas, R. Gattass, and C. Gross (eds), 21–37. Berlin/New York: Springer Verlag.
Ungerleider, L.A. and Mishkin, M. 1982. "Two cortical visual systems". In *Analysis of Visual Behavior*, D.G. Ingle, M.A. Goodale, and R.J.Q. Mansfield (eds). Cambridge: MIT Press.

CHAPTER 8

The clausal structure of linguistic and pre-linguistic behavior

Gertraud Fenk-Oczlon and August Fenk
University of Klagenfurt

1. Co-evolution of language and cognition: mutual stimulation and acceleration

The terms "evolution" and "co-evolution" imply pre-existing structures which undergo some developmental changes such as differentiation, specialization, and integration, and they imply that the scope of possible developmental directions is predisposed in these already existing structures.

Evolutionary processes are initiated or at least accelerated when the environment (e.g. the climate in a certain habitat) changes or when organisms change their environment, e.g. by conquering new habitats or niches, remigrating from land to water, etc. In the co-evolution of a system's (system 1) subsystems (1.1., 1.2. etc), the situation is similar in this respect. If we regard language (=subsystem 1.2.) as an environmental system of the cognitive system (=subsystem 1.1.), we may say that an increasing complexity of a community's language makes growing demands on relevant cognitive abilities of the "users" of this language. Individuals who are above average or extremely highly equipped with such abilities, gain higher, more "elaborated" levels in the use of this cognitive and communicative tool, and — therefore — higher chances to transmit this "equipment" to the next generation (generation $n+1$).[1] This selective mechanism is again effective in the transmission from generation $n+1$ to $n+2$, and so forth, and the more efficient, the higher the utility of the "organon" language for the individual user within his linguistic community and the higher its utility for the linguistic community as a whole.

The utility for the community (the family, the tribe, bigger societies) is immense. The activity of our cognitive apparatus is, first of all, concerned with the task of reducing uncertainty and of making events (e.g. the consequences of one own's actions) foreseeable by extracting invariants, rhythms, regularities, redundancies in the observations possible so far (Fenk and Vanoucek 1992). Anticipation allows "rational" decision-making under uncertainty in practical or even existential matters. To be able to communicate thoughts about indexical properties of events, about possibilities, conditionalities, contingencies, and probabilities, to communicate risks and chances — all that makes the social group not only fit for survival in a certain habitat, but makes it extremely fit for conquering new habitats and for adapting the habitat in an advantageous way. And the ability to succeed to meta-linguistic levels offers an extraordinarily high potential for completely new intellectual developments.

Thus, the phylogenetic and ontogenetic development of a complex language requires both: each step forward in this evolution has to allow for the actually reached "standard" of relevant cognitive capabilities. (It has to meet the demands of "cognitive ergonomics".) And each step in the complication of the "tool" language has to be accompanied (or proceeded, or followed) by an advance in the capabilities for acquisition and efficient use of this "organon". Developmental changes of language (subsystem 1.2.) stimulate and accelerate the development of cognition (subsystem 1.1.) and vice versa. This mechanism of mutual stimulation and mutual acceleration is the "motor" of the "self-acceleration" of the superordinate system 1. This view also holds, if — as in Wills (1993) or in Deacon (1997) — not cognition, but its biological substrate is taken as the language's partner in co-evolution: "The key to this is the co-evolutionary perspective which recognizes that the evolution of language took place neither inside nor outside brains, but at the interface where cultural evolutionary processes affect biological evolutionary processes". (Deacon 1997:409).

Our recent and complicated languages most probably are traced back to less complicated rudimentary predecessors in vocal, mimic and gesture communication. New steps in evolution are grounded on pre-existing structures (see first paragraph of this section), and the introduction of a "new" referential system (like "picture language" and "diagrammaticity") will always refer to already existing referential systems (Fenk 1998). This is a matter of "pragmatics" and "cognitive economy", if not of necessity.

The aim of this paper is not to reconstruct the development of "pre-language" to "language", nor to define a borderline[2] between "pre-language" and "language", nor is it concerned with possible "ritual beginnings" (Deacon 1997:401) of language and "symbolic reference", or with the changes from an initially visually-gesturally coding language to auditory coding (Givón 1998:103, in this volume). Instead, it concentrates on some cognitive pre-conditions as a "matrix" allowing for, as well as constraining the development of a complex language system.

2. The pre-linguistic matrix of language

The impulses for typological differentiation and diachronic change of languages may come from the social-communicative environment and from language contact. But the evolution and possible forms of variation of language are determined by the pressure to improve, or maintain at least, the language's efficiency as a communicative and cognitive tool in a way that accounts for the constraints of the articulatory and the cognitive system.

These goals[3] and constraints can be seen as a matrix preforming language, and this matrix will be discussed in two respects: with regard to the goal of communicating thoughts about possibilities beyond the *hic et nunc* (section 2.1), and to the "clausal" structure of cognition and psychomotility (section 2.2) predisposing the clausal structure of language.

2.1 Mental propositions

Many psycholinguistic studies (already in the 1970s) view propositions as basic semantic entities of the organization of (sensory-motor and iconic) memory. This cognitive structure cannot be described in solely linguistic terms; but when it is verbalized it is likely to take the form of a predicate-argument-structure (Fenk-Oczlon1983:30). Givón (1990:896) states that something "like a mental proposition, under whatever guise, is the basic unit of mental information storage", and in Kintsch (1974:12) propositions are characterized as "n-tuples of word concepts, one of which serves as a predicate and the others as arguments".

According to Deacon (1997:334) "this operand-operator structure (and probably subject-predicate structure) /. . ./ is the minimum requirement to make the transition from indexical to symbolic reference". He states that already the earliest symbolic systems would necessarily have been combinatorial in this sense.

Considering the universal appearence of simple declarative sentences in all languages which are apt to transcend the *hic et nunc* and to communicate about assumptions etc. (see section 1), the crosslinguistic experiments described in section 3.2. used so called "kernel sentences". A "kernel sentence" can be described as a prototypical case of clauses: a simple declarative sentence encoding **one** proposition (in the sense of a pre-linguistic entity) within **one** intonation unit in the form of a predicative syntagma (Fenk-Oczlon and Fenk 1995:231).

In written communication, where the respective text also serves as an "external memory", long, complex, and embedding sentences are much more frequently used than in everyday oral communication — even in our literate society with its well trained skills in the reception of long and complex sentences.[4] In languages without a writing system or without a longer literate tradition — like Cayuga (Sasse 1991:204) or Eipo (Heeschen 1994) — a content that we would rather encode in a complex sentence is encoded in a series of minimal-predications, each of them included in a single intonation unit.

2.2 Segmentation of extralinguistic activities and the magical number seven, plus or minus two

In the psychology of information processing the number seven is a somewhat "magical" invariant: It manifests itself as a constraint of the span of absolute judgement, the span of immediate memory, and the span of attention. But Miller (1956) warned assuming "that all three spans are different aspects of a single underlying process". This limit of about seven (plus minus two) has since figured prominently in information processing theories.[5]

Linguistic information is a special type of information processed by our cognitive apparatus. If the number seven marks some general limits of this apparatus, it should also show in languages, because languages must have developed in adaptation to the general constraints of this apparatus.[6]

A main concern of Fraisse's meanwhile classical book on the psychology of time (Fraisse 1985/1957) is the "psychological present" or "immediate memory span". He states that this span cannot be understood as a window of consciousness, on which reality moves by, but rather as a beat in rhythmically organized attentional processes. One of his arguments (Fraisse 1985:89 f.) says that we perceive the *tick-tack* of a watch always as this *tick-tack,* and not as a *tick-tack* followed by a *tack-tick* followed by a *tick-tack* etc., or that we perceive a waltz always as sequences of low-low-loud and not as a low-low-loud followed by low-loud-low followed by loud-low-low. In this context he already refers to the comprehension of sentences and states that the psychological present always corresponds to one sentence and never to the end of a sentence together with the first part of the following sentence with running elements (p 90). Regarding the **size** of the psychological present he mentions a series of strikes of a church bell: in the case of 3 or 4 strokes we know the indicated time immediately, without any counting. This is impossible with 12 strokes indicating the midnight hour. Children not yet able to count are able to correctly reproduce series of about 5 or 6 strokes.

In his experiments (Fraisse and Fraisse 1937, as discussed in Fraisse 1985:93) on the retention of sound series, he had varied the between-sound intervals: 0.17sec, 0.37 sec, 0.63 sec, 1.2 sec, 1.8 sec. For these intervals the following mean number of sounds could be reproduced: 5.7, 5.7, 5.4, 4.0, 3.3. The overall duration of the retained series was 0.8 sec in the case of the 0.17 sec interval and 4.2 sec in the case of the 1.8 sec interval. From this he concluded

- that the optimal intervals for perception are in the area betwen 1.5 and 0.7 sec;
- that the psychological present is not a content-independent scope of time;
- that a number of 5 to 6 elements marks the limit of the perception of succession; and

- that the duration of the perceived present shows a higher variability than the number of perceived elements.

Such limits regarding the **number of elements** he assumes to be not only independent of sense modality — 6 to 7 elements was also a maximum in the retention of visual stimuli (Fraisse and Fraisse 1937) — but also of the "level of intelligence": Various birds such as daws can learn to look at first in that box for food on which a certain number (let us say 5) is represented by an arbitrary array of dots. (Only that box contains food. The position of this box among other boxes representing other numbers as well as the array of dots on all of the boxes is changed from trial to trial.) This works up to about a number of 7 dots.[7]

Corresponding results regarding the relevance of a fixed number of elements showed also in non-human primates. Rhesus Macaques were the "subjects" of the two following studies:

Swartz *et al.* (2000) report on experiments concerning the monkey's ability to memorize arbitrary sequences of arbitrary stimuli. Each of their subjects mastered 4 novel 4-item lists by trial and error, and the performance of some of the subjects suggested "that, with more list-learning experience, monkeys could master longer lists and determine the ordinal position of items on new lists with fewer errors" (Swartz *et al.* 2000:284).

Experiments conducted by Brannon and Terrace (2000) "showed that monkeys trained on an ascending numerical rule spontaneously infer the ordinal values of novel numerosities when numerosity varied from 5 to 9 and "that neither language nor numerical symbols are necessary for discriminating and ordering visual stimuli on the basis of their numerosity." (Brannon and Terrace 2000:48). This reminds us, in some respects at least, of the already mentioned findings of Köhler (1952) in his experiments with daws.

Results of brain research concerning the **duration** of perceptual and memory span are in line with Fraisse's findings as well. Elbert *et al.* (1991) report on experiments in which the subjects had to reproduce the duration (1, 2, 3, 4, 6, 8 sec intervals) of the illumination of a screen.

> Reproduction was accurate for standard intervals up to 3 s but deteriorated with increasing interval length. Brain potentials during reproduction intervals of 1–3 s differed from those recorded during the longer intervals. A CNV-like slow negative shift developed during the shorter reproduction intervals. Negativity was reduced or even absent, when subjects had to reproduce standard intervals of 4 s or longer. (Elbert *et al.* 1991:648)

The authors concluded that short intervals with a maximum duration of 3–4 sec are processed primarily in left-hemispheric, frontal and temporal cortical networks and in a mode that differs qualitatively from processing events of longer duration.

Regarding the duration of the "perceived present", Fraisse (1985:96) had mentioned, among other examples, some visually ambiguous figures such as a cube with solid drawn edges throughout and with equal size of whatever one may perceive as the front plan or as the more distant back plan. The intervals in which the "foreground" subjectively becomes the "background" and then the other way round is assumed to reflect the periodicity of attentional shifts. According to more recent studies "the turnover time lies within a few seconds with a peak around 3 seconds" (Schleidt and Kien 1997:98). In these studies the respective cube representation is referred to as "Necker cube" (see also Pöppel 1986).

Rhythmical segmentation or "clausal" structure is, as we have seen, neither restricted to auditory perception nor to the species *Homo sapiens*. It is, moreover, not restricted to primarily perceptual/cognitive activities, but is a characteristic of motor activities as well. (Not surprisingly if one recalls the cognitive components of motor programming and the back-coupling of the effector-system to the cognitive system via the kinesthetic system.)

The aim of the above mentioned study of Schleidt and Kien (1997) and of Feldhütter *et al.*(1990) was an intercultural comparison of the duration and internal structure of movement patterns. A fundamental finding was that different sorts of behavior could be divided into small segments of functionally related movements. The authors' criteria for defining a sequence of movements as a separate action unit were:

> A sequence of movements was regarded as an action unit only if its beginning and end could be clearly seen, if it had an observable goal, and thus the individual movements are functionally related (e.g. wiping the nose with the finger). The consecutive movements do not belong to the same action unit if they are not functionally related (e.g. wiping the nose and then scratching it). (Schleidt and Kien 1997:79)

A further result was: from a total of 1542 action units analyzed, 93 % had a duration of 2–3 seconds.

Thus, segmentation in the range of a few seconds appears to be a universal phenomenon in human action. We find it in various cultures all over the world, not only in conscious and intentional acts like work activities, actions of communication like greeting behavior, and ritualized behavior such as dances, but also in less conscious and less culturally influenced activities like scratching one's body or playing with a baby (Schleidt and Kien 1997:81).

Such a segmentation (into somewhat shorter segments) was also observed in non-human primates, for instance in the hand movements of chimpanzees (Kien *et al.* 1991). And in Schleidt (1992) it is stated that in a higher number of movements per action unit the movements have to be shorter in order to fit into the restricted time span. (This gives the impression that the duration of the unit is

rather the invariant dimension, and not the number of elements as suggested in Fraisse 1985:93)

A central idea of the following section is that the **intonation units** in language are a special case of **action units**. To be more general: that the central nervous mechanisms underlying the segmentation of other activities are also effective in language behavior and are responsible for the clausal structure of both the perception and production of language (see Barker and Givón in this volume).

Language behavior is not only a matter of perceptual and a matter of productive activities, but a matter of intimate coordination between perception and articulation. In the role of the speaker, for instance, one is always also the hearer of one's own output. The possibility of perceiving and controlling one's own output, and the coordination and synchronization of the activities involved, must be considered as an essential characteristic of any communication system capable of development.(Similarly the possibility of the visual control of our own manual activities is often viewed as a prerequisite of tool-using and tool-manufacturing activities.)

Important indications for such a coordination between speaking and memory for the spoken come from experiments concerning the "articulatory loop hypothesis" (e.g. Baddeley 1986): One is able to recall as many words as one could pronounce in about 2 seconds. And the mapping between the psychologists' estimations of the duration of immediate memory (2 plus minus 1 sec) and the linguists' estimation of the duration of intonation units (see following section) will hardly be a mere accident.

3. Segmentation in language behavior

Spoken language is obviously segmented: "tone unit" (Quirk *et al.* 1964), "breath group" (Lieberman 1967), "tone group" (e.g. Halliday 1967) and "intonation unit" (e.g. Chafe 1994) — all these terms refer, though not always in identical way, to a prosodic unit which comprises a sequence of words spoken under a single intonation contour.

3.1 The duration of intonation units and clauses

Studies really measuring the duration of units seem to be rather rare. Määttä (1993:109) studied the "portion of speech between two measurable pauses". In Finnish he found an average length of "breath groups" in the region of 2.1 to 2.2 sec. (The mean duration of breath groups inclusive pauses comes to 3.2 to 3.3 sec.) And according to Chafe (1987:22), "new intonation units typically begin about two seconds apart. Evidently active information is replaced by other, partially different information at approximately two second intervals."

Most usually the length of intonation units is not measured in units of time but in number of words. Chafe (1994), for example, reports a mean length of 4.84 words per substantive intonation unit in English. Croft (1995), referring to Altenberg (1990) and Crystal (1969), reports similar numbers, ranging from 4 to 6 words per intonation unit. And when comparing English intonation units with intonation units in Seneca, Chafe (1994:148) found that "with respect to number of words, then, Seneca intonation units are half as long as English ones." This shows that the average length of an intonation unit, when measured in words, is highly dependent on the language in question, and especially on its morphological type. In languages with a pronounced tendency to synthetic (agglutinative or fusional) morphology we have to expect a lower number of words per intonation unit (and in polysynthetic and incorporating languages even one long word that we would encode in a sentence comprising 5 or 6 words.) This is one of many reasons why the crosslinguistic study reported in the next section took the number of syllables as an appropriate measure for the size of a basic type of intonation unit.

3.2 Segmentation within clauses and again the magical number seven, plus or minus two

In spoken language there are only two entities corresponding to **rhythmic** processing — the **syllable** as the basic element and the **clause** or **intonation unit** at a higher order level. (The unit "in between" these two levels is the word. It is the most widely used material in memory experiments and is of course interesting because of its semiotic status. But it is not the appropriate candidate in the search for elements and components of rhythmic organization.) Thus, an experimental design was applied in order to study crosslinguistic regularities of within-clause segmentation. A central assumption of this crosslinguistic study (Fenk-Oczlon 1983) is that *the number of syllables per "clause" will vary within the range of the magical number seven plus minus two.* The clauses used were of a special quality: simple declarative sentences encoding **one proposition in one intonation unit**, such as *blood is red* or *the sun is shining.* 22 German sentences of this sort were presented to native speakers of 27 different languages. (Meanwhile, the sample was extended to 34 languages, 18 Indo-European, 16 Non-Indo-European (Fenk-Oczlon and Fenk 1999)). Native speakers were asked to translate the sentences into their mother tongue and to determine the length of their translations in syllables. The mean number of syllables per clause, computed for each one of these languages, was found to be located almost exactly within Miller's (1956) often quoted range of 7 plus minus 2 elements: The lowest size was 5.05 syllables (Dutch), and only Japanese with 10.2 syllables per clause was located outside the hypothesized range of 5–9 syllables. The overall length was 6.48 syllables per simple clause.[8]

A statistical reanalysis (Fenk-Oczlon and Fenk 1985) was intended to clarify the question about the relevant factor determining the position of single languages on the continuum "mean number of syllables per sentence". The factor which came under suspicion was syllable complexity, i.e. a language's mean number of phonemes per syllable. One cause of suspicion was that the articulation of more complex syllables takes up more time. Another cause was that Dutch is known for its complex syllables and Japanese for its simple CV-syllables. The result was a highly significant negative correlation ($r=-0.77$, $p<0.1\%$) between mean number of syllables per sentence and mean number of phonemes per syllable. The interpretation therefore is that: languages with higher syllable complexity use, proportionate to the higher expenditure of time per syllable, fewer syllables for encoding a certain proposition.

In other words: for encoding a certain proposition, languages with simple syllable structure need the same time but more syllables than languages with high syllable complexity. The fact that in cross-linguistic comparison the number of syllables per simple sentence was found to be located in the area of 5–10 syllables (see Figure 1) agrees with our immediate memory span comprising about 5–9 units. And the location in this area corresponds to time-related limits, which might be operative at the level of syllable perception (and production): 200–300 milliseconds seems to be the duration necessary for auditory pattern recognition

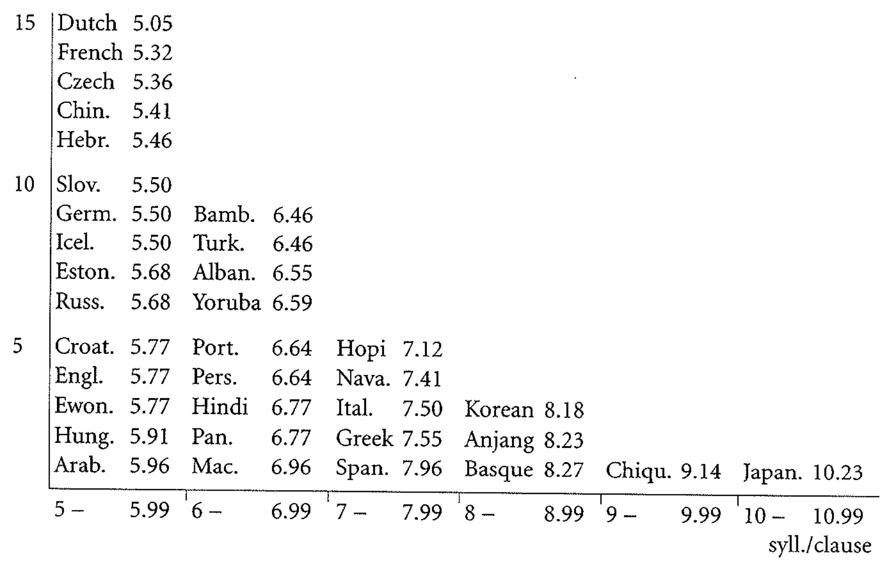

Figure 1. The frequency distribution of languages over different classes of the parameter "mean number of syllables per clause" (slightly changed from Fenk-Oczlon and Fenk 1999:158)

(Massaro 1975) and for producing the right-ear advantage in dichotic-listening experiments.

If the duration of a simple sentence coincides with our "psychological present" (ca. 2 seconds) and if the minimum duration of a syllable is estimated at ca. 200 milliseconds, then the sentence comprises 10 syllables in a "pure CV-language" (see Japanese in Figure 1) and a lower number of syllables in the case of more complex syllables (CVC, CCVC, CCVCC, . . .), proportionate to the longer duration of these more complex syllables. In this respect, at least, there seems to be nothing magical in the "magical number seven".

The upper limit (2–3 sec per clause or simple sentence) and the lower limit (200–300 millisec per syllable) are operative in the rhythmic pattern organisation, and they might be operative like set points in the self-regulation of language systems, constraining for instance the typological differentiation of languages with regard to morphosyntactic structure and complexity of syllables.

Our negative crosslinguistic correlation between number and complexity of syllables points, first of all, to the efficency of such time related constraints, and so does the whole set of correlations — e.g. the more words per sentence, the fewer syllables per word — found in a later study (Fenk and Fenk-Oczlon 1993). More generally we may conclude:

- The central unit in the rhythmic organization of language is the clause, and the clausal structure of our language is apparently preformed (phylogenetically as well as ontogenetically) by the tonal utterances that we can isolate in the (prelinguistic?) communicative behavior of our infants and of recent non-human primates.
- The variability of clauses is, first of all, restricted in terms of duration. But this relatively "constant" time interval is also filled by a rather "constant" number of elements and a rather "constant" amount of information — probably due to economy principles providing "packages" with an optimal size for cognitive "handling".
- Optimal size of packages is achieved by trade-offs between sizes of within-clause elements (see our correlations) and between the size (complexity, duration) of these elements and the information carried by them: Higher token frequency goes hand in hand with lower size of elements (e.g. Zipf 1929) and, moreover, with lower information per element. (Higher relative frequency means higher probability and corresponds with lower information per mathematical definition. Similarly, higher "subjective" probability — or higher "familiarity" — results in lower "subjective" information and lower cognitive load.)
- Quantitative relations found between relevant dimensions in recent languages — be it within a single language (e.g. Menzerath's law)[9] or in crosslinguistic computation — reflect those dynamic and economic principles that govern diachronic change and differentiation of language and probably the evolution of language out of pre-language.

4. Discussion

Dealing with the "roots" of recent human languages one probably should search for the — pre-linguistic or language-bound? — nature of propositional thinking, of man's propositional view of the world. Doing so, one might start with regularities of Gestalt-perception structuring our perceptual world by bringing about coherence between "elements" and separating groups of "elements".[10] From there one might proceed to our perceptual interpretation of — indexical, causal — relations and interactions between those objects separated by Gestalt-principles. Such "causal" interpretations on a **perceptual** level have been described by Michotte (1946). Further analyses might then be concerned with the size or duration of what can be kept within the conscious field (Mandler 1975) — be it a linguistically encoded proposition or an extralinguistic event which becomes the object of a "propositional" interpretation.

We did mention certain "positive" presuppositions of language evolution but concentrated on the "constraining" and "preforming" properties of these presuppositions. Such constraints can be observed on both levels, the **duration** of perceptual and action units and the level of the **number of elements**, and in both relevant systems, the afferent/perceptual and the efferent/motoric. Arguments were presented for an intimate perceptual/motoric coordination and back-coupling, especially in linguistic behavior.

Later studies on (a) linguistic and (b) pre-linguistic behavior correspond to the crosslinguistic and significant **negative correlation** found between number of syllables and complexity of syllables (Fenk-Oczlon and Fenk 1985):

(a) In the already mentioned linguistic study of Määttä (1993) a connection was found between syllable duration (in msec) and "length" of breath group, i.e. the number of syllables per breath group. He reports an inverse relationship between syllable duration and length of the breath group, and that the longest average syllable durations are restricted to short breath groups.

(b) Moreover, our negative correlation is in line with Feldhütter's *et al.* (1990) observation in extralinguistic and pre-linguistic action units: The higher the number of movements within the action unit, the shorter the duration of a single movement. In the case of spoken language these "single movements" are the articulatory movements forming a single syllable. Higher complexitiy of the single syllables results in longer duration of articulation and perception and in a lower number of syllables per clause. What remains relatively invariant and unaffected by this crosslinguistic within-clause variability is the duration of the clause and the information transmitted.

The underlying span of (about two seconds and) about seven syllables has, according to our findings, the appropriate size for encoding one proposition. It has the appropriate size for allowing the extraction of the meaning of a clause before moving to the following clause, or, to use Mandler's (1975:236) words again, the size "that can be kept within the conscious field, the focal attention."

Notes

1. The selective advantage of tool-using abilities and their effect on cerebral growth is not restricted to the "tool" language, but is to be assumed for *tools* in the literal sense and for cultural techniques in general (Fenk-Oczlon and Fenk 1999). Rather specific neurophysiological and -anatomical arguments (of Lieberman 1991 and Greenfield 1991) aim, according to Beaken (1996:17), at a close relationship between language and manual skills.

2. Such attempts are most commonly characterized by rather restrictive and arbitrary definitions of what is the constitutive property of "full" language. And "full" syntax which is most commonly suggested to be constitutive for "full" language, is again (e.g. Bickerton 1990) defined in a completely arbitrary way. Comprehensive criticism of Bickerton's attempt and of some other relevant approaches is presented in Sampson (1997). See also Givón's criticisms of any attempt to reduce language evolution "to a single initial cause" (Givón 1998:105).

3. The term "goal" is just a substitute for circumstantial phrases like "advantages leading via selection mechanisms to . . .", and should in no way indicate a teleologic or vitalistic position of the authors.

4. Apart from the question of, whether or not a longer literate tradition is a prerequisite for a change from paratactic to hypotactic constructions (cf. the short discussion in Sampson 1997:74), it has to be accepted as an empirical fact that in everyday oral communication hypotactical constructions are relatively rare. Impressive but unfortunately not yet published examples of an almost exclusive use of (fragments of) paratactic constructions are the dialogs recorded by Gigler (in preparation) in Carinthian.

5. One of the most recent and most interesting relevant findings is reported by Kareev (2000): small series of about 7 plus minus 2 data-pairs produce stronger correlations between the respective variables than the population. This would mean that a span of comprehension comprising about 7 elements or chunks of elements does not reflect a rather arbitrary perceptual/cognitive limit, but, furthermore, that there must have been a selective advantage and selective pressure to push up the limit to this region where minimal indications and minimal contingencies (see our section 1) can be detected with a minimum of "computational" work.

6. The present paper concentrates on the relevance of the magical number seven in the span of actual perceptual/cognitive activities. But it might also be relevant for the size of our languages' repertoire of e.g. categories (Fenk-Oczlon and Fenk 2000).

7. Fraisse's (1985:94) report on these experiments should be supplemented by their original author Köhler (1952).

8. It is worth noticing that relevant data do not only show in our experimental setting with written clauses and translations but also in recordings of oral discourse: In 9 dialogues with a total number of 1055 intonation units Gigler (in preparation) found a mean size of 6.04 syllables per unit. The lowest mean value of these dialogues (Carinthian dialect) was 4.82 syllables, the highest was 7.12 syllables per intonation unit. The mean duration per intonation unit was 1.373 seconds.

9. Menzerath (1954:100) found that German words composed of a high number of syllables tend to be composed of a "relatively" low number of phonemes.

10. Such structuring processes will be enhanced by the movements of the groups (cf. "Gesetz des gemeinsamen Schicksals") within the visual field and by eye-movements necessarily resulting in faster relative movements of those patterns which are nearer to the observer. These movements of patterns relative to the actually moving eye and/or head of the observer deliver indications regarding spatial depth as well as indications regarding coherence and separability of visual patterns.

References

Altenberg, B. 1990. *Prosodic Patterns in Spoken English: Studies in the Correlation Between Prosody and Grammar for Text-to-Speech Conversion.* [Lund Studies in English 76]. Lund: Lund University Press.

Baddeley, A.D. 1986. *Working Memory.* Oxford: Clarendon Press.

Barker, M. and Givón, T. This volume. "On the pre-linguistic origins of language processing rates".

Beaken, M. 1996. *The Making of Language.* Edinburgh: Edinburgh University Press.

Brannon, E.M. and Terrace, H.S. 2000. "Representation of the numerosities 1–9 by rhesus macaques (Macaca mulatta)". *Journal of Experimental Psychology: Animal Behavior Processes* 26:31–49.

Bickerton, D. 1990. *Language and Species.* Chicago: University of Chicago Press.

Chafe, W. 1987. "Cognitive constraints on information flow". In *Coherence and grounding in discourse,* R.S. Tomlin (ed), 21–51. Amsterdam: Benjamins.

Chafe, W. 1994. *Discourse, Consciousness, and Time.* Chicago: University of Chicago Press.

Croft, W. 1995. "Intonation units and grammatical structure". *Linguistics* 33:839–82.

Crystal, D. 1969. *Prosodic Systems and Intonation in English.* Cambridge: Cambridge University Press.

Deacon, T.W. 1997. *The Symbolic Species. The Co-evolution of Language and the Brain.* NewYork: W.W. Norton.

Elbert, T., Ulrich, R., Rockstroh, B., and Lutzenberger, W. 1991. "The Processing of Temporal Intervals Reflected by CNV-Like Brain Potentials". *Psychophysiology* 28:648–55.

Feldhütter, I., Schleidt, M. and Eibl-Eibesfeldt. 1990. "Moving in the Beats of Seconds: Analysis of the Time Structure of Human Action". *Ethology and Sociobiology* 11:1–10.
Fenk, A. 1998. "Symbols and Icons in Diagrammatic Representation". *Pragmatics and Cognition* 6:301–34. (Special Issue on the *Concept of Reference in the Cognitive Sciences*).
Fenk, A. and Fenk-Oczlon, G. 1993. "Menzerath's law and the constant flow of linguistic information". In *Contributions to Quantitative Linguistics,* R. Köhler and B. Rieger (eds), 11–31. Dordrecht: Kluwer.
Fenk, A. and Vanoucek, J. 1992. "Zur Messung prognostischer Leistung". *Zeitschrift für experimentelle und angewandte Psychologie* 39 (1):18–55.
Fenk-Oczlon, G. 1983. *Bedeutungseinheiten und sprachliche Segmentierung. Eine sprachvergleichende Untersuchung über kognitive Determinanten der Kernsatzlänge.* Tübingen: Gunter Narr.
Fenk-Oczlon, G. and Fenk, A. 1985. "The mean length of propositions is 7 plus minus 2 syllables — but the position of languages within this range is not accidental." In *Cognition, Information Processing, and Motivation,* G. d'Ydewalle (ed), 355–9. Amsterdam: North Holland.
Fenk-Oczlon, G. and Fenk, A. 1995. "Selbstorganisation und natürliche Typologie." *Sprachtypologie und Universalienforschung* 48:223–38.
Fenk-Oczlon, G. and Fenk, A. 1999. "Cognition, quantitative linguistics, and systemic typology". *Linguistic Typology* 3:151–77.
Fenk-Oczlon, G. and Fenk, A. 2000. "The magical number seven in language and cognition: empirical evidence and prospects of future research". *Papiere zur Linguistik* 62–63: 3–14.
Fraisse, P. 1957. *Psychologie du temps.* Paris: Presses Universitaires de France.
Fraisse, P. 1985. *Psychologie der Zeit.* Munich/Basel: Ernst Reinhard Verlag.
Fraisse, P. and Fraisse, R. 1937. "Études sur la mémoire immédiate. I: L'appréhension des sons". *Année psychologique* 38:48–85.
Gigler, G. In preparation. *Intonationseinheiten im Kärntnerischen.* Dissertation, University of Klagenfurt.
Givón, T. 1990. *Syntax: a Functional-Typological Introduction,* Volume II. Amsterdam: John Benjamins.
Givón, T. 1998. "On the co-evolution of language, mind and brain". *Evolution of Communication* 2:45–116.
Givón, T. This volume. "The visual information-processing system as an evolutionary precursor of human language".
Greenfield, P.M. 1991. "Language, tools and brain: The ontogeny and phylogeny of hierarchically organised sequential behaviour". *Behavioral and Brain Sciences* 14: 531–92.
Halliday, M.A.K. 1967. *Intonation and Grammar in British English.* The Hague: Mouton.
Heeschen, V. 1994. "How long are clauses and sentences in a Papuan language like Eipo?" *Semaian* 10:50–70.
Kareev, Y. 2000. "Seven (indeed, plus or minus two) and the detection of correlations". *Psychological Review* 107:397–402.

Kien, J., Schleidt, M., and Schöttner, B. 1991. "Temporal segmentation in hand movements of Chimpanzees (*Pan troglodytes*) and Comparison with Humans." *Ethology* 89:297–304.
Kintsch, W. 1974. *The Representation of Meaning in Memory.* New York: J. Wiley.
Köhler, O. 1952. "Vom unbenannten Denken". *Verhandlungen der deutschen Zoologischen Gesellschaft* 16:202–11.
Lieberman, P. 1967. *Intonation, Perception and Language.* Cambridge, Mass.: MIT Press.
Lieberman, P. 1991. *Uniquely Human: The Evolution of Speech, Thought and Selfless Behaviour.* Cambridge, Mass.: Harvard University Press.
Mandler, G. 1975. "Consciousness: Respectable, useful, and probably necessary". In *Information Processing and Cognition. The Loyola Symposium.* R. Solso (ed), 229–54. Hillsdale, NJ: Lawrence Erlbaum Associates.
Massaro, D.W. 1975. *Experimental psychology and information processing.* Chicago: Rand Mc Nally.
Määttä, T. 1993. "Prosodic structure indices in a folk tale". *Phonum* 2:107–20.
Menzerath, P. 1954. "Die Architektonik des deutschen Wortschatzes". *Phonetische Studien* 3. Bonn: Dümmler.
Michotte, A. 1946. *La perception de la causalité.* Louvain and Paris: Fondation Universitaire de Belgique.
Miller, G.A. 1956. "The magical number seven, plus or minus two: Some limits on our capacity for processing information". *Psychological Review* 63, 81–97.
Pöppel, E. 1986. *Grenzen des Bewußtseins.* Stuttgart: Deutsche Verlags-Anstalt
Quirk, R., Duckworth, A., Svartvik, J. Rusiecki, J.P., and Colin, A.J.T. 1964. "Studies in the correspondence of prosodic to grammatical features in English". In *Proceedings of the IX International Congress of Linguists,* 679–91. The Hague: Mouton.
Sampson, G. 1997. *Educating Eve: The 'Language Instinct' Debate.* London and New York: Cassell.
Sasse, H.-J. 1991. "Predication and sentence constitution in universal perspective". In *Semantic Universals and Universal Semantics,* D. Zaefferer (ed), 187–221. Berlin: Mouton de Gruyter.
Schleidt, M. 1992. "Universeller Zeittakt im Wahrnehmen, Erleben und Verhalten". *Spektrum der Wissenschaft,* December: 111–15.
Schleidt, M. and Kien, J. 1997. "Segmentation in behavior and what it can tell us about brain function". *Human Nature* 8:77–111.
Swartz, K.B., Chen, S., and Terrace, H.S. 2000. "Serial learning by Rhesus monkeys: II. Learning four-item lists by trial and error". *Journal of Experimental Psychology: Animal Behavior Processes* 26:274–85.
Wills, C. 1993. *The Runaway Brain: The Evolution of Human Uniqueness.* New York: Basic Books.
Zipf, G.K. 1929. "Relative frequence as a determinant of phonetic change. *Harvard Studies in Classical Philology* 40:1–95.