2011 September 14

Stuttering and similar features of speech
Causes and cures

Peter Naur

Begoniavej 20, DK 2820 Gentofte, Denmark, 2011

The phenomena of stuttering and similar features of speech

Everyone knows what stuttering or stammering sounds like. By the dictionary, to stutter is to keep repeating parts, esp. initial consonants, of words in effort to articulate; to utter in this way.

Stuttering is often described as a speech disorder. However, stutterers show no organic defect. Moreover about 80 percent of stutterers recover with no treatment whatsoever, usually in early adulthood or adolescence.

In a particular person the occurrence of stuttering depends both on what is spoken and on the situation in which the person finds himself or herself at the moment of speaking. In particular it is known that actors may be stutterers in their private life and yet in their acting may speak without stuttering. Likewise it is known that persons who stutter in some of their speech may sing without stuttering. Cases are known in which a person suffering heavily from stuttering in his native language and environment, after emigrating to another country with a different language speaks the new language without any stuttering.

In describing stuttering as a disorder it is ignored that there are several other very common features of much speech that, while usually being less disturbing to the listener, occur in largely the same way as stuttering, being person and situation dependent, but are not described as disorders. Most prominent among such features are insertion of hesitations, insertion of vocalizations such as 'ah' or 'eh', and repetition of phrases. Evidence of these features may be found in the transcribed samples of spontaneous speech presented in section 4.3 of the present writer's book *The neural embodiment of mental life by the synapse-state theory*. These samples record spontaneous speech by six different speakers, who are all professionally concerned with speech. The samples have durations of between 44 and 110 seconds. In each of these samples pauses and vocalization sounds are quite prominent, as, for example, in 'I think it is simply – the best – eh – love story – ah – that I know'. Repetition of phrases is also quite common, as for example 'hold it back hold it back restrain restrain'. Stuttering occurs in one place of the six samples, indicated in the transcription as 's-s-s-sesignificantly altered'.

It should further be noted that difficulty in pronunciation of particular words is quite a normal phenomenon. The Russian singer Galina Vishnevskaya in her autobiography *Galina* tells that when she first had contact with her later husband Mstislav Rostropovich she thought of him as 'the cellist whose name I couldn't pronounce', and in their very first personal contact she addressed him saying: 'Mst … Mtl… Pardon me, but it's hard to pronounce your name.'

Explaining stuttering as a disorder

In a report published in 1969 by the U.S. Public Health Service under the title *Human Communication and Its Disorders—An Overview*, stuttering appears as a speech disorder. It is said that approximately 1,400,000 persons in the United States suffer from it. It is also said that 'stuttering remains an enigma while illustrating the type of disorder which does not have either a clean-cut organic cause or a clearly habitual basis.'

In this context it is remarkable that attempts to understand how stuttering comes about are presented in contexts in which there is no understanding of how normal speech is produced. In fact, no such understanding is presented anywhere in the current literature. Thus we read in an article

from 1985 (D. B. Rosenfield and F. Boller): 'A comprehensive theory of stuttering will not emerge until we have theories of control and coordination of movement systems that permit better understanding of how developing social and communicative skills impinge on the coordination and processing of speech.'

The study by P. A. Alm (2005) in its title postulates that there is a 'causal mechanism of stuttering', but fails to find any such, and does not notice that such a causal mechanism would fail to explain that the same person may stutter in some situations but not in others, and that many stutterers recover without any treatment. The study makes no mention, neither of normal speech, nor of habits, nor of William James's explanation of speech as action chains. In its Summary it says that 'The causes of stuttering have remained obscure.'

All these considerations suggest that considering stuttering as a disorder that has a specific organic cause that may be removed or eliminated is misguided. Stuttering and other pronunciation features of speech are integral parts of certain persons' speech. Any effort to change these features in persons must be based on a thorough understanding of how any speech is generated in the nervous system.

Speech as habitual action chains

The elements of such an understanding have already been presented by William James in chapter IV, HABIT, of his *Principles of Psychology* from 1890. In this chapter James through a discussion around the theme '*our nervous system grows to the modes in which it has been exercised*' is led to state that '*habit simplifies the movements required to achieve a given result, makes them more accurate and diminishes fatigue'*. He further states that '*habit diminishes the conscious attention with which our acts are performed*'. He continues:

> One may state this abstractly thus: If an act require for its execution a chain, *A, B, C, D, E, F, G,* etc., of successive nervous events, then in the first performances of the action the conscious will must choose each of these events from a number of wrong alternatives that tend to present themselves; but habit soon brings it about that each event calls up its own appropriate successor without any alternative offering itself, and without any reference to the conscious will, until at last the whole chain, *A, B, C, D, E, F, G,* rattles itself off as soon as *A* occurs, just as if *A* and the rest of the chain were fused into a continuous stream. When we are learning to walk, to ride, to swim, skate, fence, write, play, or sing, we interrupt ourselves at every step by unnecessary movements and false notes. When we are proficient, on the contrary, the results not only follow with the very minimum of muscular action requisite to bring them forth, they also follow from a single instantaneous 'cue'…
>
> These results may be expressed as follows:
>
> In action grown habitual, what instigates each new muscular contraction to take place in its appointed order is not a thought or perception, but the *sensation occasioned by the muscular contraction just finished*. A strictly voluntary act has to be guided by idea, perception, and volition, throughout its whole course. In habitual action, mere sensation is a sufficient guide, and the upper regions of brain and mind are set comparatively free. Fig. 1 will make the matter clear:

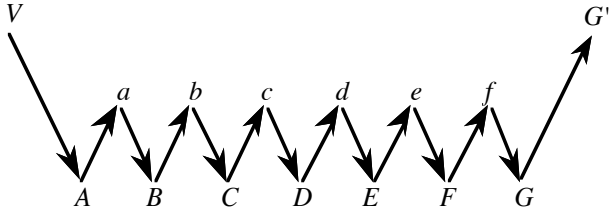

Fig. 1. Habitual chain of muscular contractions and sensations

Let *A, B, C, D, E, F, G* represent an habitual chain of muscular contractions, and let *a, b, c, d, e, f* stand for the respective sensations which these contractions excite in us when they are successively performed. Such sensations will usually be of the muscles, skin, or joints of the parts moved, but they may also be effects of the movement upon the eye or the ear. Through them, and through them alone, we are made aware whether the contraction has or has not occurred. When the series, *A, B, C, D, E, F, G,* is being learned, each of these sensations becomes the object of a separate perception by the mind. By it we test each movement, to see if it be right before advancing to the next. We hesitate, compare, choose, revoke, reject, etc., by intellectual means; and the order by which the next movement is discharged is an express order from the ideational centres after this deliberation has been gone through.

In habitual action, on the contrary, the only impulse which the centres of idea or perception need send down is the initial impulse, the command to *start*. This is represented in the diagram by *V*; it may be a thought of the first movement or of the last result, or a mere perception of some of the habitual conditions of the chain, the presence, e.g., of the keyboard near the hand. In the present case, no sooner has the conscious thought or volition instigated movement *A*, than *A*, through the sensation *a* of its own occurrence, awakens *B* reflexly; *B* then excites *C* through *b*, and so on till the chain is ended, when the intellect generally takes cognizance of the final result. The process, in fact, resembles the passage of a wave of 'peristaltic' motion down the bowels. The intellectual perception at the end is indicated in the diagram by the effect of *G* being represented, at *G'*, in the ideational centres above the merely sensational line. The sensational impressions, *a, b, c, d, e, f,* are all supposed to have their seat below the ideational lines. That our ideational centres, if involved at all by *a, b, c, d, e, f,* are involved in a minimal degree, is shown by the fact that the attention may be wholly absorbed elsewhere. We may say our prayers, or repeat the alphabet, with our attention far away.

It seems obvious that the chains of muscular contractions here described by James are what is activated at the utterance of each phrase during speech.

Details of the brain structure and action involved in chains of muscular contractions such as speech have been developed in section 3.7 of (Naur, 2008). The way they apply to speech is shown in detail in chapter 4 of the same work. By this description the relevant part of the brain is a network of neurons, nodes, and synapses. The part of the network involved in the chain of muscular contractions required in speaking the phrase 'Rostropovich' is shown in Fig. 2. The activity in the network taking place while the phrase is spoken consists of excitations of nodes, transmitted through the neurons and synapses.

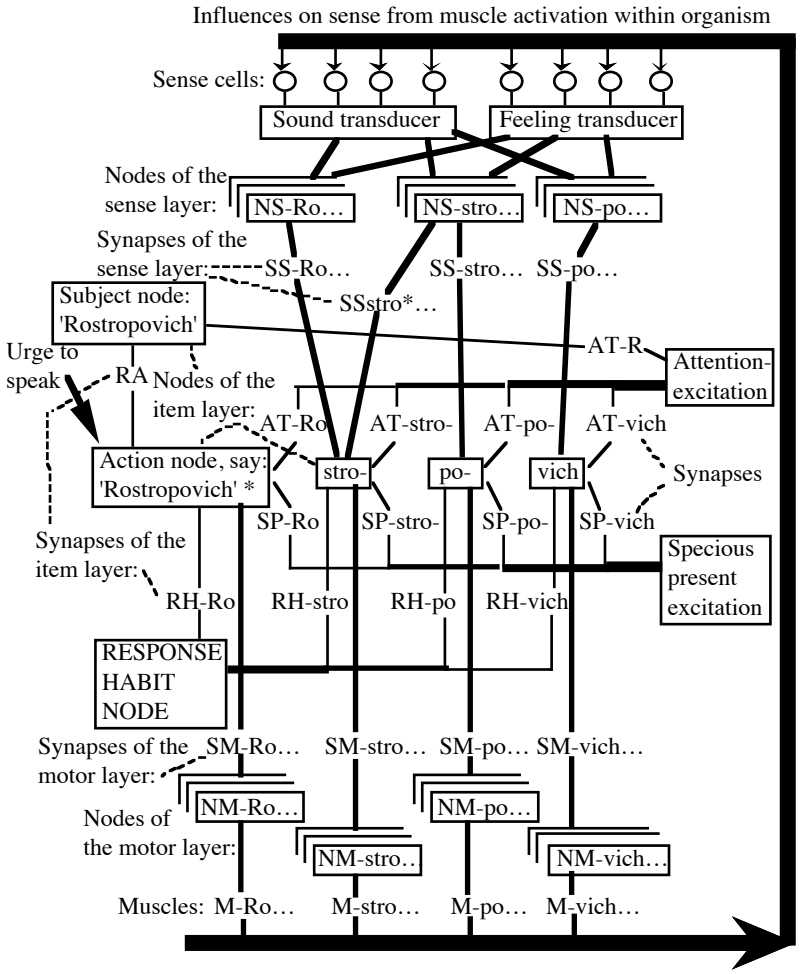

Fig. 2. Muscular action aggregate for pronouncing 'Ro-stro-po-vich' or 'Ro-stro-stro-stro…-po-vich'

In this figure designations of synapses that are shown followed by dots, such as SS.Ro… and SM.Ro…, show groups of synapses, such a group having perhaps several hundred individual members. Similarly, designations of muscles shown followed by dots, such as M.Ro… and M.vich…, show groups of muscles. Single nodes are shown as rectangular boxes. Groups of nodes are shown by a picture of a file of cards. Straight lines show neurons. Those shown with thick lines indicate groups of neurons.

Considering first the neural activity of speaking the phrase 'Rostropovich' as this happens in a well-educated neural system, the neural activity leading to excitations of muscles is prepared in the simultaneous excitation of a number of nodes of the item layer shown in Fig. 2. These nodes include, first of all, the node labelled 'Subject node: Rostropovich', which embodies the person's being acquainted with the person Rostropovich. When the person gives attention to Rostropovich this node will be excited from the Attention excitation through the synapse of the attention layer labelled AT-R. This excitation will directly be transmitted further through the synapse RA into the node 'Action node, say 'Rostropovich'*'. This node is the first of four that have to be excited so as to achieve the effect in the muscular contraction involved in speaking the four syllables of 'Rostropovich'. These four nodes are shown as 'Action node, say 'Rostropovich'*', 'stro-', 'po-', and 'vich'. While the speaking is in progress the excitations of these four nodes are formed in a number of steps, by summation in the nodes of excitation contributions coming from several sources. One constant contribution comes from the situation in which the person finds himself or herself. This is shown in the figure as the excitation of the node RESPONSE HABIT NODE. This excitation is distributed into the four muscular contraction nodes through the four synapses named RH-ro, RH-stro, RH-po, RH-vich. The significance of this contribution will be further described below.

Other contributions to the excitations arise from the person's urge to speak the word 'Rostropovich'. This urge will primarily become manifest in an excitation of the node 'Action node, say 'Rostropovich'*' coming from nodes not shown in the figure. For the speech actually to take place, the sum of these excitations must be so strong that their sum, when transmitted into the node, is sufficient to excite the synapse AT-Ro so as to make this synapse transmit a strong additional excitation from Attention excitation. This mechanism, which depends on the specific properties of the attention synapses (those named AT-ro, AT-stro-, AT-po-, AT-vich), is that by which the attention is for periods of the order of one second concentrated in one particular node. Here the strong excitation from the attention synapse AT-RO will immediately be conducted through the synapse set SM-Ro… into the set of nodes of the motor layer NM-Ro…, and from there into the set of muscles M-Ro…, which effect the sound Ro to be articulated.

The activation of the speech muscles M-Ro… will produce sensible effects in the organism that are picked up by sense cells located there. These effects may occur both in the tissue around the muscles and in the ears. This influence is shown in the figure by the thick black arrow. These effects get transformed in the Feeling and Sound transducers and produce excitations of the set of nodes of the sense layer shown as NS-Ro…. These excitations get further conducted by the set of synapses shown as SS-Ro… and so by their summation produce a strong excitation of the node stro-. This excitation acts on the synapse of the attention layer shown as AT-stro-, which reacts by adding a strong excitation into the node stro-. The resulting strong excitation of the node stro- will be transmitted through the set of synapses SM-stro… into the set of nodes of the motor layer shown as NM-stro…, which will then activate the muscles M-stro… that articulate the sound 'stro'. This excitation activity will continue in the same way until all four syllables, Ro-, stro-, po-, vich, have been articulated.

This accounts for what takes place if the person only pronounces 'Rostropovich' in one way. If the person has two different pronunciations of the word, depending on the situation, if for example the person is an actor who will speak the word stuttering in normal life but unstuttering on the stage, the neural structure will have two independent muscular action aggregates of the kind

shown in figure 2 and the choice of the one which will become excited in a particular situation depends on the excitations of the response habits nodes. The way this choice happens is shown in figure 3. Like figure 2 this shows ‘Subject node: ‘Rostropovich’’, the node that embodies that the person is acquainted with the subject Rostropovich. When the person thinks of the subject this node will be excited. This may come about for example by the person hearing the name Rostropovich spoken. This sound will be sensed in certain sense cells that through certain nodes of the sense layer, shown as NS-Rostropovich…, and a set of conductive synapses, shown as SS-Rostropovich…, will be summed into ‘Subject node: ‘Rostropovich’’.

The selection between two different styles of pronunciation depending on the situation is handled by two separate muscular action aggregates for the pronunciation, of which the action nodes are shown in Figure 3 as ‘Action node, say ‘Rostropovich’*’ and ‘Action node, say ‘Rostro-stro-stro…povich’*’. At a moment when there is attention to the subject Rostropovich both of these action nodes will receive excitations from ‘Subject node: ‘Rostropovich’’, through the synapses RA-1 and RA-2. When the urge to speak is generated in parts of the network not shown in the figure, the corresponding excitation will be sent into both of the action nodes and will be summed together with the excitations coming from the subject node. A third contribution to the excitation of the action nodes comes from whichever of the two RESPONSE HABIT NODES is excited at the moment, through one of the synapses RH-1 or RH-2. This last contribution decides which of the two action nodes will be most strongly excited, and thereby which of the two attention synapses AT-Ro1 and AT-Ro2 will become excited and by its contribution set one of the two action aggregates to be activated and thus to make either the word ‘Rostropovich’ or ‘Rostro-stro-stro…povich’ to be spoken. Precisely how many repetitions of the ‘stro’ will be spoken will depend on excitations of the action nodes from other nodes of the item layer that depend on aspects of the situation that are not taken into account here.

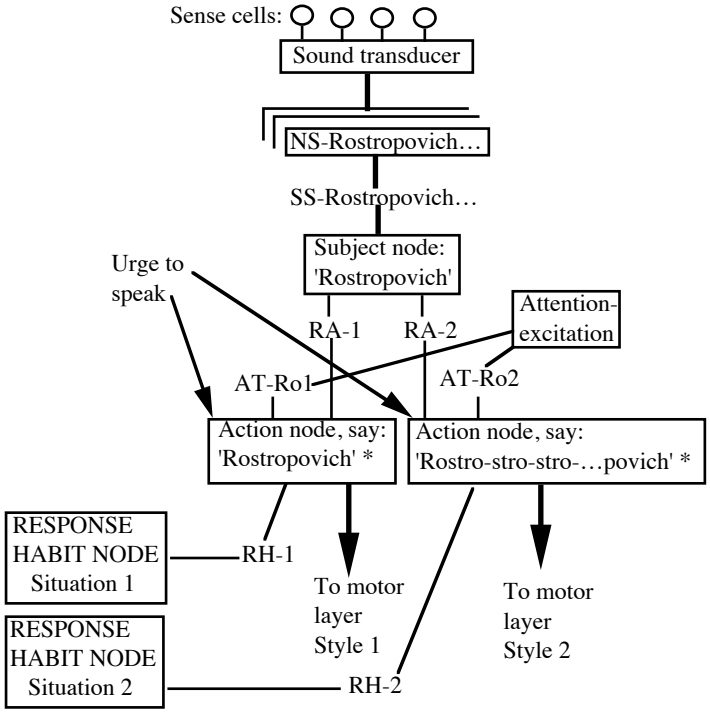

Fig. 3. Selection of speech style, by situation

The activity of pronunciation described by Fig. 2 is such that takes place in a person who can speak the name Rostropovich in a non-stuttering way. It depends on the several sets of synapses, SM-Ro…, SS-Ro…, SM-stro…, SS-stro…, SM-po…, SS-po…, SM-vich…, having been made conductive during the activity in which the person learned to pronounce the name Rostropovich.

This learning process has depended on the plasticity of the synapses with respect to their conductivity.

In its early phase this learning consisted of educating separately the parts of the neural net that enter into pronouncing the four syllables Ro-, stro-, po-, and vich. Educating such a part of the net takes place by the person's deliberate *imitation* of a sound spoken by another person. For example, when educating the part of the net for speaking stro- the person must repeatedly have in mind the intent to speak the sound spoken by someone else, which will be manifest in excitation of the node called stro-, at the same time as activating the corresponding muscles of the speech organs. These activations will be manifest in excitations of the node set NM-stro and the muscles M-stro (these excitations will come from parts of the nervous net not shown in the figure). Since at such a moment the node stro- and the node set NM-stro… are both excited the synapses connecting them, SM-stro…, will receive excitations from both of their neurons simultaneously, and so will have their conductivities increased in the plastic way characteristic of synapses.

By practising the speaking of the fours syllables a number of times the four sets of synapses, SM-ro…, SM-stro…, SM-po…, and SM-vich…, will become conductive. The result is that merely by thinking of the intent to speak 'Ro', which is manifest in the excitation of the node 'Action node, say 'Rostropovich'*', the speech muscles M-ro… will be activated.

In the next phase of the learning the person must speak the four syllables pairwise in immediate sequence many times. When first speaking the pair Ro-stro the person must attend to speaking each of the two syllables separately. However, when speaking the two syllables in immediate sequence many times this pairing of their activations will become gradually more automatic. This is a result of the synapse set SS-Ro… becoming conductive. This happens because when the first syllable Ro- is spoken, this excites a number of sense cells in the tissue around the speech muscles and in the ears. These excitations will immediately be transferred through the Feeling and Sound transducers into a certain set of nodes of the sense layer, NS-Ro… These excitations will be there when the node stro- is excited, and so the set of synapses connecting these nodes, SS-Ro…, will have their conductivities increased. During this phase the person must deliberately activate the muscles, that is excite the speaking of each of the syllables, separately many times.

Thus by the way speech is understood within the synapse-state theory of mental life, speech habits are acquired mostly by imitation of the speech of other people. The major influence is the young person's contact with parents and other older persons. This is where most of the young person's habits of perception of speech sounds and of speech production are formed.

How were the habits of stuttering acquired? In order to answer this question it must be noted, first of all, that there is a strong universal tendency in human beings that a fast movement just performed will immediately be performed again. This tendency to repeat fast movements can be observed both in infants and, as already mentioned above, in spontaneous adult speech. The tendency is readily explained by the neural mechanism of control of movements just described. Indeed, for any movement to happen there must be a certain node of the item layer, which may be called the activity node, that receives strong excitations from other nodes. If the movement is fast the excitation of the activity node will remain strong when the movement has been performed, and so will often be sufficient to release the movement again.

An habit of stuttering speech is formed when the child tries to speak (imitate) a complicated phrase, but instead of speaking it properly relapses into repeating some of its parts. This is confirmed by the fact that much stuttering is known to originate in situations in which the child is in states of stress and conflict with elders and therefore feels urged to speak quickly.

Once the person has stuttered in a phrase this will be registered in the nervous system in the form of conductivities of certain synapses. In Fig. 2 an example of such synapses are shown as

SSstro*… that will have become conductive if the person has spoken the name in a stuttering way: Ro-stro-stro-stro…

In other words, *stuttering comes into a person's speech, not together with particular sounds or letter combinations of words, but in certain phrases as the person learns to speak them by imitation of others in particular kinds of situations*. In other words, stuttering is a matter of a certain number of (bad) personal habits of speech. Like swearing to express anger or as expletive.

Against this understanding of how stuttering comes into the speech of certain persons one sometimes sees the objection that many stutterers have siblings (brothers or sisters) with perfectly normal speech. This objection is invalid, however. What matters is not the family relationship between child and parent, but the mental climate between the child and the elder from whom the child picks up speech. And the same parents may have very different mental climates in their relations to their children.

Getting rid of stuttering

Now assuming this explanation to be valid, what does it tell us about how a person can get rid of his or her stuttering? Very simply (but not easy): change your habits. The difficulty is that by its very nature any habit will get enforced, stronger, every time it is exercised. This means that in order to get rid of a particular habit, say the stuttering pronunciation of some particular word, the person must completely avoid pronouncing that word in the stuttering way. One way to achieve this, but one very unpractical for most people, would be for the person to stop talking altogether for a long period.

Alternatively, the person might train better habits. A training program along these lines will be a personal matter and will have to take up each inflicted word or phrase separately. The program will have two parts, one for determining what the inflicted words and phrases are, and one for the actual training of better pronunciations.

Determining what the inflicted words and phrases are must be done in any situation. Since the situation of speech is important, the determination should record not only the inflicted word, but that word in the context in which it is spoken in a stuttering manner. For example, if the inflicted word is *Rostropovich* this should be recorded as the phrase actually spoken in a particular situation, for example: *I have Rostropovich's recording of Haydn's concerto*.

The training must be done in the way a musician learns to play a difficult passage on his instrument, to wit, by practising, that is speaking the phrases inflicted by stuttering many times, beginning so slowly that no stuttering occurs and gradually increasing the speed, but never so fast that the stuttering happens. This training would have to be done over and over again, daily, always starting each phrase so slowly that no stutter comes forth.

If the person is incapable of speaking a particular phrase without stutter at any slow speed, a special training of that phrase by the technique of what French musicians call *decomposition* should be performed. Take as example the phrase *Rostropovich*. In decomposition this should be trained by speaking very clearly and loudly many times, never more quickly than makes the phrases quite clear, first all pairs of two syllables together:

Ro-stro, Ro-stro, Ro-stro, Ro-stro, …
stro-po, stro-po, stro-po, stro-po, …
po-vich, po-vich, po-vich, po-vich, …

Then three syllables together:

Ro-stro-po, Ro-stro-po, Ro-stro-po, Ro-stro-po, …
stro-po-vich, stro-po-vich, stro-po-vich, stro-po-vich, …

Finally all four syllables together:

Ro-stro-po-vich, Ro-stro-po-vich, Ro-stro-po-vich, Ro-stro-po-vich, …

Such a training program will make the stutterer very conscious of each inflicted phrase of his or her speech. Then hopefully when in normal conversation, the stutterer will be conscious of his or her use of each of the inflicted phrases and will be able to speak the proper version as it has been trained.

The question raised by the U.S. Public Health Service can be answered: the acquisition of speech happens by the young person's imitation of the elders. Stuttering speech is acquired when for reasons of the mental tension of the situation in which the imitation happens the young person fails to imitate properly.

Acknowledgement
In formulating the present account I have profited greatly from perceptive remarks from Erik Frøkjær and Jesper Hermann, which I gratefully acknowledge.

Literature

Alm, P. A.: 2005, *On the Causal Mechanism of Stuttering*, Lund University, Sweden.

James, W.: 1890, *The Principles of Psychology*, vol. I-II, Henry Holt, USA; reprinted in Dover, 1950.

Naur, P.: 2008, T*he neural embodiment of mental life by the synapse-state theory*, 189 pages, ISBN 87-987221-5-8, 2008. The full text is available on the internet at www.naur.com.

Rosenfield, D. B. and Boller, F.: *Stuttering*, Handbook of Clinical Neurology, Vol. 46 p. 169-172, Elsevier Science Publishers, 1985.