

## The Role of Subvocalization in Rehearsal and Maintenance of Rhythmic Patterns

Jordi Pich

University of Balearic Islands

This experiment analyzed the influence of subvocal activity in retention of rhythmical auditory patterns. Retention of sixteen percussion sequences was studied. Each sequence (a 4-s "door-knocking" pattern) was followed by one of the following six retention conditions: silence, unattended music (blocking the inner ear, i.e., Gregorian chanting), unattended music (blocking the inner ear, i.e., rock-and-roll), articulatory suppression (blocking the inner voice), tracing circles on the table with index finger (spatial task), and tapping (motor control). After silence, unattended music (chanting), or the spatial task, participants successfully reproduced most patterns. Errors increased with unattended music (rock-and-roll), but significant disruptions only occurred with tapping and articulatory suppression. Whereas the latter case supports the role of an articulatory loop in retention, the production of successive taps or syllables in both interference conditions probably relies on a general rhythm module, which disrupted retention of the patterns.

*Key words: auditory imagery, rhythm perception*

En este experimento se analizó la influencia de la actividad subvocal sobre la retención de patrones rítmicos auditivos. Se investigó la retención de dieciséis secuencias de percusión. Cada una (de 4 seg. y semejante a la típica sucesión de golpes llamando a una puerta) iba seguida por una de las siguientes seis condiciones de retención: silencio, música no atendida (bloqueo del oído interno, i.e., canto gregoriano), música no atendida (bloqueo del oído interno, i.e., "rock-and-roll"), supresión articulatoria (bloqueo de la voz interna), dibujar con el dedo índice círculos sobre la mesa (tarea espacial) y golpear regularmente con la mano sobre la mesa (control motor). En las condiciones de silencio, de música no atendida (canto gregoriano), o en la tarea espacial, los participantes reprodujeron con éxito la mayoría de las secuencias. Los errores aumentaron en la condición de música no atendida ("rock-and-roll"), pero sólo hubo interferencias que fueron estadísticamente significativas en las condiciones de control motor y supresión articulatoria. Mientras esta última condición apoya el papel del bucle articulatorio en la retención, la producción de golpes o de sílabas sucesivas en ambas condiciones experimentales depende, seguramente, de un módulo rítmico general, que perjudica la retención de los patrones.

*Palabras clave: imágenes mentales auditivas, percepción del ritmo*

Cognitive research regards various types of auditory memory representations (speech, music, or noise) as mental images that resemble the original experience of hearing a particular sound. Human beings actually appear to be able to conjure up from memory the recollections of auditory events, such as the timbre and intonation of a friend's voice, the melody, rhythm, and tempo of a symphony, or the sound of breaking glass.

On the other hand, imagery research has established a relationship between mental imagery and the working memory (WM) model originally developed by Baddeley and Hitch in 1974. Specifically, the visual component of the WM model (the visuo-spatial sketchpad) was shown to be involved in setting up, maintaining and/or transforming visual images (Baddeley, 1986; Logie, 1991, 1995). Extrapolating this argument to the auditory domain, Baddeley and Logie (1992) suggested that the phonological loop (i.e., the auditory component of the model) might represent the seat of auditory imagery. The loop comprises two components, namely a temporary phonological store linked with an articulatory rehearsal process. In a modified WM model, Logie (1995) defined these components as "Phonological Cache" and "Phonological Rehearsal."

However, this interesting suggestion raises several questions. First, the auditory capacity to recollect musical and noise materials (including spectral timbres that cannot be overtly reproduced) implies that the phonological cache should be open to sounds other than those associated with speech. For this reason, other authors (Mackay, 1992; Smith, Wilson, & Reisberg, 1995) link the articulatory loop to the inner voice phenomenon, whereas the tonal storage of the phonological cache is what permits the auditory experience within the inner ear.

Secondly, there is some experimental evidence indicating that subvocal rehearsal also seems to participate in tasks involving non-verbal tonal processing. Logie and Edworthy (1986) found that blocking subvocalization by means of concurrent articulatory suppression ("tah-tah-tah" repeated aloud by the subject) hindered recognition of a previously heard random tone sequence. In addition, Smith et al. (1995) reported that articulatory suppression impaired subjects' judgements of pitch when they were asked to decide whether a familiar melody rose or fell from the second to the third note or vice versa. This fits with Baddeley and Logie's (1992) view of the articulatory loop as an "auditory-articulatory interface." However, the concurrence of subvocalization in spectral capabilities, such as memory for the timbre of musical instruments (Crowder, 1989) or pitch discrimination (Logie & Edworthy, 1986) is not evident.

In the present study, I explored another possible role of subvocalization in the auditory memory. Specifically, I wished to examine whether subvocalization is involved in the retention of rhythmical auditory patterns. Short-term memory for percussion sequences under various types of interference during the retention period (filled with

unattended music, tapping, a spatial task, or articulatory suppression) was investigated.

In a previous study, Smith, Reisberg & Wilson (1992) showed significant disruption in the capacity to form auditory streams mentally (such as, *Anyone for tennis?* looking at letter/number strings: *NEI 4 10s?*) under articulatory suppression but not under unattended music. Here, I considered that, if subvocal activity is involved in rhythmical retention, then this retention should be more disrupted by interferences affecting the inner voice by articulatory suppression than by those that block the inner ear with musical sounds. If so, this would prove that subvocal rehearsal participates actively in maintaining the rhythmical structure of the pattern.

This musical capacity was extensively studied by the music teacher Thackray (1969), who named the capacity to reproduce temporal and intensive patterns of sounds and short melodies, while conserving a tempo, *rhythmic performance*. Fraisse (1982) also defined *rhythmic anticipation*, which implies memorizing a pattern in order to compare it with a subsequent one. In my hypothesis, I assumed that both linked activities probably depend on subvocalization.

## Method

### *Participants*

Ninety members of a university community, aged 19 to 25 years, were asked to participate in the experiment. They received a ticket for a soft drink at the bar for their participation. Fifteen individuals were randomly assigned to each of the six experimental conditions: silence, unattended music (Gregorian chant or rock-and-roll drum solo), spatial activity (tracing circles), motor activity (regular tapping), and articulatory suppression (repeating "tah-tah-tah" aloud). A between-subjects design was employed to avoid the transfer of a particular coding strategy among conditions.

### *Stimuli*

Stimuli were 16 unfamiliar percussion sequences recorded digitally with EWI software by a 32W Sound Blaster. Each sequence lasted 4 s and included a rhythmic succession of beats (ranging from 10 to 12 elements) that was repeated twice to facilitate retention. Beats consisted of a drum sound generated by the program.

### *Procedure*

The experiment was conducted in a well-isolated room, with the walls and ceiling lined with U-shaped fiber pieces for acoustic absorption. An experimenter asked the subjects to retain the sequence, performed at 75 dB(A) at subjects' ear-level. They were requested not to move during the trial. Each sequence was followed by a 10-s retention interval, the

usual period in the memory interference experimental paradigm. They then had to reproduce the sequence both vocally and by tapping, in whichever order they preferred. To become familiar with the task, there was a pre-test training trial which included three easier sequences. Afterwards, each subject was exposed to the 16 experimental sequences of increasing difficulty. Each subject's performance in this memory task (the number of these sequences—from 0 to 16—correctly reproduced -vocally and by tapping) was registered and this was the dependent variable of the experiment.

*Interference manipulation.* Musical interference during the 10-s retention interval consisted of choral music (Gregorian chant) or a percussion rock-and-roll drum solo, both played at 75 dB(A) at ear-level. Participants were instructed to treat them as disturbances, and to try not to attend to them. Nonmusical interference was provided by regular tapping and by articulatory suppression, requesting participants to perform beats (tap or say "tah") at the tempo they preferred, provided it was kept constant. In order to control the higher attentional-resource demand in these latter two conditions, which required participants to produce the taps sequentially, another group performed a spatial task consisting of tracing a series of circles on the table with their index finger, at a constant speed, during the retention interval.

## Results

After the retention period, overall results from the participants in the silent condition showed that they could correctly retain 69.6 % of the 16 experimental rhythm sequences that each of them had to memorize and reproduce ( $M = 11.13$ ,  $SD = 2.69$ ). Similar percentages were obtained by participants in the unattended Gregorian chant condition (65.8% correct reproductions:  $M = 10.53$ ,  $SD = 2.67$ ) and in the spatial task (65% correct reproductions:  $M = 10.4$ ,  $SD = 2.64$ ). The percentages of correct responses for participants in the unattended percussion music condition decreased to 55.4% ( $M = 8.86$ ,  $SD = 4.38$ ). Poorer performances were achieved under tapping and articulatory suppression conditions, in which participants' correct pattern retention only amounted to 49.1% and 35.8%, respectively ( $M = 7.86$  and  $5.73$ ,  $SD = 4.13$  and  $3.47$ , respectively). The mean percent correct responses in each condition is shown in Figure 1.

One-way ANOVA showed that the experimental condition was highly significant,  $F(5,84) = 5.359$ ,  $p < .0001$ .

A *t*-test was performed to compare the means of all the experimental conditions. The correct response mean of the silence condition was not statistically different from the means of the two musical conditions, or from the spatial task mean. On the contrary, the silence condition mean was statistically different, and higher, than the tapping condition,  $M = 11.13$  and  $7.87$ , respectively,  $t(28) = 2.562$ ,  $p < .05$ , and than articulatory suppression,  $M = 11.13$  and  $5.73$ , respectively,  $t(28) = 4.756$ ,  $p < .0001$ .

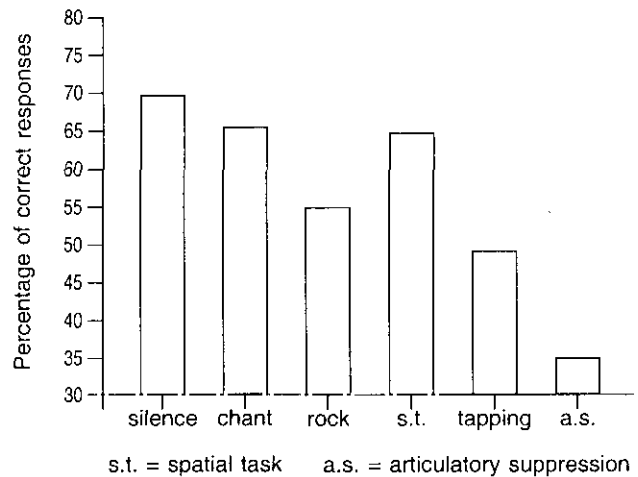


Figure 1. Percentage of correct responses in the retention of percussion sequences under the six experimental conditions: silence, unattended music (Gregorian choral chant and rock-and-roll drum solo), spatial task, tapping, and articulatory suppression.

The correct response mean of the Gregorian chant condition was also significantly different, and higher, than the tapping condition,  $M = 10.53$  and  $7.87$ , respectively,  $t(28) = 2.097$ ,  $p < .05$ , and than the articulatory suppression condition,  $M = 10.53$  and  $5.73$ ,  $t(28) = 4.244$ ,  $p < .0001$ .

The correct response mean of the other musical condition (rock-and-roll drum solo) was significantly higher than that of articulatory suppression,  $M = 8.87$  and  $5.73$ , respectively,  $t(28) = 2.168$ ,  $p < .05$ , but not higher than the tapping condition.

On the other hand, the correct response mean of the spatial task was slightly, but not significantly, higher than that of the rock-and-roll condition. However, the spatial task participants performed significantly better than those in the tapping condition,  $M = 10.4$  and  $7.87$ , respectively,  $t(28) = 1.99$ ,  $p < .05$ , and than the articulatory suppression participants,  $M = 10.4$  and  $5.73$ , respectively,  $t(28) = 4.142$ ,  $p < .0001$ .

The differences in correct responses between the tapping condition and articulatory suppression condition did not reach statistical significance,  $M = 7.87$  and  $5.73$ , respectively,  $t(28) = 1.529$ ,  $p = .137$ .

## Discussion

Retention in the control condition (silence) was good, and subjects only failed to reproduce the most difficult patterns. Theoretically, to achieve this retention, they could have kept the rhythm sequence in the auditory memory in analogical form, excluding the participation of any other type of mental activity during the retention interval. But, clearly, that did not seem to be the case. In a post-experimental questioning session, most of the participants explained that they had tried to repeat the rhythmical patterns spontaneously in their minds

by means of subvocal rehearsal. Thus, the temporal structure could be rehearsed and maintained in their short-term memory via subvocal beats. Moreover, an auditory analogue would probably have been disrupted by musical interference. Participants whose inner ear was blocked by music during the retention period reproduced the patterns as well as those in the silent condition, even when submitted to a vigorous rock-and-roll percussion solo. However, as can be seen in Figure 1, although the difference between the rock-and-roll and the silence groups did not reach statistical significance, the two groups showed a tendency to be different.

It is well established that subvocalization plays a crucial role in designing motor plans for producing and controlling speech. These "timing-motor" mechanisms could also be useful in the mental retention of the temporal structure of a sound. In fact, the strongest disruption was caused by retention under articulatory suppression, providing support to this temporal function of subvocalization. Similarly, the retention impairment caused by tapping at a regular medium tempo shows that production of subvocal beats and taps probably shares a rhythm module or resource that is necessary for mental control of rhythm structures.

Although difficult to define because of the plurality of rhythmic components and even of the related musical artistic performances, Fraisse (1982) defended the existence of a module of this sort, postulating a *general sense of rhythm*. According to this author, it could be characterized by three main factors: perceptual structuring (discrimination of temporal structures), rhythmic anticipation (memorization of a pattern), and practical-rhythmic. This last implies the coordination of alternate movements of hand and foot and the ability to adapt to changes in rhythm, both of which require voluntary control of rhythmic movement. Our results suggest that subvocalization could underlie and connect these rhythmical capacities. Moreover, from a neurophysiological point of view, there are probably strong connections between subvocalization, speech, and motor planning underlying the Broca area, which is often involved in language output and in the pre-motor organization of articulation.

However, the good performance of the participants in the spatial task condition, which required them to trace circles continuously, allows us, within reason, to reject the hypothesis that posited that poor retention of rhythmic patterns during the repetition of the vocal "tahs" or the manual tapping was exclusively due to the higher attentional demands required by these distracting tasks. Nor can I exclude the possible influence of participants' musical preferences or their general attitude towards music. I simply did not include participants with specific musical training. I also failed to control possible influences due to participants' sex. Despite these research limitations, overall results are in line with previous research showing that subvocal speech plays an important role in the production of auditory images.

Smith et al. (1992, p. 111) speculated that subvocalization might provide a "tonal frame" for the auditory image, which

is later colored to particular timbre specifications by the inner ear. In their words,

It might be impossible to hang imagined spectral characteristics onto thin air, but easily possible to assemble them onto an existing, subvocalized pitch. In that event, subvocalization would merely be the carrier wave of the auditory image, and the function of the inner ear would be to modulate that carrier in trumpet-like or flute-like ways.

Perhaps "subvocal singing" (Baddeley & Logie, 1992), "motoric control" (Crowder & Pitt, 1992), or "tonal frame" (Smith et al., 1992) are different ways of referring to the same conscious experience of this temporal dimension of auditory images, which is neither strictly "phonological," "motor," nor "tonal." The role of subvocal activity would be to provide the baseline for any specific spectro-temporal image, a sort of spectro-temporal skeleton that, in the domain of auditory imagery, would represent the same thing that the visuo-spatial sketchpad represents in the domain of spatial imagery. Moreover, research on rhythmic perception has offered evidence of the psychological reality of multi-layered rhythmic structures in which complex rhythmic patterns are overlaid on an underlying beat sequence (Dowling & Harwood, 1982). According to these authors (1982, p. 201),

The same type of mental representation has been postulated for tonal organization, namely, a solid framework (beat or tonal scale) on which more elaborate patterns (rhythm or melody) are laid. The development of these mental representations appears to follow a closely parallel course for both pitch material and rhythm.

Independently of the nature and relationships between both the above-mentioned mental representations, it is clear that humans are capable of spectral transformations (i.e., transposing an entire musical piece in a different key) and of temporal ones. Humans can retain not only a particular rhythm, but can also increase or decrease the tempo of music, speech, or regular noises, or sustain a tone indefinitely, or repeat a sequence, etc. More research on this topic will probably reveal the importance of subvocalization in all these mental transformations of auditory images.

## References

- Baddeley, A. (1986). *Working memory*. Oxford, UK: Oxford University Press.
- Baddeley, A., & Hitch, G.J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. VIII, pp. 47-90). New York: Academic Press.
- Baddeley, A., & Logie, R. (1992). Auditory imagery and working memory. In D. Reisberg (Ed.), *Auditory imagery* (pp. 179-197). Hillsdale, NJ: Erlbaum.
- Crowder, R.G. (1989). Imagery for musical timbre. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 472-478.

- Crowder, R.G., & Pitt, M.A. (1992). Research on memory/imagery for musical timbre. In D. Reisberg (Ed.), *Auditory imagery* (pp. 29-44). Hillsdale, NJ: Erlbaum.
- Dowling, W.J., & Harwood, D.L. (1982). *Music cognition*. London: Academic Press
- Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149-180). London: Academic Press.
- Logie, R. (1991). Visuo-spatial short-term memory: Visual working memory or visual buffer? In C. Cornoldi & M. MacDaniel (Eds.), *Imagery and cognition* (pp. 77-102). Berlin: Springer-Verlag.
- Logie, R. (1995). *Visuo-spatial working memory*. Hillsdale, NJ: Erlbaum.
- Logie, R., & Edworthy, J. (1986). Shared mechanisms in the processing of verbal and musical material. In D.G. Russell, D. Marks, & J. Richardson (Eds.), *Imagery 2* (pp. 33-37). Dunedin, New Zealand: Human Performance Associates.
- Mackay, D.G. (1992). Constraints on theories of inner speech. In D. Reisberg (Ed.), *Auditory imagery* (pp. 121-149). Hillsdale, NJ: Erlbaum.
- Smith, J.D., Reisberg, D., & Wilson, M. (1992). Subvocalization and auditory imagery: Interactions between the inner ear and the inner voice. In D. Reisberg (Ed.), *Auditory imagery* (pp. 95-119). Hillsdale, NJ: Erlbaum.
- Smith, J. D., Wilson, M., & Reisberg, D. (1995). The role of subvocalization in auditory imagery. *Neuropsychologia*, 33, 1433-1454.
- Thackray, R. (1969). *An investigation into rhythmic abilities*. London: Noveleo.

Received February 2, 1999

Revision received November 10, 1999

Accepted December 14, 1999