A Speech Based Connectionist Model of Human Short Term Memory

Dr. Dimitrios Bairaktaris[1] **Prof. Keith Stenning**[2]

Human Communication Research Centre,
University of Edinburgh,
2 Buccleuch Place,
Edinburgh EH8 9LW, Scotland
tel: (+44) 31 6504450, facsimile: (+44) 31 6504587,
dimitris@uk.ac.ed.cogsci, keith@uk.ac.ed.cogsci

Abstract

In recent years connectionist modelling of Short Term Memory (STM) has been a popular subject of research amongst cognitive psychologists. The direct implications in natural language generation and processing, of the speech based phenomena observed in immediate recall STM experiments, make the development of a psychologically plausible STM model very attractive. In this paper we present a connectionist Short Term Store (STS) which is developed using both traditional STM theories of interference and decay trace. The proposed store has all the essential characteristics of human short term memory. It is capable of on-line storage and recall of temporal sequences, it has a limited span, exhibits clear primacy and recency effects, and demonstrates word-length and phonological similarity effects.

Introduction

Short Term Memory (STM) has been a major subject of investigation for cognitive psychologists since the 50's. Initial experiments established that STM has a limited storage capacity, or span (Miller 1956). It was later shown that when span is exceeded, immediate STM recall performance is impaired in a very specific way. Only the first few, primacy, and the last few, recency, memory items can be recalled at some significant level of accuracy (Postman & Philips 1965). Subsequent experiments established that span is affected when memory items are phonologically similar. This feature of STM is known as the Phonological Similarity Effect (PSE) (Baddeley 1966). Later, it was also shown that the time taken to articulate a memory item has a negative effect on span. Recall performance is reduced when longer words are stored in STM (Baddeley et. al. 1975). This STM feature is known as the Word Length Effect (WLE) and together with PSE suggests that for written verbal material STM access involves some form of speech processing. Further experiments using articulatory suppression, during both STM storage and recall phases (Baddeley 1990), have shown that there exists a route for STM access which involves phonological encoding with the possible use of a lexicon. This immediately relates STM performance to language generation processing tasks and makes the modelling of STM a very attractive task indeed.
There are two major STM theories. One claims that span is limited mainly due to interference; more recent memory traces affecting earlier ones. The other claims that forgetting occurs because memory traces decay through time. Interference theory has been supported by a number of mathematical STM models such as (Murdock 1983) (Schweikert 1986) and does conform with psychological data. However, mathematical models fail to provide an account for some of the speech based characteristics of STM recall experiments (Brown & Hulme 1991). Trace decay as applied in the working memory model of STM (Baddeley 1990) provides an account to all speech based aspect of STM recall but fails in at least two ways. It does not provide an explicit computational model of STM which will facilitate testable theoretical predictions, while the articulatory loop rehearsal mechanism is not completely consistent with recent experimental evidence (Baddeley 1986) (Howard & Franklin 1990). Pioneering work on STM connectionist modelling was done by Grossberg (1976) and was further strengthened with the connectionist theory revival of the early 80's resulting in a number of connectionist STM models which conformed, to some degree, with psychological evidence. Some of these models adopted the interference theory (Wang & Arbib 1991) and others the trace decay theory (Brown 1989) (Burgess & Hitch 1992). We discuss these models and some general STM modelling issues next.

[1] Joint Council Initiative in HCI/Cognitive Science Training Research Fellow.
[2] Supported by the Economic and Social Research Council.

Connectionist Models of STM

One particular area in connectionist theory which is closely related to STM modelling is that concerned with the problem of serial order. In its simpler form the problem of serial order manifests itself in list learning and recall, a task very similar to that of immediate recall STM experiments. Is it possible for a connectionist network to learn a sequence of patterns and recall them in their original order? Various solutions have been given to this problem (Jordan 1986) (Elman 1988) (Norris 1990) (Houghton 1989) (Bairaktaris 1992). It is essential for any STM model to perform the serial order task. For example, Brown (1990) uses the solution proposed in Norris (1990) to construct his STM model, Burgess and Hitch use the solution of Houghton (1990) for their sequence storage and generation, while Wang & Arbib (1991) STM model is a improved variation of Houghton's solution.

In general most of the STM models in the literature perform well in the sequence generation task. Performance starts breaking down when they are tested against STM performance criteria such as span, primacy and recency, WLE and PSE. From all the connectionist STM models only Burgess & Hitch's (1992) makes an attempt to address all the above criteria. Their model is based on the trace decay theory of STM but a major part of its dynamical behaviour depends on the existence of random noise rather than decay. In brief, their model does well with span, primacy and WLE, but fails to demonstrate recency, PSE and it is not capable of on-line list learning.

A close study of the Burgess & Hitch model revealed that the majority of the problems with their model were due to the fact that they failed to distinguish between a Short Term Memory and Short Term Store (STS). To identify a separate STS embedded in STM is clearly consistent with the working memory (Baddeley 1986) framework. Such an approach, also favoured by Brown & Hulme (1991), provides the advantage of separating the process of generating the phonological code from the actual storage and retrieval of memory items. This paper describes an STS mechanism which has an accurately defined phonological interface to the main STM mechanism.

The model

We will divide the description of our model in two parts. First we will give an account of its static characteristics and then we will describe its dynamic behaviour as an STS module. The network model of STS is shown in Exhibit 1.

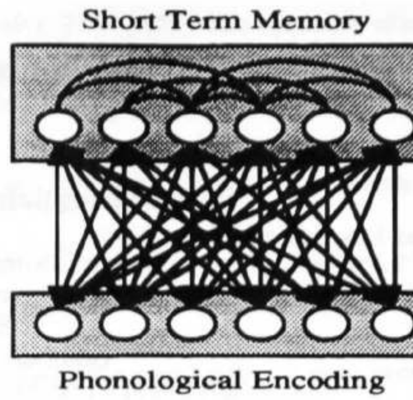

Exhibit 1. Two layer Short Term Store.

It comprises two layers of units, an input layer where the phonological code of the memory items is clamped and the STM layer where the actual memory items are stored. The two layers of units are fully connected with bidirectional connections. The STM layer nodes are also fully interconnected with excitatory connections. Similar STM models, which emphasize different aspects of STM, have been previously described in the literature (Zipser 1991) (Wang et. al 1991). The phonological encoding layer is provided with a sufficient number of nodes in order to store all the significant features of the memory item. It is assumed that a pre-processor is available for the generation of the phonological code. Such a simple pre-processor network is described in Burgess & Hitch (1992). It must be made clear however that the generation of the phonological code is a complex process which possibly requires the generation of intermediate phonetic codes (Besner 1987) as well as the use of a lexicon (Monsell 1987). It should also be mentioned that the quality of the phonological code used by the STM layer is subject to modality effects and various reception conditions such as articulatory suppression (Howard et. al. 1990). We are currently developing a model which deals explicitely with the generation of the phonological code.

Each node in the input layer stores a feature of the item's phonological code in a binary form and in such a way that phonologically similar items have similar codes. The input nodes propagate their values to the STM layer where a different node is allocated to the representation of every memory item. The connections between the input and the STM layer are modified using a hebbian learning rule in order to retain the phonological code of every memory item. It is shown in Bairaktaris (1991, 1992) that using a modifiable threshold technique, a one-to-one correspondence between STM nodes and memory items can be achieved without the use of intra layer inhibitory connections (Grossberg 1976). To avoid limiting the system's capacity artificially, STM nodes are allocated to the memory items dynamically (Bairaktaris 1991). For a network with j input nodes the STM node activation A_i and output O_i are computed as follows:

$$A_i = \sum_j W_{ij} P_j$$

$$O_i = A_i \text{ if } A_i > T_i \text{ and } O_i = 0 \text{ if } A_i \leq T_i$$

where **P** is the vector of the activations of the input nodes, **W** is the weight matrix of the connection between the input nodes and STM nodes and **T** is vector of the threshold of the STM nodes. The node allocation, threshold setting and weight modification mechanisms are described in more detail in Bairaktaris (1991, 1992). When an STM node fires its output decays through time as follows:

$$O_i(t+1) = O_i(t) \frac{\delta}{e^{(\lambda - O_i(t))}}, \text{ where } \delta, \lambda \text{ are constants}$$

The effect of the above decay rule is shown in Exhibit 2 for $O_i(0) = 1$, $\delta = \lambda = 0.6$.
The weight Zij on the connection between nodes i and j in the STM layer is modified as follows:

$$Z_{ij}(t+1) = Z_{ij}(t) + O_i(t)\, O_j(t) \quad (1)$$

During the recall phase, where there no activation propagated from the input layer to the STM layer, STM nodes compute their activation solely on feedback from other STM nodes as follows:

$$A_i = \sum_j Z_{ij} O_j$$

It is assumed that every node receives a constant amount of activation from background noise. However not all the nodes fired at the presence of background noise. It is only the node which represents the first memory item, and has the lowest threshold, which will fire thus initiating the recall phase:

$A_1(1) = \kappa$ where κ is constant; typically $\kappa = 0.15$
The training regime between the input layer and the STM layer guarantees that there is a one-to-one correspondence between the memory items and the STM nodes. In recall mode however, more than one STM node can be active at any time. This means that the system cannot decide about the exact recall sequence of the memory items. The relative output of node i (trace decay) against the sum of output of all the nodes (interference) in STM is the probability (Pi) assigned to the hypothesis of the system recalling memory item i at time t:

$$P_i(t) = \frac{O_i(t)}{\sum_j O_j(t)}$$

The proposed network architecture is very similar to the model of the articulatory loop described in Baddeley (1986). There are however two major differences between the Baddeley approach and our model.

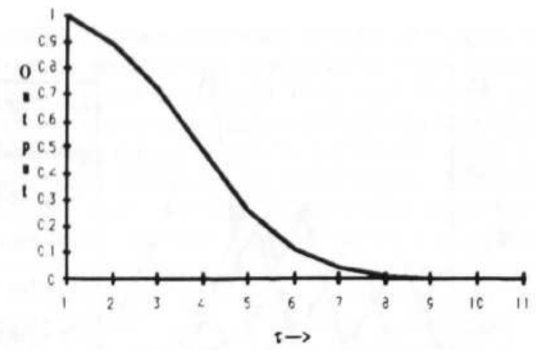

Exhibit 2. Output from STM nodes decays through time.

Baddeley proposes that memory items are rehearsed and dynamically stored in the loop between the phonological encoding and the STM layer and that span is limited purely due to trace decay. Our model stores the memory items in the STM nodes while the item sequence is stored, maintained and reproduced by the recurrent connections in the STM layer. Rehearsal is possible, but not essential (Howard et. al. 1990), via the bidirectional connections between the input layer and STM. Furthermore, recall of the memory items depends on both trace decay of individual memory nodes and interference from other memories.
We will now describe the dynamic behaviour of the proposed model. To simplify the description, we will focus on the dynamic behaviour of the STM layer assuming that the input layer provides all the appropriate phonological memory traces. As is described above when a node in the STM is allocated to the representation of a memory item its output is set to 1 and it decays thereafter. The longer it takes for the second memory item to be registered in the STM layer the weaker the recurrent output signal from the previous memory item becomes. When the second memory item is allocated a node in the STM layer, the connection between the previously active STM node and the current one is modified as shown in (1). Modifications on the STM recurrent connections occur in the same way every time a new memory item it added. A close inspection of Exhibit (3) reveals that by the time the ninth memory item is registered in STM the output of the first item has diminished to zero. This shows that the output decay mechanism applied on the STM nodes, imposes an implicit limit on span which is very close to the empirical 7±2 observation made in Miller (1956).
During recall the node representing the first memory item becomes active due to background noise and initiates the sequence generation process. Immediately after the first node all the other nodes in the layer receive varying degrees of activation depending on their original position in the sequence. Depending on whether their activation exceeds their preset threshold, they activate themselves or not. A wave of nodes firing is spread through the layer.

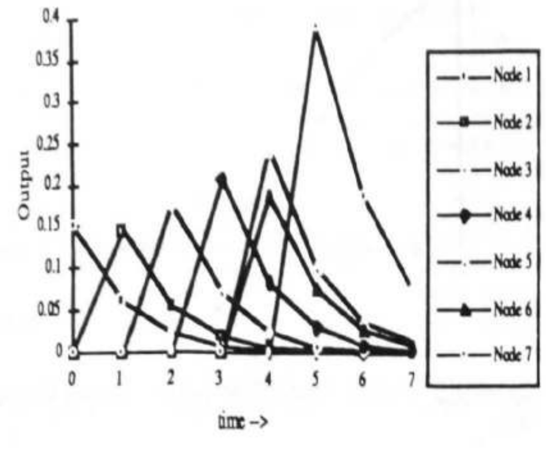

Exhibit 3. Output of STM nodes during recall.

Exhibit 3 shows the output of 7 STM layer nodes through time. The original encoding corresponds to a sequence of 7 memory items presented at equal time intervals. In the example shown $\delta = \lambda = 0.6$ and $\kappa = 0.15$. Exhibit 3 demonstrates a case where more than one node fires at the same time. At time $t = 4$, nodes 5 and 6 fire simultaneously, but node 5 has a relatively stronger output than node 6. Therefore the relative probability of the system recalling memory item 5 at time 4 is higher than the probability of recalling item 6. In general, the probability of recalling a memory item X at a particular time Y, is equivalent to the relative output of the node representing item X at time Y, over the sum of the output from all the nodes in STM at the same time Y. To place this into the context of STM immediate recall experiments, the probability of recalling item 1 at time 0, item 2 at time 1, item 3 at time 2 and so on is equivalent to the probability of correctly recalling all the memory items in their original positions in the sequence. Exhibit 4 shows the recall probabilities for every memory item, for the same sequence of items and same parameter settings of exhibit 3.

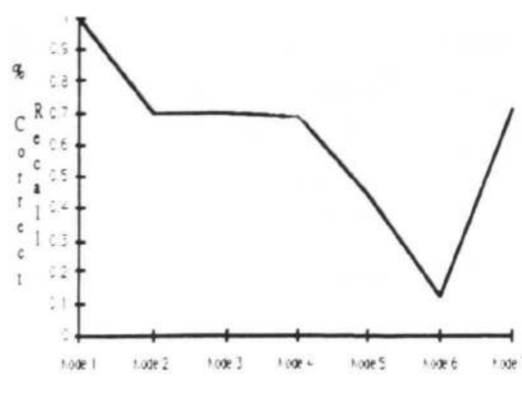

Exhibit 4. The probabilities of correctly recalling 7 memory items in their original ordering.

The above interpretation of the network's dynamics is used throughout the simulation results presented in the following section.

Simulation Results

The proposed STM store was simulated for a variety of different parameter settings before the results reported below were achieved. Setting the κ parameter of the network proved to be a difficult task, but at the same time a number of 'interesting" network behaviours emerged from the simulation process. These behaviours are currently analyzed within the context of neuropsychological, 'patient specific', STM evidence. Here we will only refer to the simulations results that are relevant to our task; to demonstrate that the proposed model conforms to psychological evidence. In all the results presented below, $\kappa = 0.15$ and $\delta = \lambda = 0.6$.

Span-Primacy - Recency

Exhibit 4 shows that the model demonstrates clear primacy and recency effects. For a list of 7 items the network is capable of recalling all the items in their original order, with greater confidence for the first and last list items and smaller confidence for the intermediate items. As is shown in Exhibit 5, the network also demonstrates primacy and recency effects for list lengths of 10 and 20. When span is exceeded, the probability of correctly recalling intermediate list items is effectively 0.

Phonological Similarity

In the introduction of the paper it was mentioned that when the memory items are phonologically similar immediate recall success rates are decreased. Phonologically similar memory items will produced phonologically similar codes at the input layer of our model and because of the hebbian learning algorithm between the input and the STM layer activation will pass not only to the node allocated to the current memory item but also to the nodes which represent phonologically similar items.

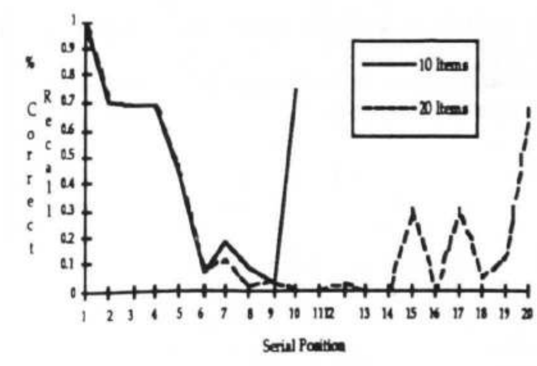

Exhibit 5. Primacy and recency effects for lists of 10 and 20 items

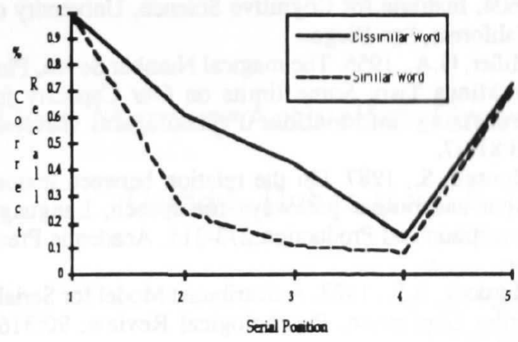

Exhibit 6. Phonological Similarity Effect when recalling a lists of 5 phonologically similar and dissimilar items.

The effect of phonologically similar memory items on the model's recall rates is shown in Exhibit 6 for a list of 5 similar and 5 dissimilar items. The phonologically similar list has lower recall rates that the dissimilar list, still both lists maintain the characteristic primacy and recency effects. An close of node outputs in the STS revealed that items with similar phonological representation are more likely to be recalled in their reverse list order.

Word Length Effect

In the description of our STS model it was mentioned that for the generation of the phonological code pre-processing of the raw data has to be made. It is reasonable to assume that the time taken to articulate a memory item is proportional to the pre-processing time required to generate the phonological code. This means that for longer words it will take longer before our model is provided with its phonological code, and for shorter words the generation of the equivalent phonological code will be shorter. In order to simulate the time taken to articulate a word in our model, we modified the number of time steps taken before two consecutive memory items are clamped at the input units. In all the simulations described above a new memory item was encoded in the STM at every time step.

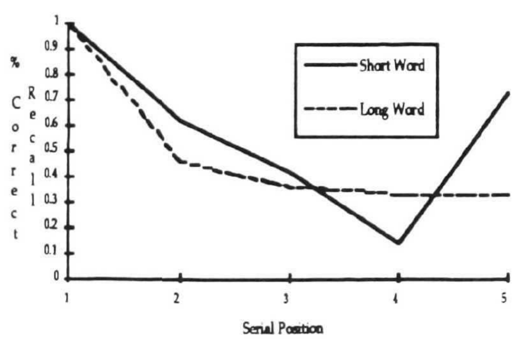

Exhibit 7. Word Length Effect when recalling lists of 5 short and long words.

Exhibit 7 shows recall rates for a list of 5 items, for both the standard case of registering a new memory item at every step (Short Word) and for the case where a new memory item was registered every three time steps instead of one (Long Word). Exhibit 7 shows that recall rates are worse for long words list than the short words list. Primacy effect is present in both cases, but recency is only present for the short word list case. Absence of primacy for the long word list case looks somewhat inconsistent with the psychological evidence. However, looking back at the original word length experiments (Baddeley et. al 1975) the recency effect in their experimental results is not very strong either. In fact the graph of Exhibit 7 is extremely similar to the equivalent word length effect graph in (Baddeley et. al. 1975).

Conclusions

A connectionist network model of a short term memory store was presented. The proposed network architecture comprises a fully interconnected layer of nodes which interacts with the core of the Short Term memory using a layer of input units where the phonological code of the memory item is clamped. The model uses a constructive learning algorithm which combined with a hebbian-type synaptic modification rule allows on-line storage of memory items. The proposed network is different to earlier STM models, in that the interpretation of its dynamics incorporates both the decay trace and the interference STM theories. Simulation results demonstrated that the model conforms to some of the major STM psychological evidence. The basic span, primacy and recency STM effects, are faithfully reproduced by the network model. These are the standard benchmark STM effects that have to be met by all STM models. It is in the interpretation of the main speech based STM effects, such as word length and phonological similarity, that our model makes a significant contribution. It provides an explicit computational account of the above effects by accurately reproducing the psychological data. Furthermore, it can explain some more subtle speech based STM effects, such as phonemic confusion, where non-adjacent phonologically similar list items are transposed during recall.
The proposed model is currently augmented with the development of a network model for the generation of the phonological code. This is intended to provide a computation account for some of the lexical access STM effects such rhyme and pseudo-homophone judgement.

References

Baddeley, A.D., 1966. Short-term memory for word sequences as a function of acoustic, semantic and

formal similarity. Quarterly Journal of Experimental Psychology, 18:302-309.
Baddeley, A.D., 1986. Working Memory. London: Oxford University Press, Oxford Psychology Series, 11:75-107.
Baddeley, A.D., Thomson, N. Buchanan, M., 1975. Word Length and the Structure of Short Term Memory. Journal of Verbal Learning and Verbal Behaviour 14:575-589.
Bairaktaris, D., 1991. Adaptive Pattern Recognition in a Real-World Environment. PhD. Thesis, Dept. of Computational Science, University of St.Andrews.
Bairaktaris, D.,1992. Discovering temporal structure using Hebbian learning. In Beale, R. and Finlay, J. (eds), Neural Networks and Pattern Recognition in Human Computer Interaction. Ellis Howard, U.K.
Besner, D., 1987. Phonology, Lexical Access and Reading and Articulatory Suppresion: A critical review. The Quarterly Journal of Experimental Psychology, 39A:467-478.
Brown, G.D.A, & Hulme, C., 1991. Connectionist Models of Human Short-Term Memory. In Progress in Neural Networks. ed.O. Omidvar, publ. Ablex, Norwood NJ, in press.
Brown, G.D.A, 1990. A Neural Net Model of Human Short-term Memory Development. In Lecture Notes in Computer Science: Proceedings of EURASIP Workshop on Neural Networks. Heidelberg: Springer Verlag.
Brown, G.D.A., 1989, A Connectionist Model of Phonological Short Term Memory, In Proceedings. of the 11th Annual Conference of the Cognitive Science Society, 572-579. Hillsdale, NJ: Lawrence Erlbaum Associates.
Burgess, N., Hitch, G., 1992. Towards a network model of the articulatory loop. Journal of Memory and Language.
Burgess, N., Moore, M.A., Shapiro, J.L.,1990. Human-Like Forgetting in Neural Network Models of Memory, Technical Report, Dept. of Theoretical Physics, University of Manchester.
Elman, J, 1998. *Finding Structure in Time*, CLR Technical Report 8801, Centre for Research in Language, University of California, San Diego.
Grossberg, S., 1976. Adaptive Pattern Classification and Universal Recoding: I Parallel Development and Coding of Neural Feature Detectors. Biological Cybernetics, 23:121-134.
Hinton, G.E., Plaut, D.C., 1987. Using fast weights to deblur old memories. In Proceedings. of the Ninth Annual Conference of the Cognitive Science Society. Hiilsdale, NJ: Lawrence Erlbaum Associates.
Houghton, G., 1989. The problem of serial order: A neural network model of sequence learning and recall. In Proceedings 2nd European Workshop on Language Generation, Edinburgh.
Howard, D., Franklin, S., 1990. Memory Without Rehearsal, In Neuropsychological Impairments of Short Term Memory:287-320. eds. Vallar & Shallice, Cambridge University Press.
Jordan, M.I., 1986. Serial Order: A Parallel Distributed Processing Approach. Technical Report 8604. Institute for Cognitive Science, University of California, San Diego.
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.
Monsell, S., 1987. On the relation between lexical input and output pathways for speech, Language Perception and Production:273-311. Academic Press Inc.
Murdock, B.B., 1983. A distributed Model for Serial-Order Generation, Psychological Review, 90:316-338.
Norris, D., 1990. Dynamic net model of human speech recognition. In G.T. Altman (Ed.) Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives, MIT Press, in press.
Postman, L., & Phillips, L.W., 1965. Short-term temporal changes in free recall. Quarterly Journal of Experimental Psychology, 17:135-.
Schweickert R., & Boruff, B., 1986. Short-Term Memory Capacity: Magic Number or Magic Spell?. Journal of Experimental Psychology, Learning, Memory, and Cognition: 419-425.
Wang, D., Arbib, A.M., 1991, A Neural Model of Temporal Sequence Generation with Interval Maintenance. In Proceedings of the 13th Annual Meeting of Cognitive Science Society, 944-948.
Waugh, N.C., & Norman, D.A., 1965. Primary Memory. Psychological Review:89-104.
Zipser, D., 1991. Recurrent Network Model of the Neural Mechanism of Short-Term Active Memory. Neural Computation, 3:179-193.