

Robotic Model of the Contribution of Gesture to Learning to Count

Marek Ruciński*, Angelo Cangelosi, Tony Belpaeme
Centre for Robotics and Neural Systems, Plymouth University
Drake Circus, PL4 8AA, Plymouth, United Kingdom
Email: marek.rucinski,angelo.cangelosi,tony.belpaeme@plymouth.ac.uk
*Telephone: +44 (0) 17525 84908

Abstract—In this paper a robotic connectionist model of the contribution of gesture to learning to count is presented. By formulating a recurrent artificial neural network model of the phenomenon and assessing its performance without and with gesture it is demonstrated that the proprioceptive signal connected with gesture carries information which may be exploited when learning to count. The behaviour of the model is similar to that of human children in terms of the effect of gesture and the size of the counted set, although the detailed patterns of errors made by the model and human children are different.

I. INTRODUCTION

It is widely accepted that gestures like pointing, touching or moving items when counting are an integral part of the development of children’s number knowledge [1]. Children use such gestures spontaneously and many studies have confirmed that it facilitates counting accuracy [1]–[8]. As number knowledge can be regarded as a major example of abstract thought, it is not surprising that counting skill has been drawing the attention of psychologists for a long time [9]. As a result, a lot of valuable experimental data on the contribution of gesture to learning to count have been gathered. For instance it is known that prevention of pointing disrupts the counting procedure; in such situations a child usually emits an indefinite stream of number words or does not count at all [2]. In addition, evidence has been provided for the importance of the actual physical contact with counted objects; counting items behind a transparent cover proves to be more difficult for children than when they are allowed to touch the objects being counted [3], [5]. It is also known that gesture plays a developmental role: it is particularly helpful for children around 4 years of age, in contrast to 2- and 6-years-olds [4]. Finally, both active gestures (a child gestures itself) and passive gestures (gesturing performed by someone else) facilitate counting accuracy, although they lead children to make different types of errors [7].

Despite all that has been learnt about the supportive role of gesture in the acquisition of the counting skill, the questions about the exact nature of this contribution remain unanswered. A number of hypotheses concerning the issue have been brought forward, but based solely on behavioural data it is hard, if not impossible, to go much further. One of the tools that may be helpful in answering the questions *how* and *why* something happens (in addition to *what* happens),

is cognitive modelling. It turns out, however, that attempts to model the contribution of gesture to learning to count are virtually nonexistent. While a number of models can be quoted which, among other numerical capabilities, look at counting [10]–[12], their focus is mostly on the distinction between sequential enumeration and so-called subitizing (i.e. the immediate visual apprehension of small numerosities) and, more importantly, they do not address in any way the relation between counting and motor capabilities.

At this point it is worth to emphasise that the contribution of gesture to learning to count is an attractive topic from the point of view of embodied cognition [13], according to which, in general terms, the functionality of the brain cannot be understood without taking into account the body. There is a growing amount of evidence that, despite its abstract appeal, numerical thinking may be to a large extent shaped by physical interactions with the environment [14]. Understanding how the body contributes to the acquisition of such a concept as the number may well shed light on how abstract representations in general are constructed, an issue that is of vital importance for cognitive science [15]. It seems likely however that it is the embodied character of the contribution of gesture to learning to count that has been putting the researchers off from trying to model it. Indeed it is difficult to imagine a solely computational model which would not impose arbitrary assumptions about representing the bodily contribution. This

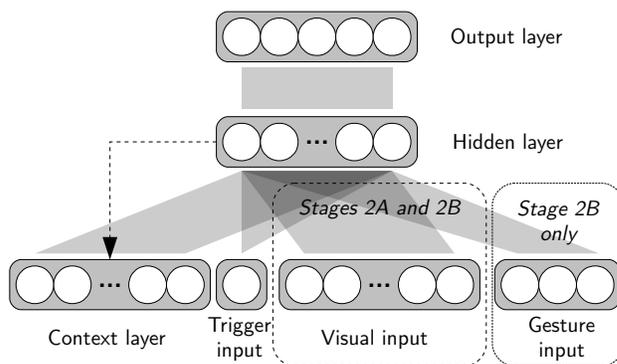


Figure 1. Architecture of the model. Gray polygons represent all-to-all connections. Activations are propagated from bottom to top.

general problem in the modelling of embodied phenomena is addressed in an elegant way by cognitive robotics [16], in which computational modelling is supplemented with an artificial, robotic body. This enables more accurate modelling with less arbitrary assumptions and the potential of this approach in the context of mathematical cognition modelling has already been demonstrated [17].

In this paper we attempt to fill-in the apparent gap mentioned above and make the first step toward understanding more specifically the contribution of gesture to learning to count. We propose a developmental robotic model of the phenomenon designed for the iCub humanoid robot platform [18], and investigate if the proprioceptive signal connected with gesture affects its counting accuracy. Furthermore we compare the behaviour of the model with data gathered in studies with human children.

The remaining of the paper is organised as follows. First, the model architecture, its development and evaluation procedures are described. Then a detailed statistical analysis of the experimental results is presented, including a comparison of the model behaviour with that of children. The paper concludes with a discussion mentioning perspectives for future work.

II. MODEL DESCRIPTION

In order to model the process of acquisition of the counting skill and the potential influence of the pointing gesture, we propose a recurrent neural network model based on the Elman architecture [19], presented in figure 1. In an Elman network, activations of the hidden units of a 3-layer artificial neural network at time $t - 1$ are available to the hidden units at time t (represented by the context layer), via connections which may be modified during training.

When formulating the model, the following assumptions about modelling the gesture have been made:

- proprioceptive information connected with gesture is an external input to the model. It is not the task of the model to *produce* gesture;
- gesture, if present, is a *correct* motor activity in the context of counting;

Motivations behind these assumptions are discussed in section IV. The following subsections describe the coding schemes adopted for model inputs and outputs.

A. Model task and output coding

The task used to assess the performance of the model throughout the experiments described in this paper is the production of a sequence of number words the length of which corresponds to the number of items presented to the visual input of the model. To this end, the output of the proposed artificial neural network is a binary vector of an a-priori chosen length (5 for the reported results), which encodes the produced words. Rather than adopting one-hot coding, output vectors are allowed to be non-orthogonal which is intended to mimic phonetic similarities present in natural languages. The particular sequence of number words to be used is a parameter of the modelling experiment and was generated randomly with

two constraints: different words need to have different vector representations and the special vector with zeros at every output unit corresponds to “silence” i.e. not producing any word. Real-life language data were not used, in order to enable the verification of stability of the model behaviour with respect to the sequence of number words. The model was trained and evaluated with numbers ranging from 1 to 10.

B. Input coding

1) *Trigger input*: The model has a 1-unit input called the trigger input. Its role is to indicate when the counting process should start. This corresponds to asking the subject in a psychological study a question (usually “How many?”) which according to the experimental protocol should encourage counting. Accordingly, the network is trained to remain silent irrespective of any additional inputs whenever the value of the trigger input is 0, and produce desired output when the trigger is 1. Trigger input activation remains fixed throughout every sequence in training and testing data sets.

2) *Visual input*: Visual input to the model is a 1-dimensional saliency map, which can be considered a simple model of a retina. Each unit of the visual input layer represents one spatial location in the visual frame of reference, and is activated or not depending on the presence of an object at this particular location. The sum of all activations over the visual input is normalised to 1 in order to eliminate the possibility of simple discrimination of cardinality based on this cue. Moreover, for a specific number of presented objects, the actual locations activated on the modelled retina are randomised between trials, so that the cardinality cannot be deduced based solely on locations of objects. The visual input consisted of 20 units to allow sufficient diversification of objects placement for the assumed maximum number (10).

3) *Gesture input*: The proprioceptive signal was obtained from a pointing gesture performed by the iCub humanoid robot [18]. As a starting point the data from joint angles of the robot right arm kinematic chain were used. This kinematic chain consisted of 6 degrees of freedom: torso yaw and pitch (roll angle was locked to eliminate unnaturally-looking postures) and the first 4 joints of the robot arm (shoulder and elbow). The robot was commanded, via its Cartesian interface, to point to 20 locations in front of it, which were assumed to correspond to the 20 locations in the visual input. These

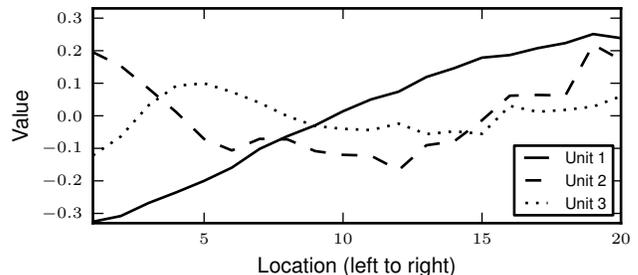


Figure 2. Values of the 3 units of the proprioceptive input (ordinate) for the 20 spatial locations (abscissa).

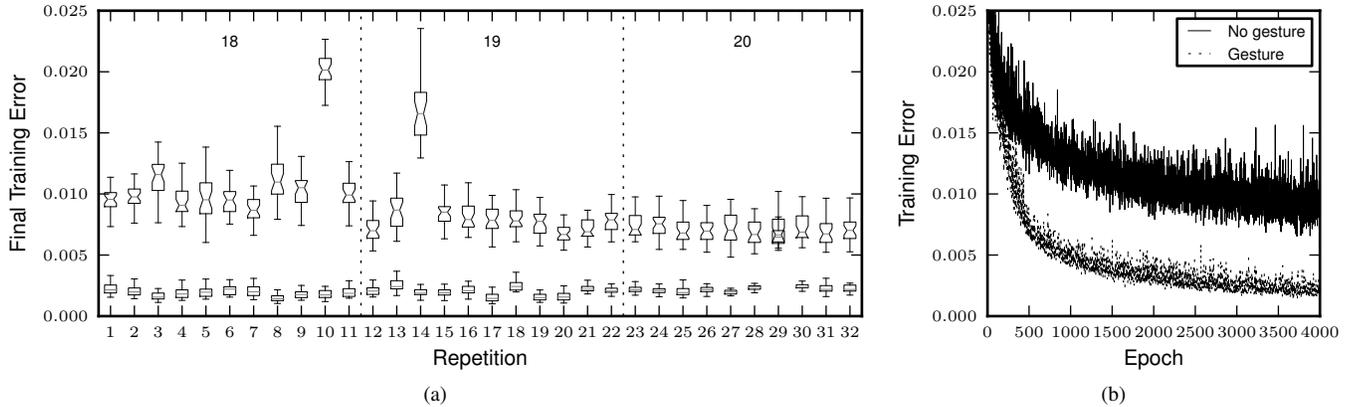


Figure 3. (a) box plot of the final training error (last 30 epochs) for each repetition in no gesture (notched boxes) and gesture (non-notched boxes) conditions. Outliers (data points outside the lower and upper quartile) were few and are omitted for clarity. Numbers near the top of the chart show the numbers of hidden units in trials; (b) typical example of full training error plots in both conditions for a selected trial (i.e. trial 1).

locations were uniformly distributed on a line placed 30cm in front of the robot, 10cm above its hip and spanned from 20cm left to 20cm right. Joint angles corresponding to pointing to each location were recorded and subsequently analysed using Principal Component Analysis which revealed that the first 3 principal components carry more than 90% of the total “statistical energy” of the original data. Therefore the dimensionality of the original signal was reduced from 6 to 3 by taking the 3 strongest principal components as the final proprioceptive input to the model. Values of these components for the considered 20 spatial locations are plotted in figure 2.

C. Model training

In order to model the development of the counting skill, the architecture of the artificial neural network (more specifically, its inputs) is adjusted throughout 3 modelling stages (see figure 1). First, the model is trained to recite a sequence of number words. Then, in order to assess the impact of the proprioceptive information connected with gesture, the training of the model branches out into two further stages, or conditions: the network is trained to count the number of objects shown to the visual input *in the absence* and *in the presence* of the proprioceptive gesture signal. All three stages use supervised learning by backpropagation through time and are described in detail below.

1) *Stage 1 – recitation of number words*: According to the findings of psychological studies with human children, early in the process of learning to count, children acquire a list of tags (i.e. number words) which is then used (more or less) consistently throughout the counting attempts [3]. In order to reflect this finding, the goal of the first stage of the model training was to teach it to produce a number word sequence, corresponding to numbers from 1 to 10. As illustrated in figure 1, at this point the only input to the model is the trigger input. The training data set consisted of two temporal sequences of the length of 20:

- the first sequence, with trigger input equal to 0 and all target outputs equal to 0 at every time step, trained the

model to remain silent when trigger input is 0;

- the second sequence, with trigger input equal to 1, trained the model to recite the 10 number words during the first 10 time steps of the sequence, and remain silent for the remaining 10 time steps;

The training lasted for 10000 epochs with learning rate 0.01 and weights updated in an on-line fashion.

2) *Stage 2A – learning to count without gesture*: In this training condition, the model is extended with the visual input, but the proprioceptive input is not added (figure 1). The task to be solved by the network is the same in both conditions (without and with gesture) and is to produce the sequence of number words with length equal to the number of objects shown to the visual input. Due to the previously mentioned need for randomisation of the positions of objects, the number of possible arrangements of objects grows exponentially with the visual input size, and thus it is highly impractical to create a data set containing all possible locations of objects for all considered numbers. In order to alleviate this problem, the network was trained in an “on-line” fashion, using smaller data sets which changed in every epoch and contained different arrangements of objects for a particular number. Each such small data set consisted of 22 sequences. For every number of objects on the retina ranging from 0 to 10 inclusive two sequences were included in the data set. For the first one, the trigger output was set to 0 and the target output consisted of 20 time steps of “silence”. For the second one, the trigger output was set to 1 and the target output contained the correct sequence of number words for the particular set of objects, followed by silence until the end of sequence. The arrangement of the objects was drawn randomly, but was the same for the two sequences in the small data set referring to a particular number.

3) *Stage 2B – learning to count with gesture*: Model training in the condition with gesture was analogous to stage 2A, main difference being the addition of the proprioceptive gesture input (figure 1). The training was also performed using

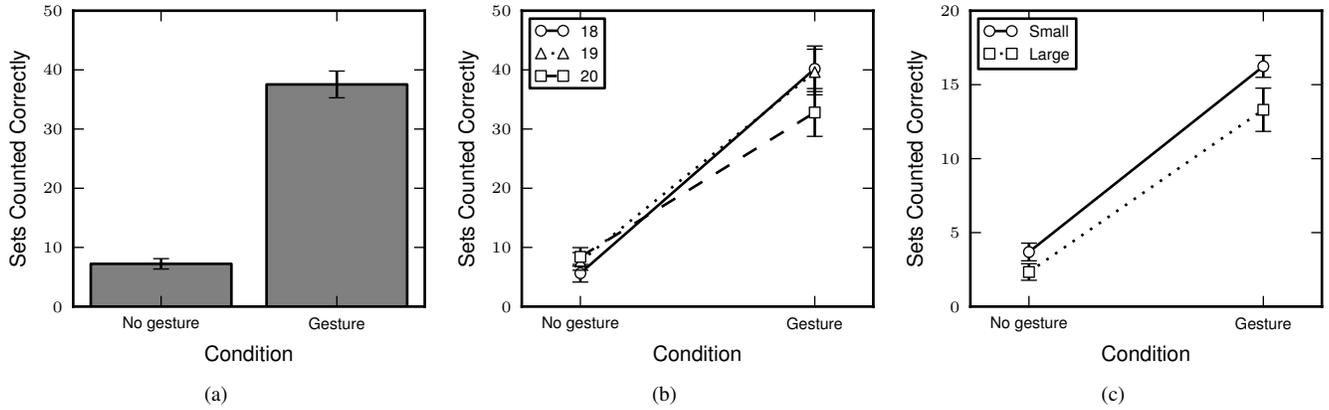


Figure 4. (a) mean number of collections from the evaluation data set counted correctly (out of 50) by the model without and with gesture; (b) profile plot of the mean number of collections counted correctly (out of 50) without and with gesture for models with 18, 19 and 20 hidden units. Visible interaction reached statistical significance (see main text); (c) profile plot of the mean number of collections counted correctly (out of 20) for small (1-4 items) and large (7-10 items) collections, without and with gesture. On all charts error bars indicate $\pm 2SEM$.

small datasets, however the sequences in the dataset contained an additional gesture signal constructed as follows. For the given arrangement of objects in the visual input, the locations of objects were considered in the left-to-right order. Assuming there were n objects on the retina, for time steps $t < n$, the proprioceptive input consisted of the joint angles (reduced to 3 dimensions as explained above) corresponding to the position of the retina at which the $(t + 1)$ -th object (in the left-to-right order) was located. For $n \leq t < 20$, proprioceptive input remained unchanged with respect to the values present for the last object. For all sequences for which the trigger input was set to 0 or the number of elements on the retina was 0, no gestural signal was provided (all 3 proprioceptive inputs set to 0). In stages 2A and 2B, training lasted for 4000 epochs using the same learning rate as stage 1.

D. Model evaluation

Evaluation of the model performance has been designed with an intention to yield maximum similarity to studies with human participants (more specifically, to [7], where authors investigate experimentally the function of gesture in learning to count with special focus on the distinction between keeping track and coordination of the recited number words with counted items), in order to allow subsequent comparison of the behavioural data. To that end, the 5-dimensional output of the model was first transformed into number words by means of nearest neighbour classification. Then, the resulting number words sequence was compared in detail with the target sequence (corresponding to the correct counting of presented items), using the same criteria of correctness and categorisation of possible errors as applied by [7]. Thus, the following counting errors were distinguished (see table 1 in [7]):

- Skip: the model does not assign a number word to an object;
- Continue: the model continues to count beyond the number of objects shown;

- Stop short: the model does not assign a number word to the last object(s);
- String error: the sequence of produced number words contains any other error;

Because of the nature of the model and the assumptions made about the gesture signal, “Double count” and “Distracted” errors described by [7] are not applicable in this study.

The data set used for model evaluation was independent from the training data sets and contained 50 sequences (5 for every number ranging from 1 to 10) with randomised locations of objects. The same evaluation data set was used to test the model in both gesture and no gesture conditions. However, consistent with the assumptions about the proprioceptive input, when the model trained in stage 2B was evaluated, sequences in the evaluation data set included the corresponding proprioceptive gesture signal.

In order to answer the question whether the addition of gesture to the designed model during training improves final counting accuracy, 32 independent repetitions of the model training were performed. In all trials the same sequence of number words was used, which is intended to correspond to the situation of testing children who use the same language. In order to verify the robustness of results with respect to the number of hidden units in the model, 18, 19 and 20 hidden units were used in 11, 11 and 10 of the trials respectively.

III. EXPERIMENTAL RESULTS

As mentioned above, the experimental set-up in this study was intentionally designed to resemble the one applied by [7]. Therefore, the analysis of the results of experiments with the proposed model presented below resembles in many aspects the one found in the quoted paper.

A. Learning the sequence of number words

In all 32 trials, the model successfully achieved the prerequisite task, i.e. has learnt to remain silent when the trigger input is 0, and to produce the correct sequence of number

words when the trigger input is 1. The final backpropagation through time training error (i.e. the mean squared error over all sequences in the training set) obtained in all trials was low, with the maximum $1.27 \cdot 10^{-4}$ (for repetition 14). In other words, in all trials the model learnt to perform the prerequisite task very well, despite differences in numbers of hidden units.

B. Gesturing and counting accuracy

Investigation of the effect of the proprioceptive gestural input on counting accuracy starts with a look at the progress of the model training in both no gesture and gesture conditions. In figure 3a a box plot of the training error for the last 30 epochs in both conditions for each trial is shown. For all but 1 repetition (29), the final training error in the gesture condition was considerably lower than in training without gesture. A typical plot of the training errors throughout the full stage 2A/B is shown in figure 3b.

Similarly to [7], we present an analysis of the correctness of the counting sequences produced by the model. In figure 4a the mean number of sets of objects from the evaluation data set counted correctly by the model in no gesture and gesture conditions is shown (this is analogous to Figure 1 in [7]). Statistical analysis of the results was performed in the form of a 3×2 (18, 19 and 20 hidden units times no gesturing and gesturing) repeated measures MANOVA with the gesturing condition as the within-subject factor (since for every experiment repetition the same model originating from the stage 1 of training was used in stage 2A and 2B) and the number of hidden units as the between-subject factor (in order to assess the stability of results with respect to the number of hidden units in the model). Comparison of the mean number of sets counted correctly for condition without and with gesture was a planned within-subject contrast. The analysis indicated strong statistical significance of the difference ($F = 633.686$, $p < 0.001$) between the gesturing conditions, meaning that models trained with proprioceptive gesture input available counted more collections in the evaluation data set correctly than those without this input. This is in perfect agreement with findings reported by [7], where children’s performance in the no gesture condition was significantly inferior to conditions with gesture. The between-subject effect of the number of hidden units in the model was not found ($F = 2.194$, $p = 0.13$), indicating that the beneficial effect of gesture was robust within the considered range of numbers of hidden units. However it has to be acknowledged that the effect of within-subject interaction between gesturing condition and the number of hidden units approached statistical significance ($F = 6.228$, $p = 0.006$). This means that in the conducted experiments the beneficial effect of the additional proprioceptive input was stronger for neural networks with less hidden units. This conclusion is illustrated in the profile plot of the estimated marginal means for gesture condition versus the number of hidden units shown in figure 4b.

Next step of the analysis presented in [7] focused on the dependence of the effect of gesturing on the size of the counted collection. In order to investigate this in the proposed model,

Table I
MEAN NUMBER OF CORRECTLY COUNTED COLLECTIONS (OUT OF 20)

Condition	Small sets	Large sets
no gesture (2A)	3.691 (0.299)	2.339 (0.280)
gesture (2B)	16.239 (0.372)	13.303 (0.731)

Table II
MODEL COUNTING ERRORS

	% of trials (out of 1600) with error made		% of models (out of 32) which made error	
	no gesture	gesture	no gesture	gesture
Skip	0.8	0.2	37.5	37.5
Continue	43.4	7.1	100.0	100.0
Stop short	39.3	14.4	100.0	100.0
String	13.1	3.9	93.8	93.8

the collections from the evaluation data set were divided into small numbers (1-4) and large numbers (7-10) and statistical analysis of the mean number of sets counted correctly within these groups was performed (meaning this time $3 \times 2 \times 2$ MANOVA). Obtained values are summarised in table I (which corresponds to Table 2 in [7]). In this table, numbers in parentheses show the standard error. Once again the behaviour of the model is in perfect agreement with the psychological study. Strong effects of gesture and set size were found ($F = 597.736$ and $F = 31.814$, respectively, $p < 0.001$ in both cases), while there was no interaction between these two factors ($F = 2.905$, $p = 0.099$). This means that the proposed model, similarly to children, counts small sets more accurately than large sets, and that gesturing improves its counting accuracy for both small and large collections of objects. This is illustrated in a profile plot in figure 4c.

C. Error patterns

Finally we look more closely at errors made by the model when counting without and with gesture. Table II reports the percent of trials with particular types of errors as well as the percentage of models which made particular kinds of errors. This is similar, but not equivalent to Table 3 published in [7], which focuses on the differences between active and passive gesture while herein such distinction is not made. When considering the general picture however, one can conclude with a fair degree of certainty that overall, patterns of errors made by the proposed model are different from the ones obtained in the study with children. While for children the most common errors are Skip and Double count, the model proposed in this paper makes Continue and Stop short errors more often. In contrast to 5% and 15% of children tested by [7] who committed a Continue and Stop short error at least once, the model made these errors at least once in every trial.

IV. CONCLUSIONS AND DISCUSSION

A recurrent artificial neural network model of the contribution of gesture to learning to count has been proposed. In an experimental set-up designed to allow comparison with human

data it has been confirmed that the proprioceptive gesturing signal enabled the model to improve its counting accuracy. The model behaviour yielded similarity to that of human children in terms of the effects of gesture and of the counted set size, although the obtained patterns of errors were different.

A few issues about the proposed model are discussed below. The first are the assumptions regarding the gesture signal (section II). In this study the gesture is an external input to the model. Although it may at first seem arbitrary and artificial, such approach is in line with a finding that children, when counting, apply the one-one correspondence principle in gesture before it is transferred to speech [1]. While designing a model which produces gesture is planned for future work, the focus of the present study was to test, from an “information-theoretic” point of view, if the proprioceptive signal carries information which may be exploited when learning to count. Obtained results confirm this is indeed the case.

Second, the proposed model makes different counting errors than human children. This may have been caused by two properties of the chosen model architecture. The assumption that the gesture is always correct affects the kinds of errors that may appear. More specifically, Double count errors do not appear at all, and Skip errors are also affected (although not ruled out). This may be seen as corresponding to the “puppet condition” in the study [7] where gesture performed by the puppet was also always correct. In addition, error patterns produced by the model may be influenced by time discretisation which is an inherent property for the Elman architecture. Here, the synchrony between gesture and number words recitation is naturally present. However, according to some hypotheses, synchronising the number words production with tagging the objects being counted may be one of the major functions of gesture [6]. Therefore, a model with continuous time would be more appropriate to investigate the importance of synchrony in the context of counting, and this is also included in the plans for future work. Error patterns would likely change as 4 out of 5 considered error types may appear as a result of problems with synchronisation.

Finally, the scalability and generalisability of the obtained results need to be addressed. As mentioned before, performed 32 experiment repetitions used the same sequence of number words. It has been confirmed however in earlier informal experiments with the model that the reported results are not due to any specific characteristic of the used number words sequence. These tests were also used to establish the training parameters (e.g. the number of epochs and the number of hidden units) for which the training of the model on the target task is successful. The crucial findings of this paper, i.e. the effect of gesture and set size on the counting accuracy, should hold for any reasonably chosen number words sequence length or retina size, as the particular values of these parameters were chosen arbitrarily and do not lead to any loss of generality. Of course this holds provided that the model architecture (most importantly the number of hidden units) and the training parameters are also adjusted accordingly.

Present study provides quantitative evidence in support of

the intuition that motor knowledge connected with pointing gesture can be transferred to a verbal and conceptual competence. Cognitive robotics approach to modelling alleviates the need for arbitrary assumptions about representing the proprioceptive signal, since a real robot which performs the actual pointing gesture is available. The analysis of the internal workings of the model and employing more sophisticated models are expected to shed even more light on the nature of the contribution of gesture to learning to count.

ACKNOWLEDGEMENT

This research has been supported by the EU project RobotDoC (235065) from the FP7 Marie Curie Actions ITN.

REFERENCES

- [1] T. A. Graham, “The role of gesture in children’s learning to count,” *Journal of Experimental Child Psychology*, vol. 74, no. 4, pp. 333–355, 1999.
- [2] B. Schaeffer, V. H. Eggleston, and J. L. Scott, “Number development in young children,” *Cognitive Psychology*, vol. 6, no. 3, pp. 357–379, 1974.
- [3] R. Gelman, “What young children know about numbers,” *Educational Psychologist*, vol. 15, no. 1, pp. 54–68, 1980.
- [4] G. B. Saxe and R. Kaplan, “Gesture in early counting: A developmental analysis,” *Perceptual and Motor Skills*, vol. 53, no. 3, pp. 851–854, 1981.
- [5] R. Gelman and E. Meck, “Preschoolers’ counting: Principles before skill,” *Cognition*, vol. 13, no. 3, pp. 343–359, 1983.
- [6] K. C. Fuson, *Children’s counting and concepts of number*, ser. Springer series in cognitive development. New York, NY, US: Springer-Verlag Publishing, 1988.
- [7] M. W. Alibali and A. A. DiRusso, “The function of gesture in learning to count: More than keeping track,” *Cognitive Development*, vol. 14, no. 1, pp. 37–56, 1999.
- [8] R. A. Carlson, M. N. Avraamides, M. Cary, and S. Strasberg, “What do the hands externalize in simple arithmetic?” *Journal of Experimental Psychology-Learning Memory and Cognition*, vol. 33, no. 4, pp. 747–756, 2007.
- [9] J. Piaget, *The child’s conception of number*. Oxford, England: W. W. Norton & Co., 1952.
- [10] P. Rodriguez, J. Wiles, and J. L. Elman, “A recurrent neural network that learns to count,” *Connection Science*, vol. 11, no. 1, pp. 5–40, 1999.
- [11] S. A. Peterson and T. J. Simon, “Computational evidence for the subitizing phenomenon as an emergent property of the human cognitive architecture,” *Cognitive Science*, vol. 24, no. 1, pp. 93–122, 2000.
- [12] K. Ahmad, M. Casey, and T. Bale, “Connectionist simulation of quantification skills,” *Connection Science*, vol. 14, no. 3, pp. 165–201, 2002.
- [13] R. Pfeifer, M. Lungarella, and F. Iida, “Self-organization, embodiment, and biologically inspired robotics,” *Science*, vol. 318, no. 5853, pp. 1088–1093, 2007.
- [14] G. Lakoff and R. Núñez, *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York, NY: Basic Books, 2000.
- [15] L. Barsalou, “Perceptual symbol systems,” *Behavioral and brain sciences*, vol. 22, no. 04, pp. 577–660, 1999.
- [16] M. Asada, K. MacDorman, H. Ishiguro, and Y. Kuniyoshi, “Cognitive developmental robotics as a new paradigm for the design of humanoid robots,” *Robotics and Autonomous Systems*, vol. 37, no. 2-3, pp. 185–193, 2001.
- [17] M. Ruciński, A. Cangelosi, and T. Belpaeme, “An embodied developmental robotic model of interactions between numbers and space,” in *33rd Annual Meeting of the Cognitive Science Society*, L. Carlson, C. Hoelscher, and T. F. Shipley, Eds., 2011, pp. 237–242.
- [18] G. Metta, L. Natale, F. Nori, G. Sandini, D. Vernon, L. Fadiga, C. von Hofsten, K. Rosander, M. Lopes, J. Santos-Victor, A. Bernardino, and L. Montesano, “The iCub humanoid robot: An open-systems platform for research in cognitive development,” *Neural Networks*, vol. 23, no. 8-9, pp. 1125–1134, 2010.
- [19] J. Elman, “Finding structure in time,” *Cognitive Science*, vol. 14, no. 2, pp. 179–211, 1990.