

# Vowel recognition with four coupled spin–torque nano–oscillators

Miguel Romera<sup>1,5</sup>, Philippe Talatchian<sup>1,5</sup>, Sumito Tsunegi<sup>2</sup>, Flavio Abreu Araujo<sup>1,4</sup>, Vincent Cros<sup>1</sup>, Paolo Bortolotti<sup>1</sup>, Juan Trastoy<sup>1</sup>, Kay Yakushiji<sup>2</sup>, Akio Fukushima<sup>2</sup>, Hitoshi Kubota<sup>2</sup>, Shinji Yuasa<sup>2</sup>, Maxence Ernout<sup>1,3</sup>, Damir Vodenicarevic<sup>3</sup>, Tifenn Hirtzlin<sup>3</sup>, Nicolas Locatelli<sup>3</sup>, Damien Querlioz<sup>3\*</sup> & Julie Grollier<sup>1\*</sup>

In recent years, artificial neural networks have become the flagship algorithm of artificial intelligence<sup>1</sup>. In these systems, neuron activation functions are static, and computing is achieved through standard arithmetic operations. By contrast, a prominent branch of neuroinspired computing embraces the dynamical nature of the brain and proposes to endow each component of a neural network with dynamical functionality, such as oscillations, and to rely on emergent physical phenomena, such as synchronization<sup>2–6</sup>, for solving complex problems with small networks<sup>7–11</sup>. This approach is especially interesting for hardware implementations, because emerging nanoelectronic devices can provide compact and energy-efficient nonlinear auto-oscillators that mimic the periodic spiking activity of biological neurons<sup>12–16</sup>. The dynamical couplings between oscillators can then be used to mediate the synaptic communication between the artificial neurons. One challenge for using nanodevices in this way is to achieve learning, which requires fine control and tuning of their coupled oscillations<sup>17</sup>; the dynamical features of nanodevices can be difficult to control and prone to noise and variability<sup>18</sup>. Here we show that the outstanding tunability of spintronic nano-oscillators—that is, the possibility of accurately controlling their frequency across a wide range, through electrical current and magnetic field—can be used to address this challenge. We successfully train a hardware network of four spin-torque nano-oscillators to recognize spoken vowels by tuning their frequencies according to an automatic real-time learning rule. We show that the high experimental recognition rates stem from the ability of these oscillators to synchronize. Our results demonstrate that non-trivial pattern classification tasks can be achieved with small hardware neural networks by endowing them with nonlinear dynamical features such as oscillations and synchronization.

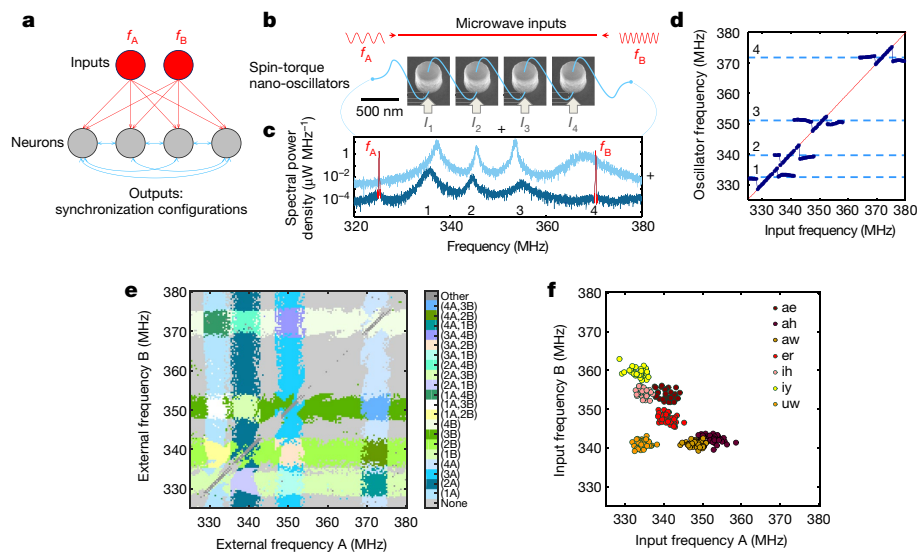
Spin-torque nano-oscillators are natural candidates for building hardware neural networks made of coupled nanoscale oscillators<sup>8–10,13,15,18,19</sup>. These nanoscale magnetic tunnel junctions emit microwave voltages when they are driven by direct-current injection in a regime of sustained magnetization precession through the effect of spin torque. In addition, they have exceptional capacities to synchronize their rhythms to periodic electric and magnetic input signals and to other spin-torque nano-oscillators<sup>20–24</sup>. This property originates from the high tunability of their frequency, in other words, the large frequency changes induced by applied d.c. currents and magnetic fields. Single spin-torque nano-oscillators can achieve impressive cognitive computations<sup>25</sup>. However, it has not been shown experimentally that a coupled network of spin-torque nano-oscillators can learn to perform computational tasks through synchronization. Here, we use the ability of spin-torque nano-oscillators to modify their frequency in response to injected direct currents to train in real-time a network of coupled oscillators to categorize different input patterns into different synchronization configurations<sup>2,17,18</sup>.

We transpose to hardware the neural network illustrated in Fig. 1a<sup>17</sup> with the set-up illustrated in Fig. 1b. The four neurons in Fig. 1a are experimentally implemented with four spin-torque nano-oscillators (Fig. 1b), in our case circular magnetic tunnel junctions with 375 nm diameter and an FeB free layer with a vortex as ground state (see Methods)<sup>26</sup>. The double arrow connections between neurons (blue in Fig. 1a) indicate that the output of neuron *i* influences the behaviour of neuron *j*, and vice versa. We implement these symmetric neural interconnections by connecting electrically the four oscillators using millimetre-long wires as schematized in Fig. 1b: in this configuration, the microwave current generated by each oscillator propagates in the electrical microwave loop and in turn influences the dynamics, and in particular the frequency, of the other oscillators through the microwave spin-torques it creates<sup>24</sup>. The sum of all microwave emissions is detected by a spectrum analyser. Importantly, we can control the frequency of each oscillator by adjusting the direct current flowing through each (see Methods and Extended Data Fig. 1). Here, for computing, we choose direct currents leading to close but not identical frequencies. The light blue curve in Fig. 1c shows a four-peak spectrum typical of this regime of moderate coupling where the dynamics of the oscillators are correlated but do not lead to mutual synchronization.

The inputs to the neural network are encoded in the frequencies  $f_A$  and  $f_B$  of two fixed-amplitude microwave signals. Injected in a strip line fabricated above the active magnetic layers, they modify the dynamics of the oscillators through the radiofrequency magnetic fields they generate. Figure 1d shows that when the frequency of one of the microwave sources is swept, each oscillator synchronizes to the source in turn. Indeed, when the frequency of the source gets close to the frequency of one of the oscillators, the strong signal of the source pulls the adaptable frequency of the oscillator towards its own. In the locking range, the frequency of the oscillator becomes equal to the frequency of the source<sup>27</sup>. The dark blue curve in Fig. 1c shows an example of spectrum measured when the two microwave inputs are injected simultaneously. Two peaks (in red) appear at frequencies  $f_A$  and  $f_B$  owing to capacitive coupling with the strip line. In comparison to the spectrum without inputs (light blue curve), the emission peaks of oscillators 1 and 2 are pulled towards  $f_A$ , whereas oscillator 4 is phase-locked to input B (its emission peak merges with the one of input B at  $f_B$ ). We label this synchronization configuration as (4B).

The possible outputs of the neural network, represented in different colours in Fig. 1e, are the different synchronization configurations that appear for different frequencies of the two input signals, keeping the direct currents through the oscillators fixed. Depending on the frequencies of inputs, zero (grey regions), one, or two oscillators are phase-locked. For example, in the petrol-blue region labelled (2A), oscillator 2 is synchronized to input A. In the white region labelled (1A,3B), oscillators 1 and 3 are synchronized to inputs A and B, respectively.

<sup>1</sup>Unité Mixte de Physique, CNRS, Thales, Université Paris-Sud, Université Paris-Saclay, Palaiseau, France. <sup>2</sup>National Institute of Advanced Industrial Science and Technology (AIST), Spintronics Research Center, Tsukuba, Ibaraki, Japan. <sup>3</sup>Centre de Nanosciences et de Nanotechnologies, CNRS, Université Paris-Sud, Université Paris-Saclay, Orsay, France. <sup>4</sup>Present address: Institute of Condensed Matter and Nanosciences, UC Louvain, Louvain-la-Neuve, Belgium. <sup>5</sup>These authors contributed equally: Miguel Romera, Philippe Talatchian. \*e-mail: damien.querlioz@u-psud.fr; julie.grollier@cnrs-thales.fr



**Fig. 1 | Approach for pattern classification with coupled spin-torque nano-oscillators.** **a**, Schematic of the emulated neural network. **b**, Schematic of the experimental set-up with four spin-torque nano-oscillators electrically connected in series and coupled through their own emitted microwave currents. Two microwave signals encoding information in their frequencies  $f_A$  and  $f_B$  are applied as inputs to the system through a strip line, which translates into two microwave fields. The total microwave output of the oscillator network is recorded with a spectrum analyser. **c**, Microwave output emitted by the network of four oscillators without (light blue) and with (dark blue) the two microwave signals applied to the system. The two curves have been shifted vertically for clarity. The four peaks in the light blue curve correspond to the emissions of the four

oscillators. The two narrow red peaks in the dark blue curve correspond to the external microwave signals with frequencies  $f_A$  and  $f_B$ . **d**, Evolution of the four oscillator frequencies when the frequency of external source A is swept. One after the other, the oscillators phase-lock to the external input when the frequency of the source approaches their natural frequency. In the locking range, the oscillator frequency is equal to the input frequency. **e**, Experimental synchronization map as a function of the frequencies of the external signals  $f_A$  and  $f_B$ . Each colour corresponds to a different synchronization state. **f**, Inputs applied to the system, represented in the  $(f_A, f_B)$  plane. Each colour corresponds to a different spoken vowel, and each data point corresponds to a different speaker.

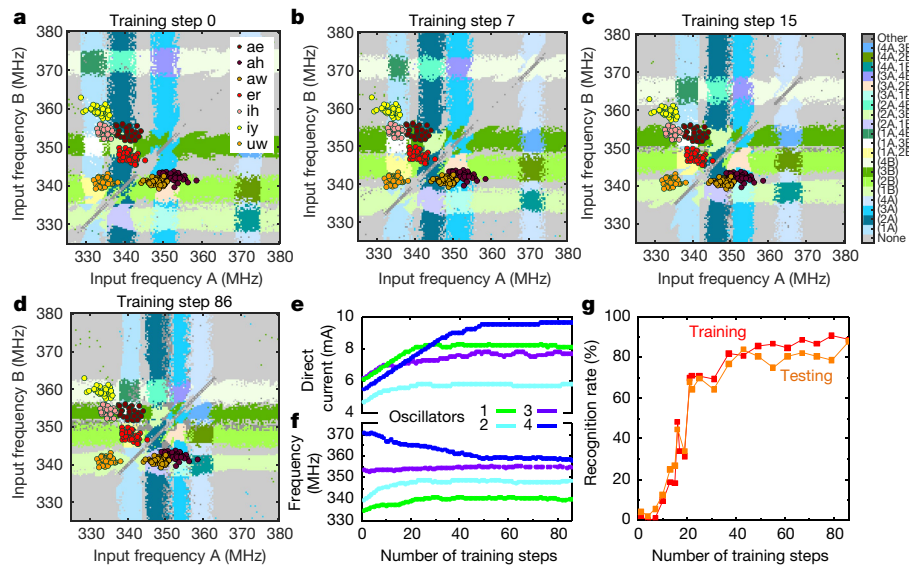
We now describe how this neural network can recognize patterns by classifying spoken vowels, which are naturally characterized by frequencies called formants<sup>28</sup>. We use as input data a subset of the Hillenbrand database (available at <https://homepages.wmich.edu/~hillenbr/voweldata.html>; see Supplementary Information) comprising seven vowels pronounced by 37 different female speakers, where each vowel is characterized by 12 different frequencies. Formant frequencies are typically in the range between 500 Hz and 3,500 Hz, so a transformation is needed to obtain input frequencies  $(f_A, f_B)$  in the range of operation of our oscillators, between 325 MHz and 380 MHz. As detailed in Methods, we obtain  $f_A$  and  $f_B$  through two different linear combinations of the 12 formant frequencies that fit the grid-like geometry of the oscillator synchronization maps. In the resulting map shown in Fig. 1f, each point corresponds to one speaker. The spread in frequency for each vowel indicates that each speaker has a different pronunciation. Our goal is to recognize the vowel presented as input to the oscillator network independently of the speaker. For this purpose, the scattered points corresponding to each vowel pronounced by different speakers should all be contained inside a different region of the oscillator synchronization map in Fig. 1e.

As can be seen from Fig. 2a, in which the input vowel map and the oscillator synchronization map are superposed, initially they do not coincide: the initial oscillator frequencies have been set randomly and are not adequate to solve the problem. The oscillatory neural network must learn to perform the classification properly. During this training stage, the internal parameters of the network need to be finely tuned until each synchronization region encompasses the cloud of points corresponding to the vowel that it has been assigned. For this purpose, we take advantage of the highly tunable nature of spin-torque nano-oscillators to modify the synchronization map by tuning the direct current through each oscillator, adapting a training algorithm first proposed in ref.<sup>17</sup>. We have developed an automatic real-time learning procedure involving a feedback loop between the experimental setup and the computer that controls it (see Methods). At each training step, we consecutively apply seven inputs  $(f_A, f_B)$  to the oscillators, one

for each vowel, randomly picked between the different speakers. The oscillator emissions corresponding to each of the seven input microwave signals are recorded with a spectrum analyser. A computer identifies the corresponding synchronization states (see Methods). If all the seven vowels have been correctly classified in their assigned synchronization regions of the map  $(f_A, f_B)$ , the direct currents are not changed. If one or several vowels have not been correctly classified, direct currents in the oscillators are modified to bring the assigned synchronization regions closer to the corresponding input frequency pairs  $(f_A, f_B)$  and thus reduce the classification error (see Methods). In the next learning step, another set of seven vowels is applied, and so on.

Figure 2 shows synchronization maps obtained at different stages of the training process (Fig. 2a–d), together with the evolution of the direct currents applied to the oscillators (Fig. 2e), their frequencies (Fig. 2f) and the average recognition rates for the seven vowels (Fig. 2g) (for a short video (20 s), see Supplementary Information or <https://youtu.be/bbRqQcxc-po>; for a longer video (3 min 30 s), see <https://youtu.be/IHYnh0oJgOA>). After 48 training steps, an optimum is found, direct currents and frequencies stop evolving, and the recognition rates stop increasing, signifying that the training process can be stopped. During training, we do not use all the vowels in the database. We always retain 20% of the vowels to test the ability of the system to recognize unknown data. The final recognition rates on the training and testing datasets reach values up to 89% and 88%, respectively (Fig. 2g).

We now interpret these experimental recognition rates by comparing them to the performances that can be achieved with ideal oscillators trained on the same task with the same learning process. For this purpose, we model the oscillator dynamics with coupled van der Pol equations accounting for their collective magnetization coordinates (see Supplementary Information)<sup>20</sup>. The simulated oscillators are noiseless and differ only by a 2% mismatch in their natural frequencies, analogous to the one observed experimentally. We first vary their ability to synchronize by modifying their frequency tunability (see Supplementary Information). Black circles in Fig. 3a show the recognition rate of the ideal simulated network as a function of the



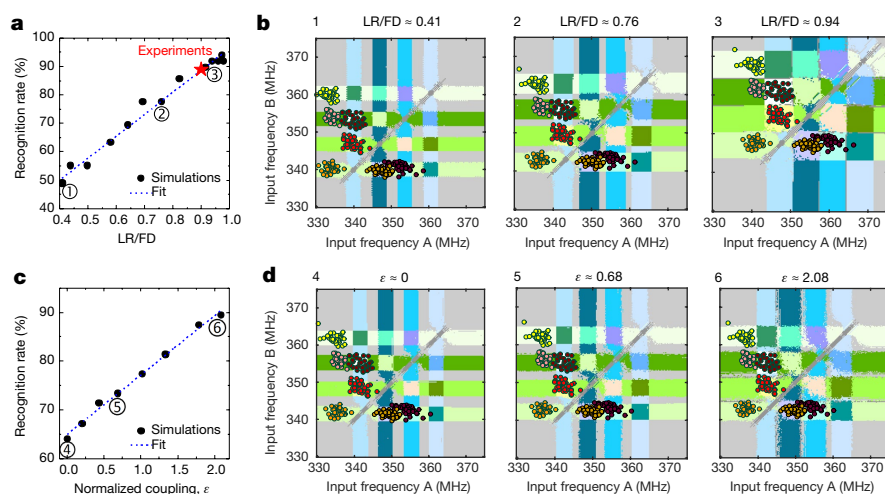
**Fig. 2 | Learning to classify patterns by tuning the frequencies of oscillators.** **a–d**, Experimental synchronization map as a function of the frequencies of the external signals, at different steps of the training procedure: **a**, step 0; **b**, step 7; **c**, step 15; and **d**, step 86. The coloured dots represent the inputs applied to the oscillatory network: vowels pronounced by different speakers. Different vowels are shown in different colours.

average locking range of the oscillators normalized by their frequency difference. The recognition rate increases linearly with the oscillator locking ranges (see dotted blue linear fit in Fig. 3a). Indeed, as shown in the simulated maps of Fig. 3b, when the oscillator locking ranges increase, the regions of synchronization grow, thus encompassing and classifying an increasing number of points in each of the different vowel clouds. As shown in Fig. 3c, d, the mutual coupling between oscillators also enhances their locking ranges<sup>27</sup>, leading to increased recognition rates when the mutual interactions increase. The red star in Fig. 3a pinpoints where the experimental result features in this graph. The experimental vowel recognition rate of 89% is close to the maximum recognition rate of 94% that can be achieved with the same neural

A video is provided as Supplementary Information. **e**, Direct current applied through each oscillator as a function of the number of training steps. **f**, Frequency of each oscillator as a function of the number of training steps. **g**, Recognition rates obtained with the sets of data points used for training and for testing, as a function of the number of training steps.

network composed of ideal, noiseless oscillators. This high performance is due to the large experimental locking ranges resulting from the high tunability, coupling and low noise of the hardware spin-torque nano-oscillators.

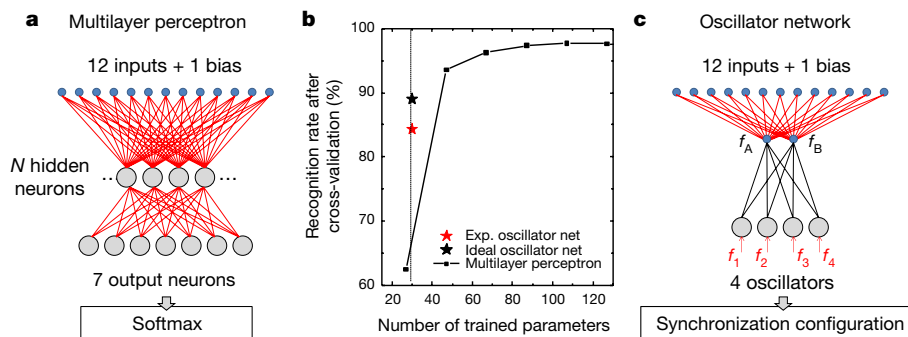
We then compare the dynamical oscillator-based neural network studied in this paper to more conventional forms of neural networks. For this purpose, we first extract a reference value for the experimental recognition rate by repeating the training procedure experimentally several times with different combinations of training and testing sets (see Methods). This cross-validation technique yields an average value of 84.3% for the experimental recognition rate on the testing set that we can compare to other neural networks performances. First, we



**Fig. 3 | Comparing the recognition rates of experimental and ideal oscillators.** Simulations of vowel recognition with a network of four identical oscillators trained with the same procedure as in the experiments are illustrated, in the absence of noise. The simulated oscillators differ only by a 2% mismatch in their natural frequencies. **a**, Recognition rate on the training set (black circles) as a function of the average oscillator locking range normalized by the frequency difference between oscillators (LR/FD). The locking range is varied by modifying the tunability of the oscillator frequency. The blue dotted line is a linear

fit to the simulation results. The red star indicates where experimental oscillators feature in this graph. **b**, Synchronization maps simulated with the network of oscillators used in **a**, for three different values of the normalized locking range. **c**, Recognition rate on the training set (black circles) as a function of the mutual coupling between oscillators normalized by their coupling to the microwave inputs. The blue dotted line is a linear fit to the simulation results. **d**, Synchronization maps simulated with the network of oscillators used in **c**, for three different values of the normalized coupling  $\epsilon$ .





**Fig. 4 | Benchmarking performances with classical neural networks.** **a**, Flow chart of the simulated multilayer perceptron. The trained parameters are indicated in red. **b**, Recognition rate obtained through cross-validation versus the total number of trained parameters for the neural network in **a**, in which the number of hidden neurons is varied.

consider a conventional, static, multi-layer neural network. This kind of network can achieve better-than-human recognition rates at complex tasks, such as image classification. This performance, however, comes at the expense of the large number of parameters that need to be trained, a major hurdle for hardware implementation. Figure 4b shows the recognition rate of a multilayer perceptron, trained in software through backpropagation on the same database as the experimental neural network, with 30,000 vowel presentations (see Methods). As illustrated in Fig. 4a, this network, composed of static neurons, takes as inputs the 12 formant frequencies characterizing each pronounced vowel. The hidden layer neurons receive a weighted sum of these inputs (plus a bias term). The output layer, with softmax activation functions, has seven neurons, one for each vowel class (see Methods). As can be seen in Fig. 4b, the recognition rate is excellent, reaching 97% when the number of trained parameters is large (synaptic weights illustrated in red in Fig. 4a). However, the performance rapidly degrades for small numbers of trained parameters, diving below 65% for 27 trained parameters. This result is quite general: as can be seen from Extended Data Fig. 2, state-of-the-art networks with feedback such as standard recurrent neural networks or long short-term memory networks have limited performance when the number of trained parameters is small. In contrast, the recognition rate of our experimental oscillatory neural network is over 84% for only 30 trained parameters: as illustrated in red in Fig. 4c, the 26 weights converting formants to inputs, and the currents through the oscillators. For an ideal, noiseless, oscillatory network, the success rate reaches 89% after cross-validation. The network also learns rapidly (350 vowel presentations are used). This high performance with a small number of trained parameters comes from the combination of two phenomena: as shown in Fig. 3c, the oscillatory network can do better than the sum of its individual components, owing to its complex, coupled, dynamical features, and in addition, the oscillators collectively contribute to pattern recognition by synchronizing to the inputs. This result shows that the performance of hardware neural networks can be boosted by enhancing neuron functionalities beyond simple nonlinear activation functions, through oscillations and synchronization.

In the future, such dynamical neural networks will have to be scaled up to solve challenging classification problems on software-benchmarked databases. Spin-torque nano-oscillators offer numerous advantages towards this goal. Their energy consumption is comparable to or lower than complementary metal-oxide-semiconductor (CMOS) oscillators, and contrary to the latter, their lateral dimensions can be scaled down to a few nanometres in diameter (a detailed comparison is presented in Extended Data Table 2). Their quality factor can exceed several thousands<sup>26</sup>, and their natural frequency can be controlled by the aspect ratio of the magnetic dot from hundreds of megahertz to several gigahertz in small pillars, opening the path to nano-oscillators assemblies with a wide range of natural frequencies<sup>19</sup>. In addition, their simple structure is similar to spin-torque magnetic random access memory cells, which means that they can be produced by billions

The red star corresponds to the experimental results with the network of spin-torque nano-oscillators. Exp., experimental. **c**, Flow chart of the experimental oscillatory neural network. The trained parameters are indicated in red.

on top of CMOS. Finally, their synchronization can be detected with CMOS circuits that count the number of oscillations<sup>29</sup> or measure the additional d.c. voltages produced by the oscillators when they phase-lock (see Methods and Extended Data Fig. 3)<sup>30</sup>. Therefore, the wide variety of possible magnetic and electric couplings offered by spintronics<sup>21–24</sup>, and the different ways of driving and controlling magnetization dynamics (spin torques, spin-orbit torques, electric fields), could be exploited in the future to implement large-scale hardware neural networks<sup>15</sup>.

### Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at <https://doi.org/10.1038/s41586-018-0632-y>.

Received: 24 November 2017; Accepted: 31 July 2018;

Published online 29 October 2018.

- Silver, D. et al. Mastering the game of Go without human knowledge. *Nature* **550**, 354–359 (2017).
- Borisov, R., Denham, M., Hoppensteadt, F., Kazanovich, Y. & Vinogradova, O. An oscillatory neural network model of sparse distributed memory and novelty detection. *Biosystems* **58**, 265–272 (2000).
- Jaeger, H. & Haas, H. Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. *Science* **304**, 78–80 (2004).
- Rabinovich, M., Huerta, R. & Laurent, G. Transient dynamics for neural processing. *Science* **321**, 48–50 (2008).
- Sussillo, D. Neural circuits as computational dynamical systems. *Curr. Opin. Neurobiol.* **25**, 156–163 (2014).
- Pikovsky, A. & Rosenblum, M. Dynamics of globally coupled oscillators: progress and perspectives. *Chaos* **25**, 097616 (2015).
- Kumar, S., Strachan, J. P. & Williams, R. S. Chaotic dynamics in nanoscale NbO<sub>2</sub> Mott memristors for analogue computing. *Nature* **548**, 318–321 (2017).
- Csaba, G. & Porod, W. Computational study of spin-torque oscillator interactions for non-Boolean computing applications. *IEEE Trans. Magn.* **49**, 4447–4451 (2013).
- Yogendra, K., Fan, D., Jung, B. & Roy, K. Magnetic pattern recognition using injection-locked spin-torque nano-oscillators. *IEEE Trans. Electron Dev.* **63**, 1674–1680 (2016).
- Macià, F., Kent, A. D. & Hoppensteadt, F. C. Spin-wave interference patterns created by spin-torque nano-oscillators for memory and computation. *Nanotechnology* **22**, 095301 (2011).
- Fang, Y., Yashin, V. V., Levitan, S. P. & Balazs, A. C. Pattern recognition with “materials that compute”. *Sci. Adv.* **2**, e1601114 (2016).
- Pickett, M. D., Medeiros-Ribeiro, G. & Williams, R. S. A scalable neuristor built with Mott memristors. *Nat. Mater.* **12**, 114–117 (2013).
- Pufall, M. R. et al. Physical implementation of coherently coupled oscillator networks. *IEEE J. Explor. Solid-State Comput. Devices Circuits* **1**, 76–84 (2015).
- Sharma, A. A., Bain, J. A. & Weldon, J. A. Phase coupling and control of oxide-based oscillators for neuromorphic computing. *IEEE J. Explor. Solid-State Comput. Devices Circuits* **1**, 58–66 (2015).
- Grollier, J., Querlioz, D. & Stiles, M. D. Spintronic nanodevices for bioinspired computing. *Proc. IEEE* **104**, 2024–2039 (2016).
- Parihar, A., Shukla, N., Jerry, M., Datta, S. & Raychowdhury, A. Computational paradigms using oscillatory networks based on state-transition devices. In *2017 International Joint Conference on Neural Networks (IJCNN)* 3415–3422 (IEEE, 2017).
- Vassiliev, E., Pinto, G., de Barros, J. A. & Suppes, P. Learning pattern recognition through quasi-synchronization of phase oscillators. *IEEE Trans. Neural Netw.* **22**, 84–95 (2011).



18. Vodenicarevic, D., Locatelli, N., Araujo, F. A., Grollier, J. & Querlioz, D. A nanotechnology-ready computing scheme based on a weakly coupled oscillator network. *Sci. Rep.* **7**, 44772 (2017).
  19. Locatelli, N., Cros, V. & Grollier, J. Spin-torque building blocks. *Nat. Mater.* **13**, 11–20 (2014).
  20. Slavin, A. & Tiberkevich, V. Nonlinear auto-oscillator theory of microwave generation by spin-polarized current. *IEEE Trans. Magn.* **45**, 1875–1918 (2009).
  21. Kaka, S. et al. Mutual phase-locking of microwave spin torque nano-oscillators. *Nature* **437**, 389–392 (2005).
  22. Mancoff, F. B., Rizzo, N. D., Engel, B. N. & Tehrani, S. Phase-locking in double-point-contact spin-transfer devices. *Nature* **437**, 393–395 (2005).
  23. Houshang, A. et al. Spin-wave-beam driven synchronization of nanocontact spin-torque oscillators. *Nat. Nanotech.* **11**, 280–286 (2016).
  24. Lebrun, R. et al. Mutual synchronization of spin torque nano-oscillators through a long-range and tunable electrical coupling scheme. *Nat. Commun.* **8**, 15825 (2017).
  25. Torreon, J. et al. Neuromorphic computing with nanoscale spintronic oscillators. *Nature* **547**, 428–431 (2017).
  26. Tsunegi, S. et al. High emission power and  $Q$  factor in spin torque vortex oscillator consisting of FeB free layer. *Appl. Phys. Express* **7**, 063009 (2014).
  27. Romera, M. et al. Enhancing the injection locking range of spin torque oscillators through mutual coupling. *Appl. Phys. Lett.* **109**, 252404 (2016).
  28. Hillenbrand, J., Getty, L. A., Wheeler, K. & Clark, M. J. Acoustic characteristics of American English vowels. *J. Acoust. Soc. Am.* **97**, 3099–3111 (1994).
  29. Vodenicarevic, D., Locatelli, N., Grollier, J. & Querlioz, D. Synchronization detection in networks of coupled oscillators for pattern recognition. In *2016 International Joint Conference on Neural Networks (IJCNN)* 2015–2022 (IEEE, 2016).
  30. Fang, B. et al. Giant spin-torque diode sensitivity in the absence of bias magnetic field. *Nat. Commun.* **7**, 11259 (2016).
- Acknowledgements** This work was supported by the European Research Council ERC under grant bioSPiNspired 682955, the French National Research Agency (ANR) under grant MEMOS ANR-14-CE26-0021, and a public grant overseen by the ANR as part of the ‘Investissements d’Avenir’ programme (Labex NanoSaclay, reference ANR-10-LABX-0035).
- Reviewer information** *Nature* thanks F. Hoppensteadt, A. Kent and the other anonymous reviewer(s) for their contribution to the peer review of this work.
- Author contributions** The study was designed by J.G. and D.Q. Samples were optimized and fabricated by S.T. and K.Y. The main experiments were performed by M.R. and P.T. Spin diode experiments were performed by P.T. and J.T. Numerical simulations were realized by P.T., M.E., M.R., T.H. and D.V. All authors contributed to analysing the results and writing the paper.
- Competing interests** The authors declare no competing interests.
- Additional information**  
**Extended data** is available for this paper at <https://doi.org/10.1038/s41586-018-0632-y>.  
**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-018-0632-y>.  
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.  
**Correspondence and requests for materials** should be addressed to D.Q. or J.G.  
**Publisher’s note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## METHODS

**Samples.** Magnetic tunnel junction (MTJ) films with a stacking structure of buffer/PtMn(15)/Co<sub>71</sub>Fe<sub>29</sub>(2.5)/Ru(0.9)/Co<sub>60</sub>Fe<sub>20</sub>B<sub>20</sub>(1.6)/Co<sub>70</sub>Fe<sub>30</sub>(0.8)/MgO(1)/Fe<sub>80</sub>B<sub>20</sub>(6)/MgO(1)/Ta(8)/Ru(7) (thicknesses in nm) were prepared by ultrahigh-vacuum (UHV) magnetron sputtering. After annealing at 360 °C for 1 h, the resistance–area product was  $RA \approx 3.6 \Omega \mu\text{m}^2$ . Circular-shaped MTJs with a diameter of about 375 nm were patterned using Ar ion etching and e-beam lithography. The resistance of the samples is close to 40  $\Omega$ , and the magneto-resistance ratio is about 100% at room temperature. The FeB layer presents a structure with a single magnetic vortex as the ground state for the dimensions used here. In a small region called the vortex core (of about 12 nm diameter at remanence for our materials), the magnetization spirals out of plane. Under direct current injection and the action of the spin transfer torques, the core of the vortex steadily gyrates around the centre of the dot with a frequency in the range of 150 MHz to 450 MHz for the oscillators we used here.

**Database and inputs.** In this study, we classify seven spoken vowels with the oscillatory network. Spoken vowels are characterized by a set of frequencies called formants, which we obtain from a subset of the Hillenbrand database (<https://homepages.wmich.edu/~hillenbr/voweldata.html>) given in Supplementary Information. We use the first three formants ( $F_1$ ,  $F_2$  and  $F_3$ ) sampled at four different times of the duration of the spoken vowel: at the steady state and at 20%, 50% and 80% of the vowel duration (that is, 12 parameters in total). When one of these 12 parameters could not be measured, or when irresolvable formants mergers occurred, Hillenbrand et al.<sup>28</sup> put a zero in this parameter in the database. For our study, we have removed the vowel utterances whose corresponding set of formants is not complete. Moreover, we use the same number of speakers for each vowel. The resulting formant database comprising 37 female speakers that we used is provided as Supplementary Data.

We perform two linear combinations of these formants to obtain two characteristic frequencies ( $f_A$  and  $f_B$ ) in the range of operation of the spin-torque nano-oscillators (between 325 MHz and 380 MHz for the applied field value that we are using):

$$\begin{aligned} f_A &= A_1 F_1^{\text{steady\_state}} + B_1 F_2^{\text{steady\_state}} + C_1 F_3^{\text{steady\_state}} + D_1 F_1^{20\%} \\ &\quad + E_1 F_2^{20\%} + G_1 F_3^{20\%} + H_1 F_1^{50\%} + I_1 F_2^{50\%} + J_1 F_3^{50\%} + K_1 F_1^{80\%} \\ &\quad + L_1 F_2^{80\%} + M_1 F_3^{80\%} + N_1 \\ f_B &= A_2 F_1^{\text{steady\_state}} + B_2 F_2^{\text{steady\_state}} + C_2 F_3^{\text{steady\_state}} + D_2 F_1^{20\%} \\ &\quad + E_2 F_2^{20\%} + G_2 F_3^{20\%} + H_2 F_1^{50\%} + I_2 F_2^{50\%} + J_2 F_3^{50\%} \\ &\quad + K_2 F_1^{80\%} + L_2 F_2^{80\%} + M_2 F_3^{80\%} + N_2 \end{aligned}$$

To choose the coefficients of the two linear combinations, we first record an experimental synchronization map that is used as a calibration of the network. The calibration map allows to assign a synchronization pattern to each vowel. Then, the linear transformation of the formants that best matches the data points of each vowel with its associated synchronization pattern is determined through fitting by least-square regression. The coefficients used in the two linear combinations and the two frequencies  $f_A$  and  $f_B$  corresponding to each vowel are provided as Supplementary Data.

Once this calibration is done and the coefficients and characteristic frequencies are calculated, the direct currents are reset to random values to begin the learning experiment. Two fixed-amplitude microwave signals with frequencies  $f_A$  and  $f_B$  are used as inputs to the experimental network of coupled nano-oscillators.

**Experimental set-up.** Extended Data Fig. 1 shows a schematic of the experimental set-up with the four coupled vortex nano-oscillators. A magnetic field of  $\mu_0 H = 530$  mT is applied perpendicularly to the oscillator layers to get an efficient spin transfer torque acting on the oscillator vortex core. A direct current is injected into each oscillator to induce vortex dynamics, which leads to periodic oscillations of the magnetoresistance, giving rise to an oscillating voltage at the same frequency than the vortex core dynamics. The four oscillators are electrically connected in series by millimetre-long wires. They are therefore coupled through the microwave currents they emit, and too far away to be coupled through the magnetic dipolar fields that they radiate. Four direct currents ( $I_{DC1}$ ,  $I_{DC2}$ ,  $I_{DC3}$ ,  $I_{DC4}$ ) are supplied to the circuit by four different sources, allowing an independent control of the current flowing through each oscillator. The actual current flowing through each spin-torque oscillator is given by  $I_{STO1} = I_{DC1}$ ,  $I_{STO2} = I_{DC2} + I_{DC1}$ ,  $I_{STO3} = I_{DC3} + I_{DC2} + I_{DC1}$  and  $I_{STO4} = I_{DC4} + I_{DC3} + I_{DC2} + I_{DC1}$ , respectively, where  $I_{STOi}$  corresponds to the current flowing through the  $i$ th oscillator. Two microwave sources are used to inject two external microwave signals with frequencies  $f_A$  and  $f_B$  and power  $P = -9$  dBm through a strip line, creating two microwave fields as inputs to the oscillator network. The amplitude of the generated magnetic field, set by Ampere's law, depends only on the cross-section of the antenna (in addition to the distance between the strip line and the active magnetic layer of the oscillators). Therefore, the length of the antenna is only set by the number of oscillators it should cover. In our case, the strip line has a width of 2.5  $\mu\text{m}$  and is fabricated 370 nm above the pillar (separated by an insulating layer). The resulting input microwave fields have an amplitude of

0.1 mT. They strongly affect the magnetization dynamics of the four oscillators, and thus the total microwave output emitted by the network. The microwave emissions are recorded with a spectrum analyser. As can be seen in Fig. 1d, the input signals from the antenna can be detected in addition to the oscillator emissions due to capacitive coupling between the strip line antenna and the metallic electrodes connecting the oscillator. The analysis of the output, which depends on the frequencies of the microwave inputs, can therefore easily be used to classify the spoken vowels.

Each spectrum recorded with the spectrum analyser is sent to the computer, where it is analysed by a program in real time. The information we use as input to this program is: (1) the value of the two frequencies of the external microwave signals ( $f_A$ ,  $f_B$ ) and (2) the oscillator frequencies at each direct current value in the absence of external microwave signals ( $f_1^0$ ,  $f_2^0$ ,  $f_3^0$ ,  $f_4^0$ ). The output data that we extract from each spectrum analysis are the four values of the oscillator frequencies in the presence of microwave inputs. Then, another program takes these oscillator frequencies to calculate the synchronization states and check whether the applied vowel was properly recognized, as follows. If one of the detected frequencies coincides with the frequency of one of the external signals ( $\pm 0.5$  MHz), we consider that the oscillator is synchronized to it. From this analysis, the synchronization pattern that corresponds to the input vowel is calculated. This is compared to the synchronization pattern initially assigned to that specific vowel to check whether it was successfully classified.

If we are in the training procedure and the vowel is not properly classified, the online learning algorithm calculates how the four direct currents should be modified to reduce the recognition error, as described in 'Real-time learning algorithm' below. This information is then sent back to the experimental set-up, where the currents are automatically modified.

**Real-time learning algorithm.** In this section, we present the supervised learning procedure that was applied to our spin-torque nano-oscillator network to learn to recognize different classes of input stimuli. Here these classes correspond to seven different spoken English vowels: ae, ah, aw, er, ih, iy and uw (see ref.<sup>28</sup> for details; the sounds can be heard at <https://homepages.wmich.edu/~hillenbr/voweldata.html>). Initially, we assign a synchronization pattern to each class of vowel (column 2 in Extended Data Table 1).

For a perfect recognition of one class of vowel, all data points in the frequency input map that corresponds to this vowel (Fig. 1f) must be contained in their assigned synchronization pattern in the experimental map (Fig. 1e). If this is not the case, for each association spoken vowel–synchronization pattern we define a frequency difference vector with four components (one for each oscillator; see third column in Extended Data Table 1) that will be used in the learning procedure.

Starting from a random map configuration (Fig. 1e), the automatic learning rule that we developed allows us to converge to a configuration where most data points for each vowel class are contained in their respective assigned synchronization pattern. The learning rule works in the following way.

(1) We present to the network a randomly chosen input data point  $i$  belonging to one vowel class, by sending two microwave inputs with frequencies  $f_A^i$  and  $f_B^i$ .

(2) From the resulting spectra, we extract the frequencies of the four spin-torque oscillators ( $f_1, f_2, f_3, f_4$ ) in presence of the microwave inputs.

(3) We determine the resulting synchronization configurations by comparing the oscillator frequencies to the input frequencies  $f_A^i$  and  $f_B^i$ . Then, we compare the obtained synchronization configuration with the one assigned to this vowel.

(4) For each vowel presented to the network, we define an associated frequency difference vector, which describes the frequency distance between the applied input and the assigned synchronization region. For instance, if the presented data point belongs to the vowel class 'ae', we compute  $\mathbf{d}_{ae} = [(f_A^i - f_1), 0, (f_B^i - f_3), 0]^T$ .

If one of the two synchronization events assigned to 'ae' has occurred, we only compute the frequency difference that corresponds to the other event. For instance, if oscillator 1 is correctly synchronized to external source  $f_A^i$ , then we compute only  $\mathbf{d}_{ae} = [0, 0, (f_B^i - f_3), 0]^T$ .

(5) We repeat steps (1) to (4) for all seven vowel classes.

(6) We compute the sign of the vector sum of all seven associated frequency difference vectors  $\mathbf{D}: \mathbf{D} = \text{sgn}(\mathbf{d}_{ae} + \mathbf{d}_{ah} + \mathbf{d}_{aw} + \mathbf{d}_{er} + \mathbf{d}_{ih} + \mathbf{d}_{iy} + \mathbf{d}_{uw}) = (D_1, D_2, D_3, D_4)^T$ .

(7) We then compute the new direct current set ( $I_1', I_2', I_3', I_4'$ ), which will be applied to the four oscillators:

$$\begin{pmatrix} I_1' \\ I_2' \\ I_3' \\ I_4' \end{pmatrix} = \begin{pmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{pmatrix} + \mu \begin{pmatrix} D_1 \text{sgn}\left(\frac{\partial \omega_1}{\partial I}\right)_{I=I_1} \\ D_2 \text{sgn}\left(\frac{\partial \omega_2}{\partial I}\right)_{I=I_2} \\ D_3 \text{sgn}\left(\frac{\partial \omega_3}{\partial I}\right)_{I=I_3} \\ D_4 \text{sgn}\left(\frac{\partial \omega_4}{\partial I}\right)_{I=I_4} \end{pmatrix}$$

In this equation,  $\mu = 0.1$  mA is the learning rate of our algorithm. At each step, the applied direct current through each oscillator can be modified only by  $\pm\mu$ . Here  $\text{sgn}[(\partial f_k / \partial I)_{I=I_k}]$  represents the sign of the frequency evolution versus injected direct current of the  $k$ th oscillator at the value of current  $I_k$ . For this, the frequency–current dependence of each independent oscillator has been previously characterized.

Upon modifying the direct currents following this learning procedure, the oscillator frequencies change. This translates into a displacement of the synchronization patterns in the experimental synchronization map (Fig. 2a–d).

(8) We repeat all previous steps (steps (1) to (7))  $N$  times, where  $N$  is the total number of training steps. At each iteration, the synchronization map evolves towards an optimal configuration where the global frequency difference vector  $\mathbf{d}_{\text{tot}} = \mathbf{d}_{\text{ae}} + \mathbf{d}_{\text{ah}} + \mathbf{d}_{\text{aw}} + \mathbf{d}_{\text{er}} + \mathbf{d}_{\text{ih}} + \mathbf{d}_{\text{iy}} + \mathbf{d}_{\text{uw}}$  is minimized. On increasing the number of training steps, we observe an increase of the recognition rate until it saturates after step 48, reaching a value of 89% (Fig. 2f). In our training experiment, we set the maximum number of training steps to  $N = 87$ , which corresponds to applying three times each of the 29 data points of the training database.

**Cross-validation procedure.** Training was realized using 80% of the total number of vowels in the database. The testing procedure was done using the remaining 20% data points. The cross-validation technique allows estimating accurately the recognition performances of the network by repeating the training/testing procedure five times over distinct data point samples. Each time, the selected data points used for testing are different: in the first (respectively second, third, fourth and fifth) cross-validation period, we use the first (respectively second, third, fourth and fifth) quintile (20%) of the data points for testing. The final recognition rate was obtained by averaging the testing recognition rates of the five cross-validation experiments. The same cross-validation procedure is used for all the neural networks (experimental and simulated).

**Comparison of spin-torque nano-oscillators to CMOS oscillators.** Extended Data Table 2 compares features of CMOS and spin-torque nano-oscillators. ‘Vortex spin-torque oscillators’ refer to the magnetic tunnel junctions used in this study; ‘10 nm spin-torque oscillators’ refer to state-of-the-art magnetic tunnel junctions currently used as memory cells.

**Comparison with a multilayer perceptron.** To benchmark the results of the experimental oscillatory network, we first ran a standard multi-layer perceptron, schematized in Fig. 4a, on the same vowel database.

The network takes as inputs the 12 formants of a given vowel in a database and has seven outputs, one for each vowel class. We have varied the number of hidden neurons between 1 and 20 to evaluate the recognition rate as a function of the number of trained parameters. More precisely, each formant has been rescaled between  $-1$  and  $1$  before being fed into the first layer of neurons. The neuron activation functions are tanh functions at the hidden layer, and softmax at the output layer: the outputs  $z_i$  ( $i = 1$  to 7) are defined as  $z_i = e^{y_i} / \sum_{j=1}^7 e^{y_j}$ , where  $y_j$  is the input to the output neuron  $j$ . The output with the largest  $z_i$  is taken as the vowel class corresponding to the input. We also tried ReLU activation functions, but they performed worse than tanh on this task.

For training the network we performed backpropagation, that is, gradient descent over the negative log-likelihood (or cross entropy).

As in the experimental conditions, the samples are picked and presented randomly to the network. One learning iteration corresponds to one forward pass of a given sample through the network, its subsequent gradient evaluation and weight update. The learning rate has been tuned to obtain the best result. Weights and biases before learning were randomly sampled from a Gaussian of mean 0 and variance 0.01.

For each trial, we ran training over 100,000 iterations to ensure convergence with a learning rate of 0.05. In practice, optimization techniques such as root-mean-square propagation or adaptive moment estimation could be used to accelerate training. All results are reported in Fig. 4b, where we show the recognition rate after cross validation as a function of the number of trained parameters.

**Comparison with RNNs.** In addition to the multilayer perceptron (Extended Data Fig. 2b), we also ran, on the same vowel database, a perceptron (Extended Data Fig. 2c), as well as a recurrent neural network (RNN; Extended Data Fig. 2d) and a long short-term memory network (LSTM) recurrent neural network (Extended Data Fig. 2e) with four hidden units. The procedure is similar to the multilayer perceptron. Formants are presented sequentially to the network which outputs a vowel once all of them have been swept through. Softmax activation functions were used at the output layer and tanh elsewhere. Outputs are encoded in a ‘one-hot’ fashion: for example, the ae vowel (out of the seven in total) is encoded by (1,0,0,0,0,0). We take the maximum activation value as the classification result. As in the experimental conditions, the samples are picked and presented randomly to the network. One learning iteration corresponds to one forward pass of a given sample through the network, its subsequent gradient evaluation and weight update. For each architecture, the choice of the learning rate has been tuned to obtain the best result. Weights and biases before learning were randomly sampled from a

Gaussian of mean 0 and variance 0.01. No gradient inertia or learning rate adaptation technique was used. For the LSTM and the RNN, we ran training over 500,000 and over 1,000,000 iterations to ensure convergence with a learning rate of 0.01 and 0.0005, respectively. If needed, optimization techniques such as root-mean-square propagation or adaptive moment estimation could be used to accelerate training. Owing to the mini-batch size, gradient descent is highly stochastic, and we average the test and training rates over the last 5,000 iterations to obtain reliable training and error rate for a given trial. All results are reported in Extended Data Fig. 2a where we show the cross-validation success as a function of the number of parameters learnt.

**Synchronization detection through oscillator rectified voltages.** In the present work, synchronization of the oscillators is detected using a spectrum analyser, allowing a comprehensive understanding of the systems and of the physics of the oscillators. In a final integrated system, simpler techniques could be used to detect synchronization of oscillators. A possibility is given in ref. 29. Another method, involving less energy overhead, consists in exploiting the spin diode effect<sup>31</sup>, which causes synchronized oscillators to generate a supplementary direct voltage<sup>32</sup>. Extended Data Fig. 3a and b illustrates this effect in one of our oscillators. The appearance of a rectified voltage measured between the oscillator electrodes (Extended Data Fig. 3a) coincides with the locking range (Extended Data Fig. 3b). The generated rectified voltage is proportional to the fraction of the external microwave current  $I_{\text{ext}}$  flowing through the oscillator<sup>30,32</sup>. In our experiments,  $I_{\text{ext}}$  is small: the input microwave signals are sent through a strip line isolated from the oscillators, in a geometry minimizing by design the capacitive coupling between oscillator and strip line ( $I_{\text{ext}} = 7.5 \times 10^{-3} I_{\text{stripline}}$ ). As a result, the measured rectified voltages are small (approximately 0.5 mV). In the future, these values can be increased up to several tens of millivolts by optimizing the coupling between oscillator and strip line. Indeed, as demonstrated experimentally, rectification effects due to oscillator phase locking can be large, with sensitivities reaching 75.4 mV for the generated d.c. voltage per microwatt of injected microwave power<sup>30</sup>.

We now present how synchronization detection through the resulting rectified voltages may be implemented in a final integrated circuit, using a differential method. We propose to use four reference resistors with the same resistance as the mean resistance of the nano-oscillators and polarized in the same manner. Comparing the voltage across a nano-oscillator and the corresponding reference resistance then allows detection of whether the oscillator is experiencing synchronization (Extended Data Fig. 3c). We designed a simple two-stage CMOS circuit to perform this comparison (Extended Data Fig. 3d,e). The first stage is composed of two differential amplifiers (voltage to current) in parallel. It is followed by a gain stage (current to voltage amplifier). The mismatch between the two amplifiers, a standard design technique, allows high gain. The output of the circuit is therefore a binary voltage, high if the oscillator is synchronized to the input signal, low otherwise. This voltage can be used directly by standard CMOS digital circuit to obtain the class of the input. In the circuit, bias voltages ( $V_{\text{bias1}}$  and  $V_{\text{bias2}}$ ) can be adjusted to vary the speed and power consumption of the circuit.

We simulated this circuit in transient operation using the Cadence Spectre SPICE simulator, a standard tool in commercial integrated circuit design, with the design kit of a 28-nanometre commercial CMOS technology, and optimized the bias voltages for minimal energy consumption, while retaining a response time of the circuit below 600 ns. Extended Data Fig. 3f shows the energy consumed by the detection circuit as a function of the rectified direct voltage due to synchronization, taking into account the whole transient of the detection. This energy can be low: it is below 200 fJ for rectified direct voltages above 50 mV, which can be achieved in structures optimized for spin diode effect<sup>30</sup>. For a full system, this detection must be performed twice (we send two input signals), for the four oscillators, leading to a detection energy of  $2 \times 4 \times 200 \text{ fJ} = 1.6 \text{ pJ}$ .

Using our current oscillators, this energy would be smaller than the energy dissipated by the oscillators and the reference resistors. By contrast, with scaled nano-oscillators (see Extended Data Table 2), this 1.6 pJ detection energy would become dominant.

It is interesting to compare this quantity with the energy consumption of a purely CMOS neural network, implementing the multilayer perceptron of Fig. 4a. Optimized CMOS neural networks compute in reduced precision, usually 8-bit integers, which allows low energy consumption<sup>33</sup>. Taking into account the arithmetic operations (sum and multiplications), in the same commercial 28-nanometre technology as the detection circuit that we implemented, we calculated that an 8-bit integer neural network implementing the second layer of the neural network of Fig. 4a consumes 2.2 pJ. We only took into account the second layer of the neural network, as it is the part implemented by the nano-oscillators. To obtain the energy estimation, we synthesized a Verilog description of a multiply and accumulate block and computed its energy consumption with the Cadence encounter tools using appropriate value change dump files generated by the Cadence ncsim simulator.

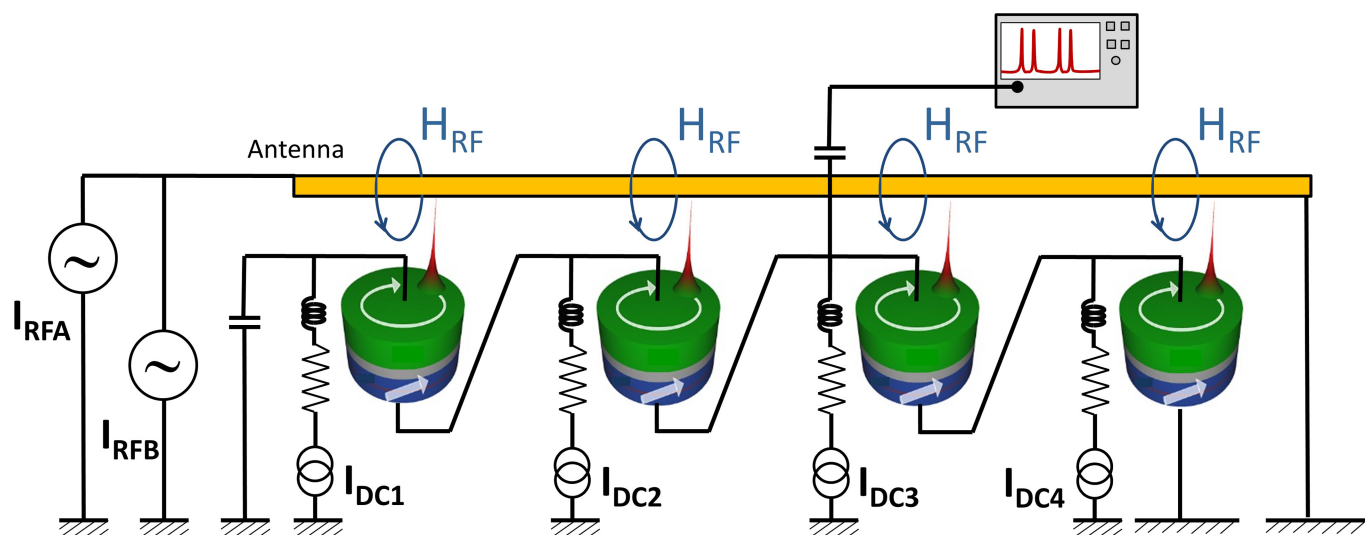


These energy considerations show that on our tiny control system, a nano-oscillator-based solution would provide an energy consumption slightly smaller than an optimized CMOS-based solution. We expect that the full benefit of the oscillator system will appear in deep networks composed of many layers of spin-torque nano-oscillators. Indeed, cascading the synchronization states from one layer to the next can be achieved directly through oscillatory interlayer coupling and does not require synchronization detection. Only at the last layer will detection circuits be required to communicate their state to other circuits. Therefore, we expect that in a deep network of oscillators, the energy consumption will be largely dominated by the oscillator energy consumption, which can be low for a scaled-down oscillator, as can be seen from Extended Data Table 2.

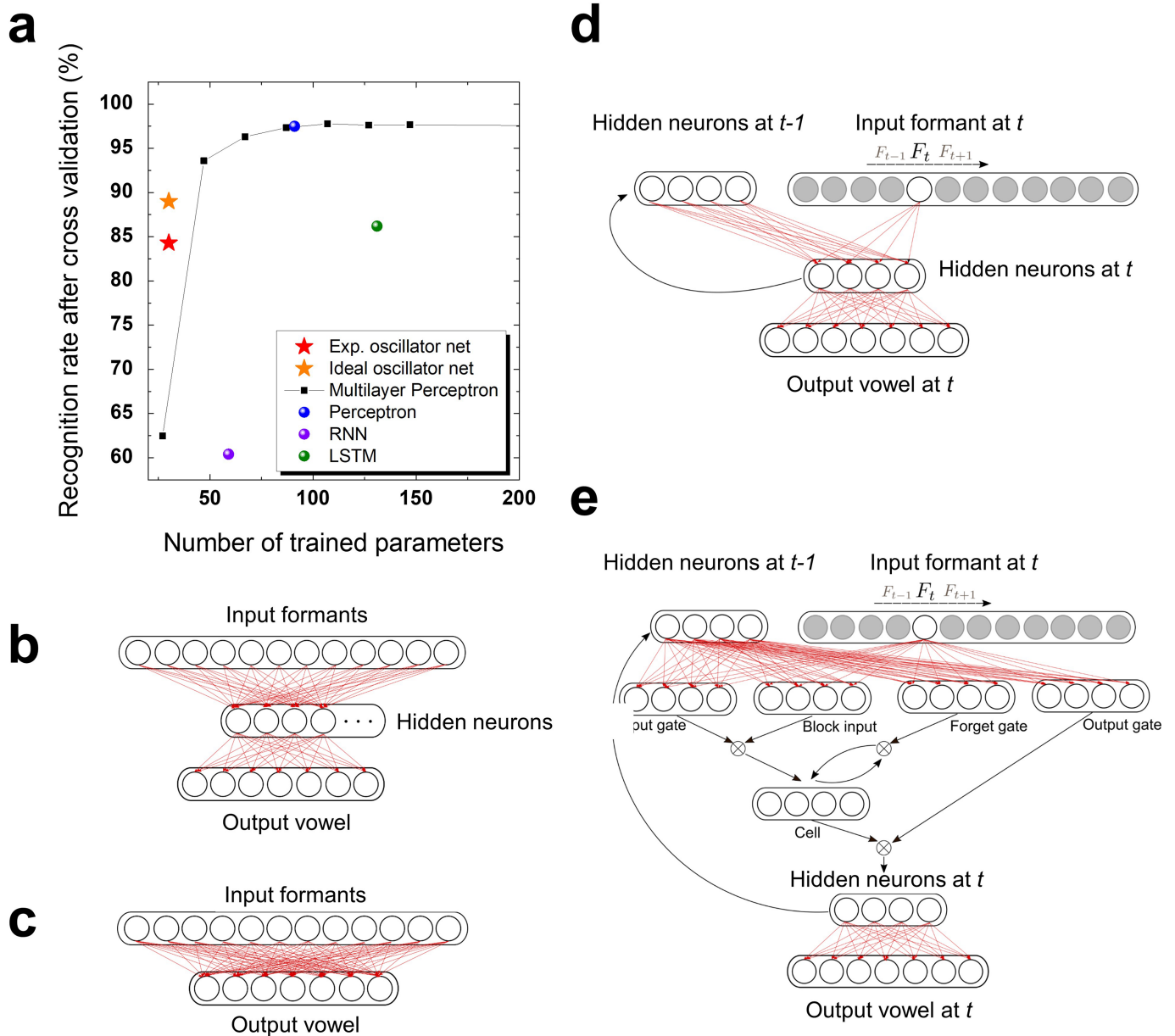
### Data availability

The datasets generated and analysed during this study are available from the corresponding authors on reasonable request.

31. Tulapurkar, A. A. et al. Spin-torque diode effect in magnetic tunnel junctions. *Nature* **438**, 339–342 (2005).
32. Louis, S. et al. Low power microwave signal detection with a spin-torque nano-oscillator in the active self-oscillating regime. *IEEE Trans. Magn.* **53**, 1–4 (2017).
33. Jouppi, N. P. et al. Datacenter performance analysis of a tensor processing unit. In *Proc. 44th Annual International Symposium on Computer Architecture* 1–12 (ACM, 2017).
34. Livi, P. & Indiveri, G. A current-mode conductance-based silicon neuron for address-event neuromorphic systems. In *2009 IEEE International Symposium on Circuits and Systems* 2898–2901 (IEEE, 2009).
35. Qiao, N. & Indiveri, G. Scaling mixed-signal neuromorphic processors to 28 nm FD-SOI technologies. In *2016 IEEE Biomedical Circuits and Systems Conference (BioCAS)* 552–555 (IEEE, 2016).
36. Wijekoon, J. H. B. & Dudek, P. Compact silicon neuron circuit with spiking and bursting behaviour. *Neural Netw.* **21**, 524–534 (2008).
37. Tran, D. X. & Dang, T. T. An ultra-low power consumption and very compact 1.49 GHz CMOS voltage controlled ring oscillator. In *2014 International Conference on Advanced Technologies for Communications (ATC 2014)* 239–244 (IEEE, 2014).
38. Tomita, Y. et al. An 8-to-16 GHz 28 nm CMOS clock distribution circuit based on mutual-injection-locked ring oscillators. In *2013 Symposium on VLSI Circuits* C238–C239 (IEEE, 2013).
39. Gajek, M. et al. Spin torque switching of 20 nm magnetic tunnel junctions with perpendicular anisotropy. *Appl. Phys. Lett.* **100**, 132408 (2012).



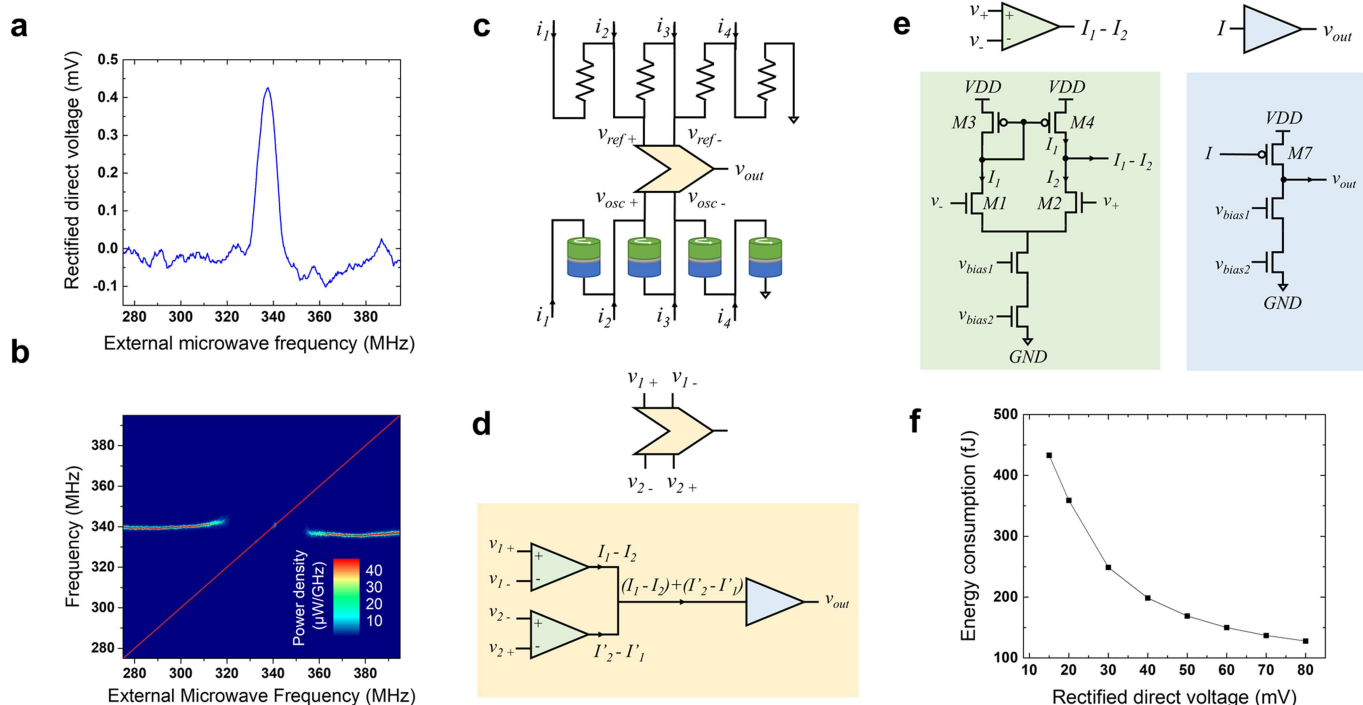
**Extended Data Fig. 1 | Schematic of the experimental set-up.** The four coupled vortex nano-oscillators are shown.  $I_{RFA}$  and  $I_{RFB}$  are the microwave currents injected in the strip line by the two microwave sources.  $H_{RF}$  is the resulting microwave field.  $I_{DC1-4}$  are the applied direct currents.



**Extended Data Fig. 2 | Recognition rates obtained by different neural networks on the formant database. a**, Recognition rates of different neural networks on the formant database as a function of the number of

trained parameters. **b–e**, Schematics of the simulated neural networks: **b**, multi-layer perceptron; **c**, perceptron; **d**, RNN; and **e**, LSTM.





**Extended Data Fig. 3 | Synchronization detection by the spin diode effect.** **a**, Rectified direct voltage measured between oscillator electrodes when the external microwave signal is injected in the strip line above the oscillator and its frequency is swept. Here, the direct current through the oscillator is 5 mA, the magnetic field is 585 mT and the injected microwave power is +1 dBm. **b**, Oscillator spectrum emission measured during the same frequency sweep as **a**. **c**, Proposed differential measurement

configuration for CMOS-based detection of synchronization-induced rectified voltages. **d**, Two-stage CMOS circuit. **e**, The first stage, composed of two differential amplifiers (green), is followed by a gain stage (blue). VDD, supply voltage; GND, ground. **f**, Energy consumption of the CMOS circuit for one synchronization detection event, as a function of the amplitude of the generated rectified direct voltages.

Extended Data Table 1 | Learning rule

7 different spoken vowel classes	Synchronization pattern	Associated frequency difference vector
<b>ae</b>	(1A, 3B)	$\mathbf{d}_{ae} = \begin{pmatrix} f_A^i - f_1 \\ 0 \\ f_B^i - f_3 \\ 0 \end{pmatrix}$
<b>ah</b>	(3A, 1B)	$\mathbf{d}_{ah} = \begin{pmatrix} f_B^i - f_1 \\ 0 \\ f_A^i - f_3 \\ 0 \end{pmatrix}$
<b>aw</b>	(2A, 1B)	$\mathbf{d}_{aw} = \begin{pmatrix} f_B^i - f_1 \\ f_A^i - f_2 \\ 0 \\ 0 \end{pmatrix}$
<b>er</b>	(1A, 2B)	$\mathbf{d}_{er} = \begin{pmatrix} f_A^i - f_1 \\ f_B^i - f_2 \\ 0 \\ 0 \end{pmatrix}$
<b>ih</b>	(3B)	$\mathbf{d}_{ih} = \begin{pmatrix} 0 \\ 0 \\ f_B^i - f_3 \\ 0 \end{pmatrix}$
<b>iy</b>	(4B)	$\mathbf{d}_{iy} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ f_B^i - f_4 \end{pmatrix}$
<b>uw</b>	(1B)	$\mathbf{d}_{uw} = \begin{pmatrix} f_B^i - f_1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

Column 1, spoken vowel class; column 2, synchronization pattern assigned to each vowel; column 3, frequency difference vector between the spoken vowels and their associated patterns. The index  $i$  refers to the  $i$ th data point of a vowel class ( $i$ th speaker).

Extended Data Table 2 | Comparison of CMOS and spin-torque nano-oscillators for neuromorphic computing

	Lateral dimensions	Energy / oscillation	Frequency	Power consumption	Ability to synchronize	References
<b>CMOS neuron</b>	> 30 $\mu\text{m}$	265 pJ	10 Hz	2.65 nW	Yes	(34)
<b>Scaled CMOS neuron</b>	$\approx 7 \mu\text{m}$	50 pJ	30 Hz	1.5 nW	Yes	(35)
<b>Accelerated CMOS neuron</b>	$\approx 10 \mu\text{m}$	8.5 pJ	1 MHz	8.5 $\mu\text{W}$	Yes	(36)
<b>CMOS ring oscillator</b>	6 $\mu\text{m}$	6 fJ	200 KHz	1.2 nW	Unknown	(37)
<b>CMOS ring oscillator</b>	6 $\mu\text{m}$	33 fJ	1.5 GHz	50 $\mu\text{W}$	Unknown	(37)
<b>CMOS ring oscillator</b>	$\approx 300 \mu\text{m}$	1.4 pJ	16 GHz	23 mW	Yes	(38)
<b>Vortex spin-torque oscillator</b>	300 nm	3 pJ	300 MHz	1 mW	Yes	(24)
<b>10 nm spin-torque oscillator (projection)</b>	10 nm	100 aJ	10 GHz	1 $\mu\text{W}$	Yes	(39)

Data from refs <sup>24,34–39</sup>.