Bi-weekly Random Bits from the Internet

2015-06-06

(CANNOT FIND A GOOD STAND-UP COMMEDY SHOW)

California Dreaming

P2, Chiara Barzini, Vogue, June 2015 Issue

Deep Learning

P6, Yann LeCun, Yoshua Bengio and Geoffrey Hinton, Nature 521, May 28, 2015 Issue

How to Really Motivate Salespeople

P21, Doug J. Chung, Harvard Business Review, April 2015 Issue

Richard Pevear & Larissa Volokhonsky, The Art of Translation No. 4

P30, Susannah Hunnewell, The Paris Review 213

California Dreaming

Chiara Barzini, Vogue, June 2015 Issue

We were shooting a commercial for an Italian version of Spam when my father told us that our family was moving to Hollywood. “Don’t you want to go where it’s always summer?” he asked my brother and me between takes. We smiled for the camera as we ate salad topped with fluorescent red meat, duly molded and sprayed by a food stylist to deliver what the director called “a glorious, meaty glow.”

FROM LEFT: BRANDI QUINONES, NIKI TAYLOR, KRISTEN KLOSTERMAN, KATE MOSS, AND SHALOM HARLOW, PHOTOGRAPHED BY ELLEN VON UNWERTH FOR *VOGUE*, 1993.

It was the spring of 1994. Berlusconi was entering politics, and my father, a director when he wasn’t appearing in his friends’ commercials, predicted that Italy was about to go down the drain. The best solution, he announced, was to leave, and conquer Los Angeles I was fifteen, enrolled in one of Rome’s most prestigious classical-studies high schools. My twelve-year-old brother had just discovered the joys of taking public transportation unaccompanied by an adult. Neither Latin declensions nor subway lines would be useful where we were going.

I didn’t like jogging, I hated the idea of having to drive everywhere, and I was unsettled by the existence of an artificial place like Disneyland. I spent the next month sulking. And yet, I was a little curious about America. My grandfather had lived in New York City for many years and made his big break as a writer with a caustic book about our national character called The Italians. His last girlfriend, 21 Washington, D.C.—born countess called Viviana “Vivi” Crespi, who spoke with a Southern drawl and was intimate with the Kennedy family, used to appear at our Roman

doorstep brimming with gifts from the legendary American toy store FAD. Schwarz. My aunt Benedetta, a stunning woman who had lived in Manhattan before settling in Milan, told us stories about posing for Richard Avedon and I1ving Penn in the sixties, dating Salvador Dali, and attending Truman Capote's Black and White Ball. My thirteen-year-old cousin Anna, who lived in Washington, DC, with her journalist parents, sent me mix tapes introducing me to pivotal American music, including an essential Nirvana baptism in 1991.

I prepared by spending my afternoons watching films about L.A. (Chinatown, The Graduate, Sunset Boulevard, and the seminal surf movies Point Break and Big Wednesday - and became intrigued by the city's beach life. I took out the American Vognes I'd collected through the years, sat on my wintlowsill overlooking Saint Peter's Basilica, and sought out clues about my imminent future. I stopped at a photo of Kate Moss and other models lying in a row on the sand. and gave my first sigh of relief. So this was L.A.: a land when: girls were heart-shaped sunglasses, bikinis, and lip gloss. They seemed happy and bronzed, not burned like the sun-crazed Italian girls of the nineties They puckered their lipkall of them, except for Kate Moss, the most beautiful and mysterious of the bunch. This was what reassured me. I didn't have to smile in California. There were options.

Two months later I stood with my brother next to a SHARK ALERT sign on Malibu's El Matador Beach, staring at ferocious, tall waves creaming on the sand. I closed my eyes and felt the cold wind scratch against my naked thighs: Nature in California was hostile and unforgiving. This didn't stop my mother from getting naked and sprawling next to my father on the sand. My parents were nudists. All it took was a ray of sunshine and a body of water, and their clothes came undone in an instant.

A helicopter roared in the sky. It was a cop screaming at my parents from a megaphone to put their bathing suits back on. When they got a ticket for indecent exposure, I saw their folly clearly: They were treating Los Angeles as if it were a beach town on the Mediterranean. They had no clue where they were, and, worse still, our father had miscalculated the family budget. Instead of living in Beverly Hills, as my brother and I had fantasized, we'd landed in the scorching San Fernando Valley basin, in Van Nuys, where gunshots, police sirens, and the roar of the 405 Freeway became our new sound track. After picking us up from school, our parents drove us around the Valley in a 1973 convertible Cadillac. My mother played the part of the Hollywood diva in her cat-eye sunglasses while my father told tales of the early 1900s, when Van Nuys and the southeastern part of the San Fernando Valley were prime neighborhoods for movie shoots "Stars lived here!" he insisted. "Did you know Marilyn Monroe went to Van Nuys High in 1941'?"

“Exactly—1941,”I shot back. The LA. riots’ aftershocks were still fresh, and the city was rife with racial tension. The epicenter of the Northn’dge earthquake, which had occurred on my fifteenth birthday (another terrible piesage), was just a few miles away from our new house. The forces were speaking clearly to us: Go back to your country.

My parents, however, loved the anonymity their new life alforded them. Nobody judged them on the basis of their last name or the reviews of my father’s latest film; they could go grocery shopping in their pajamas. But I was infuriated. How was I supposed to be enriched by a labyrinthine mass of freeway overpasses? My few friends all had cars, but when we went out, they proposed afternoon visits to the mall. Back in Rome, being underage meant very little if you had a Vespa. I felt trapped.

American school was nothing like what I was used to. I had moved from a high school of 300 students to one of 3,500, where I was ordered not to wear red or blue so as to not be affiliated with the Bloods or the Crips. Even though I spoke English, I was automatically placed in an ESL class. My teacher kept thinking I came from Rome, Georgia. I complained to my parents every chance I got, and befriended a group of Salvadoran kids who went to raves, the only parties a teenager could get into without
I.D. On Sundays, I would walk aimlessly up and down the grid of the San Fernando Valley, collecting objects at yard sales. Something about claiming other people’s histories made me feel like I was validating my present. It was a way to accept a city that seemed to have, and to want, no memory of its past.

For all my homesickness, when I returned to Italy the following summer, I soon felt as out of place as I had in LA. My old schoolmates seemed infantile with their faded political debates that hadn‘t evolved since the protests of 1968. My best friends listened to Inti-Illimani and peaceful reggae music, but I was secretly falling in love with Trent Reznor’s dark and gutsy voice, looping the uncensored version of “Closer” on my yellow Sony Sports Walkman while fantasizing about impossible loves I had transformed into a strange hybrid: an Italian Valley Girl.

After two years of “taking meetings” in LA, my father was finally gaining entry to the world that had drawn us away from home. It helped that my mother, who was horrified by the juice culture that was beginning to emerge in California, was an amazing cook. Valeria Golino, Greta Scacchi, Carla Gugino, Danny Huston, and Martin Donovan (a co-screenwriter of Death Becomes Her) all came down from “the hills,” lured by my mother’s arrosto alla genovese.

In 1998, my father finally made his debut behind the camera in LA: an Italian-American TV series to be shot at the Alexandria. The downtown hotel had been one of the most luxurious venues in the city in the early 1920s, with gold-leaf detailing and a ballroom with stained-glass ceilings Rudolph Valentino once had a suite there, but by this point it was a rundown, semisquatted mass of apartment buildings My brother and I ended up writing and starring in one of the episodes, with my nerdy high school friends volunteering as crew members. My mother produced and worked in the catering department.

Right when things seemed to be picking up, my brother fell in with the wrong crowd. He is younger than I, and was more vulnerable. Some of his friends were involved in gangs and drugs, and after a few graffiti-related experiences with the LAPD, it became clear that Los Angeles was not the right city for him at that moment. So in 1999, five years after moving out west, my parents decided to head back to the comforts of the Eternal City. “Aren’t you sick of living where it’s always summer.” my father asked now. By this point, I wasn’t.

When I said goodbye to my family, I felt as though I was betraying them, but I knew I had to stay. I went on to live in Santa Cruz, the Bay Area, and eventually Brooklyn. When I finally made my way back to Rome ten years later, I found myself longing for wildness and violent winds. I missed the snow in New York and the Topanga Canyon sunset. I forced my boyfriend, Luca, to move to a remodeled barnyard close to the sea: an imaginary Malibu. We had to drive everywhere I kept writing in English. I even started jogging. I have my own family now and a rich life in Rome, but my heart still aches for a life that keeps living itself without me.

Deep Learning

Yann LeCun, Yoshua Bengio and Geoffrey Hinton, Nature 521, May 28, 2015 Issue

Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. These methods have dramatically improved the state-of-the-art in speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics. Deep learning discovers intricate structure in large data sets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer. Deep convolutional nets have brought about breakthroughs in processing images, video, speech and audio, whereas recurrent nets have shone light on sequential data such as text and speech.

Machine-learning technology powers many aspects of modern society: from web searches to content filtering on social networks to recommendations on e-commerce websites, and it is increasingly present in consumer products such as cameras and smartphones. Machine-learning systems are used to identify objects in images, transcribe speech into text, match news items, posts or products with users' interests, and select relevant results of search. Increasingly, these applications make use of a class of techniques called deep learning.

Conventional machine-learning techniques were limited in their ability to process natural data in their raw form. For decades, constructing a pattern-recognition or machine-learning system required careful engineering and considerable domain expertise to design a feature extractor that transformed the raw data (such as the pixel values of an image) into a suitable internal representation or feature vector from which the learning subsystem, often a classifier, could detect or classify patterns in the input.

Representation learning is a set of methods that allows a machine to be fed with raw data and to automatically discover the representations needed for detection or classification. Deep-learning methods are representation-learning methods with multiple levels of representation, obtained by composing simple but non-linear modules that each transform the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level. With the composition of enough such transformations, very complex functions can be learned. For classification tasks, higher layers of representation amplify aspects of the input that are important for discrimination and suppress irrelevant variations. An image,

for example, comes in the form of an array of pixel values, and the learned features in the first layer of representation typically represent the presence or absence of edges at particular orientations and locations in the image. The second layer typically detects motifs by spotting particular arrangements of edges, regardless of small variations in the edge positions. The third layer may assemble motifs into larger combinations that correspond to parts of familiar objects, and subsequent layers would detect objects as combinations of these parts. The key aspect of deep learning is that these layers of features are not designed by human engineers: they are learned from data using a general-purpose learning procedure.

Deep learning is making major advances in solving problems that have resisted the best attempts of the artificial intelligence community for many years. It has turned out to be very good at discovering intricate structures in high-dimensional data and is therefore applicable to many domains of science, business and government. In addition to beating records in image recognition and speech recognition, it has beaten other machine-learning techniques at predicting the activity of potential drug molecules, analysing particle accelerator data, reconstructing brain circuits, and predicting the effects of mutations in non-coding DNA on gene expression and disease. Perhaps more surprisingly, deep learning has produced extremely promising results for various tasks in natural language understanding, particularly topic classification, sentiment analysis, question answering and language translation.

Supervised learning

The most common form of machine learning, deep or not, is supervised learning. Imagine that we want to build a system that can classify images as containing, say, a house, a car, a person or a pet. We first collect a large data set of images of houses, cars, people and pets, each labelled with its category. During training, the machine is shown an image and produces an output in the form of a vector of scores, one for each category. We want the desired category to have the highest score of all categories, but this is unlikely to happen before training. We compute an objective function that measures the error (or distance) between the output scores and the desired pattern of scores. The machine then modifies its internal adjustable parameters to reduce this error. These adjustable parameters, often called weights, are real numbers that can be seen as ‘knobs’ that define the input–output function of the machine. In a typical deep-learning system, there may be hundreds of millions of these adjustable weights, and hundreds of millions of labelled examples with which to train the machine.

To properly adjust the weight vector, the learning algorithm computes a gradient vector that, for each weight, indicates by what amount the error would increase or

decrease if the weight were increased by a tiny amount. The weight vector is then adjusted in the opposite direction to the gradient vector.

The objective function, averaged over all the training examples, can be seen as a kind of hilly landscape in the high-dimensional space of weight values. The negative gradient vector indicates the direction of steepest descent in this landscape, taking it closer to a minimum, where the output error is low on average.

In practice, most practitioners use a procedure called stochastic gradient descent (SGD). This consists of showing the input vector for a few examples, computing the outputs and the errors, computing the average gradient for those examples, and adjusting the weights accordingly. The process is repeated for many small sets of examples from the training set until the average of the objective function stops decreasing. It is called stochastic because each small set of examples gives a noisy estimate of the average gradient over all examples. This simple procedure usually finds a good set of weights surprisingly quickly when compared with far more elaborate optimization techniques. After training, the performance of the system is measured on a different set of examples called a test set. This serves to test the generalization ability of the machine — its ability to produce sensible answers on new inputs that it has never seen during training.

Many of the current practical applications of machine learning use linear classifiers on top of hand-engineered features. A two-class linear classifier computes a weighted sum of the feature vector components. If the weighted sum is above a threshold, the input is classified as belonging to a particular category.

Since the 1960s we have known that linear classifiers can only carve their input space into very simple regions, namely half-spaces separated by a hyperplane19. But problems such as image and speech recognition require the input–output function to be insensitive to irrelevant variations of the input, such as variations in position, orientation or illumination of an object, or variations in the pitch or accent of speech, while being very sensitive to particular minute variations (for example, the difference between a white wolf and a breed of wolf-like white dog called a Samoyed). At the pixel level, images of two Samoyeds in different poses and in different environments may be very different from each other, whereas two images of a Samoyed and a wolf in the same position and on similar backgrounds may be very similar to each other. A linear classifier, or any other ‘shallow’ classifier operating on raw pixels could not possibly distinguish the latter two, while putting the former two in the same category. This is why shallow classifiers require a good feature extractor that solves the selectivity–invariance dilemma — one that produces representations that are selective to the aspects of the image that are important for discrimination,

but that are invariant to irrelevant aspects such as the pose of the animal. To make classifiers more powerful, one can use generic non-linear features, as with kernel methods, but generic features such as those arising with the Gaussian kernel do not allow the learner to generalize well far from the training examples21. The conventional option is to hand design good feature extractors, which requires a considerable amount of engineering skill and domain expertise. But this can all be avoided if good features can be learned automatically using a general-purpose learning procedure. This is the key advantage of deep learning.

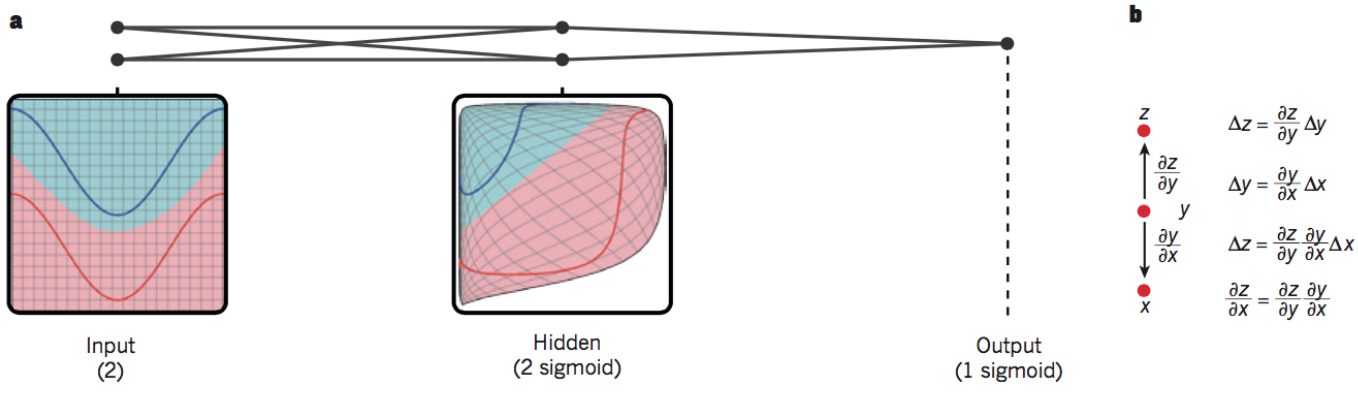
A deep-learning architecture is a multilayer stack of simple modules, all (or most) of which are subject to learning, and many of which compute non-linear input–output mappings. Each module in the stack transforms its input to increase both the selectivity and the invariance of the representation. With multiple non-linear layers, say a depth of 5 to 20, a system can implement extremely intricate functions of its inputs that are simultaneously sensitive to minute details — distinguishing Samoyeds from white wolves — and insensitive to large irrelevant variations such as the background, pose, lighting and surrounding objects.

Backpropagation to train multilayer architectures

From the earliest days of pattern recognition, the aim of researchers has been to replace hand-engineered features with trainable multilayer networks, but despite its simplicity, the solution was not widely understood until the mid 1980s. As it turns out, multilayer architectures can be trained by simple stochastic gradient descent. As long as the modules are relatively smooth functions of their inputs and of their internal weights, one can compute gradients using the backpropagation procedure. The idea that this could be done, and that it worked, was discovered independently by several different groups during the 1970s and 1980s.

The backpropagation procedure to compute the gradient of an objective function with respect to the weights of a multilayer stack of modules is nothing more than a practical application of the chain rule for derivatives. The key insight is that the derivative (or gradient) of the objective with respect to the input of a module can be computed by working backwards from the gradient with respect to the output of that module (or the input of the subsequent module) (Fig. 1). The backpropagation equation can be applied repeatedly to propagate gradients through all modules, starting from the output at the top (where the network produces its prediction) all the way to the bottom (where the external input is fed). Once these gradients have been computed, it is straightforward to compute the gradients with respect to the weights of each module.

Many applications of deep learning use feedforward neural network architectures (Fig. 1), which learn to map a fixed-size input (for example, an image) to a fixed-size output (for example, a probability for each of several categories). To go from one layer to the next, a set of units compute a weighted sum of their inputs from the previous layer and pass the result through a non-linear function. At present, the most popular non-linear function is the rectified linear unit (ReLU), which is simply the half-wave rectifier f(z) = max(z, 0). In past decades, neural nets used smoother non-linearities, such as tanh(z) or 1/(1 + exp(−z)), but the ReLU typically learns much faster in networks with many layers, allowing training of a deep supervised network without unsupervised pre-training28. Units that are not in the input or output layer are conventionally called hidden units. The hidden layers can be seen as distorting the input in a non-linear way so that categories become linearly separable by the last layer (Fig. 1).



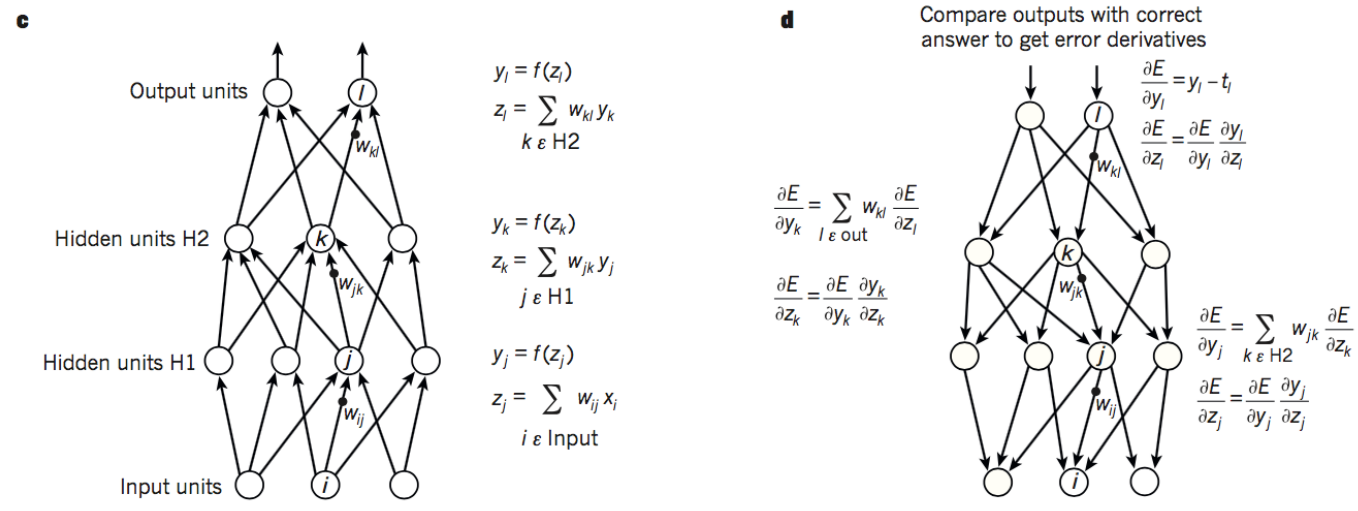


Figure 1 | Multilayer neural networks and backpropagation. **a**, A multilayer neural network (shown by the connected dots) can distort the input space to make the classes of data (examples of which are on the red and blue lines) linearly separable. Note how a regular grid (shown on the left) in input space is also transformed (shown in the middle panel) by hidden units. This is an illustrative example with only two input units, two hidden units and one output unit, but the networks used for object recognition or natural language processing contain tens or hundreds of thousands of units. Reproduced with permission from C. Olah (http://colah.github.io/). **b**, The chain rule of derivatives tells us how two small effects (that of a small change of x on y, and that of y on z) are composed. A small change Δx in x gets transformed first into a small change Δy in y by getting multiplied by $\partial y/\partial x$ (that is, the definition of partial derivative). Similarly, the change Δy creates a change Δz in z. Substituting one equation into the other gives the chain rule of derivatives — how Δx gets turned into Δz through multiplication by the product of $\partial y/\partial x$ and $\partial z/\partial x$. It also works when x, y and z are vectors (and the derivatives are Jacobian matrices). **c**, The equations used for computing the forward pass in a neural net with two hidden layers and one output layer, each constituting a module through which one can backpropagate gradients. At each layer, we first compute the total input z to each unit, which is a weighted sum of the outputs of the units in the layer below. Then a non-linear function $f(.)$ is applied to z to get the output of the unit. For simplicity, we have omitted bias terms. The non-linear functions used in neural networks include the rectified linear unit (ReLU) $f(z) = \max(0,z)$, commonly used in recent years, as well as the more conventional sigmoids, such as the hyberbolic tangent, $f(z) = (\exp(z) - \exp(-z))/(\exp(z) + \exp(-z))$ and logistic function logistic, $f(z) = 1/(1 + \exp(-z))$. **d**, The equations used for computing the backward pass. At each hidden layer we compute the error derivative with respect to the output of each unit, which is a weighted sum of the error derivatives with respect to the total inputs to the units in the layer above. We then convert the error derivative with respect to the output into the error derivative with respect to the input by multiplying it by the gradient of $f(z)$. At the output layer, the error derivative with respect to the output of a unit is computed by differentiating the cost function. This gives $y_l - t_l$ if the cost function for unit l is $0.5(y_l - t_l)^2$, where t_l is the target value. Once the $\partial E/\partial z_k$ is known, the error-derivative for the weight w_{jk} on the connection from unit j in the layer below is just $y_j\,\partial E/\partial z_k$.

In the late 1990s, neural nets and backpropagation were largely forsaken by the machine-learning community and ignored by the computer-vision and speech-recognition communities. It was widely thought that learning useful, multistage, feature extractors with little prior knowledge was infeasible. In particular, it was commonly thought that simple gradient descent would get trapped in poor local minima — weight configurations for which no small change would reduce the average error.

In practice, poor local minima are rarely a problem with large networks. Regardless of the initial conditions, the system nearly always reaches solutions of very similar quality. Recent theoretical and empirical results strongly suggest that local minima are not a serious issue in general. Instead, the landscape is packed with a combinatorially large number of saddle points where the gradient is zero, and the surface curves up in most dimensions and curves down in the remainder. The analysis seems to show that saddle points with only a few downward curving directions are present in very large numbers, but almost all of them have very similar values of the objective function. Hence, it does not much matter which of these saddle points the algorithm gets stuck at.

Interest in deep feedforward networks was revived around 2006 (refs 31–34) by a group of researchers brought together by the Canadian Institute for Advanced Research (CIFAR). The researchers introduced unsupervised learning procedures that could create layers of feature detectors without requiring labelled data. The objective in learning each layer of feature detectors was to be able to reconstruct or model the activities of feature detectors (or raw inputs) in the layer below. By 'pre-training' several layers of progressively more complex feature detectors using this reconstruction objective, the weights of a deep network could be initialized to sensible values. A final layer of output units could then be added to the top of the network and the whole deep system could be fine-tuned using standard backpropagation. This worked remarkably well for recognizing handwritten digits or for detecting pedestrians, especially when the amount of labelled data was very limited.

The first major application of this pre-training approach was in speech recognition, and it was made possible by the advent of fast graphics processing units (GPUs) that were convenient to program and allowed researchers to train networks 10 or 20 times faster. In 2009, the approach was used to map short temporal windows of coefficients extracted from a sound wave to a set of probabilities for the various fragments of speech that might be represented by the frame in the centre of the window. It achieved record-breaking results on a standard speech recognition benchmark that used a small vocabulary and was quickly developed to give record-breaking results on a large vocabulary task. By 2012, versions of the deep net from 2009 were being developed by many of the major speech groups and were

already being deployed in Android phones. For smaller data sets, unsupervised pre-training helps to prevent overfitting, leading to significantly better generalization when the number of labelled examples is small, or in a transfer setting where we have lots of examples for some 'source' tasks but very few for some 'target' tasks. Once deep learning had been rehabilitated, it turned out that the pre-training stage was only needed for small data sets.

There was, however, one particular type of deep, feedforward network that was much easier to train and generalized much better than networks with full connectivity between adjacent layers. This was the convolutional neural network (ConvNet). It achieved many practical successes during the period when neural networks were out of favour and it has recently been widely adopted by the computer-vision community.

Convolutional neural networks

ConvNets are designed to process data that come in the form of multiple arrays, for example a colour image composed of three 2D arrays containing pixel intensities in the three colour channels. Many data modalities are in the form of multiple arrays: 1D for signals and sequences, including language; 2D for images or audio spectrograms; and 3D for video or volumetric images. There are four key ideas behind ConvNets that take advantage of the properties of natural signals: local connections, shared weights, pooling and the use of many layers.

The architecture of a typical ConvNet (Fig. 2) is structured as a series of stages. The first few stages are composed of two types of layers: convolutional layers and pooling layers. Units in a convolutional layer are organized in feature maps, within which each unit is connected to local patches in the feature maps of the previous layer through a set of weights called a filter bank. The result of this local weighted sum is then passed through a non-linearity such as a ReLU. All units in a feature map share the same filter bank. Different feature maps in a layer use different filter banks. The reason for this architecture is twofold. First, in array data such as images, local groups of values are often highly correlated, forming distinctive local motifs that are easily detected. Second, the local statistics of images and other signals are invariant to location. In other words, if a motif can appear in one part of the image, it could appear anywhere, hence the idea of units at different locations sharing the same weights and detecting the same pattern in different parts of the array. Mathematically, the filtering operation performed by a feature map is a discrete convolution, hence the name.

Although the role of the convolutional layer is to detect local conjunctions of fea-

tures from the previous layer, the role of the pooling layer is to merge semantically similar features into one. Because the relative positions of the features forming a motif can vary somewhat, reliably detecting the motif can be done by coarse-graining the position of each feature. A typical pooling unit computes the maximum of a local patch of units in one feature map (or in a few feature maps). Neighbouring pooling units take input from patches that are shifted by more than one row or column, thereby reducing the dimension of the representation and creating an invariance to small shifts and distortions. Two or three stages of convolution, non-linearity and pooling are stacked, followed by more convolutional and fully-connected layers. Backpropagating gradients through a ConvNet is as simple as through a regular deep network, allowing all the weights in all the filter banks to be trained.

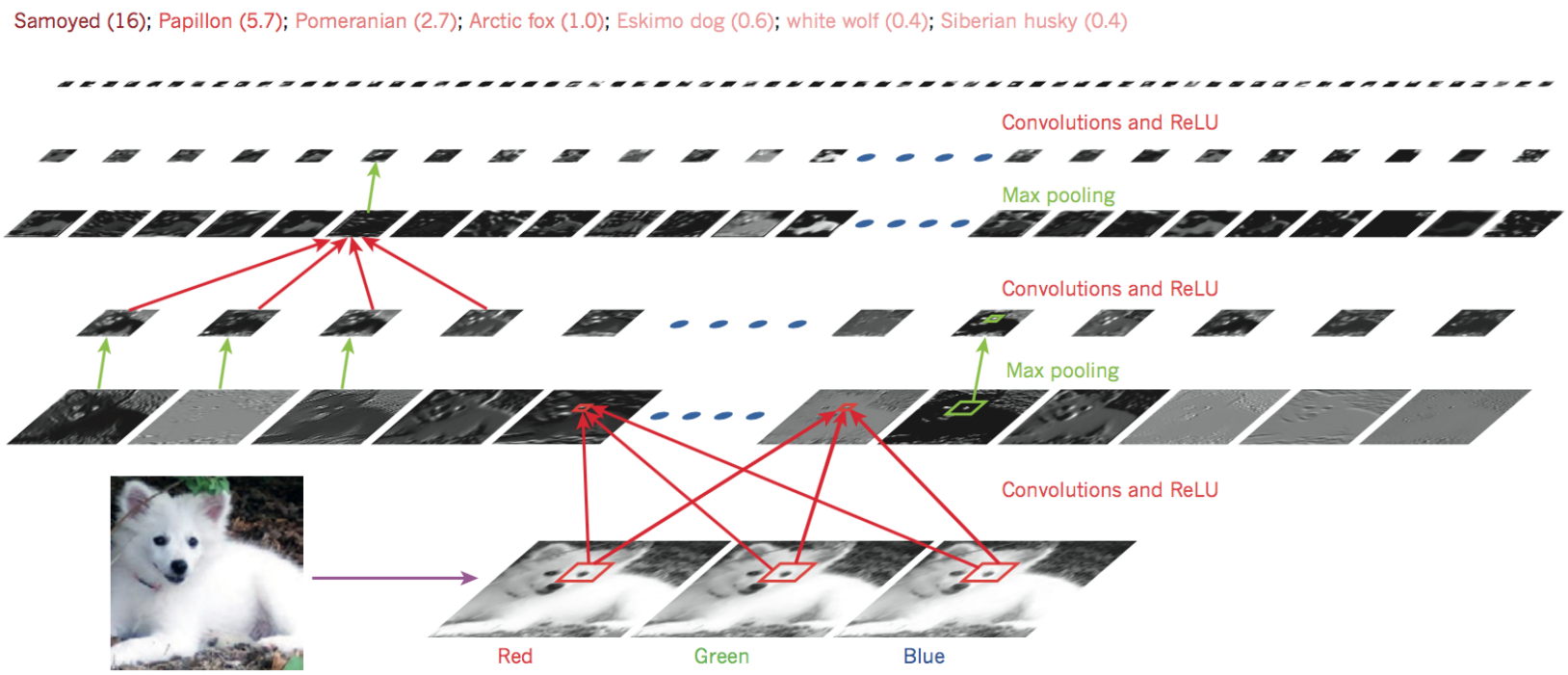


Figure 2 | Inside a convolutional network. The outputs (not the filters) of each layer (horizontally) of a typical convolutional network architecture applied to the image of a Samoyed dog (bottom left; and RGB (red, green, blue) inputs, bottom right). Each rectangular image is a feature map corresponding to the output for one of the learned features, detected at each of the image positions. Information flows bottom up, with lower-level features acting as oriented edge detectors, and a score is computed for each image class in output. ReLU, rectified linear unit.

Deep neural networks exploit the property that many natural signals are compositional hierarchies, in which higher-level features are obtained by composing lower-level ones. In images, local combinations of edges form motifs, motifs assemble into parts, and parts form objects. Similar hierarchies exist in speech and text from sounds to phones, phonemes, syllables, words and sentences. The pooling allows representations to vary very little when elements in the previous layer vary in position and appearance.

The convolutional and pooling layers in ConvNets are directly inspired by the classic notions of simple cells and complex cells in visual neuroscience, and the overall architecture is reminiscent of the LGN–V1–V2–V4–IT hierarchy in the visual cortex ventral pathway. When ConvNet models and monkeys are shown the same picture, the activations of high-level units in the ConvNet explains half of the variance of random sets of 160 neurons in the monkey's inferotemporal cortex. ConvNets have their roots in the neocognitron, the architecture of which was somewhat similar,

but did not have an end-to-end supervised-learning algorithm such as backpropagation. A primitive 1D ConvNet called a time-delay neural net was used for the recognition of phonemes and simple words.

There have been numerous applications of convolutional networks going back to the early 1990s, starting with time-delay neural networks for speech recognition and document reading. The document reading system used a ConvNet trained jointly with a probabilistic model that implemented language constraints. By the late 1990s this system was reading over 10% of all the cheques in the United States. A number of ConvNet-based optical character recognition and handwriting recognition systems were later deployed by Microsoft. ConvNets were also experimented with in the early 1990s for object detection in natural images, including faces and hands, and for face recognition.

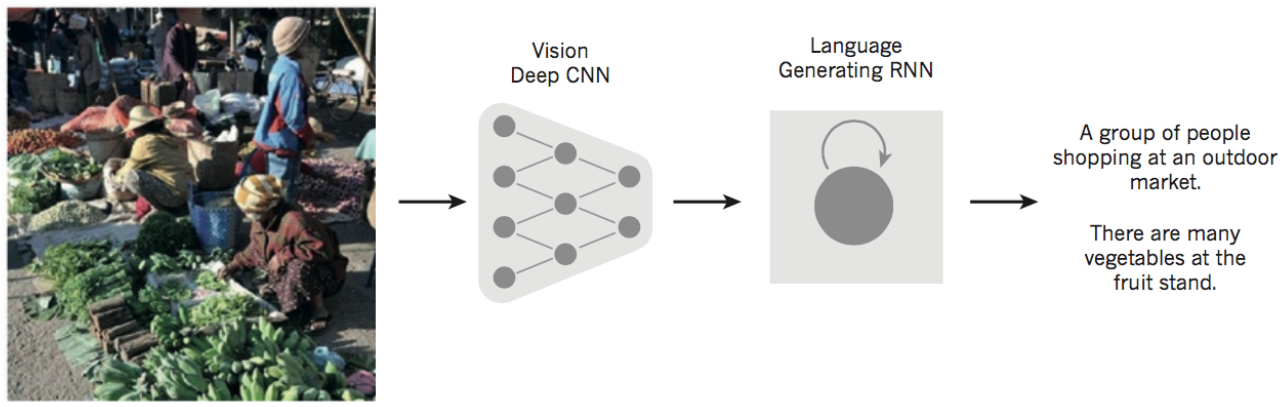
Image understanding with deep convolutional networks

Since the early 2000s, ConvNets have been applied with great success to the detection, segmentation and recognition of objects and regions in images. These were all tasks in which labelled data was relatively abundant, such as traffic sign recognition, the segmentation of biological images particularly for connectomics, and the detection of faces, text, pedestrians and human bodies in natural images. A major recent practical success of ConvNets is face recognition.

Importantly, images can be labelled at the pixel level, which will have applications in technology, including autonomous mobile robots and self-driving cars. Companies such as Mobileye and NVIDIA are using such ConvNet-based methods in their upcoming vision systems for cars. Other applications gaining importance involve natural language understanding and speech recognition.

Despite these successes, ConvNets were largely forsaken by the mainstream computer-vision and machine-learning communities until the ImageNet competition in 2012. When deep convolutional networks were applied to a data set of about a million images from the web that contained 1,000 different classes, they achieved spectacular results, almost halving the error rates of the best competing approaches. This success came from the efficient use of GPUs, ReLUs, a new regularization technique called dropout, and techniques to generate more training examples by deforming the existing ones. This success has brought about a revolution in computer vision; ConvNets are now the dominant approach for almost all recognition and detection tasks and approach human performance on some tasks. A recent stunning demonstration combines ConvNets and recurrent net modules for the generation of image captions (Fig. 3).

Recent ConvNet architectures have 10 to 20 layers of ReLUs, hundreds of millions of weights, and billions of connections between units. Whereas training such large networks could have taken weeks only two years ago, progress in hardware, software and algorithm parallelization have reduced training times to a few hours.



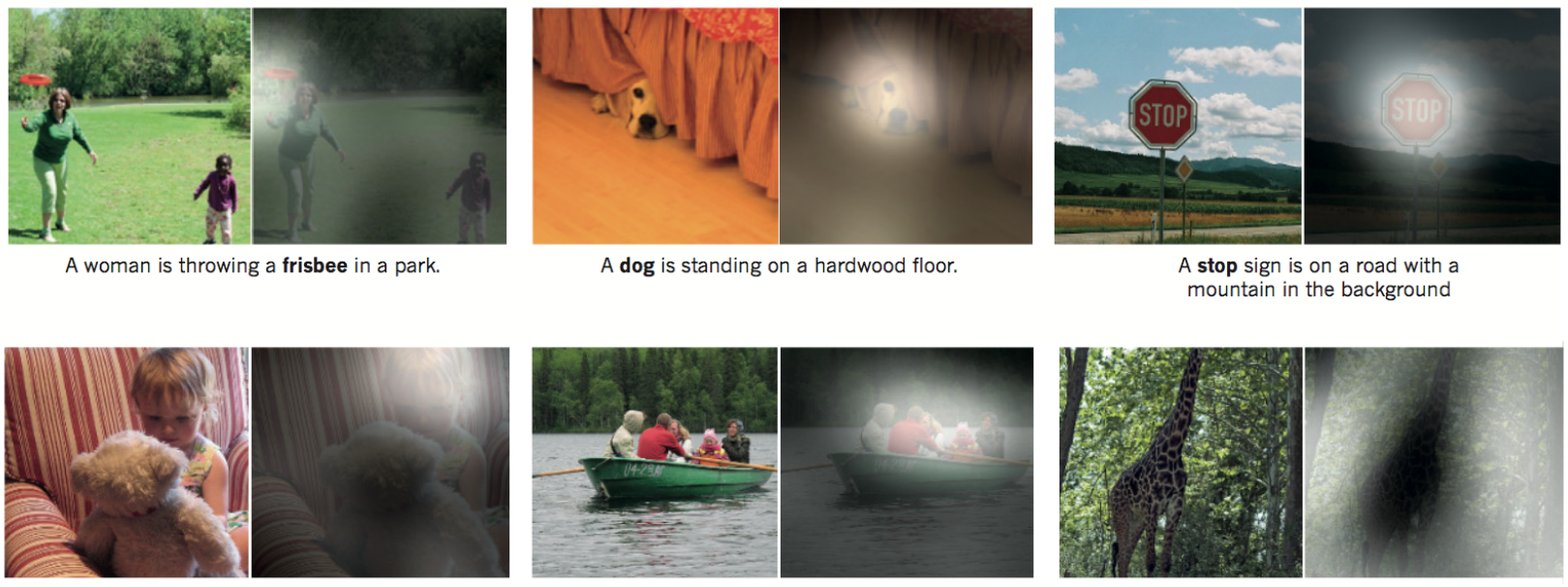


Figure 3 | From image to text. Captions generated by a recurrent neural network (RNN) taking, as extra input, the representation extracted by a deep convolution neural network (CNN) from a test image, with the RNN trained to 'translate' high-level representations of images into captions (top). Reproduced with permission from ref. 102. When the RNN is given the ability to focus its attention on a different location in the input image (middle and bottom; the lighter patches were given more attention) as it generates each word (bold), we found[86] that it exploits this to achieve better 'translation' of images into captions.

The performance of ConvNet-based vision systems has caused most major technology companies, including Google, Facebook, Microsoft, IBM, Yahoo!, Twitter and Adobe, as well as a quickly growing number of start-ups to initiate research and development projects and to deploy ConvNet-based image understanding products and services.

ConvNets are easily amenable to efficient hardware implementations in chips or field-programmable gate arrays. A number of companies such as NVIDIA, Mobileye, Intel, Qualcomm and Samsung are developing ConvNet chips to enable real-time vision applications in smartphones, cameras, robots and self-driving cars.

Distributed representations and language processing

Deep-learning theory shows that deep nets have two different exponential advantages over classic learning algorithms that do not use distributed representations. Both of these advantages arise from the power of composition and depend on the underlying data-generating distribution having an appropriate componential structure. First, learning distributed representations enable generalization to new combinations of the values of learned features beyond those seen during training (for example, 2n combinations are possible with n binary features). Second, composing layers of representation in a deep net brings the potential for another exponential advantage (exponential in the depth).

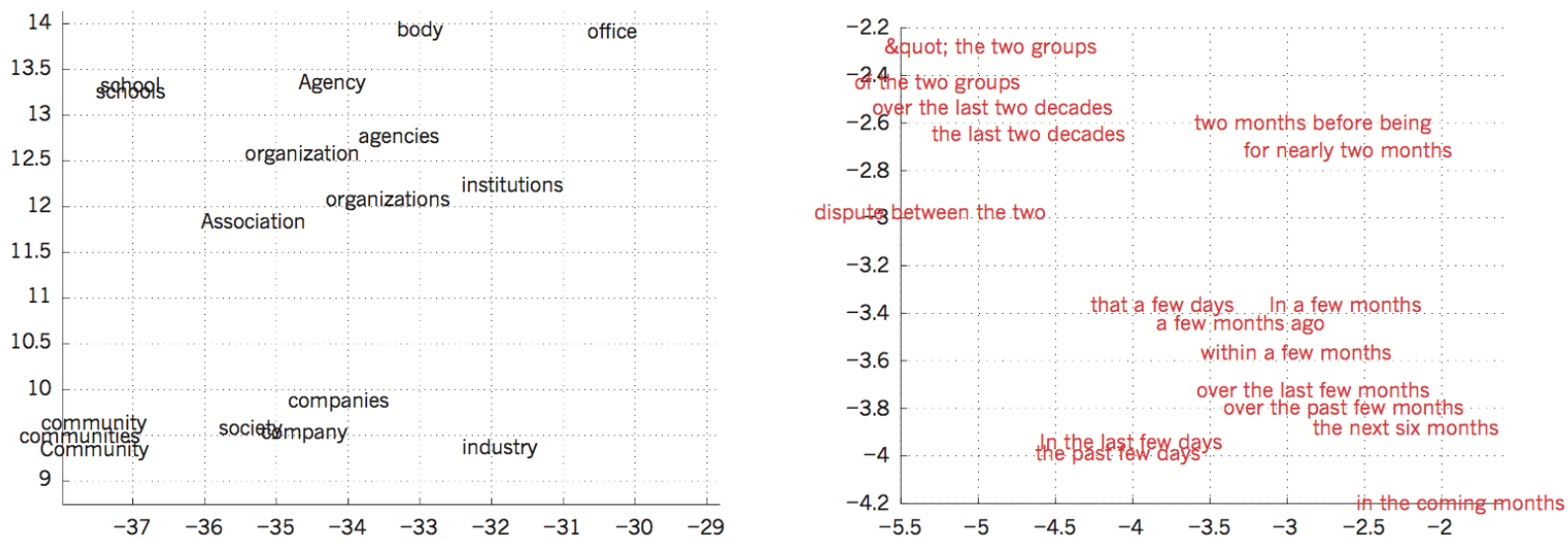


Figure 4 | Visualizing the learned word vectors. On the left is an illustration of word representations learned for modelling language, non-linearly projected to 2D for visualization using the t-SNE algorithm[103]. On the right is a 2D representation of phrases learned by an English-to-French encoder–decoder recurrent neural network[75]. One can observe that semantically similar words or sequences of words are mapped to nearby representations. The distributed representations of words are obtained by using backpropagation to jointly learn a representation for each word and a function that predicts a target quantity such as the next word in a sequence (for language modelling) or a whole sequence of translated words (for machine translation)[18,75].

The hidden layers of a multilayer neural network learn to represent the network's inputs in a way that makes it easy to predict the target outputs. This is nicely demonstrated by training a multilayer neural network to predict the next word in a sequence from a local context of earlier words71. Each word in the context is presented to the network as a one-of-N vector, that is, one component has a value of 1 and the rest are 0. In the first layer, each word creates a different pattern of activations, or word vectors (Fig. 4). In a language model, the other layers of the network learn to convert the input word vectors into an output word vector for the predicted next word, which can be used to predict the probability for any word in the vocabulary to appear as the next word. The network learns word vectors that contain many active components each of which can be interpreted as a separate feature of the word, as was first demonstrated in the context of learning distributed representations for symbols. These semantic features were not explicitly present in the input. They were discovered by the learning procedure as a good way of factorizing the structured relationships between the input and output symbols into multiple 'micro-rules'. Learning word vectors turned out to also work very well when the word sequences come from a large corpus of real text and the individual micro-rules are unreliable. When trained to predict the next word in a news story, for example, the

learned word vectors for Tuesday and Wednesday are very similar, as are the word vectors for Sweden and Norway. Such representations are called distributed representations because their elements (the features) are not mutually exclusive and their many configurations correspond to the variations seen in the observed data. These word vectors are composed of learned features that were not determined ahead of time by experts, but automatically discovered by the neural network. Vector representations of words learned from text are now very widely used in natural language applications.

The issue of representation lies at the heart of the debate between the logic-inspired and the neural-network-inspired paradigms for cognition. In the logic-inspired paradigm, an instance of a symbol is something for which the only property is that it is either identical or non-identical to other symbol instances. It has no internal structure that is relevant to its use; and to reason with symbols, they must be bound to the variables in judiciously chosen rules of inference. By contrast, neural networks just use big activity vectors, big weight matrices and scalar non-linearities to perform the type of fast 'intuitive' inference that underpins effortless commonsense reasoning.

Before the introduction of neural language models, the standard approach to statistical modelling of language did not exploit distributed representations: it was based on counting frequencies of occurrences of short symbol sequences of length up to N (called N-grams). The number of possible N-grams is on the order of V^N, where V is the vocabulary size, so taking into account a context of more than a handful of words would require very large training corpora. N-grams treat each word as an atomic unit, so they cannot generalize across semantically related sequences of words, whereas neural language models can because they associate each word with a vector of real valued features, and semantically related words end up close to each other in that vector space (Fig. 4).

Recurrent neural networks

When backpropagation was first introduced, its most exciting use was for training recurrent neural networks (RNNs). For tasks that involve sequential inputs, such as speech and language, it is often better to use RNNs (Fig. 5). RNNs process an input sequence one element at a time, maintaining in their hidden units a 'state vector' that implicitly contains information about the history of all the past elements of the sequence. When we consider the outputs of the hidden units at different discrete time steps as if they were the outputs of different neurons in a deep multilayer network (Fig. 5, right), it becomes clear how we can apply backpropagation to train RNNs.

RNNs are very powerful dynamic systems, but training them has proved to be problematic because the backpropagated gradients either grow or shrink at each time step, so over many time steps they typically explode or vanish.

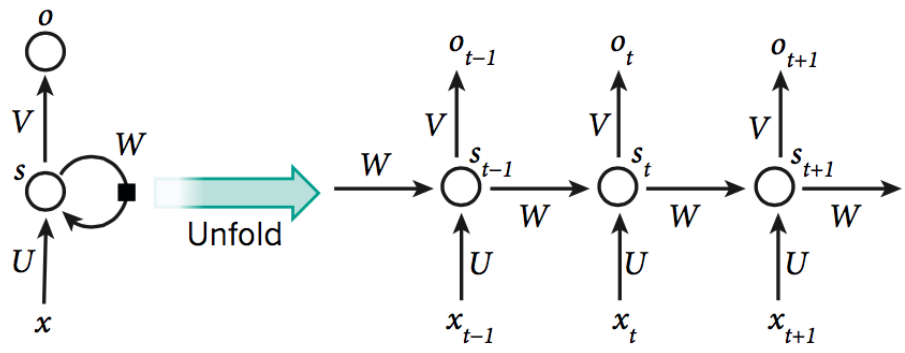


Figure 5 | A recurrent neural network and the unfolding in time of the computation involved in its forward computation. The artificial neurons (for example, hidden units grouped under node s with values s_t at time t) get inputs from other neurons at previous time steps (this is represented with the black square, representing a delay of one time step, on the left). In this way, a recurrent neural network can map an input sequence with elements x_t into an output sequence with elements o_t, with each o_t depending on all the previous $x_{t'}$ (for $t' \leq t$). The same parameters (matrices U, V, W) are used at each time step. Many other architectures are possible, including a variant in which the network can generate a sequence of outputs (for example, words), each of which is used as inputs for the next time step. The backpropagation algorithm (Fig. 1) can be directly applied to the computational graph of the unfolded network on the right, to compute the derivative of a total error (for example, the log-probability of generating the right sequence of outputs) with respect to all the states s_t and all the parameters.

Thanks to advances in their architecture and ways of training them, RNNs have been found to be very good at predicting the next character in the text or the next word in a sequence, but they can also be used for more complex tasks. For example, after reading an English sentence one word at a time, an English 'encoder' network can be trained so that the final state vector of its hidden units is a good representation of the thought expressed by the sentence. This thought vector can then be used as the initial hidden state of (or as extra input to) a jointly trained French 'decoder' network, which outputs a probability distribution for the first word of the French translation. If a particular first word is chosen from this distribution and provided as input to the decoder network it will then output a probability distribution for the second word of the translation and so on until a full stop is chosen. Overall, this process generates sequences of French words according to a probability distribution that depends on the English sentence. This rather naive way of performing machine translation has quickly become competitive with the state-of-the-art, and this raises serious doubts about whether understanding a sentence requires anything like the internal symbolic expressions that are manipulated by using inference rules. It is more compatible with the view that everyday reasoning involves many simultane-

ous analogies that each contribute plausibility to a conclusion.

Instead of translating the meaning of a French sentence into an English sentence, one can learn to 'translate' the meaning of an image into an English sentence (Fig. 3). The encoder here is a deep ConvNet that converts the pixels into an activity vector in its last hidden layer. The decoder is an RNN similar to the ones used for machine translation and neural language modelling. There has been a surge of interest in such systems recently (see examples mentioned in ref. 86).

RNNs, once unfolded in time (Fig. 5), can be seen as very deep feedforward networks in which all the layers share the same weights. Although their main purpose is to learn long-term dependencies, theoretical and empirical evidence shows that it is difficult to learn to store information for very long.

To correct for that, one idea is to augment the network with an explicit memory. The first proposal of this kind is the long short-term memory (LSTM) networks that use special hidden units, the natural behaviour of which is to remember inputs for a long time. A special unit called the memory cell acts like an accumulator or a gated leaky neuron: it has a connection to itself at the next time step that has a weight of one, so it copies its own real-valued state and accumulates the external signal, but this self-connection is multiplicatively gated by another unit that learns to decide when to clear the content of the memory.

LSTM networks have subsequently proved to be more effective than conventional RNNs, especially when they have several layers for each time step, enabling an entire speech recognition system that goes all the way from acoustics to the sequence of characters in the transcription. LSTM networks or related forms of gated units are also currently used for the encoder and decoder networks that perform so well at machine translation.

Over the past year, several authors have made different proposals to augment RNNs with a memory module. Proposals include the Neural Turing Machine in which the network is augmented by a 'tape-like' memory that the RNN can choose to read from or write to, and memory networks, in which a regular network is augmented by a kind of associative memory. Memory networks have yielded excellent performance on standard question-answering benchmarks. The memory is used to remember the story about which the network is later asked to answer questions.

Beyond simple memorization, neural Turing machines and memory networks are being used for tasks that would normally require reasoning and symbol manipulation. Neural Turing machines can be taught 'algorithms'. Among other things, they

can learn to output a sorted list of symbols when their input consists of an unsorted sequence in which each symbol is accompanied by a real value that indicates its priority in the list88. Memory networks can be trained to keep track of the state of the world in a setting similar to a text adventure game and after reading a story, they can answer questions that require complex inference90. In one test example, the network is shown a 15-sentence version of the The Lord of the Rings and correctly answers questions such as “where is Frodo now?”.

The future of deep learning

Unsupervised learning had a catalytic effect in reviving interest in deep learning, but has since been overshadowed by the successes of purely supervised learning. Although we have not focused on it in this Review, we expect unsupervised learning to become far more important in the longer term. Human and animal learning is largely unsupervised: we discover the structure of the world by observing it, not by being told the name of every object.

Human vision is an active process that sequentially samples the optic array in an intelligent, task-specific way using a small, high-resolution fovea with a large, low-resolution surround. We expect much of the future progress in vision to come from systems that are trained end-to-end and combine ConvNets with RNNs that use reinforcement learning to decide where to look. Systems combining deep learning and reinforcement learning are in their infancy, but they already outperform passive vision systems at classification tasks and produce impressive results in learning to play many different video games.

Natural language understanding is another area in which deep learning is poised to make a large impact over the next few years. We expect systems that use RNNs to understand sentences or whole documents will become much better when they learn strategies for selectively attending to one part at a time.

Ultimately, major progress in artificial intelligence will come about through systems that combine representation learning with complex reasoning. Although deep learning and simple reasoning have been used for speech and handwriting recognition for a long time, new paradigms are needed to replace rule-based manipulation of symbolic expressions by operations on large vectors.

How to Really Motivate Salespeople

Doug J. Chung, Harvard Business Review, April 2015 Issue

Before I became a business school professor, I worked as a management consultant. One engagement in particular had a profound influence on my career. The project involved working with the Asia-based sales force of a global consumer products company. This company practiced "route sales," which meant reps spent their days visiting mom-and-pop convenience stores, servicing accounts. One thing about the organization surprised me: Its sales managers spent inordinate time listening to the reps complain about their compensation.

The complaints were based on what the reps saw as a myriad of problems. Their quotas were set too high, so they couldn't possibly reach them. Or their territory was subpar, limiting their ability to sign new accounts. Sometimes the complaints focused on fairness: A rep who was hitting his quotas and making decent money would want a manager to do something about a "lazy" colleague who was earning outsize pay simply because he had a good territory. Imagine any conceivable complaint a salesperson might have about pay, and I guarantee that sales managers at my client's company had heard it.

The reps weren't the only ones obsessed with the compensation system. The company liked to play around with the system's components to try to find better ways to motivate reps and boost revenue, or to increase the return on the money it spent paying salespeople—a large part of its marketing budget. This company's sales comp system was fairly basic: Reps earned a salary and a commission of around 1% of sales. The company worried that the system was too focused on outcomes and might over- or under-reward reps for factors outside their control. So it began basing compensation on their effort and behavior, not just on top-line sales. For instance, under the new system, a portion of compensation was based on customer satisfaction surveys, the number of prospective accounts visited (even if they didn't buy), and the retention of existing accounts.

Largely because of this consulting assignment, I became so curious about the best ways to compensate salespeople that I began reading academic articles on the subject. Eventually I pursued a PhD in marketing at Yale, where I studied the theory and practice of how companies can and should manage and pay salespeople—research I now continue at Harvard Business School.

Although there are fewer academics studying sales force compensation and management than researching trendy marketing subjects, such as the use of social me-

dia or digital advertising, in the past decade it's become a fast-moving field. While some of the basic theories established in the 1970s and 1980s still apply, academics have begun testing those theories using two methods new to this area of research: empirical analysis of companies' sales and pay data, and field experiments in which researchers apply various pay structures to different groups of salespeople and then compare the groups' effort and output.

This new wave of research is already providing evidence that some standard compensation practices probably hurt sales. For instance, the research suggests that caps on commissions, which most large companies use, decrease high-performing reps' motivation and effort. Likewise, the practice of "ratcheting" quotas (raising a salesperson's annual quota if he or she exceeded it the previous year) may hurt long-term results. Research based on field experiments (as opposed to the lab experiments academics have been doing for many years) is also yielding new insight into how the timing and labeling of bonuses can affect salespeople's motivation.

In this article I will take readers through the evolution of this research and suggest the best ways to apply it. With luck, this knowledge not only will help companies think about better ways to compensate salespeople, but also might mean that their managers spend fewer hours listening to them gripe about unfair pay.

The Dangers of Complex Compensation Systems

Researchers studying sales force compensation have long been guided by the principal-agent theory. This theory, drawn from the field of economics, describes the problem that results from conflicting interests between a principal (a company, for instance) and an agent hired by that principal (an employee). For example, a company wants an employee's maximum output, but a salaried employee may be tempted to slack off and may be able to get away with it if the company can't observe how hard the employee is working. Most incentive or variable pay schemes—including stock options for the C-suite—are attempts to align the interests of principals and agents. Commission-based plans for salespeople are just one example.

Salespeople were paid by commission for centuries before economists began writing about the principal-agent problem. Companies chose this system for at least three reasons. First, it's easy to measure the short-term output of a salesperson, unlike that of most workers. Second, field reps have traditionally worked with little (if any) supervision; commission-based pay gives managers some control, making up for their inability to know if a rep is actually visiting clients or playing golf. Third, studies of personality type show that salespeople typically have a larger appetite for risk than other workers, so a pay plan that offers upside potential appeals to them.

During the 1980s several important pieces of research influenced firms' use of commission-based systems. One, by my Harvard colleague Rajiv Lal and several coauthors, explored how the level of uncertainty in an industry's sales cycle should influence pay systems. They found that the more uncertain a firm's sales cycle, the more a salesperson's pay should be based on a fixed salary; the less uncertain the cycle, the more pay should depend on commission. Consider Boeing, whose salespeople can spend years talking with an airline before it actually places an order for new 787s. A firm like that would struggle to retain reps if pay depended mostly on commissions. In contrast, industries in which sales happen quickly and frequently (a door-to-door salesperson may have a chance to book revenue every hour) and in which sales correlate more directly with effort and so are less characterized by uncertainty, pay mostly (if not entirely) on commission. This research still drives how companies think about the mix between salaries and commissions.

Another important study, from the late 1980s, came from the economists Bengt Holmstrom and Paul Milgrom. In their very theoretical paper, which relies on a lot of assumptions, they found that a formula of straight-line commissions (in which salespeople earn commissions at the same rate no matter how much they sell) is generally the optimal way to pay reps. They argue that if you make a sales comp formula too complicated—with lots of bonuses or changes in commission structure triggered by hitting goals within a certain period—reps will find ways to game it. The most common method of doing that is to play with the timing of sales. If a salesperson needs to make a yearly quota, for instance, she might ask a friendly client to allow her to book a sale that would ordinarily be made in January during the final days of December instead (this is known as "pulling"); a rep who's already hit quota, in contrast, might be tempted to "push" December sales into January to get a head start on the next year's goal.

While a very simple comp plan such as the one advocated by Holmstrom and Milgrom can be appealing (for one thing, it's easier and less costly to administer), many companies opt for something more complex. They do so in recognition that each salesperson is unique, with individual motivations and needs, so a system with multiple components may be more attractive to a broad group of reps. In fact, to get the optimal work out of a particular salesperson, you should in theory design a compensation system tailored to that individual. For instance, some people are more motivated by cash, others by recognition, and still others by a noncash reward like a ski trip or a gift card. Some respond better to quarterly bonuses, while others are more productive if they focus on an annual quota. However, such an individualized plan would be extremely difficult and costly to administer, and companies fear the "watercooler effect": Reps might share information about their compensation with

one another, which could raise concerns about fairness and lead to resentment. So for now, individualized plans remain uncommon.

Concerns about fairness create other pressures when designing comp plans. For instance, companies realize that success in any field, including sales, involves a certain amount of luck. If a rep for a soft-drink company has a territory in which a Walmart is opening, her sales (and commission) will increase, but she's not responsible for the revenue jump—so in essence the company is paying her for being lucky. But when a salesperson's compensation decreases owing to bad luck, he or she may get upset and leave the firm. That attrition can be a problem. So even though there are downsides to making a compensation system more complex, many companies have done so in the hope of appealing to different types of salespeople and limiting the impact of luck by utilizing caps or compensating people for inputs or effort (such as number of calls made) instead of simply for closing sales.

Using Real Company Data to Build Understanding

The big difference between earlier research on sales compensation and the research that's come out in the past decade is that the latter is not based just on theories. Although companies tend to be very secretive about their pay plans, researchers have begun persuading them to share data. And companies have been opening up to academics, partly because of the attention being given to big data; managers hope that allowing researchers to apply high-powered math and estimation techniques to their numbers will help them develop better tools to motivate their workforce. Indeed, these new empirical studies have revealed some surprises, but they have also confirmed some of what we already believed about the best ways to pay.

Tom Steenburgh, a professor at the University of Virginia's Darden School of Business, published one of the first of these papers, in 2008. He persuaded a B2B firm selling office equipment to give him several years of sales and compensation information. This unique data set allowed Steenburgh to look at sales and pay data for individual salespeople and use it to make assumptions about how pay influences behavior. The company had a complex compensation plan: Reps earned a salary, commissions, quarterly bonuses based on hitting quotas, an additional yearly bonus, and an "overachievement" commission that kicked in once they passed certain sales goals. He focused on the issue of timing games: Was there evidence that salespeople were pushing or pulling sales from one quarter to another to help them hit their quotas and earn incentive pay? That's a really important question, because pushing and pulling don't increase a firm's revenue, and so paying salespeople extra for doing that is a waste.

Even though the salespeople in the study could receive (or miss out on) substantial bonuses for hitting (or missing) quotas, Steenburgh found no evidence of timing games. He concluded that the firm's customers required sales to close according to their own needs (at the end of a quarter or a year, say) and that the firm's managers were able to keep close enough tabs on the reps to prevent them from influencing the timing of sales in a way that would boost their incentive payments. That finding was significant, because quotas and bonuses are a large part of most sales compensation plans.

In 2011 Sanjog Misra, of UCLA, and Harikesh Nair, of Stanford, published a study that analyzed the sales comp plan of a Fortune 500 optical products company. In contrast with the firm Steenburgh studied, this company had a relatively simple plan: It paid a salary plus a standard commission on sales after achieving quota, and it capped how much a rep could earn in order to prevent windfalls from really big sales. Such caps are relatively common in large companies.

As they analyzed the data, Misra and Nair concluded that the cap was hurting overall sales and that the company would be better off removing it. They also determined that many reps' motivation was hurt by the firm's practice of ratcheting. Setting and adjusting quotas is a very sensitive piece of the sales compensation formula, and there's disagreement over ratcheting: Some feel that if you don't adjust quotas, you're making it too easy for reps to earn big commissions and bonuses, while others argue that if you raise a person's quota after a very strong year, you're effectively penalizing your top performers.

Misra and Nair estimated that if this firm removed the cap on sales reps' earnings and eliminated quotas, sales would increase by 8%. The company implemented those recommendations, and the next year companywide revenue rose by 9%.

A third empirical study of sales rep pay, on which I am the lead author, was published in Marketing Science in 2014. Like Steenburgh, we utilized data from a B2B office equipment supplier with a complex compensation plan. We examined how the components of the plan affected various kinds of reps: high performers, low performers, and middle-of-the-road performers.

We found that although the salary and straight commission affected the three groups in similar ways, the other components created different incentives that appealed to certain subsets of the sales force. For instance, overachievement commissions were important for keeping the highest performers motivated and engaged after they'd hit their quotas. Quarterly bonuses were most important for the lower performers: Whereas the high performers could be effectively incentivized by

a yearly quota and bonus, more-frequent goals helped keep lower performers on track. Some people compare the way people compensate a sales force to the way teachers motivate students: Top students will do fine in a course in which the entire grade is determined by a final exam, but lower-performing students need frequent quizzes and tests during the semester to motivate them to keep up. Our study showed that the same general rule applies to sales compensation.

How to Create a Sales Comp Plan

STEP 1 SET THE PAY LEVEL

This is crucial for attracting and retaining talent.

STEP 2 BALANCE SALARY AND INCENTIVES

The proportion of earnings that comes from salary and from incentives determines the riskiness of the plan. The proper balance varies by industry and is often based on the degree of certainty that a salesperson's efforts will directly influence sales.

STEP 3 DESIGN THE PLAN

Metrics Most companies still pay salespeople a commission based on gross revenue, although some companies pay on the basis of profitability of sales.

Plan Type Many companies supplement salary and commissions with bonuses based on exceeding quotas or reaching other goals.

Payout Curve Caps on earnings limit the pay of top performers and flatten the payout curve (or make it "regressive"); accelerators or overachievement commissions ramp up the pay of top performers, creating a "progressive" structure.

STEP 4 CHOOSE PAYOUT PERIODS

Companies can set quotas and bonus structures to cover periods ranging from a single week to an entire year. Research shows that shorter payout periods help keep low performers motivated and engaged.

STEP 5 CONSIDER ADDITIONAL ELEMENTS

Many companies use nonmonetary incentives, such as contests or recognition programs.

SOURCE ADAPTED FROM *THE POWER OF SALES ANALYTICS*, BY ANDRIS A. ZOLTNERS, PRABHAKANT SINHA, AND SALLY E. LORIMER

FROM "HOW TO *REALLY* MOTIVATE SALESPEOPLE," APRIL 2015

© HBR.ORG

Our research also suggested that the firm would benefit if it shifted from quarterly bonuses to cumulative quarterly bonuses. For example, say a salesperson is supposed to sell 300 units in the first quarter and 300 units in the second quarter. Under a regular quarterly plan, a salesperson who misses that number in the first quarter but sells 300 units in the second quarter will still get the second-quarter bonus. Under a cumulative system, the rep needs to have cumulative (year-to-date) sales of 600 units to get the second-quarter bonus, regardless of his first-quarter performance. Cumulative quotas do a better job of keeping reps motivated during periods in which they're showing poor results, because reps know that even if they're going to miss their number, any sales they can squeeze out will help them reach their cumulative number for the next period. In fact, even before we made our recommendations to the company in our study, managers there decided to move to cumulative quotas.

Out of the Lab, Into the Field

In addition to sharing sales and compensation data with academics, companies in the past several years have been allowing controlled, short-term field experiments in which researchers adjust reps' pay and measure the effects. Prior to the use of field experiments, most academic experiments regarding sales force compensation took place in labs and involved volunteers (usually undergraduates) rather than real salespeople. Shifting from this artificial setting into actual companies helps make the results of these studies more practical and convincing.

As an example of one such experiment, consider recent work my colleague Das Narayandas and I did with a South Asian company that has a retail sales force for its consumer durable products. The company uses a simple system of linear commissions—reps earn a fixed percentage of sales, with no quotas, bonuses, or overachievement commissions. Managers were interested in seeing how instituting bonuses would affect the reps' performance, so over six months we tested various ways to frame and time bonuses—always comparing results against a control group.

For one of our experimental groups, we created a bonus that was payable at the end of the week if a rep sold six units. For another group, we framed the bonus differently, using the well-known concept of loss aversion, which posits that the pain people feel from a loss exceeds the happiness they feel from a gain. Instead of telling reps they would receive a bonus if they sold six units, we told them they would receive a bonus unless they failed to sell at least six units. To test the concept even further, the company's managers suggested another experiment in which we paid

the bonuses at the beginning of the week and then had the reps return the money if they missed the goal.

The results showed that all three types of bonuses exerted similar effects and that in every case the group receiving the bonus generally outsold the control group. Loss aversion didn't have much effect. We believe that's partly because we were using cash, which is liquid and interchangeable; in the future we might experiment with noncash rewards, such as physical objects.

We also tried to measure the impact on sales reps' effort of cash payments that were framed as gifts (as opposed to bonuses). Whereas bonuses are viewed as transactional, research shows that framing something as a gift creates a particular form of goodwill between the giver and recipient. In our study we used cash but told employees it was a gift because there were no strings attached—they didn't have to meet a quota to receive it. We found that the timing of a gift directly influences how reps respond: If you give the gift at the beginning of a period, they view it as a reward for past performance and tend to slack off. If you tell them they will receive a gift at the end of a period, they work harder. We concluded that if companies want to encourage that kind of reciprocity, they need to pay careful attention to timing.

Other researchers are using field experiments to better understand how salespeople react to changes in payment schemes, but most of this work is so new that it hasn't been published yet. One paper presented at a conference in 2014 showed that if salespeople receive cash incentives for passing tests about the product they are selling, they will sell more. (This is an example of sales compensation based on effort as opposed to results.) Another recent field experiment found that sales reps valued noncash incentives (such as points that could be used for vacations or for items such as televisions) more than the actual monetary cost of the good the points could purchase. As more researchers and companies embrace the use of field experiments, sales managers will learn even more about the best ways to motivate their teams.

It Pays to Experiment

After spending a decade in academia studying sales force compensation, I sometimes wonder what would happen if I were transported back into my job as a management consultant. What would I tell sales force managers to do differently?

Some of my advice would be straightforward: I would urge managers to remove the caps on commissions or, if they have to retain some ceiling for political reasons, to set it as high as possible. The research is clear on this point: Companies sell more when they eliminate thresholds at which salespeople's marginal incentives are re-

duced. There might be problems if some reps' earnings dramatically exceed their bosses' or even rival a C-suite executive's compensation, but the evidence shows that firms benefit when these arbitrary caps are removed.

I would tell sales managers to be extremely careful in setting and adjusting quotas. For instance, the research clearly shows that ratcheting quotas is detrimental. It's tempting to look at a sales rep who blows through her yearly number and conclude that the quota must be too low—and quotas do need to be adjusted from time to time. But in general it's important to prevent reps from feeling that unfairness or luck plays a part in compensation, and resetting quotas can contribute to that perception. And if something outside the salesperson's control—such as an economic downturn—made it more difficult to hit a goal, I would consider reducing the quota in the middle of the year. It's important to keep quotas at the right level to properly motivate people.

On the basis of my own research, I would advocate for a pay system with multiple components—one that's not overly complicated but has enough elements (such as quarterly performance bonuses and overachievement bonuses) to keep high performers, low performers, and average performers motivated and engaged throughout the year.

Finally, I would urge my client companies to consider experimenting with their pay systems. Over the past decade managers have become attuned to the value of experimentation (A/B testing, in particular); today many consumer goods companies experiment constantly to try to optimize pricing. There are important lessons to be learned from doing controlled experiments on sales reps' pay, because the behaviors encouraged by changes in incentives can exert a large influence on a firm's revenue, and because sales force compensation is a large cost that should be managed as efficiently as possible. Involving academic researchers in these experiments can be beneficial: Having a trained researcher take the lead generally will result in a more controlled environment, a more scientific process, and more-robust findings. These studies also help the world at large, because research that improves how companies motivate salespeople will result in better and more-profitable businesses for employees and shareholders.

Richard Pevear & Larissa Volokhonsky, The Art of Translation No. 4

Susannah Hunnewell, The Paris Review 213

Credited with starting a "quiet revolution," Larissa Volokhonsky and Richard Pevear have joined the small club of major translators whose interpretation of a master-piece displaces the one read by generations before. Volokhonsky, who is Russian, and Pevear, who is American, have been married thirty-three years. In that time, they have translated much of Russian literature as we know it. Their thirty or so translations include The Brothers Karamazov, Crime and Punishment, Demons, The Idiot, Notes from Underground, War and Peace, Anna Karenina, Hadji Murat, The Death of Ivan Ilyich and Other Stories, The Master and Margarita, Doctor Zhivago, Gogol's Collected Tales, Dead Souls, he Enchanted Wanderer and Other Stories by Nikolai Leskov, and Chekhov's Selected Stories.

Until their translation of The Brothers Karamazov was published in 1990, the English-speaking world got its Dostoevsky (their preferred spelling— with one y) Sent from my iPad her translations of Turgenev and Chekhov are generally considered virtuosic, her versions of Dostoevsky, Gogol, and Tolstoy have drawn criticism for Victorian elision. Her Gogol translations are "dry and lat, and always unbearably demure," complained Nabokov. "The reason English-speaking readers can barely tell the difference between Tolstoy and Dostoevsky is that they aren't reading the prose of either one," grumbled Joseph Brodsky. The critic Korney Chukovsky summed it up best and most brutally when he wrote, "Who does not feel the convulsions, the nervous trembling of Dostoevsky's style? ... But with Constance Garnett it becomes a safe bland script: not a volcano, but a smooth lawn mowed in the English manner— which is to say a complete distortion of the original." For her part, Garnett once wrote, "Dostoievsky is so obscure and so careless a writer that one can scarcely help clarifying him."

Pevear and Volokhonsky's translations have been lauded for restoring the idiosyncrasies of the originals—the pagelong sentences and repetitions of Tolstoy, the cacophonous competing voices of Dostoevsky. hough almost unanimously praised by reviewers and Slavic scholars, they have a few critics who accuse them, in fierce blog posts, of being too literal or prone to unidiomatic turns of phrase. Pevear, who is sometimes drawn into the online jousting, never apologizes for erring on the side of the unfamiliar sounding over muting the original.

In 2004, the translators were propelled to commercial success when Oprah Win-

frey chose their translation of Anna Karenina for her book club, making the 137-year-old book an instant best seller. (he Moscow Times called it "the greatest promotion of Russian literature since Omar Sharif cantered across the steppe in a fur hat as Doctor Zhivago.")

Pevear and Volokhonsky have won the pen Translation Prize twice, for The Brothers Karamazov and Anna Karenina. Pevear, who has also translated French and Italian works, is Distinguished Professor Emeritus of Comparative Literature at the American University of Paris. In addition to translating Russian contemporary poets, Volokhonsky, who attended Yale Divinity School, has translated theological texts into Russian. They have two trilingual children.

The interview took place in January over two long afternoons in their groundloor apartment in the 15th arrondissement of Paris, where they have lived since 1998. Volokhonsky is warm and reserved. She has strong opinions, sometimes delivered bluntly. She doesn't like facile answers. She accepts praise with sincere embarrassment and pleasure. She speaks with a thick Russian accent, which adds to her considerable charm. Pevear looks like a New England ship captain, bearded and with an excellent head of hair. He has a slow, easygoing manner which belies his precise tastes. He enjoys puns and repartee. In the beginning, the couple took turns speaking, listening respectfully in mortuary silence as the other spoke. But soon, they were interrupting each other, finishing the other's sentence, prodding the other to speak, teasing or correcting each other, though they were always in general agreement. At the end of the interview, which, like a nine-hundred-page Russian novel, seemed to contain all subjects simultaneously, we opened a half bottle of champagne. Pevear had bought it to celebrate New Year's Eve, but the two had fallen asleep before midnight. — Susannah Hunnewell

INTERVIEWER

How did you meet?

PEVEAR

We actually met because of Russian literature. I had written an essay on the Soviet dissident and writer Andrei Sinyavsky. It was published in The Hudson Review in 1972. I remarked ironically that the poet Yevtushenko was giving readings in Madison Square Garden—among his translators were John Updike and Richard Wilbur—while Sinyavsky was in a Soviet labor camp. I received a letter from Irene Kirk, a professor at the University of Connecticut. She told me he wasn't in prison, he had been released but also stripped of his citizenship and deported. She had helped him

and his family leave for France.

VOLOKHONSKY

When I arrived in the United States, I stayed for a while with this professor, and she started matchmaking. Succeeded after a while. Not immediately.

INTERVIEWER

Who was resistant?

PEVEAR

Circumstances. Irene told me there was someone I should meet, and she invited me down to Connecticut. I was very surprised. I lived in Maine and worked in a boat-yard as a woodworker—boats were still made of wood back then. I took a little time of and drove down. It happened that Larissa had to renew her visa, which meant she had let for England just as I arrived.

INTERVIEWER

You missed each other.

PEVEAR

It took a few more years. By then we were in Manhattan, both of us.

INTERVIEWER

You lived on West 107th Street.

PEVEAR

Yes. And by some miracle I found Larissa an apartment on the same street.

VOLOKHONSKY

It was convenient.

PEVEAR

Larissa always says that if it hadn't been for 107th Street, we'd never have been married. When I moved to New York, I took up cabinetmaking. That's how I earned a living.

VOLOKHONSKY

Yes. We were neighbors with a wonderful, crotchety woman, an old translator from Russian, Mirra Ginsburg. She was a very good translator. We liked her. When we started to try to translate The Brothers Karamazov, we showed her samples. By then, Richard had translated some Russian children's poetry. Richard's very good at jingles.

PEVEAR

Who's that knocking at my door?
His badge is stamped with number four. His shoulder bag is big and fat.
His coat is blue, so is his hat.

VOLOKHONSKY

She read our samples and said, You can't do it. Just stick to your cabinetry and these children's poems. You're so good at these children's poems. And she said one phrase that sent me through the roof. She repeated it several times—she said, I adore the smell of wood shavings.

INTERVIEWER

Your first critic.

VOLOKHONSKY

Yes. We had yet another critic, at the very beginning, an old Russian emigrelady. When we first told her we were translating The Brothers Karamazov, she said, Oh, Dostoevsky, I hope you correct his awful style. I said, No, that is precisely what we're going to keep.

INTERVIEWER

How did you come to translate The Brothers Karamazov?

PEVEAR

I had read it for a summer course I took at Harvard in Russian literature. I happened to have a wonderful professor, Vladimir Markov. The course transformed me. I loved these books. I read Constance Garnett's translation. After we were married, I thought I'd try David Magarshack's translation. I started reading it. hen Larissa got curious.

VOLOKHONSKY

I had my Russian edition of Dostoevsky, and I decided to read along. Dostoevsky had always really gripped me. Usually if you read in your native tongue, unless you're either a scholar or an especially curious and attentive reader, you just read. You follow the plot, the characters, you hope maybe this time this one won't murder that one! But now I started actually looking at the language. I said, How is Magarshack going to translate this? And lo and behold, he didn't. It wasn't there. he jokes, or the unusualness, just disappeared.

INTERVIEWER

What was there instead?

VOLOKHONSKY

Something very bland. Something tame, not right. he meaning is there, but the style, the tone, the humor are gone. For example, there is a character, Mr. Miusov. He's a secondary character, but he's important because this particular scene is seen through his eyes. Mr. Miusov has just come from abroad. He's a liberal, he's cultivated, he's reined. Describing him, Dostoevsky adds this sarcastic touch—he says Miusov is "столичный, заграничный." It has the same jingle as hoity-toity. English kindly gave us "metropolitan, cosmopolitan." We were lucky. We're not always so lucky.

INTERVIEWER

What had the other translators said?

VOLOKHONSKY

"Who had been in capitals and abroad." hey would give the information but not the voice. his is the kind of thing I began to notice throughout the novel. Sometimes three times, five times on a page.

PEVEAR

And I discovered during our work together on Dostoevsky that he was not a brooding, obsessed man, but a very playful, free spirit. You see it in his style. he style of Dostoevsky is extremely varied. He would practice writing pages in different voices. He shows characters through the voice, through the way they use or misuse language. Which meant a lot of people used to say that he didn't write very well! For example, there is a little note at the beginning of Karamazov, "From the Author," about how he came to write the book. he "author" is not Dostoevsky—he makes that perfectly clear—although everybody seems to think that Dostoevsky is the narrator. But the narrator isn't a writer at all. He just happens to live in the town where the novel is set. He got interested in the story of the Karamazov brothers and the murder of their father and wanted to record it. he whole point of this preface is to introduce all possible voicings of this narrator, who writes absurd things like, "Being at a loss to resolve these questions, I am resolved to leave them without any resolution." And of course all the translators vary the words, because Flaubert said you should never use the same word twice on the same page. Finally he says, "Well, that is the end of my introduction. I quite agree that it is superluous, but since it is already written, let it stand." Dostoevsky gets you into the entire question of whether this man is trustworthy. Does he know what he's talking about? he uncertainty surrounding this narrator is very important, and all of that is introduced just by the way it's written. So the light suddenly went on.

VOLOKHONSKY

I said to Richard, You are reading a different book.

PEVEAR

It occurred to us that there was a whole other register to Dostoevsky, and the translators hadn't translated it. here was something to be done there.

INTERVIEWER

So what happened exactly?

VOLOKHONSKY

We had no names as translators. Richard had some kind of name as an essayist and as a poet.

PEVEAR

I had published one book, a translation of the French philosopher Alain, with New Directions.

VOLOKHONSKY

So we prepared four passages representative of four different kinds of narrative and dialogue. We sent these samples to five of the most prominent Dostoevsky scholars, and they all sent us very positive responses.

PEVEAR

We had this package of samples and letters, and I started mailing it out to major publishers, who all turned it down.

INTERVIEWER

What was their reason?

PEVEAR

There was no need for a new translation.

VOLOKHONSKY

And then, finally, there was this wonderful small press.

PEVEAR

I sent it to Jack Shoemaker at North Point Press. His assistant called us on the phone from California. He said they wanted to publish it, and with regard to an advance, asked, How does a thousand dollars sound? I said, Very small. He said, I'll get back to you. Which usually means you'll never hear from them again. But he actually called the next day and said, How does six thousand sound? I said, Much better than one thousand. They put everything they had into it. They really did a beautiful job. They made a press kit—I wish I could show it to you. It was a double portfolio with samples and letters. They sent it all over. We got wonderful reviews in smalltown newspapers because they didn't have to read the book. They just read the press kit. My favorite one was from the Wichita Eagle, which did a full-page review, with a

full-page photograph. he title was “Karamazov Still Leads Creative Way,” and the photograph was—Tolstoy! With his big beard and scowling face ...

VOLOKHONSKY

And big eyebrows.

INTERVIEWER

You translated four Dostoevsky works in a row—The Brothers Karamazov, Crime and Punishment, Notes from Underground, and Demons. What is it like living in a world of toxic narrators and tortured murderers for five years? Does it affect you personally?

VOLOKHONSKY

No. It’s a professional thing.

PEVEAR

Oh, it does. I think it affects me, certainly.

VOLOKHONSKY

I never noticed.

PEVEAR

You didn’t see me twitching?

VOLOKHONSKY

We had two small children in a row. I had my blind old mother living with us. I had my own solid reality right there.

PEVEAR

I do live in the book, in the voice or voices. If you don’t enter into it, you can’t really translate it. But there is also a certain detachment. You keep having to step back and think, How do I say that in English? Translation isn’t done by principle or by a machine. he only way you can judge what you’re doing is by how it feels to you. Is that

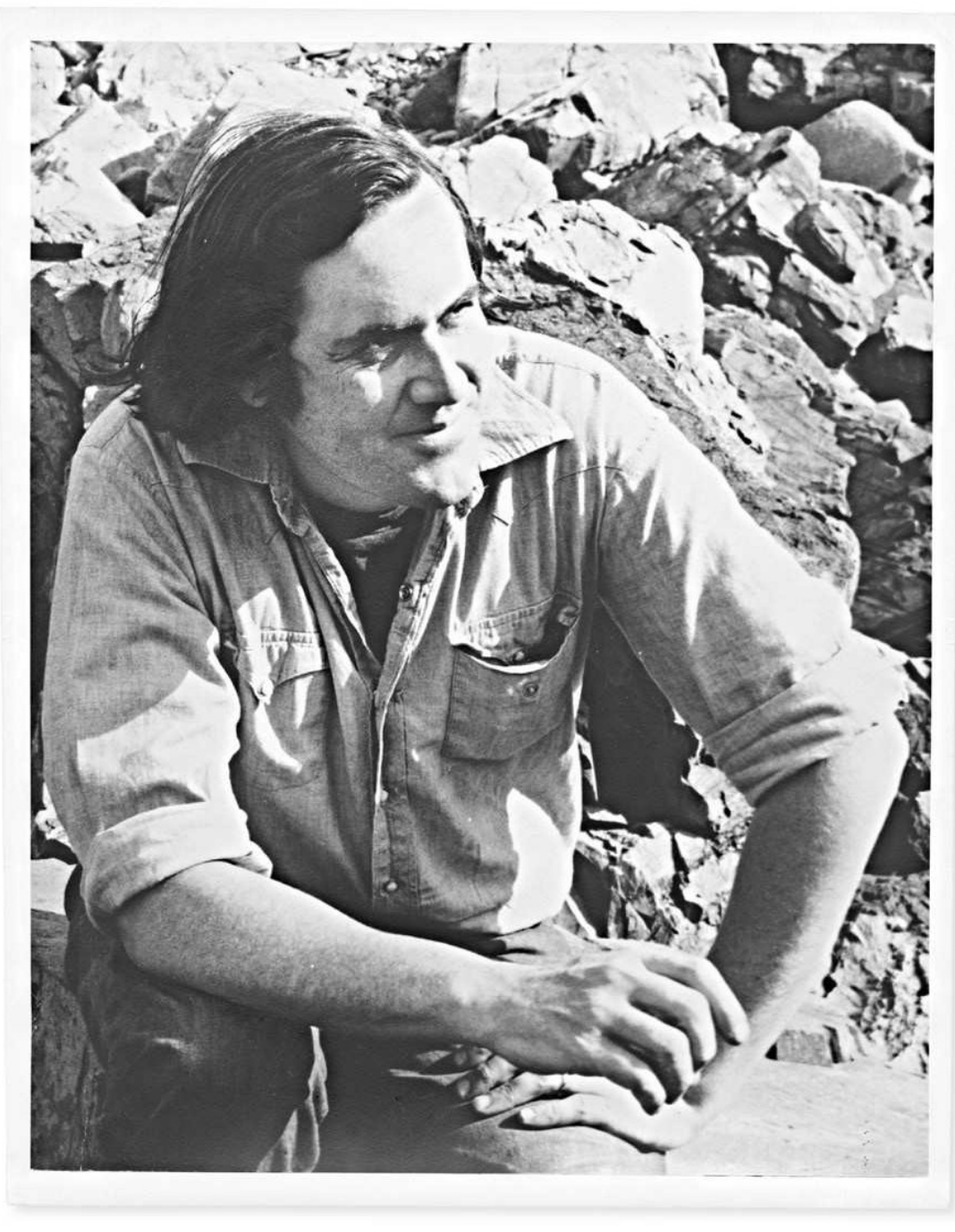
the life of it? And for that there has to be a lot of identiication—not with the characters but with the art of the book, the art that went into it. You have to have that in order to choose your words. They have to feel right. It's impossible to deine. Writers know this feeling.

INTERVIEWER

You write poetry. Is it similar to writing your own work?

PEVEAR

Yes, it's very close. Because there's a kind of singing that goes on in your head. And you realize, hat phrasing didn't work. If I invert it, it would be better. Robert Frost used to say, "It goes right." In a way, translating prose is closer to poetic writing than it is to real prose writing. Prose writers have a lot of dirty work to do. A good prose writer has to make a house, put furniture in it, open doors, bring people in, give them hair and eyes and clothes. They have to make a world and populate it. For a translator, that's all been done.



INTERVIEWER

Since you've started translating, have you written less poetry?

PEVEAR

Yes. When I got married, I stopped writing poetry.

VOLOKHONSKY

Ah, I'm not inspiring.

PEVEAR

Poets always write about longing. Dante is looking for Beatrice in heaven because he couldn't have her on earth.

INTERVIEWER

What was your original interest in writing?

PEVEAR

I fell in love with poetry when I was about fourteen. I picked up a couple of the little paperback anthologies of English poetry edited by Oscar Williams and used to carry them around with me everywhere. A little later I acquired E.E. Cummings's Poems 1923–1954, my first real book of poetry. Just recently I reread Cummings's Eimi, the novelistic journal of his trip to Russia in 1931, and was struck not only by the liveliness of the writing, but also by his grasp of what was actually happening in the Soviet Union under Stalin, when so many enthusiasts failed to see it. I read Yeats constantly, and Blake and Pound and Frost and Donne and Hopkins and Chaucer and Shakespeare of course. And the other Elizabethans. And John Skelton, whom I learned about from Robert Graves's Oxford lectures on poetry. I went on to write my senior thesis in college on Skelton. And I taught myself Italian in order to read Dante. Kaka, Laurence Sterne, and Flann O'Brien are among the prose writers I read and reread. And my teaching of Homer brought me to Simone Weil's essay "The Iliad, or the Poem of Force," and from that to all of her writings, which have been of the greatest importance to me.

INTERVIEWER

When did you first start translating?

PEVEAR

In college, I was writing poetry, but pretty soon I started translating, too. My high school language was Spanish, and I've translated a lot of Spanish—Antonio Machado, Borges—just for my desk drawer. But the first thing I

really seriously translated was "La jolie rousse" by Apollinaire. I loved the poem. I loved the liveliness of the poet—the freedom, the life, and the wit.

INTERVIEWER

Tell me about your background, Richard.

PEVEAR

My ancestors were Huguenots from Bordeaux who let France after the repeal of the Edict of Nantes, in 1685. They felt uncomfortable because they were Protestant. So they moved to the island of Guernsey, of the coast of Normandy, and then thought they weren't far enough away, so they moved to the Massachusetts Bay Colony in the early eighteenth century, like Paul Revere's father, Apollos Rivoire. He was a Huguenot in the same colony. My paternal grandfather worked in the family business, a coal company. My grandparents on my mother's side were teachers. My grandmother taught Latin at the Boston Latin School. My grandfather taught high school science. My father studied art at the School of the Museum of Fine Arts in Boston. He was a great craftsman. He could paint and draw, but it was the Depression so he decided to be practical and study commercial art. He went to work as a designer for Lever Brothers in Cambridge. He became their packaging manager when they moved to the new Lever House in New York in 1952. I grew up in Albertson, on Long Island.

INTERVIEWER

Now from Long Island to Leningrad!

VOLOKHONSKY

Yes. I grew up in Leningrad. My family was Jewish. Not intellectual but there was some intellectual pretense. My mother was a housewife, but she read a lot and she wanted us to be educated. So she had us study languages. She invited private teach-

ers to tutor us, my brother and me, at home. So I have known English since I was little. My father worked as an administrator in a factory.

INTERVIEWER

What kind of factory was it?

VOLOKHONSKY

Production to keep deaf and mute people occupied. My parents survived the thirties and Stalin somehow. They had a difficult and poor life. They never talked about it. They were very frightened people. They weren't dissidents, but I remember my mother on the day Stalin died. I was very little. My mother and I went out and there were big red lags with broad black trimming everywhere. And I said, Mama, what are these lags? And my mother said, Stalin's no more. Just very curtly. Nobody knew what was going to happen, who would come to power, what would be the new rules.

I owe a lot to my older brother, Anri Volokhonsky. Very early, he educated himself at home in humanities and in classics, in history, literature, and philosophy. It was an eclectic kind of education, but it was much more than a normal Soviet school could provide. He was also trained in mathematics and chemistry. He's an interesting man and a very good poet. I love him and have had a great respect for him all my life. So because of him I read a lot. We had a very lively literary life in Leningrad. Joseph Brodsky was slightly younger than my brother, but they knew each other. When I was in high school, we ran to a poet's café to listen to Brodsky read his early poetry. I still remember how he read. In two notes. Dadadada high, dadadada low. People formed groups and fought with each other, wrote epigrams about each other, and read to each other. Everybody knew each other. his is what foreigners liked when they came to Russia. They would come and the next day they would know everybody.

PEVEAR

Also, you could actually live as a translator.

VOLOKHONSKY

he poets and writers who could not publish their own things for ideological reasons could do translation either for money or simply as something else to do. So everybody translated. And that is why we had such good translators. Pasternak did several plays by Shakespeare. It was a really respectable occupation.

INTERVIEWER

You have a degree in mathematical linguistics. What does that mean exactly?

VOLOKHONSKY

I wanted to study Greek and Latin, but my parents just were dead against it. Not practical. Mathematical linguistics was the beginning of computer science—mathematics as applied to language, programming.

INTERVIEWER

Which is a form of translation.

VOLOKHONSKY

Yes, it is a language.

INTERVIEWER

Has any of that come to your aid as a translator?

VOLOKHONSKY

No. his is the first time that this has occurred to me, but I wrote my university thesis on machine translation. My professor asked me to analyze formally certain German constructions for the purposes of machine translation. It was before the digital age, so it was still very far from any practical realization, but it was then that I understood that machine translation is not possible. I can formalize all these syntactic connections and relationships until I'm blue and gray and dead in my grave, and still I will never convey the taste, the flavor, the rhythm, the smell, the music that I hear. It is really a matter of having a worldview.

PEVEAR

Machines don't have a worldview.

VOLOKHONSKY

But I did somehow get this degree, and I never did anything with it.

INTERVIEWER

In fact, you went to work for an institute of marine biology.

Yes, evasion after evasion!

VOLOKHONSKY

In the Soviet Union, after university, you would be sent to certain institu tions to work, and you had to go. They wanted to assign me to the socalled postoice boxes, which means military institutions—ship and aircrat build ing. My whole group went, but I refused to. I wasn't particularly conscious of what I was doing, but something told me not to do it. So I found a job in marine biology. They needed someone to translate from English.

INTERVIEWER

Your first translating job. What was the subject?

VOLOKHONSKY

The genetics of Pacific salmon. I went to Vladivostok, Sakhalin Island, and Kamchatka. I saw the most beautiful places in the world.

PEVEAR

They skinned bears.

VOLOKHONSKY

I didn't skin bears. You imagined that.

PEVEAR

Your friends skinned bears.

VOLOKHONSKY

We were allowed to hunt, you see. They were very wild places. But I didn't hunt. I could have told you I did, but I didn't.

INTERVIEWER

Then you emigrated to the United States, via Israel.

VOLOKHONSKY

It was the early seventies, Brezhnev just came to power. It felt bad. It stank. We weren't dissidents in the sense that we would go to Red Square and pro test, but we didn't like it.

I was in Israel for two years. It's a very interesting country, but I never wanted to go to Israel. I knew very little about it. I went because my brother went. I wasn't really happy there, because I liked to be in a big country. You know how it is—you feel good in some places. Emigration also is a very strange thing. You start from zero, you don't know anybody, you don't know the language. It's very difficult.

INTERVIEWER

Once in the United States, you decided to go to divinity school.

VOLOKHONSKY

I don't know why. I decided—I think correctly, without knowing much about it—that the way we are is because of our religion. I studied the Bible, the Old Testament, the New Testament, the history of Christianity, the history of religions. I didn't have any practical purpose in mind.

INTERVIEWER

Though it may have come in handy when you and Richard translated Dostoevsky. Tell me how your collaboration process works.

VOLOKHONSKY

See this lamp and that ugly little board? I sit there and I write the first draft with a pencil. I try to be very close to the text, which is sometimes more possible, sometimes less possible. In the margins, I point out everything I see. For example, his is a Russian proverb that rhymes. Or, his is a cliché, totally banal. Or, his is a cliché, but not quite a cliché. Like, "to defend something to the last drop of his blood" would be a Russian cliché, but Dostoevsky just says, "To the last drop." I will note if it is a

repetition of the same word, if it's almost the same word, or if it's a parallel construction.

PEVEAR

Or a biblical quotation.

VOLOKHONSKY

Or a distorted biblical quotation, quoted from memory. Or there is a four foot trochaic meter, but it's hidden. Is it important? A scholar would say yes—scholars are usually so happy to discover such things—but it may not be important. What is more important is to make a decent English sentence out of it. So these are the kind of things that emerge. his nineteenth-century literature is illed with hidden quotations, references, allusions to poetry or songs which I must indicate.

INTERVIEWER

Do you recognize it yourself, or is it noted in your Russian text?

VOLOKHONSKY

A lot of it I recognize myself. I will get a feeling that something is there. But we also have very good Russian annotated editions, and that helps.

PEVEAR

So Larissa produces this pile of manuscript. It's a very good way to use old proofs. We take old page proofs and turn them over and there's the next translation. It's ecological.

INTERVIEWER

You're doing it by hand?

PEVEAR

Right, Larissa does it by hand, and then I start going through it. I have the original, I have earlier translations, and I start listening to it.

INTERVIEWER

Richard, you once wrote that rumors of your ignorance of Russian are somewhat exaggerated. What is the actual state of your Russian?

PEVEAR

I can't really speak Russian. But I can hear it, and I understand quite a lot. I look at the text all the time as I translate. I don't just use Larissa's translated manuscript. She even sometimes gets very angry. She says, Where did you get that? You must have looked in the dictionary!

VOLOKHONSKY

Richard has something better than the knowledge of Russian. He has intuition and literary style.

INTERVIEWER

When your first draft goes to Richard, what does it look like? Is it close to what it might become?

VOLOKHONSKY

You mean, how bad is it? How bad is it, Richard? Tell me.

PEVEAR

She makes it as bad as possible so that I have something to do.

VOLOKHONSKY

I could produce something more literary, but I deliberately don't do it, because this would eliminate possibilities. I want to give Richard as many possibilities as I can. And besides, how does one choose? For instance, you can say proud or haughty, arrogant, supercilious, down the nose, snooty... I cannot and don't want to make these decisions, because first of all, it depends on previous choices and previous sentences. It is up to Richard to make these choices.

INTERVIEWER

If Richard changes your suggested translation, is it hurtful?

VOLOKHONSKY

No. I never wanted to be a writer.

PEVEAR

But I sometimes get very firmly told to put something back that I changed.

INTERVIEWER

So you sometimes have to control the poet...

PEVEAR

That's the idea. his tandem work is very useful for us because I might start inventing and she can always bring me back. Each of us needs what the other one has.

VOLOKHONSKY

I have no ambition to be a writer. I have had occasions in my life when I needed to write some article or memoir, and I did it with great, great difficulty, very slowly and painfully.

PEVEAR

Don't you know that every writer does?

VOLOKHONSKY

Yes, but a real writer wants to do it again, and I never did.

INTERVIEWER

Back to your process.

VOLOKHONSKY

I look at Richard's version and read it very attentively, against the original, trying to respond to what he asks and asking my own questions about what he did. hen we sit down together, with reference books and dictionaries, and work on it, word by

word, phrase by phrase.

PEVEAR

At that point, we can sometimes spend an hour on one word. hen I take this mass of paper and produce the third complete version.

VOLOKHONSKY

And then Richard reads it aloud in English, and I follow with the Russian text. We make some changes, but few. And then it goes to the publisher.

INTERVIEWER

How quickly do you translate?

PEVEAR

In one day, I could do five pages of Dostoevsky. I discovered I could do nine or ten of Tolstoy. But two, maybe three of Gogol.

VOLOKHONSKY

You were younger then. You've become slower and lazier.

PEVEAR

Well, now that we've established that! We do shorter and shorter things, and we take longer and longer to do them.

INTERVIEWER

Richard, you have translated French and Italian authors. How is it different to translate by yourself?

PEVEAR

It's a lot easier.

VOLOKHONSKY

No one bothers you.

INTERVIEWER

You translated Alexandre Dumas, another nineteenth-century giant.

PEVEAR

Larissa was starting War and Peace and I had to give her a year's start, so I had nothing to do. Dumas was fun to translate. I also discovered, as usual, that it's not as obvious as you'd think. You think that it's all sword fights, but there are hardly any. he book is ninety percent dialogue. Much more important is what they say than what they do. Dumas's prose is very direct, very simple. I noticed that the translators into English tend to add rhetoric to it because they think there should be swashbuckling. But in fact, Dumas is very plain. He's reporting what happened. It has to move quickly. It has to be alert.

VOLOKHONSKY

And there is a story behind it. In Russia, he Three Musketeers is very popular, much better known than in the Anglophone world. And when we were first married, I discovered that Richard hadn't read he Three Musketeers. And I said, What? I'm not going to be married to a man who has never read he Three Musketeers. I went to the French bookstore and bought a copy, and Richard read he Three Musketeers.

PEVEAR

I loved it. his story has a very funny sequel to it, too, because at that time I was translating the French poet Yves Bonnefoy. So I went up to Yale, where he was teaching, to go over the translations with him. He was wonderful to work with, very restrained.

INTERVIEWER

How's his English?

PEVEAR

Perfect. But he let me do what I wanted. He found a few things to say, but he gave me great freedom. Anyway, when we finished working, we had din ner and were talking. I said, By the way, I've just read he Three Musketeers, it's a great book. He

looked at me very sadly and said, I have never read it.

INTERVIEWER

Back to your collaboration. How was it going from Dostoevsky to Gogol? Was it a relief ?

PEVEAR

No, it was much harder. Gogol is quietly wild, I would say. People think he's rhetorical, and some people say our translations are too simple. But we translate what he wrote, and he didn't use rhetoric. He narrates, with a very lat voice, all kinds of extraordinary things. Everything he says is unexpected. And of course it's wildly comic. Dostoevsky is funny, but not in that way.

INTERVIEWER

And you're getting this from Larissa's first manuscript? Her translation and her commentary are giving you Gogol?

PEVEAR

That's what's going on.

INTERVIEWER

A great responsibility for Larissa.

VOLOKHONSKY

I am beginning to tremble in my shoes. Belated trembling. I don't know what to say.

PEVEAR

We're two very simpleminded people.

VOLOKHONSKY

I'm not a scholar. I just read the text as a reader, but a very attentive reader. But yes, it's a great responsibility. Maybe someone else would read it differently. For instance, in Leskov's short story whose title I translated as "A Flaming Patriot," which

is literally what it means in Russian, this very simple lady is a companion to a Russian princess abroad in Austria. They're promenading in a park in a carriage. They see the emperor, Franz Joseph, leave his royal carriage to have a beer with the local workers. he princess wants to see him better, so she orders the driver to move the horses forward. he simple woman is terribly embarrassed, so much so that in the end she abandons her position as a companion.

PEVEAR

She has her pride as a Russian. She is the flaming patriot.

VOLOKHONSKY

Of course, it's ironic.

PEVEAR

But one critic said the title should have been "A Woman Who Loves Her Country Ardently." I wrote and asked him, Are you serious? I even think Leskov may have borrowed it from the English cliché. He'd heard the phrase—it was used just around that time, in the mid-nineteenth century, in America and England.

VOLOKHONSKY

But all that is to say that my responsibility is great. I have to decide, is it ironic or not? here is another wonderful story by Leskov, "The Man on Watch," which all other translators have translated as "The Sentry," which is nice as a title. In Russian, it's "Человек на часах," which means "The Man Who Stands Watch." It's about a sentry who has a heart, who has feelings. He is a human being, first of all. And he abandons his watch in order to save another human being, a drowning drunkard, and is punished for it. And so this is my choice. he title does mean "sentry," but I think it's better to say "The Man on Watch," as Leskov did. He was perfectly capable of saying "Часовой," "The Sentry," but he didn't.

INTERVIEWER

What do you do when you have a disagreement?

PEVEAR

We don't disagree.

VOLOKHONSKY

We're boring. The thing is that our realms of competence overlap, but they don't coincide fully. In English, Richard always has the final word because it's his language. Let's try an experiment. "Mortally in love." Can you say that in English?

INTERVIEWER

It's unusual.

PEVEAR

Some readers would pause and be angry because they don't want to pause. "Why do they say 'mortally'? We don't say that!"

INTERVIEWER

What is it from?

VOLOKHONSKY

It's from a story called "The Blizzard," by Pushkin. It's one of the most charming stories in the world. here is a young girl who becomes delirious. Her parents figure out that she is "mortally in love." In Russian, it is now a standard phrase. Yet when I asked my Russian friends, Do we say it or do we not? I got different answers. his is what happens when you start mulling things over in your head. But Pushkin was one of the first to use it in fiction, and it entered the language. Whether it was unusual before him is impossible to know. It is such a wonderful phrase that we decided to use it. And even if it makes you pause, it is immediately clear what it means.

INTERVIEWER

Fatal is often used with love.

PEVEAR

But that suggests fate, while mortally means "until death."

VOLOKHONSKY

So this is the kind of discussion we have. How can we quarrel?

INTERVIEWER

The subject of translation can be very emotional.

PEVEAR

Yes, once people realize that there's more than one translation. Very naive readers think you take the Russian and you put it in English, and then you're done. Why would there be two translations? People still ask us that—hasn't it been translated already? Some people feel very uneasy once they discover that there might be two or three different ways of translating something. How do I know which one to read? And that part is very emotional.

VOLOKHONSKY

It's like with editing. You are given a text to edit. Immediately the instinct would be, I would write it differently. It's the same with the critique of translation. When I'm given a text to read and to comment on, I always say to myself, Just calm down, this is another person's style, have some respect.

There is also such a thing as the first translation you read. For instance, I read Proust in a very good Russian translation. For me, that was Proust. I never aspired to read Proust in French. And then, in the seventies, another translation began to appear. And the whole world just split in two. here were those who loved the first one and those who loved the second one. I preferred the first one but then I thought, here are very interesting things in the second one. here must be something in the original that he's trying to convey. his is why I don't like this phrase that publishers use to advertise a book, "the deinitive translation." hey shouldn't say that. It's a little bit embarrassing. here is always something else to say. But I know from my own experience that you are comfortable with the first translation you read. I still think it is a great miracle that our translations were so well received, because it is very hard to introduce a new translation.

PEVEAR

I understand when people say, I still prefer the Maude or the Garnett. I understand it completely.

INTERVIEWER

I have to ask—what do you think of Constance Garnett?

VOLOKHONSKY

She's a good translator. It's just that we discovered she omitted certain registers of Dostoevsky's work. But otherwise she was a very good translator.

PEVEAR

Very often, when there are a number of translations, hers is the closest to the original. She translated an enormous amount. People say about us, Look at them, they pour it out! We've done about thirty translations. She did more than seventy. She did almost all of Chekhov—seventeen volumes. Almost all of Dostoevsky. Almost all of Tolstoy and Turgenev. And more besides.

INTERVIEWER

When you translate, how much do you consult other translations?

PEVEAR

I usually have at least one. If there's one that I'm interested in, or if I want to control or check myself, I keep it open.

INTERVIEWER

There must be passages in the Russian that are very hard to understand for whatever reason. Is that also a moment when you look at other translations?

VOLOKHONSKY

Pasternak has a lot of that. His poetic light takes him so high that you no longer know where you are, and some of it is deliberate, God knows. And we usually try to keep the same kind of difficulty. his gives, of course, a lot of material to people who would like to criticize. But if there is something that is simple but that I don't understand, I will look at other translations. Or ask somebody. But it usually doesn't help because you see that they also don't understand.

INTERVIEWER

You once said Chekhov is especially hard to translate.

VOLOKHONSKY

I say that about everything. he thing is that with Dostoevsky, it's complex, it's rich, it's interesting, it's rough. You have a bump, you translate a bump. You know what you're doing. But with Chekhov, it's an even road. It's simple, direct. And yet it isn't. here is no excess in his prose. When you're on a vast, smooth plain, you have to ind your way. It's hard.

PEVEAR

he words seem simple, but when you start looking, there's enormous life underneath. And that's what you have to try to catch. here's a wellknown Russian poet and translator named Olga Sedakova. he thing about transla tion, she says, is that it's a matter of both the separate words and the whole, which struck me as absolutely true. here are these two things to consider simultaneously. You look at each word

and you think about it, but there is also the whole. Otherwise it's just piecemeal. here's no movement, and movement is the hardest thing to deine, but it's the most important. In a review, critics have to quote something. And you can't quote what Olga calls "the whole." So you pick something and say, I think this is wrong. Or, Look at that compared to this. But in all these cases, you miss the overall quality, which makes all the difference.

VOLOKHONSKY

Now really everybody has become an expert at translation. here is a lot of talk about foreignization, familiarization, domestication! All this incredible terminology that doesn't help much. In the end, it's a general impression. Sedakova says, How do you judge or deine it? Alas, only by intuition. And that's true— it's very hard to deine really what happens, why one translation achieves something that the other doesn't. It's not just the knowledge of one language and the knowledge of another. As we know, scholars are often not the best translators. It's not anything speciic. It's partly your sensibility, your reading experience.

INTERVIEWER

Tell me about your reading experience.

VOLOKHONSKY

As an adolescent, I read a lot of classic ancient authors. Homer had been translated by very great translators, the poets Gnedich and Zhukovsky, in the nineteenth century. his was my older brother's inluence. He also taught me to love Dante, whose Divine Comedy we read in an excellent translation by Mikhail Lozinsky. Russian classics, as a matter of course, the preferred ones being Pushkin, Gogol, Dostoevsky, and Leskov. English classics—Shakespeare, Dickens, Jane Austen. I love them and reread them regularly. At the age of sixteen, I discovered Russian twentieth-century poetry—Gumilev, Mandelstam, Akhmatova, Tsvetayeva, and many others. Mandelstam's poetry still moves me as if I'm reading it for the first time. So does Akhmatova's. Approximately at the same time, in the sixties, a lot of translations of Western writers became available in Russia—Kaka, Faulkner, Melville, Salinger. I later reread most of the English ones in the original. When already living in the U.S., I discovered Flannery O'Connor. I think she is a great writer, and her letters published by Sally Fitzgerald under the title he Habit of Being make up one of my favorite books. I enjoy tremendously everything written by Isak Dinesen. At some point after I emigrated, I read the memoirs of Nadezhda Mandelstam, the widow of the poet Osip Mandelstam. They had been translated by Max Hayward and published in English

in two volumes, Hope Against Hope and Hope Abandoned. If anyone wants to understand what the Soviet regime was like, they should read these books. I read everything written by the recently deceased Russian scholar, thinker, poet, and translator Sergei Averintsev. But I do not want to discuss living writers. I love some of Dorothy Sayers's murder mysteries. But she did not write many, and those I prefer I know almost by heart.

When I was ten years old, my mother went to the lea market and bought a very beautiful old edition of the Bible. It contained the books of the Old and New Testament and was abundantly illustrated by Gustave Doré. My agnostic mother's thinking was that an educated person should have read the Bible. hat it was impossible to understand our world, its art and literature, without a knowledge of the Bible. his book impressed me very much, first as an object, and then, as I began to read it, by what it says. I have been reading it ever since, although the beautiful edition had to be let in Russia when we emigrated.

I knew I would forget someone very important, and I did. When Mikhail Bulgakov's The Master and Margarita was published in the late sixties, we all lost our minds over it. We did not know one could write like that in Soviet Russia. It was published in a magazine in two big installments, and somehow everybody managed to read it overnight. We knew it almost by heart and talked in quotations. Many think now that the book has been overrated, but I disagree and consider it one of the best Russian books of the twentieth century.

INTERVIEWER

And you subsequently translated it.

PEVEAR

Yes. It was our second translation for Penguin. We got a very nasty review in the Times Literary Supplement.

INTERVIEWER

What was the criticism?

PEVEAR

Rather blunt—that ours was the worst translation of the book. he reviewer is a scholar of Bulgakov. She studied with the man who did the first translation and

finds his version still the best.

VOLOKHONSKY

Maybe she sincerely hated ours. She quoted only one sentence, one that we were particularly proud of, because it's a very important sentence. It marks the change from one part of the book to another, from modern Moscow to biblical Jerusalem. And the style changes abruptly, becomes almost rhythmical prose, epic. here is this very spectacular entry of Pontius Pilate in the novel, and it was very important to end the phrase with the words "Pontius Pilate." Because the phrase gets repeated three times in the novel and also it's the last words of the novel. So we felt it was very, very important.

PEVEAR

It's a beautifully suspended sentence. "In a white cloak with blood-red lining, with the shuffling gait of a cavalryman, early in the morning of the fourteenth day of the spring month of Nisan, there came out to the covered colonnade between the two wings of the palace of Herod the Great the procurator of Judea, Pontius Pilate."

VOLOKHONSKY

Our critic thought it was all wrong. Because there's an inversion. People say it's archaic. We think that sometimes inversions are very expressive and necessary.

PEVEAR

When we were translating Anna Karenina, Penguin had an in-house reader who said, I want my Tolstoy to read smoothly.

VOLOKHONSKY

To be smooth and reader friendly.

PEVEAR

And I said, Smooth translations slide smoothly into oblivion.

INTERVIEWER

I read you had trouble with the editing of the British Penguin edition of Anna

Karenina.

VOLOKHONSKY

They hated what we did.

PEVEAR

It was quite something. For example, Kitty meets Levin at the skating rink. She asks him, “Did you come recently?” And the copy editor wrote a comment which said, I’m not sure if you’re aware of it, but this word has now acquired different meanings. And there is a better example! Kitty is discussing the upcoming ball. Seventeen--year-old, completely innocent Kitty says, “I do like balls.” Again the copy editor wrote, I’m not sure if you’re aware... hen the editor had this other problem. I had written that Anna “got into the carriage.” And the editor said, his is the American usage of the word got. We can’t do this in a British edition. You should say Anna went into the carriage. I wrote back, I’m not sure if you’re aware of it, but this word has now acquired different meanings . . .

VOLOKHONSKY

They blue-penciled everything.

PEVEAR

Even the preface. here’s a famous story about a little fragment by Pushkin that got Tolstoy started on Anna Karenina. He read this little fragment and he loved it so much. I quoted the beginning of it in my preface—“he guests arrived at the summer house.” One of their inhouse readers commented, I’m not sure if it’s “arrived” or “had arrived.” Perhaps it should be “were arriving.” I told them it was impossible to work this way. It’ll take thirty years! hey finally gave up. They even agreed to pay us a hundred percent of the royalties for the paperback because they were so sure it wouldn’t sell.

INTERVIEWER

What happened?

PEVEAR

The British hardcover edition sold fewer than six hundred copies, but the American

hardcover sold twenty thousand copies in the first three months.

VOLOKHONSKY

We won the PEN Translation Prize.

INTERVIEWER

And four years later, Oprah Winfrey propelled you to international stardom by choosing Anna Karenina for her book club.

PEVEAR

I remember it very well. We were at our place in Burgundy, and I answered the phone. It was our American editor. She said, I want to ask you a question. Does the name Oprah Winfrey mean anything to you? And I said I thought she might be a countryandwestern singer. And she laughed her head of.

VOLOKHONSKY

Because we let New York in the late eighties, and we never had any televi sion any-way.

PEVEAR

I said, What happens as a result? She said they had just shipped nine hun dred thousand copies.

VOLOKHONSKY

I overheard the number. I thought they were talking about some best seller. I shouted from the other room, Nine hundred thousand copies of what? Richard said, Of Anna Karenina! I thought it was some kind of joke.

INTERVIEWER

How did that change your life?

VOLOKHONSKY

My mother always told me not to discuss money.

PEVEAR

So did mine.

INTERVIEWER

So tell me, Dostoevsky or Tolstoy?

VOLOKHONSKY

I always preferred Dostoevsky. When I was young, Dostoevsky's world, his heroes, interested me more. But now I have changed. I stopped comparing. They're completely different writers. I have enormous admiration for Tolstoy as a writer. Working with Tolstoy's text, I feel that I'm in good hands.

INTERVIEWER

What do you mean?

VOLOKHONSKY

I don't know. Richard, help me.

PEVEAR

There is a whole, very solid world in Tolstoy.

VOLOKHONSKY

There is death. here is suffering. But there is also stability. You belong.

PEVEAR

You're at home.

VOLOKHONSKY

You enter it and you live in it. And of course, he belongs to the nineteenth century. He doesn't, like Dostoevsky, make these incredible leaps into modernity. And yet, with all this nineteenth-century quality of his prose, these characters and

their problems and their lives and their relations, they are also ours. You recognize them. You recognize them as your friends, your neighbors, your aunt. But there is a strange misconception—the widespread opinion is that Dostoevsky wrote badly and Tolstoy wrote well.

PEVEAR

Good Russian prose, that's what they say, that Tolstoy wrote good Russian prose.

VOLOKHONSKY

True, he wrote good Russian prose. But not in the sense it's meant. Tolstoy was very aware of the means of Russian language. Rhetorical means. He used chiasmus, parallel constructions, repetition. He wrote long, long sentences. In one passage, Anna is seen through the eyes of Kitty. Seven times in one scene, Tolstoy uses the word прелестна to describe Anna. It can mean "lovely," "charming," "enchanting." But we had to choose. In Russian the word has a Slavonic etymology, which has the spiritual meaning of a seduction, of a magical spell. And it's repeated seven times, very rhetorically. "Enchanting her firm neck with its string of pearls, enchanting her curly hair in disarray." So Anna almost becomes like a witch.

PEVEAR

He liked to attach certain epithets to certain people.

VOLOKHONSKY

Anna has a firm neck. Vronsky has even teeth, a wall of teeth. You don't see the separate teeth. It's very hard to translate. He's very precise describing people. For instance, in the very beginning, Prince Oblonsky is in disgrace, his wife threw him out. He is in his study, on his leather sofa. He wakes up, he's forgotten everything, and it all comes back to him. He puts on his dressing gown and throws a knot in it. hen he takes a deep breath, or inhales deeply. And that is how other translators have translated it. What Tolstoy actually says is, "Drawing a goodly amount of air into the broad box of his chest."

INTERVIEWER

So when you read it in Russian, it's striking.

VOLOKHONSKY

It is completely striking. It's unusual.

PEVEAR

In another instance, introducing a minor character named Madame Stahl, he writes that Kitty pointed to "a bathchair in which something lay, dressed in something grey and blue."

VOLOKHONSKY

We immediately see that Tolstoy is completely contemptuous of this Madame Stahl, who is some kind of whining invalid.

PEVEAR

This is one difference between Dostoevsky and Tolstoy. In Dostoevsky, the narrator would be a person, a character. In Tolstoy, it's Tolstoy. And it's his emotion, his con-

tempt. He inhabits these people, even the worst of them, and they become alive.

VOLOKHONSKY

These characters come out of the page at you. They are outside the frame.

INTERVIEWER

Tolstoy famously said that “War and Peace is not a novel, still less an historical chronicle, but what the author wanted and was able to express in the form in which it is expressed.” What do you think he meant?

PEVEAR

What on earth does he mean? But I think he’s right. It isn’t a novel in the ordinary sense. He breaks every possible rule. He comments on his own characters, he digresses, he philosophizes. There’s his polemical view of the history of the Napoleonic Wars. And then there are the lives of these people who have nothing to do with any of that. They’re riding through the snow on the sleigh. You can hear the runners running and the horses galloping.

 he night when Petia Rostov is about to be killed, it’s dark and there’s no visual image. It’s all aural. Tolstoy says, “Капли капали.” Simply that— “Drops dripped.” We had a discussion about that with a well-known French scholar and translator. We asked him, Could you do that in French? And he said, No, you can’t do it in French. We were lucky, because it works perfectly in English just as it does in Russian.

INTERVIEWER

Translator’s luck! Why did you translate Tolstoy’s essay What Is Art? first?

PEVEAR

We were asked. his was when we ran into a wall with Random House. We finished the three Dostoevsky books in our contract. I sent our editor a list of suggestions for further translations, and she wrote back and said, I don’t ind anything of interest. he list included Anna Karenina and Gogol!

VOLOKHONSKY

We panicked. Because by then we'd already acquired a taste for it.

PEVEAR

So I wrote another publisher, Penguin. And they suggested What Is Art? and we accepted. In fact, it's a polemic against art. Boneheaded. But Tolstoy was very good at being boneheaded, and he loved doing it. Tolstoy writes about art that he never saw. He sent one of his daughters to Europe. She would look at artworks and describe them to him, and then he would denounce them.

VOLOKHONSKY

Tolstoy loved music. He wept when he heard Beethoven. But then he would say, All this Mozart and Beethoven, it's elitist! He was all for folk art.

PEVEAR

He would gather his peasants and make them sing. There's a beautiful symbolist painting by Odilon Redon in the British Museum entitled he Golden Cell, with a woman's profile in blue. Tolstoy's daughter had described it to him, and he had duly denounced it. I was very pleased when I got Penguin to use it as the book's cover!

INTERVIEWER

I notice you always have a collaborator, Richard Nelson, when you translate plays. Why is that?

PEVEAR

It's a very deliberate thing. I read an article by the French translator André Markowicz about his first translation for the theafter. He was commissioned by the head of the ComédieFrançaise to translate he Inspector, by Gogol, and he was very excited. He finished it and the director said it was very good but started asking him a theafter director's questions. And Markowicz real ized he had never thought of it from that point of view. He redid the whole thing, working with the director. When I read that, I decided that if I ever did a play, I would want to do it with a playwright. I'd want it to be for the stage, not just to read.

VOLOKHONSKY

Therefore, each time I raised the question of translating plays, because I like trans-

lating dialogue—it's just fun—Richard said no. Always a firm Anglo Saxon no. hen some five years ago—I still remember the moment—Richard said, Your prayers have been answered! He read me a letter from the play wright Richard Nelson asking if we would be interested in collaborating with him on translating Russian plays.

INTERVIEWER

So what is the process with him?

VOLOKHONSKY

We do our version, and he then scribbles all over it, and we discuss it.

PEVEAR

And we do it face-to-face, not by mail. He'll say, My actors won't be able to say that.

VOLOKHONSKY

He always pulls it in the direction of practical staging—"My actors won't know what it means"—and he's right. But in the end, we work it out. For instance, in Chekhov's Cherry Orchard, there is the word недотёпа. It is repeated four or five times, and it is the last word of the play. It means someone who bungles things, someone who is not quite a failure but who stumbles. When I did the research, I discovered Chekhov was the first to use it, then it became an ordinary Russian word. So how to translate it? "Nincompoop"? We weren't terribly pleased with that. It's a silly word. hen we looked at other translations. Someone translated it as "half-baked bungler," which is too much. We settled for "good-for-nothing."

PEVEAR

The most simple, plain, ordinary...

VOLOKHONSKY

Banal, totally uninteresting, and not quite right. And we lived with it for a year, and then a few months ago, I said no. Good-for-nothing, I won't have it. We went to our good old Roget's hesaurus and came up with...

PEVEAR

Blunderhead!

INTERVIEWER

Speaking of Roget's hesaurus, what are your other reference tools?

PEVEAR

he Oxford English Dictionary, which, in addition to thorough definitions, tells me when a word entered the English language, an old Russian–English dictionary and a newer one. Eric Partridge's dictionary of word origins is sometimes useful.

VOLOKHONSKY

I use a Russian–English dictionary edited by O.S. Akhmanova. To verify the meanings of Russian words, I also use the tolkovyi slovar of V. Dal and other Russian dictionaries. For the terms of clothing, food, coats of dogs and horses, details of harness, various hunting and card-playing terms, there are now all sorts of sites on the Internet that can be consulted. We occasionally consult friends, scholars, and simple mortals about the use of a word or a phrase. Here we discover that people rarely want to commit themselves or else feel the same hesitation we do. In any case, it helps to simply discuss a problem with a disinterested person.

INTERVIEWER

You have taught translation. How do you teach it?

PEVEAR

I didn't give the students passages to work on. I'd have them choose a project and hand out samples from their work on it. Then we'd go over them in class. Translation is a very broad and loose subject. It has no real rules or principles. It all depends on what the translator wants to do. Louis Zukofsky "translated" Catullus into English words that mimicked the sounds of the Latin, often regardless of sense and grammar, so that

Furi et Aureli, comites Catulli,
sive in extremos penetrabit Indos,
litus ut longe resonante Eoa tunditur unda

becomes

Furius, Aurelius: comities—Catullus.
If he penetrate most remote India,
lit as with the long resonant coast East's wave thundering under—

Marvelous things! And they do make poetic sense after all—because Zukofsky was a true poet, of course. I would try to get my students to see the rich possibilities and, at the same time, to respect the original, as Zukofsky obviously does. To move between languages, rather than from one to another.

INTERVIEWER

Richard, you once said you had never been to Russia and you weren't curious to see it. Is that really true?

VOLOKHONSKY

You did say something very snooty.

PEVEAR

It was something very stupid. And of course they printed it. Because the implication was you had to go to Russia in order to translate Russian, and I wanted to make the point that they're different things. I've never been to Spain, but I've translated a lot of Spanish.

INTERVIEWER

When you went to Russia, did you recognize it from what you'd been translating?

PEVEAR

Yes, especially Dostoevsky's Petersburg. We went to the Haymarket, to the bridge where Raskolnikov stood contemplating, to the canals...

INTERVIEWER

Was it hard to go back to what you knew as Leningrad, Larissa?

VOLOKHONSKY

When I emigrated, I said to myself, I'm not going to live in a Russian ghetto. You know what happens when you turn back.

PEVEAR

You turn to salt.

VOLOKHONSKY

Yes. Anyway, I forbade myself even to be nostalgic. But when I went back after twenty-seven years, I almost fell apart. I stood on the Troitsky Bridge and wept hot tears. I realized how I missed the city. It's a beautiful city with its own life, but it's changed now. It is like in "Rip Van Winkle"— you can't return to your past. We only went three times to Russia. We went to Gogol's birthplace in Ukraine, which was paradise, and Tolstoy's estate, Yasnaya Polyana.

INTERVIEWER

How do you explain the extraordinary richness and variety of nineteenth century Russian literature?

PEVEAR

The unusual thing about Russia is that it reached cultural maturity in the nineteenth century. Russia didn't have the Middle Ages of Dante and Chaucer, the Renaissance of the Italians, or the Elizabethan age of the British. They weren't even sure what language to write in. Pushkin more or less created the Russian literary language, and Pushkin was born in 1799. They were doing for the first time what other cultures had been doing for hundreds of years.

INTERVIEWER

What about the idea of the so-called Russian soul?

VOLOKHONSKY

he Russian soul is a myth. Some people believe in it, and some don't. It is very difficult to describe what it is. he words mysterious Russian soul don't mean anything. It is nothing but a convention. We don't know where our own personal soul is, still less where a nation's soul is. It is an idea that is pleasing to nationalists, I guess. As translators, we don't live with these ideas. We live with words.

INTERVIEWER

You once said that one of your subliminal aims as a translator was “to help energize English itself.” Can you explain what you mean?

PEVEAR

It seemed to me that American fiction had become very bland and mostly self--centered. I thought it needed to break out of that. One thing I love about translating is the possibility it gives me to do things that you might not ordinarily do in English. I think it’s a very important part of translating. he good effect of translating is this crosspollination of languages. Sometimes we get criticized—this is too literal, this is a Russianism—but I don’t mind that. Let’s have a little Russianism. Let’s use things like inversions. Why should they be eliminated? I guess if you’re a contemporary writer, you’re not supposed to do it, but as a translator I can. I love this freedom of movement between the two languages. I think it’s the most important thing for me—that it should enrich my language, the English language. And I hope that this is the positive effect of our work generally.

INTERVIEWER

Lydia Davis wrote that “we must get to know our own language even better when we are translating.”

VOLOKHONSKY

I like finding out etymologies of words that have rarely needed translation. I love words. I love my dictionaries. I literally love what I am doing. And there is this feeling of having a mission. I cherish the thought that we supply something that has not been done. Some of these writers were very well translated. But with Leskov particularly, and with Dostoevsky, I was so happy to have done this work. his sense that it was the right thing to do lasted, it stayed with us, no matter what people say about our translations.

INTERVIEWER

You are a flaming patriot.

VOLOKHONSKY

Thank you.

INTERVIEWER

One last question, which was once asked of Robert Fitzgerald. What are the peculiar satisfactions of translation?

PEVEAR

I could answer that by paraphrasing what Gregory Rabassa said in the first Translation Review interview in the seventies. Translators are the only ones who live in this place between languages, both as reader and writer. You move back and forth from being one to being the other, and you never stop. No one else is quite in that same ... whatever it is ... hiatus.