Lexical Plasticity in Early Bilinguals Does Not Alter Phoneme Categories: I. Neurodynamical Modeling

J. P. Larsson[1], Fátima Vera Constán[2], Núria Sebastián-Gallés[2], and Gustavo Deco[1,3]

Abstract

■ Sebastián-Gallés et al. [The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language, 52,* 240–255, 2005] contrasted highly proficient early Spanish–Catalan and Catalan–Spanish bilinguals, using Catalan materials in a lexical decision task (LDT). They constructed two types of experimental pseudowords, substituting Catalan phoneme /e/ for Catalan /ɛ/, or vice versa. Catalan-dominant bilinguals showed a performance asymmetry across experimental conditions, making more mistakes for /ɛ/→/e/ changes, than for /e/→/ɛ/ ones. This was considered evidence of a developed acceptance of mispronounced Catalan /ɛ/-words, caused by exposure to a bilingual environment where mispronunciations by Spanish-dominant bilinguals using their /e/-category abound. Although this indicated modified or added lexical representations, an open issue is whether such lexical information also modifies phoneme categories. We address this using a biophysically realistic neurodynamic model, describing neural activity at the synaptic and spiking levels. We construct a network of pools of neurons, representing phonemic and lexical processing. Carefully analyzing the dependency of network dynamics on connection strengths, by first exploring parameter space under steady-state assumptions (mean-field scans), then running spiking simulations, we investigate the neural substrate role in a representative LDT. We also simulate a phoneme discrimination task to address whether lexical changes affect the phonemic level. We find that the same network configuration which displays asymmetry in the LDT shows equal performance discriminating the two modeled phonemes. Thus, we predicted that the Catalan-dominant bilinguals do not alter their phoneme categories, although showing signs of having stored a new word variation in the lexicon. To explore this prediction, a syllable discrimination task involving the /e/–/ɛ/ contrast was set up, using Catalan-dominants displaying performance asymmetry in a repetition of the original LDT. Discrimination task results support the prediction, showing that these subjects discriminate both categories equally well. We conclude that subjects often exposed to dialectal word variations can store these in their lexicons, without altering their phoneme representations. ■

INTRODUCTION

What is the architecture of the speech processing system? How do the different types of lexical and sublexical information interact? For many years it has been accepted that the way the speech perception system deals with the problem of variability is that of ''filtering it out,'' thus representing information (phonemes, syllables, words) in an abstract code. One problem with such an approach is that of how the linguistic representations can be modified once they are established. Indeed, everyday evidence shows that, for instance, some time after having moved to a new region, where a different dialect is spoken, human beings modify their language production and their speech becomes accented with the new sounds. If the speech perception system filters out all the variability not conforming to the stored abstract representations, how can this occur?

Furthermore, in the last years, a series of studies have shown that the way speech information is stored in the brain is highly dynamic. In a seminal paper, Norris, McQueen, and Cutler (2003) showed that listeners can calibrate their phoneme representations to the particular characteristics of the speech they hear in a surprisingly short interval. In that study, participants were exposed either to some properly pronounced words (containing either /s/ or /f/) or to words with ambiguous sounds (e.g., halfway between /s/ and /f/). The results showed that listeners shifted their phonetic boundaries to accommodate these ''variations.'' The results of that study have been replicated several times in a number of studies. Interestingly, some of the studies indicate that these changes are not transient in nature, but relatively long-lasting. After 25-min postexposure periods, containing either silence or different types of intervening auditory stimuli, Kraljic and Samuel (2005) observed numerically larger shifts than those measured immediately after the main stimuli were presented. Eisner and McQueen (2006) have reported convergent results with

[1]Universitat Pompeu Fabra, Barcelona, Spain, [2]Universitat de Barcelona, Barcelona, Spain, [3]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

12-hr intervals (even after sleep). All these results indicate that exposure to variability may induce changes in the speech perception system that remain stable over time and despite new exposures. The changes studied in these investigations have mostly consisted of the modification of speech categories (phonemes) induced by exposure to nonstandard utterances. However, an open question is to which extent the lexicon is modified and whether variation is also represented at this level.

As suggested, one way of studying this issue is analyzing word recognition in listeners exposed to dialectal variation. Most dialects are characterized by changes at the phonological (and, specifically, phonetic) level. The literature about adaptation of the speech perception system to dialectal variability has not been focused on how dialectal variability is encoded in the lexicon, but mostly on more broad aspects of perception of accented speech (see, for instance, Clopper & Pisoni, 2005). Two studies are worth mentioning here, both addressing the core issue of how different pronunciations of the same words are stored in the mental lexicon. In the first one, Connine (2004) investigated the recognition of a frequently heard spoken word form variant in American English (flapping, i.e., when the tongue makes a flap against the palate to produce the sound replacing the full /t/ sound in words such as the example ''pretty'' below). She presented words and pseudowords including either the standard /t/ or the flap sound (e.g., /pridi/–/bridi/ or /priti/–/briti/; English word: pretty). Participants were asked to identify the initial segment (/b/ or /p/). The results showed better performance in identifying the target sound when it occurred in the more frequently heard flap carrier. The author considered that the results showed that listeners stored in their mental lexicon the flap variant of the words together with the standard form. The second study where evidence about the representation of dialectal variation of spoken words was obtained was that of Sebastián-Gallés, Echeverría, and Bosch (2005) (later replicated in Sebastián-Gallés, Rodríguez-Fornells, Diego-Balaguer, & Diaz, 2006). We will describe that work in some detail because it constitutes the basis of the present work. Before that, however, a few words concerning the linguistic properties of Catalan and Spanish as well as the sociolinguistic situation of the participants of that study are in order.

Catalan and Spanish are two Romance languages, co-official in Catalonia (northeast of Spain, Barcelona being the largest city). Although both languages are similar at different levels of linguistic description, critically they differ at the level of the phoneme repertoire. For instance, although Spanish has no voiced fricatives, Catalan does. Central to the present study, Catalan has eight vowels, and Spanish only five. In particular, Catalan has two mid-front vowels (/e/ and /ɛ/), whereas Spanish has only one, falling roughly midway between the two Catalan ones. This phoneme distribution is very similar to the one of the English liquid contrast /r/–/l/ and the Japanese liquid /l/. Like Japanese natives learning English with the /r/–/l/ contrast, Spanish natives have great difficulties in perceiving (and producing) the Catalan-specific /e/–/ɛ/ contrast.

Because Spanish and Catalan are co-official in Catalonia, they are both widely used, in formal as well as in informal contexts. According to official statistics (www.idescat.net), about 95% of the population living in Catalonia understand Catalan and about 75% declare they can speak it. In the Barcelona area, more than 40% of the people declare that Catalan is their ''usual'' language of use (''llengua habitual''). Illustratively, at the University of Barcelona, more than 60% of the courses are taught in Catalan and students are free to use any of the languages to write their essays and exams. In fact, students entering Catalan universities must show a written and oral command of both Spanish and Catalan equivalent to that which students elsewhere in Spain must show for Spanish. Thus, both Catalan and Spanish are frequently used. Still, as a series of studies have shown (Sebastián-Gallés et al., 2006; Pallier, Colomé, & Sebastián-Gallés, 2001; Bosch, Costa, & Sebastián-Gallés, 2000; Sebastián-Gallés & Soto-Faraco, 1999; Pallier, Bosch, & Sebastián-Gallés, 1997), most individuals (even those raised in Catalonia) born within monolingual Spanish families fail to perceive (and produce) many Catalan-specific sounds; in particular, the Catalan /e/–/ɛ/ contrast. Instead, they use the Spanish five-vowel system when speaking in Catalan, thus producing Catalan with a Spanish accent. Because of the bilingual nature of Catalan society, it is very common to hear both native and accented Catalan. That way, listeners are frequently exposed to different varieties of the same words. For instance, the Catalan word ''galleda'' (pronounced /gallɛda/[1]) is pronounced with the Spanish phonemes /galleda/. In the same way as American natives consider the liquid sound produced by Japanese speakers as an /l/, Catalan natives assimilate the Spanish mid-front /e/ vowel to the Catalan /e/. Thus, for words such as ''galleda,'' they hear two different acoustic varieties, the one produced by Catalan natives, including vowel /ɛ/, and another produced by Spanish natives, including vowel /e/. The exposure to different forms does not occur for Catalan words containing vowel /e/, like ''finestra'' (pronounced /finestra/), as both Catalan and Spanish natives produce words that to Catalan natives sound equivalent.

In the most relevant study preceding this work (Sebastián-Gallés et al., 2005), Catalan participants were asked to do an auditory lexical decision task (LDT). Pseudowords were created by changing one vowel of real words. There were three types of pseudowords: Control pseudowords involved a vowel change existing in Spanish (for instance, the Catalan word ''cadira'' was transformed into /cadura/, /i/–/u/ being a vowel contrast common to both Spanish and Catalan). /ɛ/-type pseudowords consisted of pseudowords created from Catalan words including vowel /ɛ/, which was replaced by vowel /e/ (like

the ''galleda'' example just described). The /e/-type pseudowords were created from Catalan words including vowel /e/, being replaced by vowel /ɛ/ (like the ''finestra'' example just described). As can be seen, /ɛ/-type pseudowords corresponded in fact to the way Spanish natives mispronounce some Catalan words. The results of the experiment showed that Catalan natives had great difficulties in rejecting the /ɛ/-type pseudowords as such (the percentage of correct responses in this category was less than 60%, whereas for the other two pseudoword categories they were close to 90%; see Table 1). Therefore, an asymmetry was observed in the percentage of correct rejections of both types of experimental pseudowords. The authors interpreted these results by postulating that the repeated exposure to the dialect that Spanish natives use when speaking in Catalan created multiple representations for /ɛ/-type words, for instance, in the ''galleda'' case, they had stored in their lexicon both forms, /gallɛda/ and /galleda/. Thus, when confronted with the ''pseudoword'' /galleda/, they often failed to reject it as a pseudoword. However, this would not happen for /e/-type stimuli. As we have stated, Spanish natives never mispronounce these items, for which reason Catalan natives would never have stored /finɛstra/ in their lexicon. These results were considered to indicate that participants held in their mental lexicon both the correct (native) pronunciation of words and the incorrect (nonnative accented) ones.

The present work aims to elucidate possible mechanisms for and consequences of the asymmetry in the performance of Catalan natives, observed by Sebastián-Gallés et al. (2005), exploring the implications of such mechanisms for the architecture of models of spoken word recognition. To this end, a modeling approach is proposed, using a biologically plausible mathematical formulation of the functioning of single neurons as a foundation and then constructing a network of these neurons, organized in functional populations (henceforth, pools), whose interactions can account for the behavior observed. We propose two competing network architectures, each representing a specific cognitive hypothesis. One model supposes an interaction at the phonemic level as the main mechanism behind the observed results, whereas the other model explores whether the context of a complete word is necessary for the asymmetry to be accounted for, in which case it would rather be a lexical effect. The pools in a given model are interconnected and each represents one or more phonemes. The weights of certain connections between pools are the parameters of these models, whose values will determine the behavior of each model. The two models differ in terms of which of all possible interpool connection weights are considered parameters. We present these models with stimuli and quantify their responses using neuronal firing rates. Upon the presentation of a stimulus, the persistent coactivation of pools, which together represent a full word, is taken to be word recognition. Thus, comparing the modeled responses with the behavior of real experimental subjects, we are able to identify the best model. To this end, a careful analysis of the network is performed to establish the appropriate parameter values. To mitigate the computational burden of parameter space exploration, we first use a steady-state mean-field approximation to the full spiking neuron model, to perform so-called mean-field scans. Thereby, we identify plausible parameter values, for which we then simulate the full models over multiple trials, obtaining time-dependent neuronal activity, which is then compared to real subjects' performance. After simulating the LDT and finding the best model, we make a final modeling effort, where we explore the behavior of our network in response to a single-phoneme input, simulating a discrimination task, in order to further clarify which of the proposed models is more appropriate.

In these models, a particular parameter configuration will be considered to be the result of plasticity in the brain, having occurred on the long-term due to exposure to ambient language. Such plasticity is commonly considered to be the result of Hebbian learning (Hebb, 1949), in which the connections between coactivated neurons, such as those receiving the same stimulus, strengthen. Similarly, lack of common input can lead to the weakening of connections between neurons, the so-called anti-Hebbian learning.

It is stressed here that this approach makes explicit use of the neurodynamics underlying any operation carried out by the human cortex, and the stochastic nature of the fluctuations inherent in the cortex as well as in the model makes a decisive contribution to the solution of the problem. Specifically, with a proper description of

Table 1. Different Types of Stimuli, Native and Dialectal Pronunciations, and Percentage of Correct Responses to Pseudoword Stimuli Used in Sebastián-Gallés et al. (2005)

Stimulus Type	*English Translation*	*Word*	*Pseudoword*	*Native Pronunciation*	*Dialectal Pronunciation*	*% Correct Pseudoword Responses*
/ɛ/-type	Bucket	Gallɛda	Galleda	/ɛ/	/e/	58
/e/-type	Window	Finestra	Finestra	/e/	/e/	89
Control	Chair	Cadira	Cadura	/i/	/i/	90

the nonstationary transients, using a spiking neuron model and realistic synaptic dynamics, one can account for response times in psychophysical experiments, as well as make use of fluctuations to get a probabilistic response behavior (basis for decision making). As models of this kind have already been successfully employed modeling working memory (e.g., Brunel & Wang, 2001) and its role in the discrimination between two sequential sensory stimuli (Deco & Rolls, 2006; Machens, Romo, & Brody, 2005; Miller & Wang, 2002) as well as in other decision-making tasks (Wong & Wang, 2006; Wang, 2002), we believe they can make an important contribution to elucidating the workings of language processing as well.

Although there exist models of phoneme perception, notably Norris, McQueen, and Cutler (2000) and McClelland and Elman (1986), to our knowledge, an approach with such a solid basis in neuroscience as ours (time course and probabilistic behavior) has not been previously applied to this kind of problem. Although it is acknowledged that other approaches include input with a closer correspondence to real sequential sound streams, this is not of crucial importance, as we see this work as a proof-of-principle of the applicability of such more neurodynamical modeling schemes to behavioral data from psychology.

In addition to the modeling, with the aim of exploring the prediction of perfect phoneme discrimination based on the model results, we designed and conducted a psychophysical experiment to investigate whether phoneme categories are modified in the Catalan-dominant early bilinguals, or whether the effect is purely lexical, as predicted if, indeed, phoneme discrimination is perfect in these individuals.

METHODS

The model is based on the single-neuron level of description, using *integrate-and-fire* neurons (cf. The Integrate-and-Fire Model section). This allows for the use of realistic biophysical constants obtained in neurophysiological studies, (e.g., synaptic conductances and delays), which enables the study of time scales and neural firing rates involved in shaping the neural activity underlying cognitive processes. Benefits of using such a detailed model include, as emphasized in the Introduction, the possibility of studying the time course of the neural activity, including transient responses, and using the model's inherent stochasticity to explain variability in experimental data. Specifically, the former can be used to obtain reaction times from the simulation data, whereas the latter can be used to obtain bistable response patterns, for instance, corresponding to a certain distribution of responses in data from a binary choice experimental task. Also, obtaining such elaborate model data allows us to compare at a detailed level with neurophysiological data, if available.

Using the said single-neuron model as a building block, we construct networks of neurons, organized in interconnected functional pools. The neurons of such a pool share common input and are thus coactivated. This is assumed to have set the *intrinsic* connections of that pool higher than network average through a Hebbian learning process (Hebb, 1949). In the framework of biased competition and cooperation (e.g., Deco & Rolls, 2004, 2005a, 2005b; Rolls & Deco, 2002), such a pool typically represents a certain aspect of the external input, such as visual or auditory object identities. External input to a network of such pools can bias the internal activity of the network through intrapool cooperation and interpool competition. With a sufficiently strong synaptic connection binding two or more pools, one can also get cooperation between pools. These connection strengths or weights describe relative deviations of the synaptic conductivities from their average value across the network. Stronger intrinsic weights implement reverberation of a pool's neuronal activity, which can underlie the formation of working memory, that is, the appearance of sustained high activity coding for a previously presented stimulus, which is absent at present. This concept of an attractor in the system's dynamics is used to encode presented stimuli, which are then considered to have been recognized by the system.

The modeling scheme using integrate-and-fire neurons grouped in such functional pools has the advantage that it can easily be analyzed using the mean-field approach, which is a theoretical tool that uses the pool organization of the network to calculate approximate solutions of the stationary state of the dynamics. This will be elaborated on in the Simulations and Analysis section.

The Integrate-and-Fire Model

The model used is based on nonlinear leaky integrate-and-fire neurons, which are coupled together to form a network of neurons. This neural network basis is adapted from Brunel and Wang (2001), and we refer to that article for a more detailed description of the neurodynamical properties. In Appendix A, the mathematical model for a single neuron is defined. Essentially, it consists of an equation governing the evolution of the neuron's subthreshold membrane potential. Via synapses, excitatory and inhibitory input currents effect an increase or decrease of the membrane potential, respectively. In the absence of such input, the membrane potential decays exponentially over time. When the membrane potential reaches the threshold potential, an action potential (a spike) is emitted. Spikes in the membrane potential are in the model treated as unitary events of equal magnitude and duration. An emitted spike propagates to connected neurons via explicitly modeled synaptic mechanisms, namely, α-Amino-3-hydroxy-5-Methyl-4-isoxazole Propionic Acid (AMPA), N-Methyl-D-Aspartate (NMDA), and Gamma-AminoButyric Acid A ($GABA_A$) receptors on the postsynaptic neurons.

More specifically, the membrane potential of a single neuron is modeled by an electrical circuit, which consists of a capacitor C_m connected in parallel with a resistor R_m. This circuit describes how the membrane potential $V(t)$ evolves in time depending on external currents entering from other neurons. When the threshold membrane potential V_{thr} is reached, an action potential is emitted and propagates to other neurons, whereas the membrane potential of the neuron that spiked is set to the reset potential V_{reset}, at which it is kept for a refractory period τ_{ref}. Both excitatory and inhibitory neurons have a resting potential $V_L = -70$ mV, a firing threshold $V_{thr} = -50$ mV, and a reset potential $V_{reset} = -55$ mV. The membrane parameters are different for excitatory and inhibitory neurons. The former type of neuron is modeled with membrane capacitance $C_m = 0.5$ nF, leak conductance $g_m = 25$ nS, membrane time constant $\tau_m = 20$ msec, and refractory period $t_{ref} = 2$ msec, whereas the inhibitory neurons have the corresponding values $C_m = 0.2$ nF, $g_m = 20$ nS, $\tau_m = 10$ msec, and $t_{ref} = 1$ msec. Values are taken from McCormick, Connors, Lighthall, and Prince (1985).

Incoming synaptic influences are both excitatory and inhibitory. Excitatory neurons transmit action potentials via the glutamatergic receptors AMPA and NMDA, both modeled with exponential terms. We neglect the rise time of the current mediated by the AMPA channel because it is typically very short ($\ll$1 msec), and just model the decay period with $\tau_{AMPA} = 2$ msec (Spruston, Jonas, & Sakmann, 1995; Hestrin, Sah, & Nicoll, 1990). The NMDA channel is modeled with a rise term, $\tau_{NMDA,rise} = 2$ msec; a decay term, $\tau_{NMDA,decay} = 100$ msec (Spruston et al., 1995; Hestrin et al., 1990); and an extra voltage dependence controlled by the extracellular magnesium concentration, $[Mg^{2+}] = 1$ mM (Jahr & Stevens, 1990). The inhibitory postsynaptic potential is mediated by the $GABA_A$ receptor and is described by a decay term with time constant $\tau_{GABA} = 10$ msec (Xiang, Huguenard, & Prince, 1998; Salin & Prince, 1996). The $GABA_A$ receptors mediate global shunting inhibition, a negative feedback mechanism preventing runaway neuronal activity, which can otherwise arise due to recurrent excitation. The inclusion of the slow, voltage-gated NMDA receptors is necessary for maintaining stability of the persistent activity state of the network when it has reached an attractor (e.g., Wang, 1999). Specifically, if only fast AMPA receptors are used for excitatory synaptic transmission, the interplay with the slower time-constants involved in negative feedback mechanisms can disrupt persistent activity. This effect is countered by the slower positive feedback introduced through the NMDA receptors.

Network Setup and Architecture

Our network contains a total of 2000 neurons, all connected to each other. Of these, $N_E = 1600$ are excitatory and $N_I = 400$ are inhibitory. This proportion of inhibitory to excitatory neurons is consistent with the observed proportion of interneurons to pyramidal neurons in the cerebral cortex (Abeles, 1991; Braitenberg & Schütz, 1991). These neurons are subdivided into six pools of three types. Four of the pools are selective (S) pools, each consisting of 150 excitatory neurons. These pools are activated in association with the input, in this case, one or many phonemes. In addition, there is a nonselective (NS) pool consisting of 1000 excitatory neurons, which are not activated in association with the input. Finally, there is a pool of 400 inhibitory neurons (I pool), which provides global inhibition, thereby helping control network activity. Figure 1 schematically illustrates the basic network structure.

However, before creating this pool structure by differentially setting connection strengths through the modification of synaptic *weights*, all connections' synaptic *conductances* are initially given values, which ensure a spontaneous activity of 3 Hz and 9 Hz for the excitatory and inhibitory neurons, respectively (for this procedure, see Brunel & Wang, 2001). These values approximately correspond to those obtained in neurophysiological studies (Wilson, Scalaidhe, & Goldman-Rakic, 1994; Koch & Fuster, 1989).

Each *S pool* is then characterized by a higher *intrinsic* connection strength, designated by the weight w_+. This is in accordance with the assumption that shared input leads to correlated activity, which in turn strengthens the connections between these neurons according to the concept of Hebbian learning (Hebb, 1949). Here, it is assumed that weights have achieved the constant values they have in a certain simulation through such a Hebbian learning process, possibly over a very large time span as in long-term plasticity.

The connection strength between different S pools, w_-, has a value of less than 1, which on the network level compensates for the effect of the high value of w_+. In specific, $w_- = 1 - f(w_+ - 1)/(1 - f)$, where f is the fraction of all excitatory neurons in each selective pool (i.e., here $f = 150/1600$). This relationship ensures that the overall recurrent excitatory drive in the spontaneous state remains constant as w_+ varies (Brunel & Wang, 2001). This recurrent excitation, mediated by the AMPA and NMDA receptors, is assumed to be dominated by the NMDA current to provide a more robust behavior during the period after stimulus offset.

Connections from the *I pool* (GABAergic connections) to all pools, including itself, are set to 1. The connections from the *NS* pool to itself are set to 1, as its neurons do not share selectivity. The connections from the NS pool to S pool j is given by $w_{nj} = (1 - \sum_{\forall i} f w_{ij})/f_n$, where f is as before, f_n is the fraction of all excitatory neurons in the NS pool, w_{ij} is the weight value from pool i to pool j and the sum goes over all S pools. This setting of the weights ensures that the average input to each S pool is 1.

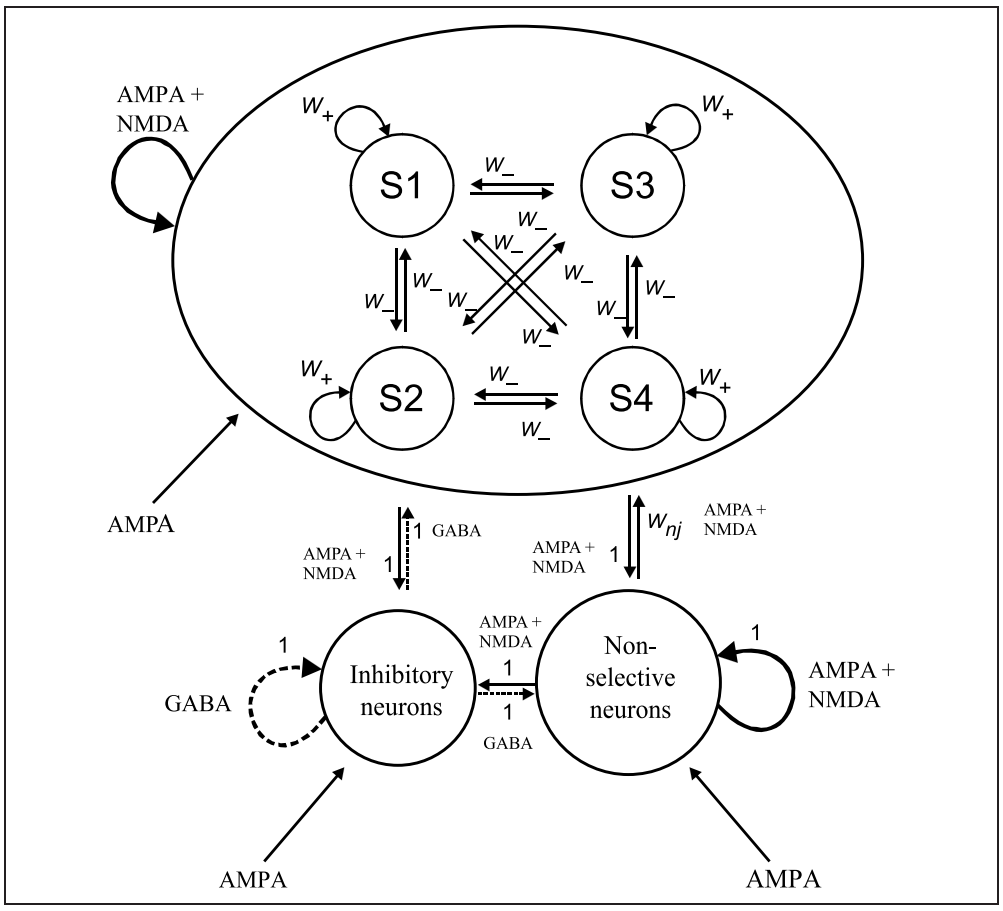

Figure 1. Network architecture. The network has five pools of excitatory neurons, four selective (S1–S4) pools and one nonselective pool, as well as one pool of inhibitory neurons. Dashed lines indicate inhibitory synaptic connections, solid lines indicate excitatory synaptic connections. For each connection, the type of neurotransmitter(s) employed is (are) indicated. See main text for explanation of synaptic weights w_+, w_-, and w_{nj}.

This brings us to the definition of the two specific architectures here considered (Figures 2 and 3).

In both models, there are two types of S pools, phoneme pools and word context pools, which together can represent a word. A phoneme pool simply codes for a specific phoneme, whereas a word context pool can be imagined as a conglomeration of single phoneme pools which together create the context. Alternatively, one could model each of a word's constituent phonemes by its own S pool, but this is not computationally feasible. Thus, we chose the intermediate representation of a context pool, while explicitly modeling the crucial critical contrast phonemes of the stimuli. A specific model architecture is defined by letting certain interpool connections have a weight value exceeding w_-. These values are then the parameters of the model in question, along with w_+.

In Model 1 (Figure 2), parameters w_{pp1} and w_{pp2} interconnect the two phoneme (p) pools, which opens for the possibility of a change having occurred in the representation of phoneme categories. The parameters w_{w1} and w_{w2} each help define a complete word (w), by

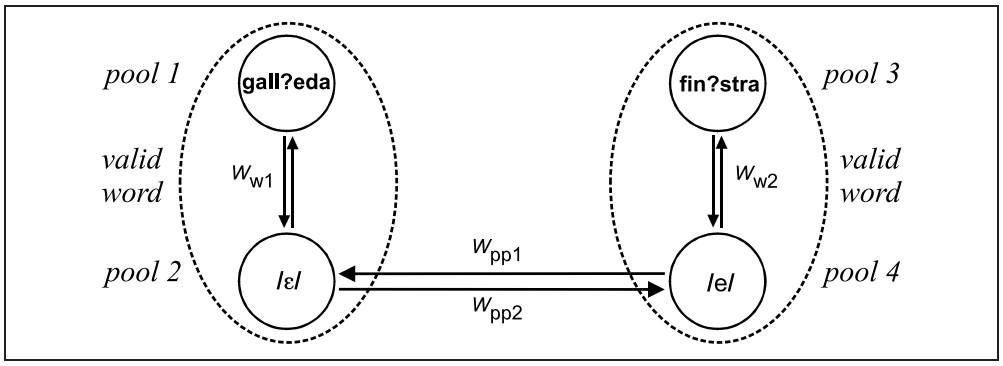

Figure 2. Model 1. The figure shows the parameters which, apart from w_+, are varied in the search for specific phenomena in the network behavior. In the figure, connections with the value w_- are omitted. This model was designed to look for an explanation of the experimental data at the phonemic level, represented by the interconnected Pools 2 and 4. The four pools illustrated here correspond to the selective pools in Figure 1.

binding a context pool with a phoneme pool to enable the representation and recognition of a *valid* (see figure) word, as opposed to a pseudoword. Values w_{w1} and w_{w2} are assigned to their respective connections bidirectionally.

Figure 3 is a schematic drawing of Model 2. Here, w_{w1} and w_{w2} have the same function as in Model 1. However, instead of parameters connecting the phoneme pools, we have the bidirectional w_{pc1} and w_{pc2}, which each interconnect a phoneme (p) pool with a context (c) pool. These connections open for the possibility that a new lexical representation may have evolved over time, thus enabling the recognition of a pseudoword, such as /galleda/ (enabled by w_{pc1}), as a word.

It is stressed here that, although the pools in Figures 2 and 3 are marked by the critical phonemes and their corresponding word contexts, which together constitute the two stimuli which were used as representatives in the discussion of the behavioral experiment we are modeling, the results obtained should be equally valid for other stimuli used in the underlying experimental study. Henceforth, the S pools will be referred to by their respective numbers (see Figures 2 and 3). The networks have no spatial structure or extension.

All neurons in the network are exposed to an external Poisson spike train input through 800 excitatory connections, mediated by AMPA receptors at a rate of ν_{ext} = 3 Hz per synapse, corresponding to a typical value for spontaneous activity in the cerebral cortex (Rolls & Treves, 1998; Wilson et al., 1994). This results in a total external input of 2.4 kHz to each neuron in the network, representing noise input from surrounding areas in the cortex not explicitly modeled.

The network is presented with a stimulus by increasing the external input of the pools corresponding to the stimulus with λ_{stim} = 0.1 Hz, yielding $\nu_{stim} = \nu_{ext} + \lambda_{stim}$ = 3.1 Hz. For example, to present the input /galleda/, this higher external input is applied to Pools 1 and 4, adding a total of 80 Hz to the total external input to each neuron in those pools. Input is always maintained for 500 msec, regardless of the stimulus presented, a value which approximately corresponds to the durations of the four 3-syllable stimuli used as representatives in our discussion, the Catalan words ''galleda'' and ''finestra'' and their corresponding pseudowords.

The stimulus input influences the inherent competition and cooperation between areas in the network and has the capacity to drive the activity of the pools given the input into an attractor. Setting interpool *parameters* higher than the surrounding interpool *connections*, combined with the higher value of w_+ (creating so-called pool cohesion), can create competition between distinct groups of pools, whereas pools constituting a specific group instead display cooperation. An example would be the tendency to coactivate in the group consisting of Pools 3 and 4, representing /finestra/, because of setting w_{w2} higher than w_-.

Simulations and Analysis

We used spiking simulations to analyze the time course of network activity and the influence of fluctuations. Because these are computationally demanding, they cannot be run for all parameter configurations of the network. Here the mean-field approximation is useful for assessing network behavior. It assures that the dynamics of the network will converge to a stationary attractor which is consistent with the asymptotic behavior of an asynchronously firing network of integrate-and-fire neurons (Brunel & Wang, 2001). In the standard mean-field approach, the network is partitioned into populations of neurons which share the same statistical properties of the afferent currents, and fire spikes independently at the same rate. The essence of the mean-field approximation is to simplify the integrate-and-fire equations by replacing, in accordance with the diffusion approximation (Tuckwell, 1988), the sum of the synaptic currents by an average component (constant under steady-state assumptions) and a fluctuation term. The stationary dynamics of each population can be described by a

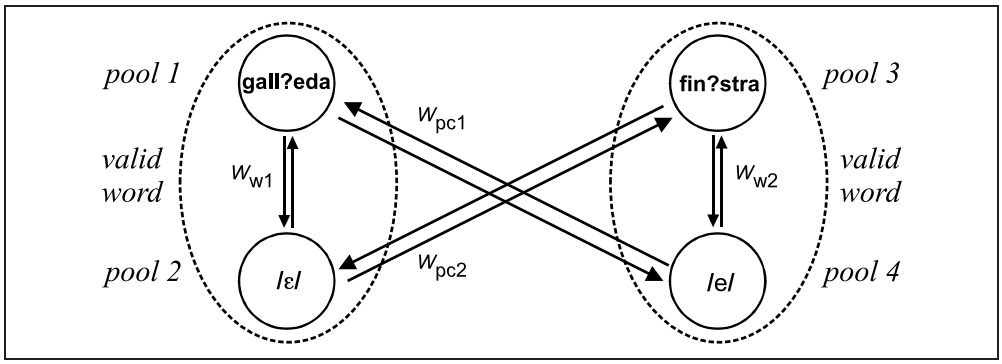

Figure 3. Model 2. The figure shows the parameters which, apart from w_+, are allowed to vary in the search for specific phenomena in the network behavior. In the figure, connections with the value w_- are omitted. This model was designed to look for an explanation of the experimental data at the lexical level, interconnecting a context pool with a different phoneme pool than its normal critical phoneme constituent. The four pools illustrated here correspond to the selective pools in Figure 1.

population transfer function, which provides the average population rate as a function of the average input current (for further details, see Appendix B).

We used mean-field simulations to calculate the attractor states of the network in a large number of points in parameter space, thereby assessing which points displayed the behavior of interest. Although the mean-field attractor states reflect the true behavior of the network, the fluctuations and external changes added in the spiking simulations can significantly alter the behavior. Therefore, and because we need the temporal dimension to compare simulation results with experimental data, spiking simulations were then run for the identified interesting parameter settings.

At a first stage, we ran extensive mean-field scans for both models over their respective five network parameters, namely, the intrinsic pool cohesion parameter, w_+, and the four interpool weights (lateral, word-cohesion weights w_{w1} and w_{w2} for both models, plus w_{pp1} and w_{pp2} for Model 1 or w_{pc1} and w_{pc2} for Model 2). In the mean-field scans, we looked for parameter space points in which the behavior of the network be the following. Upon the presentation of *words*, for example, /gallɛda/ and /finestra/, the average firing rates of their corresponding pools should be above threshold, whereas the other two pools of the network should display below-threshold activity in the mean (a ''yes'' response); upon presentation of unlikely *pseudowords*, for example, /finɛstra/, all pools should show below-threshold firing rates (a ''no'' response); however, upon presentation of *pseudowords* more likely to occur, for instance, /galleda/, the pools receiving input (Pools 1 and 4) should show an above-threshold mean firing rate, indicating that in this point we might find the bistable behavior we are looking for in the network's response to this kind of stimulus. That is, we expected this mean behavior to fluctuate when adding the noise of the full spiking simulations, yielding different results in different trials, just as in the experiments.

Specifically, typical values for the mean-field scans of both models were: $w_+ \in [1.9, 2.5]$, while w_{w1}, w_{w2}, w_{pc1}, w_{pc2}, w_{pp1}, and w_{pp2} typically varied between 1.0 and 1.2. Step sizes varied depending on the interval being searched for each parameter, with the smallest used being 0.01 and the largest 0.1. To keep down computation time, step sizes were only gradually decreased during the search. Important constraints in finding interesting points were the interrelations between weights. Specifically, the value of w_+ should always be greater than all other weights in the network, thus maintaining individual pool reverberatory activity. This constraint only aided in the initial choice of parameter intervals which were scanned for the other parameters. Two further constraints were, however, used to discard points which otherwise displayed wanted behavior in the mean-field simulations. One of them, $w_{w1} < w_{w2}$ reflected the fact that the frequency of occurrence of /gallɛda/ is lower in the environment than ditto for /finestra/, which, through Hebbian learning, should have led to the connection binding the constituent pools for the latter word being stronger than the one for the former word. The second additional constraint was $w_{pc1} > w_{pc2}$, which was added to increase the likelihood of bistability in spiking trials following the presentation of /galleda/, while decreasing the likelihood of /finɛstra/ being recognized as a word. By varying these weights, we predicted that one could modify overall network behavior in these two crucial pseudoword cases.

Spiking simulations start with a 300-msec prestimulus period, time during which the asterisk is shown to participants in the LDT experiment in Sebastián-Gallés et al. (2005). We present the stimulus for 500 msec to the corresponding context and phoneme pools, together constituting the stimulus word or pseudoword, by increasing the external input as described before. Afterward, a period of another 1000 msec is simulated, corresponding approximately to a typical reaction time in the LDT task. This yields a total of 1800 msec for each trial simulated. In this context, a point in parameter space corresponds to an individual participant and a trial corresponds to one trial for that participant. We typically ran 100 trials per point. The analysis of the data thus obtained (a 1800-msec time course of the firing rate of each pool, averaged over its neurons and sampled every 20 msec) was analyzed as follows: First, an average of the last 500 msec simulated was calculated for each trial (leaving time for network activity to stabilize somewhat after stimulus offset). Then, if this average firing rate exceeded a threshold of 10 Hz, the pool was taken to be in an attractor state, meaning it had retained a representation of the stimulus it codes for. Thus, we obtained either attractor states or spontaneous states for all the 100 trials, for each pool in the network. These data are obtained for each stimulus given to the network, for a total of four different such datasets to analyze (one for each word and one for each pseudoword), for *each* point/individual we chose to simulate.

For each point, once we had the outcome of the 100 trials with respect to the threshold, for each of the four stimuli, we checked if the outcome corresponded to a ''yes'' or a ''no'' response, that is, if the stimulus was taken to be a word or not. Thus, for each stimulus, using the 100 different trials, we obtained a percentage of *correct* trials using a binary test based on the definition of a ''yes'' (or ''no'') response. Then, we used this percentage in the comparison of model data with experimental data, either directly or computing the A' statistic for comparison with the corresponding values in the experimental paper. The A' values were calculated using the following formula (Macmillan & Creelman, 1991):

$$A' = 1/2 + \frac{(\text{Hit} - \text{FA})(1 + \text{Hit} - \text{FA})}{4\text{Hit}(1 - \text{FA})}, \text{Hit} \geq \text{FA},$$

where FA is the fraction of false alarms, that is, the response ''yes'' to the presentation of a pseudoword, and Hit is the fraction of correct responses, that is, the response ''yes'' when the stimulus presented was, indeed, a word. For the case when Hit < FA, Hit and FA change places in this equation.

Experimental Methods

In the Sebastián-Gallés et al. (2005) study, participants were only tested with an LDT. In the present work, inspired by our modeling results, the very same participants are tested in a discrimination task and in an LDT. The methods of the LDT were exactly the same as those employed in the study just mentioned.

Lexical Decision Task[2]

Thirty-two native Catalan participants (8 men) took part in this experiment (age average: 22.7 years, *SD* = 9.4). They were students from the University of Barcelona and they were born in Barcelona or its metropolitan area. They participated in exchange for course credits. No auditory or language learning problems were reported. Materials were exactly the same as in the Sebastián-Gallés et al. (2005) experiment. One pair of stimuli was not included in the final analysis because it elicited a very high error rate. As in the previous study, experimental stimuli were words and pseudowords containing either vowel /e/ or vowel /ɛ/. To create pseudowords, both vowels were exchanged (see Table 1). Participants were seated in a soundproof booth in front of a personal computer screen where instructions were displayed. Instructions emphasized that pseudowords could be very similar to real words, and that the participants should pay attention to vowels because pseudowords had been created by replacing a single vowel. It was stressed that, in many cases, the replacement involved the exchange of vowels /ɛ/ and /e/. Responses were made by pressing one of two labeled buttons (''yes'' for words, ''no'' for pseudowords) with their dominant hand. Participants were encouraged to keep their response fingers over the response buttons to respond as quickly as possible. Stimuli were presented in two lists. Each member of each word–pseudoword pair appeared in only one list. Half of the participants were tested with one list, and the other half with the other list. Presentation order was fully randomized for each subject. Reaction times were measured from stimulus onset. Total duration of this experiment was approximately 15 min.

Discrimination Task

Three different female Catalan native speakers recorded several tokens of syllables /de/ and /dɛ/. The most representative and prototypical tokens from each category were selected. There were no statistical differences between the two types of stimuli, neither in length nor in amplitude [average lengths: /e/-type stimuli: 390 msec (*SD* = 49.3); /ɛ/-type stimuli: 412 msec (*SD* = 38.5); average amplitudes: /e/-type stimuli: 69.8 dB; /ɛ/-type stimuli: 69.7 dB].

Participants were tested in individual soundproof booths and they were seated in front of a personal computer screen where instructions were displayed. Stimuli were delivered binaurally through headphones. Participants were asked to press a designated key whenever they detected a change in category, not in voice. Stimuli were presented in a pseudorandom order with at least three tokens of the same category and a maximum of eight tokens before a change was produced. As there were two categories, two changes of direction were possible: from /dɛ/ to /de/ or vice versa (ɛ → e and e → ɛ, respectively). There was a training phase in which participants were asked to do the same task with the same procedure, but only two tokens—one from each category—from one speaker were played. In this phase, visual cues were used to indicate that there were two categories. Also, when a change took place, another cue appeared, indicating that they should press the key. These cues were not presented in the experimental phase, during which only a fixed asterisk was shown as a fixation point. The training phase consisted of 25 trials. The experimental phase consisted of 400 trials, divided into two blocks. Subjects could rest between these two blocks.

After extensive piloting, stimulus onset asynchrony was fixed to 800 msec. The experiment lasted for about 15 min.

RESULTS

The model results are presented in two stages. First, we focus on the main modeling study, the LDT described in the Introduction. Second, we describe the discrimination task modeling and the consequences its results had as to the interpretation of results obtained in the LDT simulations, as well as regarding the implications for model structure and the problem studied.

Looking for the desired behavior of the network using mean-field scans (see Methods) yielded many feasible points in Model 2 in a fairly contiguous parameter space, which were candidates for full spiking simulations. However, in Model 1, only a single parameter configuration was found which displayed this behavior in the mean-field simulations, after *very* extensive scans, indicating the intractability of that model for explaining the experiments, already at an early stage of simulations.

The single feasible parameter configuration found in mean-field simulations for Model 1, shown in Table 2, was used in a spiking simulation of 100 trials.

These spiking simulations did not show satisfactory results. Specifically, this network configuration showed

Table 2. The Only Feasible Parameter Configuration, used in Spiking Trials Run for Model 1

	w_+	w_{w1}	w_{w2}	w_{pp1}	w_{pp2}
Point 1	2.10	1.01	1.02	1.11	1.08

a breakdown in identifying /finɛstra/ as a pseudoword, with an almost chance performance of 54% correct trials. In Figure 4, a spiking trial representative of the 46% erroneous trials is shown, with activity high in Pools 2 and 3, signaling word recognition. These are not the data we wish to reproduce, the real Catalan-dominant bilinguals having a far higher score on these stimuli.

These results indicate that Model 1 was not feasible and all our analytical efforts thereafter focused on Model 2.

As concerns Model 2, having decided on the points which should be interesting behaviorally, we consequently ran full spiking simulations with 100 trials per point, for 11 parameter configurations in total. These configurations can be seen in Table 3 (*point* refers to a point in the 5-D parameter space, i.e., a specific parameter configuration).

Table 3. Parameter Configurations used in Spiking Trials Run for Model 2

	w_+	w_{w1}	w_{w2}	w_{pc1}	w_{pc2}
Point 1	2.02	1.06	1.09	1.06	1.00
Point 2	2.02	1.07	1.09	1.06	1.00
Point 3	2.02	1.07	1.09	1.05	1.00
Point 4	2.02	1.08	1.09	1.06	1.00
Point 5	2.02	1.09	1.09	1.06	1.00
Point 6	2.03	1.06	1.09	1.06	1.00
Point 7	2.03	1.07	1.09	1.06	1.00
Point 8	2.03	1.08	1.09	1.06	1.00
Point 9	2.02	1.08	1.09	1.07	1.00
Point 10	2.02	1.09	1.09	1.07	1.00
Point 11	2.02	1.09	1.09	1.04	1.00

The results of the spiking simulations of Model 2 looked very promising when going through the pool activity time courses for each stimulus and point. As an

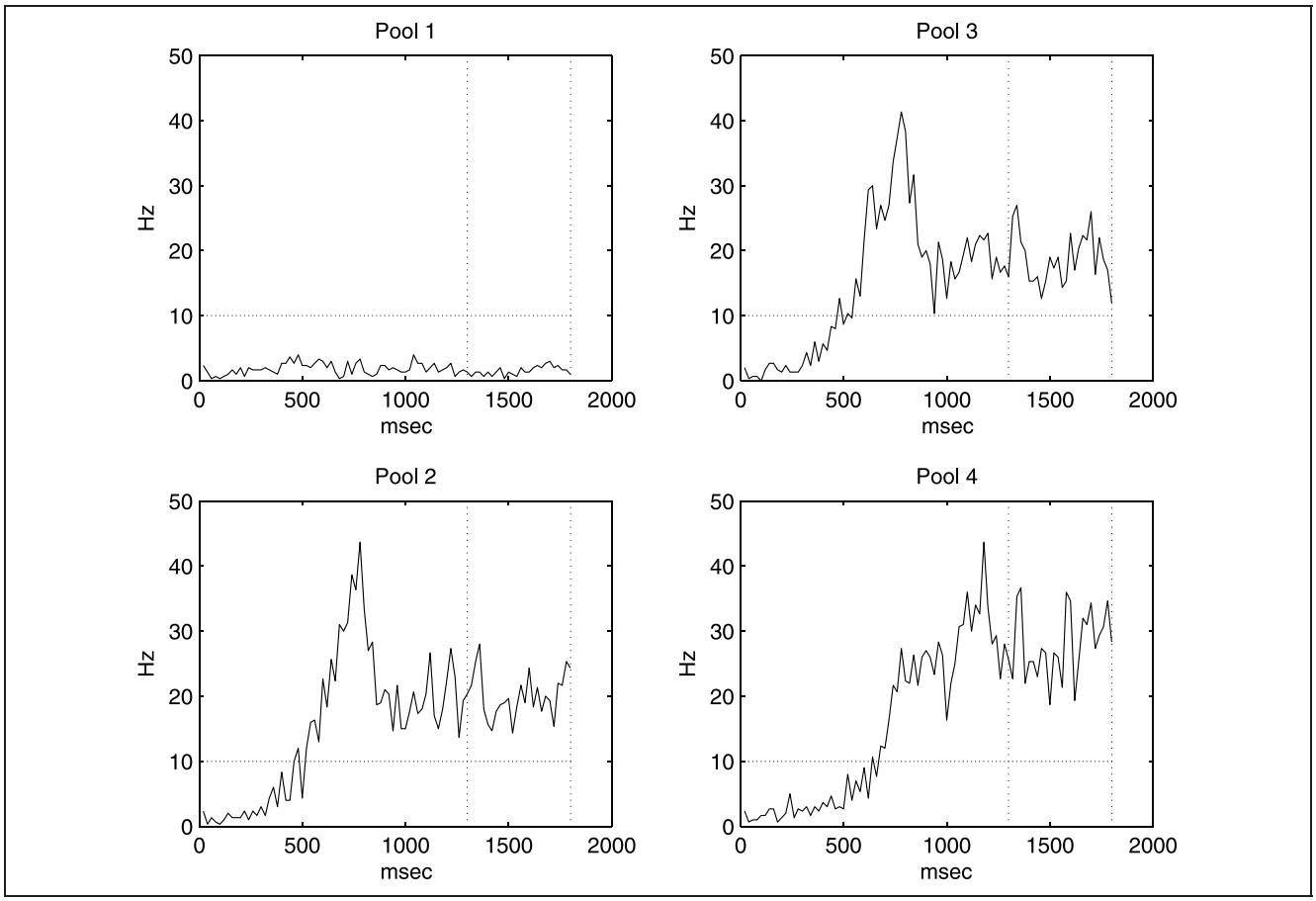

Figure 4. Model 1 spiking simulation for the input /finɛstra/. The figure shows the activity of the four pools, organized in relation to each other in the same way as in the model architecture. As can be seen in this trial, the pseudoword input drives the corresponding pools into an attractor and is thus recognized as a word, something which only very rarely would happen in reality. The horizontal line indicates the threshold; the two vertical lines enclose the time interval used for averaging the signal to obtain a response.

illustration, we here include a spiking simulation with the input /galleda/, presented to Pools 1 and 4, showing the sought behavior, that is, taking the pseudoword stimulus to be a word (Figure 5).

However, as it is hard to get an overview here of all the data obtained in these simulations by presenting the exact time course, we had to find a good measure which, at the same time as having a clear-cut correspondence in the experimental data we model, was easily presentable in a written form. To this end, we calculated the percentages ''yes'' and ''no'' responses over trials to each stimulus presentation (using the response definitions recounted before) and thereafter calculated the A' statistic for each network configuration simulation, using the different outcomes of the distinct trials to obtain the percentages. Using this approach, it was found that the experimental data were readily reproduced in the mean, with a variability of responses corresponding roughly to the distribution of behavior in the experimental paper (Sebastián-Gallés et al., 2005). In Table 4, you can see the percentage *correct* responses (''yes'' response to a word presentation, ''no'' response to a pseudoword presentation) to the presentation of the four stimuli we have considered in our modeling, together with the corresponding A' scores. The average values are in the bottom of the table, and correspond well to the average values obtained for Catalan-dominant early bilinguals in the experimental study (see Table 1).

To facilitate comparison of average values, we have generated figures. Figure 6 shows a comparative graph of average percentage correct responses, from experiment and model. In Figure 7 there is a comparative graph of the corresponding average A′ values, from experiment and model.

Based on the results obtained in the modeling of the LDT, we predicted that in order for the effect to manifest itself in our simulations, the connections between the two phoneme pools (Pools 2 and 4) needed to be ''symmetric'', that is, their weights' values needed to be equal. In Model 2, whose strong results indicate its suitability for the LDT task modeling, these weights are indeed equal, as they are both set to w_{-}. In the only point found in Model 1, these weights (w_{pp1} and w_{pp2}) were ''asymmetric,'' (i.e., differed in value). This led us

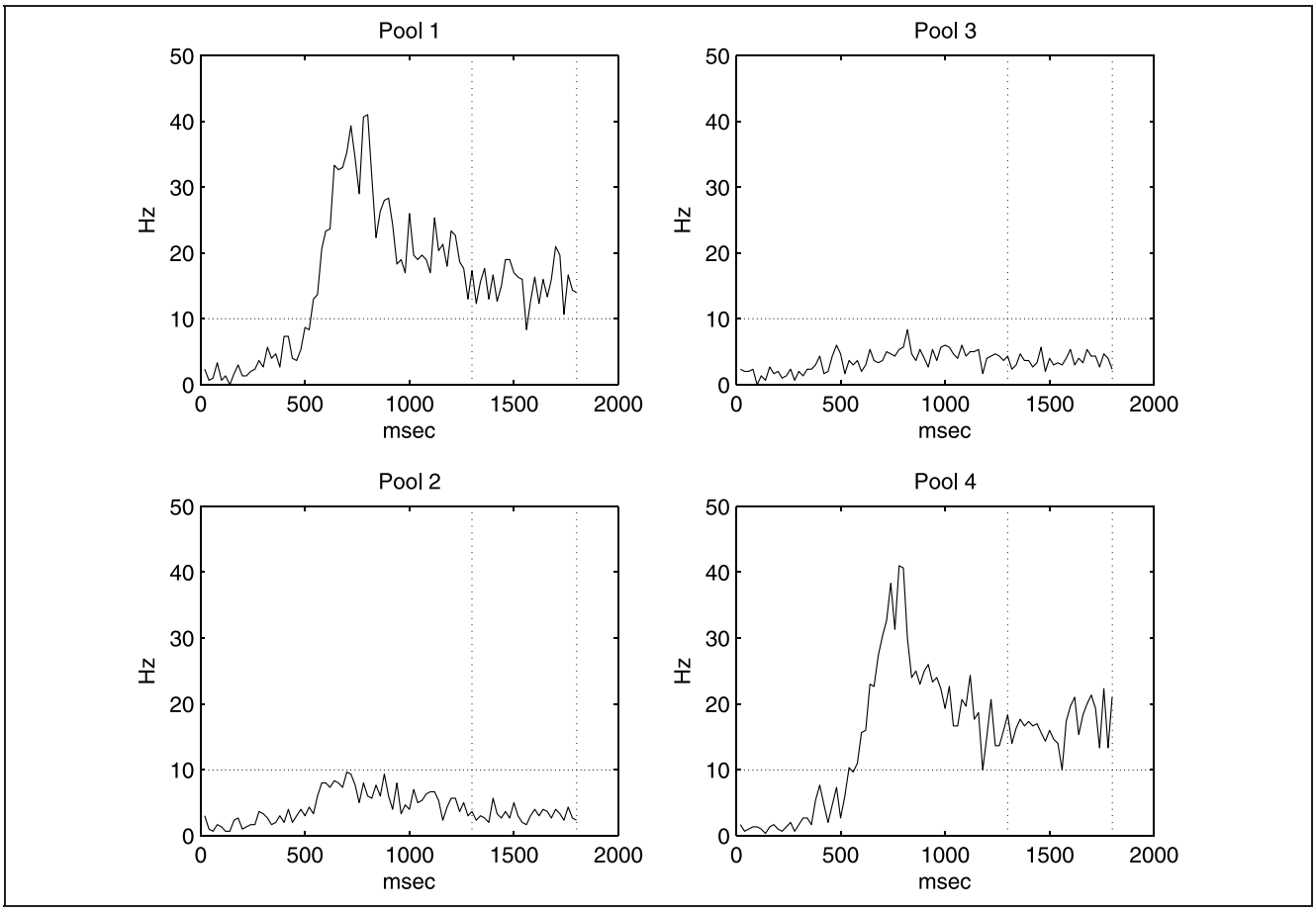

Figure 5. Model 2 spiking simulation for the input /galleda/. The figure shows the activity of the four pools, organized in relation to each other in the same way as in the model architecture. As can be seen in this trial, the pseudoword input drives the corresponding pools into an attractor, and is thus recognized as a word, which also happens frequently in reality for individuals showing the asymmetry in performance. The horizontal line indicates the threshold; the two vertical lines enclose the time interval used for averaging the signal to obtain a response.

Table 4. Individual and Mean Percent Correct Responses and Corresponding A' Values for All Stimuli, Obtained in Model 2 Spiking Simulations (W = word; PW = pseudoword)

	% Correct Responses					
	/ɛ/-Type		*/e/-Type*		A'	
Point	*W*	*PW*	*W*	*PW*	*/ɛ/-Type*	*/e/-Type*
1	57	70	92	94	0.715	0.962
2	76	75	91	97	0.838	0.969
3	81	87	91	94	0.905	0.960
4	89	77	92	93	0.900	0.960
5	94	77	92	90	0.919	0.951
6	78	54	99	87	0.751	0.964
7	89	55	98	88	0.824	0.964
8	94	60	96	90	0.869	0.963
9	86	62	89	93	0.833	0.960
10	93	64	89	90	0.876	0.941
11	100	85	100	81	0.963	0.953
Mean	85.18	69.64	93.55	90.64	0.854	0.959

to run further simulations, this time of a ''discrimination task,'' in which the network response must reflect a change in phoneme stimuli. The motivation for this was our prediction that in a discrimination task using the critical phonemes as stimuli, the Catalan-dominant bilinguals should show equally perfect performance regardless of the direction of the change in stimuli in a sequential presentation. That is, we suspected they would not show the asymmetry displayed in the LDT (i.e., in this case, they would *not* perform worse on /ɛ/ stimuli than on /e/ stimuli). Linguistically, this would be due to the lexical context not being present in such a task, in which case performance would not be affected by long-term changes in the lexicon. What we hoped to show with the modeling of this task was that, using the same points which showed the asymmetry in the LDT, we would obtain equal or nearly equal, very high A' values on the discrimination task, but *only* if there was not a significant asymmetry in weight values between the phoneme pools.

In this task, we simulated 100 trials for each of the 11 points already run for the LDT task, but with a different stimulus setup and duration. The initial period of zero input was maintained at 300 msec. Thereafter, the stimulus /ɛ/ or /e/ was presented to its corresponding pool, for a duration of 400 msec in both stimulus cases. This was followed by an interstimulus interval of 400 msec, after which either the same phoneme was presented again, or the other phoneme was presented, representing a switch in stimuli. Finally, the same period of 1000 msec as in the LDT task was simulated after the last stimulus offset, for a total of 2500 msec simulation for the discrimination task.

In all four cases (two switches and two nonswitches of stimuli), if activation showed that the last phoneme input was still retained by the network, the corresponding trial was deemed a correct response. That is, for the critical changes to be detected, the last stimulus presented to the network in a ''change'' trial had to be the winner in the competition induced by the switch. Using this paradigm, we again calculated both percentages and A' values for all 11 parameter configurations, but not only for these exact points. We also did the same for points in which an asymmetry in the phoneme-to-phoneme weights, w_{pp1} and w_{pp2} (Model 1 terminology),

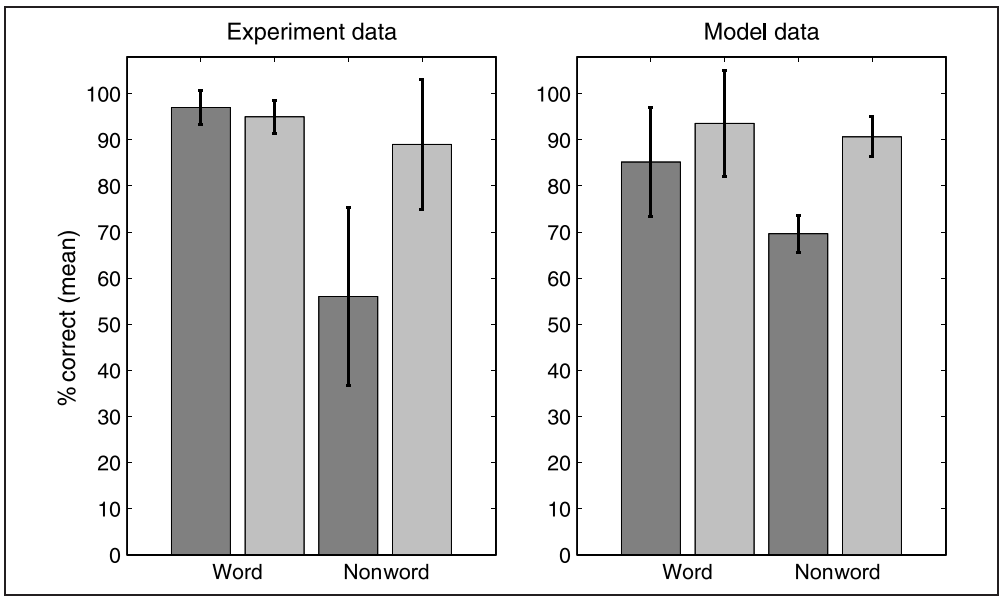

Figure 6. Shown are the average percentages of correct responses obtained in the experiment (left) and with Model 2 (right), for all four stimuli. The experimental values are for Catalan-dominant early bilinguals, taken from Sebastián-Gallés et al. (2005). Color coding: dark gray for /ɛ/-type stimuli, light gray for /e/-type stimuli. Error bars represent standard deviations.

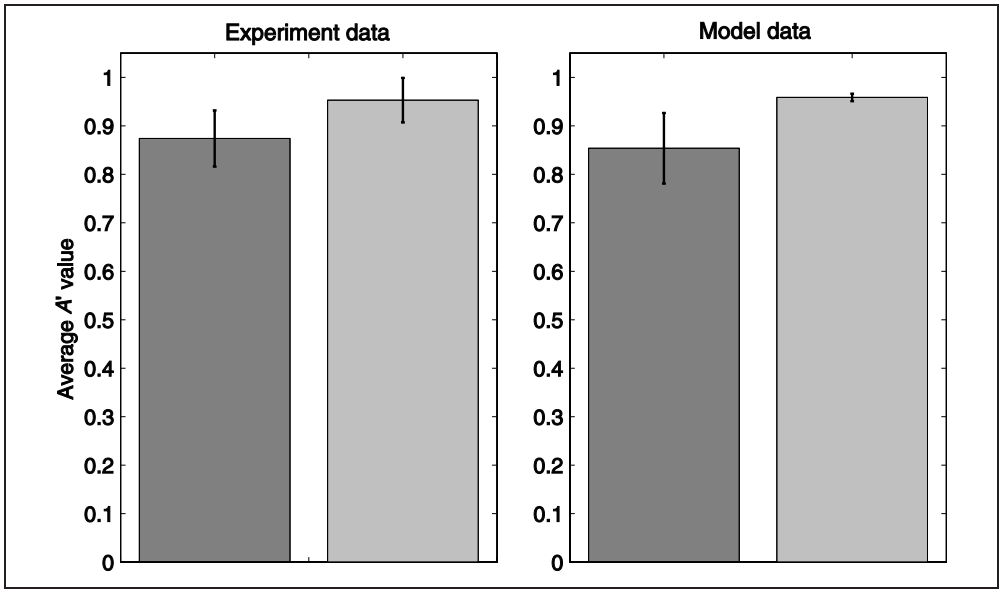

Figure 7. Shown are the average A' values obtained in the LDT experiment (left) and with Model 2 (right). The experimental values are for Catalan-dominant early bilinguals, taken from Sebastián-Gallés et al. (2005). Color coding: dark gray for /ɛ/-type stimuli, light gray for /e/-type stimuli. Error bars represent standard deviations.

was introduced, one of the weights set to a greater-than-w_- value and the other retained at w_-. Specifically, the values were $w_{pp1} = w_- = 0.89$ and $w_{pp2} = 0.97$, or vice versa, in order to investigate the effect of an asymmetry in both directions.

The results largely corroborated our prediction. In Table 5, we show the mean A' values obtained from the spiking simulations for the symmetric case and for the two asymmetric cases, calculated from the individual A' values of all 11 points (data not shown).

As can be seen in Table 5, in the symmetric weight case, A' values differ by a mere 2.5%, whereas in the asymmetric cases, they differ by 14% and 8%, respectively, in the order of presentation of the table. This indicates that there is a sensitive balance to be maintained, in which these interphoneme pool weights should be kept equal or may at least not differ by a significant amount, for discrimination performance not to show asymmetry.

The results of our simulations of the discrimination task strengthened our prediction that plasticity at the lexical level does not alter phoneme categories' representations and that the performance on a phoneme categorization task using the critical phonemes here discussed should approach perfect discrimination. Subsequently, as described in the Methods section, we conducted behavioral experiments to investigate the prediction, including a syllable discrimination task.

Table 5. Mean A' Values Obtained for the Different Spiking Simulations of the Discrimination Task

	Direction of Change	
Weight (A)symmetry	/ɛ/→/e/	/e/→/ɛ/
$w_{pp1} = w_{pp2}$	0.958	0.934
$w_{pp1} < w_{pp2}$	0.986	0.866
$w_{pp1} > w_{pp2}$	0.908	0.982

Experimental Results

Lexical Decision Task

As in Sebastián-Gallés et al. (2005), /ɛ/-type pseudowords yielded high error rates. The percentage of correct responses for each type of stimulus is shown in Table 6.

Discrimination Task

To obtain the A' statistics,[3] the percentages of correct responses (Hits) and false alarms (FAs) were calculated for each change of direction. Participants could give their response whenever during the entire trial duration. A trial in which the presented stimulus belonged to a different category than the previous one was considered a ''change trial.'' A Hit was scored when the response was given during change trials. Any response given at any other time was considered a false alarm. Participants showed very high performance, independently of the direction of change ($A'_{\varepsilon\to e} = 0.945$; $A'_{e\to\varepsilon} = 0.936$). In Table 7, the percentages of Hits and FAs are given.

Table 6. Percentage of Correct Responses and A' Scores for Each Condition (Lexical Decision Task)

	Word	*Pseudoword*	A'
/ɛ/-stim	95.1	56.9	0.865
/e/-stim	95.5	85.2	0.946
Control	98.2	95.1	0.982

Table 7. Percentage of Correct Responses (Hits) and False Alarms (FAs) for Each Direction of Change (Discrimination Task)

	Correct Responses		*False Alarms*	
	/ɛ/→/e/	/e/→/ɛ/	/ɛ/→/e/	/e/→/ɛ/
Mean	86.4	81.8	5.3	4.4
SD	12.8	15.7	4.2	3.0

Comparison of Results

As the modeling predicted, the Catalan-dominant early bilinguals in the discrimination task, indeed, showed virtually equal performance on both directions of change of critical vowel. In Figure 8 we show the average A' values attained by participants in the task, together with the average A' values obtained in the simulation of a phoneme discrimination task using the model.

Like the results of the modeling (right part of Figure 8), the performance in the experiment (left part of Figure 8) shows practically equal values on both directions of change, indicating that the performance is indeed equal at the phoneme level for the same kind of bilinguals who displayed asymmetry across stimuli in the LDT.

DISCUSSION

We take the modeling results to suggest that Model 2 (depicted in Figure 3) is the appropriate architecture for showing the asymmetry effect found in Sebastián-Gallés et al. (2005). The failure of Model 1 combined with the results of our discrimination task simulations substantiate the appropriateness of Model 2 by showing that as soon as an asymmetry occurs, neither an LDT nor a discrimination task shows the appropriate behavior. Furthermore, the results of the discrimination task simulations strengthen our belief that the asymmetry in the LDT task is a lexical effect, and thus, does not reflect an alteration of representations at the phoneme level, because such an alteration (weight asymmetry) deteriorated performance in the simulation of the LDT.

Based on the results obtained in our simulations, we maintain that to explain the effect observed in Sebastián-Gallés et al. (2005), the context of the word is needed, that is, it is not merely a modification of phoneme categories that underlies this behavior (as indicated in Norris et al., 2003), but a specific modification in the connections between phoneme pools which together represent a new word in the lexicon. Hence, the acceptance of the mispronounced word as a word is a lexical-level phenomenon. Furthermore, based on our modeling of a discrimination task, we expected the performance on a real discrimination task of subjects with the same background as those showing the asymmetry in the LDT in Sebastián-Gallés et al., to tend to be perfect for both directions of change.

In the subsequent psychophysical experiments, we showed that the prediction from the modeling results is, indeed, reasonable, as the very same kind of subjects as in Sebastián-Gallés et al. (2005) showed the asymmetry found earlier in an LDT, but practically equally high performance on a syllable discrimination task involving the same critical phonemes as in the lexical task. This indicated that a new word form stored in the lexicon could account for the asymmetry in the results, without the information having affected phoneme category representations, as performance at the phonemic level was practically equally high for both categories.

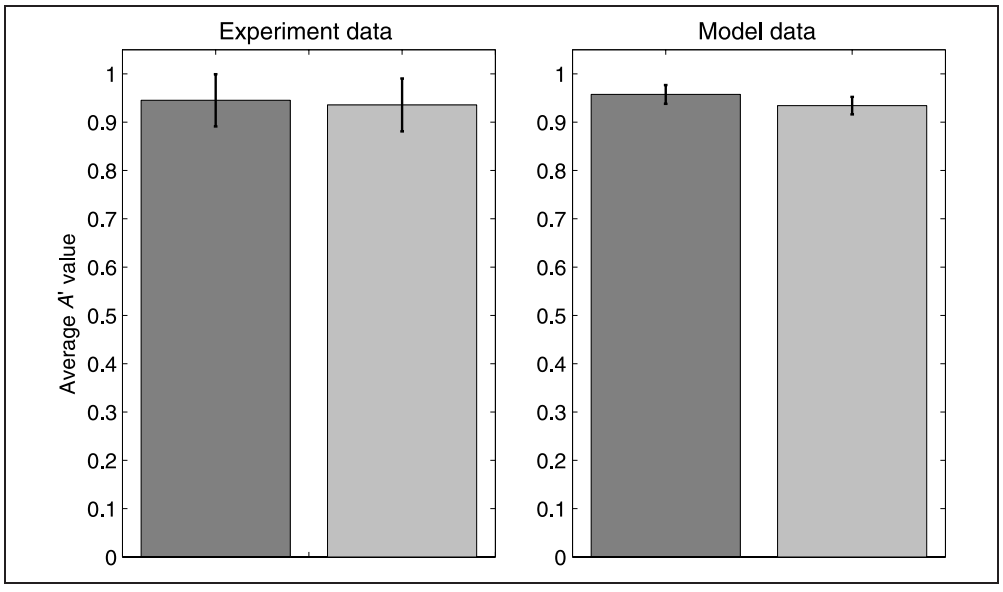

Figure 8. Shown are the average A' values obtained in the discrimination experiment (left) and with the model (right). The experimental values are for Catalan-dominant early bilinguals. Color coding: dark gray for /ɛ/–/e/ change, light gray for /e/–/ɛ/ change. Error bars represent standard deviations.

In a recent experimental paper mentioned in the Introduction, we found evidence in this direction (Connine, 2004). The author of that work used a phoneme identification experiment to examine the representation of phonological variants, specifically, the flap variant of spoken words which according to spelling are pronounced with a /t/, such as /pridi/ instead of /priti/. She found that her results provided ''strong evidence for the claim that representation of auditory form includes explicit representations of the frequently heard variant'' (p. 1088). She claimed that ''words with highly frequent auditory forms (e.g., containing a flap) that differ from their orthographic forms may develop parallel representations in the auditory lexicon'' (p. 1088) and her results indicated that listeners did not recode the flap variant into an underlying /t/ version, but rather recognized it via a preexisting lexical representation of the word containing this variant. Thus, those results are strongly supportive of our claim that the frequently heard variant /galleda/ could have induced in the subjects an explicit new lexical entry through long-term plasticity (Hebbian learning), rather than a change in phoneme categories. Consistent with these results, Cutler, Weber, and Otake (2006) recently used an eye-tracking paradigm to demonstrate an asymmetric mapping from phonetic to lexical representations in Japanese listening to English (their L2). Specifically, the authors found that although such subjects fail to make a distinction between the isolated English phonemes /r/ and /l/, they still have lexical entries for words containing both phoneme categories. In the article, the purported explanation for these seemingly paradoxical results was that the lexical entries had been stored through direct language instruction, although it was recognized that this would not explain if similar phenomena were to be found in subjects who had not acquired their L2 through explicit teaching. Although the situation in Cutler et al. (2006) is distinct from the one we have investigated, seeing as we focus on L1 listening in bilinguals, it can, nevertheless, be argued that the false-alarm recognition of, for instance, the pseudoword /galleda/ in our case has similar root causes, namely, two minimal-pair lexical entries /gallɛda/ and /galleda/ having been stored, although through environmental exposure instead of explicit language instruction, especially in the case of /galleda/, which is never taught as an actual word at school in Catalonia.

Although we in this work have shown a ''proof of principle'' of the feasibility of using detailed neurodynamical modeling for higher cognitive phenomena, such as language perception, a possible extension of this model would be to incorporate lower-level auditory processing, thus connecting our more representative firing rate input with the underlying stages, such as frequency mapping and spectro-temporal response fields in the auditory cortex (Kowalski, Depireux, & Shamma, 1996) or even models of the basilar membrane distributed frequency mapping in the cochlea. One good model of peripheral auditory processing is the Development System for Auditory Modelling (DSAM; O'Mard & Meddis, 1997), whose output is the actual spike trains of auditory nerve fibers. This kind of lower-level processing output could then be fed to higher-level networks, such as ours. One interesting possibility would be to use the approach found in Drew and Abbott (2003), where the authors also use an integrate-and-fire model and show that in order to get the model to process more than two sequential (birdsong) syllables, all one needs to do is to expand the number of pools in the network from the original four by an additional two per syllable. This would be interesting to explore in combination with a more realistic input (e.g., a filter bank or the DSAM model), presented with the real sound files of three-syllable words, which are then processed by a network able to handle three syllables. The principle on which the work by Drew and Abbott is based is a general one, so although the spectral complexity of human language syllables is higher than more harmonic birdsong syllables, this should be feasible. One could also look at more technical or mathematical extensions of the model, including explicit incorporation of plasticity in order to study how the resultant changes we here hypothesize come about through experience.

For the representation of a single phoneme used in our model, we find support in an interesting study by Houde and Jordan (1998), which looks at adaptation in vowel production following a training phase with real-time feedback of the subject's own transformed speech. The transformation corresponded to a shift in the vowel formant space, and the adaptation in the production of test words went in the other direction, so as to counter the transformation. The adaptation of the vowel /ɛ/ generalized to different word contexts than the ones presented during training, something taken by the authors as evidence for a common representation of the production of the vowel /ɛ/ in the brain, shared by the production representations of the training words. Although the study of Houde and Jordan focused on speech production, we argue that their results indicate that it is reasonable to assume a representation of words composed of representations of its constituent phonemes, also for speech perception. As for the neural substrate of this representation, the results of a study by Obleser, Lahiri, and Eulitz (2004), using magnetoencephalography to measure the neural activity evoked by tokens of German vowels, point to the auditory association cortices, that is, areas of the auditory cortex peripheral to the primary auditory cortex, AI.

Concerning the discrimination experiment, it could be possible that the lack of differences between the two directions of change be due to a ceiling effect (although extensive piloting was done to avoid it). Although this explanation cannot be totally discarded, further experimental results, both with behavioral and electrophysiological measurements, speak against such a possibility.

For instance, the results of Sebastián-Gallés, Vera-Constán, Larsson, Costa, and Deco (submitted), measuring the MMN in an oddball paradigm, show no differences between the two directions of change, thus confirming the present behavioral study and supporting the modeling.

APPENDIX A: NEURAL AND SYNAPTIC DYNAMICS

In this work, we use the mathematical formulation of the nonlinear leaky integrate-and-fire neuron and its modeled synaptic input currents, as described in Brunel and Wang (2001). What follows is a presentation of the essentials in that formulation.

The evolution in time of the subthreshold membrane potential $V(t)$ of a neuron is given by the equation

$$C_{\mathrm{m}}\frac{\mathrm{d}V(t)}{\mathrm{d}t} = -g_{\mathrm{m}}(V(t)-V_{\mathrm{L}}) - I_{\mathrm{syn}}(t), \qquad (1)$$

where C_{m} is the membrane capacitance set to 0.5 nF for excitatory neurons and 0.2 nF for inhibitory neurons; g_{m} is the membrane leak conductance, with the values 25 nS for excitatory neurons and 20 nS for inhibitory neurons; V_{L} is the resting potential of -70 mV and I_{syn} the synaptic input current. The firing threshold is $V_{\mathrm{thr}} = -50$ mV and the reset potential $V_{\mathrm{reset}} = -55$ mV.

The synaptic current is the sum of recurrent excitatory currents ($I_{\mathrm{AMPA,rec}}$ and $I_{\mathrm{NMDA,rec}}$), an external excitatory current ($I_{\mathrm{AMPA,ext}}$) and an inhibitory current (I_{GABA}):

$$I_{\mathrm{syn}}(t) = I_{\mathrm{AMPA,ext}}(t) + I_{\mathrm{AMPA,rec}} + I_{\mathrm{NMDA,rec}}(t) + I_{\mathrm{GABA}}(t) \qquad (2)$$

The currents are defined by

$$I_{\mathrm{AMPA,ext}}(t) = g_{\mathrm{AMPA,ext}}(V(t)-V_{\mathrm{E}})\sum_{j=1}^{N_{\mathrm{ext}}} s_j^{\mathrm{AMPA,ext}}(t) \qquad (3)$$

$$I_{\mathrm{AMPA,rec}}(t) = g_{\mathrm{AMPA,rec}}(V(t)-V_{\mathrm{E}})\sum_{j=1}^{N_{\mathrm{E}}} w_j^{\mathrm{AMPA}} s_j^{\mathrm{AMPA,rec}}(t) \qquad (4)$$

$$I_{\mathrm{NMDA,rec}}(t) = \frac{g_{\mathrm{NMDA}}(V(t)-V_{\mathrm{E}})}{1+[\mathrm{Mg}^{2+}]\exp(-0.062V(t))/3.57} \times \sum_{j=1}^{N_{\mathrm{E}}} w_j^{\mathrm{NMDA}} s_j^{\mathrm{NMDA}}(t) \qquad (5)$$

$$I_{\mathrm{GABA}}(t) = g_{\mathrm{GABA}}(V(t)-V_{\mathrm{I}})\sum_{j=1}^{N_{\mathrm{I}}} w_j^{\mathrm{GABA}} s_j^{\mathrm{GABA}}(t) \qquad (6)$$

where $V_{\mathrm{E}} = 0$ mV, $V_{\mathrm{I}} = -70$ mV, w_j are the synaptic weights, s_j are the fractions of open channels for the different receptors, g_X designates the synaptic conductance for channel X and j indexes presynaptic neurons providing input to the neuron under consideration. The NMDA synaptic current is dependent on the potential and controlled by the extracellular concentration of magnesium ($[\mathrm{Mg}^{2+}] = 1$ mM). The values for the synaptic conductances for excitatory neurons are $g_{\mathrm{AMPA,ext}} = 2.08$ nS, $g_{\mathrm{AMPA,rec}} = 0.052$ nS, $g_{\mathrm{NMDA}} = 0.1635$ nS and $g_{\mathrm{GABA}} = 0.625$ nS and for inhibitory neurons $g_{\mathrm{AMPA,ext}} = 1.62$ nS, $g_{\mathrm{AMPA,rec}} = 0.0405$ nS, $g_{\mathrm{NMDA}} = 0.129$ nS and $g_{\mathrm{GABA}} = 0.4865$ nS. These values are obtained from the ones used in Brunel and Wang (2001) by correcting for the different number of neurons in our model. In their work, the conductances were calculated so that in an unstructured network the excitatory neurons have a spontaneous spiking rate of 3 Hz and the inhibitory neurons a spontaneous rate of 9 Hz. The fractions of open channels are governed by

$$\frac{\mathrm{d}s_j^{\mathrm{AMPA,ext}}(t)}{\mathrm{d}t} = -\frac{s_j^{\mathrm{AMPA,ext}}(t)}{\tau_{\mathrm{AMPA}}} + \sum_k \delta\left(t-t_j^k\right) \qquad (7)$$

$$\frac{\mathrm{d}s_j^{\mathrm{AMPA,rec}}(t)}{\mathrm{d}t} = -\frac{s_j^{\mathrm{AMPA,rec}}(t)}{\tau_{\mathrm{AMPA}}} + \sum_k \delta\left(t-t_j^k\right) \qquad (8)$$

$$\frac{\mathrm{d}s_j^{\mathrm{NMDA}}(t)}{\mathrm{d}t} = -\frac{s_j^{\mathrm{NMDA}}(t)}{\tau_{\mathrm{NMDA,decay}}} + \alpha x_j(t)\left(1-s_j^{\mathrm{NMDA}}(t)\right) \qquad (9)$$

$$\frac{\mathrm{d}x_j(t)}{\mathrm{d}t} = -\frac{x_j(t)}{\tau_{\mathrm{NMDA,rise}}} + \sum_k \delta\left(t-t_j^k\right) \qquad (10)$$

$$\frac{\mathrm{d}s_j^{\mathrm{GABA}}(t)}{\mathrm{d}t} = -\frac{s_j^{\mathrm{GABA}}(t)}{\tau_{\mathrm{GABA}}} + \sum_k \delta\left(t-t_j^k\right), \qquad (11)$$

where decay time constants are $\tau_{\mathrm{NMDA,decay}} = 100$ msec for NMDA synapses, $\tau_{\mathrm{AMPA}} = 2$ msec for AMPA synapses and $\tau_{\mathrm{GABA}} = 10$ msec for GABA synapses. $\tau_{\mathrm{NMDA,rise}} = 2$ msec is the rise time for NMDA synapses (the rise times for AMPA and GABA are in reality very short and are therefore neglected in the model) and $\alpha = 0.5$ msec^{-1}. Each sum over k represents a sum over spikes in the form of δ-peaks ($\delta(t)$), emitted by presynaptic neuron j at time $t = t_j^k$.

The equations were in our implementation integrated numerically using a second order Runge–Kutta method with step size 0.02 msec. The Mersenne Twister algorithm was used as a random number generator for the

external Poisson spike trains and each of the typically 100 different trials for one specific parameter configuration differed in its random seed from the other trials.

APPENDIX B: THE MEAN-FIELD FORMULATION

The mean-field approximation used in the present work was derived in Brunel and Wang (2001). Its basic assumption is that the network of integrate-and-fire neurons has reached a stationary state. In this formulation the potential of a neuron is calculated as

$$\tau_x \frac{\mathrm{d}V(t)}{\mathrm{d}t} = -V(t) + \mu_x + \sigma_x \sqrt{\tau_x}\eta(t), \tag{12}$$

where $V(t)$ is the membrane potential, and x labels the different populations. τ_x is the effective membrane time constant, μ_x is the mean value which the membrane potential would have in the absence of spiking and fluctuations, σ_x measures the magnitude of the fluctuations and η is a Gaussian process with an exponentially decaying correlation function and the time constant τ_{AMPA}. The quantities μ_x and σ_x^2 are given by

$$\mu_x = \frac{(T_{\mathrm{ext}}\nu_{\mathrm{ext}} + T_{\mathrm{AMPA}}n_x^{\mathrm{AMPA}} + \rho_1 n_x^{\mathrm{NMDA}})V_{\mathrm{E}} + \rho_2 n_x^{\mathrm{NMDA}}\langle V_x\rangle + T_{\mathrm{I}}n_x^{\mathrm{GABA}}V_{\mathrm{I}} + V_{\mathrm{L}}}{S_x} \tag{13}$$

$$\sigma_x^2 = \frac{g_{\mathrm{AMPA,ext}}^2(\langle V_x\rangle - V_{\mathrm{E}})^2 N_{\mathrm{ext}}\nu_{\mathrm{ext}}\tau_{\mathrm{AMPA}}^2\tau_x}{g_{\mathrm{m}}^2\tau_{\mathrm{m}}^2} \tag{14}$$

where $\nu_{\mathrm{ext}} = 3$ Hz $(+\lambda_{\mathrm{stim}})$, ν_{I} is the spiking rate of the inhibitory pool, $\tau_{\mathrm{m}} = C_{\mathrm{m}}/g_{\mathrm{m}}$ with the values for the excitatory or inhibitory neurons depending on the pool considered. The other quantities are given by

$$S_x = 1 + T_{\mathrm{ext}}\nu_{\mathrm{ext}} + T_{\mathrm{AMPA}}n_x^{\mathrm{AMPA}} + (\rho_1 + \rho_2)n_x^{\mathrm{NMDA}} + T_{\mathrm{I}}n_x^{\mathrm{GABA}} \tag{15}$$

$$\tau_x = \frac{C_{\mathrm{m}}}{g_{\mathrm{m}}S_x} \tag{16}$$

$$n_x^{\mathrm{AMPA}} = \sum_{j=1}^{p} f_j w_{jx}^{\mathrm{AMPA}}\nu_j \tag{17}$$

$$n_x^{\mathrm{NMDA}} = \sum_{j=1}^{p} f_j w_{jx}^{\mathrm{NMDA}}\psi(\nu_j) \tag{18}$$

$$n_x^{\mathrm{GABA}} = \sum_{j=1}^{p} f_j w_{jx}^{\mathrm{GABA}}\nu_j \tag{19}$$

$$\psi(\nu) = \frac{\nu\tau_{\mathrm{NMDA}}}{1+\nu\tau_{\mathrm{NMDA}}} \times \left(1 + \frac{1}{1+\nu\tau_{\mathrm{NMDA}}}\sum_{n=1}^{\infty}\frac{(-\alpha\tau_{\mathrm{NMDA,rise}})^n T_n}{(n+1)!}\right) \tag{20}$$

$$T_n(\nu) = \sum_{k=0}^{n}(-1)^k\frac{n!}{(n-k)!k!} \times \frac{\tau_{\mathrm{NMDA,rise}}(1+\nu\tau_{\mathrm{NMDA}})}{\tau_{\mathrm{NMDA,rise}}(1+\nu\tau_{\mathrm{NMDA}}) + k\tau_{\mathrm{NMDA,decay}}} \tag{21}$$

$$\tau_{\mathrm{NMDA}} = \alpha\tau_{\mathrm{NMDA,rise}}\tau_{\mathrm{NMDA,decay}} \tag{22}$$

$$T_{\mathrm{ext}} = \frac{g_{\mathrm{AMPA,ext}}\tau_{\mathrm{AMPA}}}{g_{\mathrm{m}}} \tag{23}$$

$$T_{\mathrm{AMPA}} = \frac{g_{\mathrm{AMPA,rec}}N_{\mathrm{E}}\tau_{\mathrm{AMPA}}}{g_{\mathrm{m}}} \tag{24}$$

$$\rho_1 = \frac{g_{\mathrm{NMDA}}N_{\mathrm{E}}}{g_{\mathrm{m}}J} \tag{25}$$

$$\rho_2 = \beta\frac{g_{\mathrm{NMDA}}N_{\mathrm{E}}(\langle V_x\rangle - V_{\mathrm{E}})(J-1)}{g_{\mathrm{m}}J^2} \tag{26}$$

$$J = 1 + \gamma\exp(-\beta\langle V_x\rangle) \tag{27}$$

$$T_{\mathrm{I}} = \frac{g_{\mathrm{GABA}}N_{\mathrm{I}}\tau_{\mathrm{GABA}}}{g_{\mathrm{m}}} \tag{28}$$

$$\langle V_x\rangle = \mu_x - (V_{\mathrm{thr}} - V_{\mathrm{reset}})\nu_x\tau_x, \tag{29}$$

where p is the number of excitatory pools, f_x the fraction of neurons in the excitatory pool x, $w_{j,x}$ the weight of the connections from pool x to pool j, ν_x the spiking rate of the excitatory pool x, $\gamma = [\mathrm{Mg}^{2+}]/3.57$ and $\beta = 0.062$.

The spiking rate of a pool as a function of the defined quantities can then be described by:

$$\nu_x = \phi(\mu_x, \sigma_x), \tag{30}$$

where

$$\phi(\mu_x, \sigma_x) = \left(\tau_{rp} + \tau_x \int_{\beta(\mu_x,\sigma_x)}^{\alpha(\mu_x,\sigma_x)} du\sqrt{\pi}\exp(u^2)[1 + \mathrm{erf}(u)]\right)^{-1} \quad (31)$$

$$\alpha(\mu_x, \sigma_x) = \frac{(V_{thr} - \mu_x)}{\sigma_x}\left(1 + 0.5\frac{\tau_{AMPA}}{\tau_x}\right) + 1.03\sqrt{\frac{\tau_{AMPA}}{\tau_x}} - 0.5\frac{\tau_{AMPA}}{\tau_x}x \quad (32)$$

$$\beta(\mu_x, \sigma_x) = \frac{V_{reset} - \mu_x}{\sigma_x} \quad (33)$$

$\mathrm{erf}(u)$ is the error function and τ_{rp} is the refractory period, which is considered to be 2 msec for excitatory neurons and 1 msec for inhibitory neurons. To solve the equations defined by Equation 30 for all x, we numerically integrate Equation 29 and the differential Equation 34 below, whose fixed point solutions correspond to solutions to Equation 30:

$$\tau\frac{d\nu_x}{dt} = -\nu_x + \phi(\mu_x, \sigma_x). \quad (34)$$

The equations were integrated using the Euler method with step size 0.2 and 4000 iterations, enough for sufficient convergence to be attained.

Acknowledgments

This study was supported by the James S. McDonnell Foundation (JMSF 20002079), the European research project EmCAP (FP6-IST, contract 013123) and the Spanish Ministerio de Educación y Ciencia (with EC Fondos FEDER SEJ2004-07680-C02-01/PSIC). J. L. was also supported by Generalitat de Catalunya through an IGSOC/IQUC grant and F. V. C. was supported by a fellowship from the Spanish Ministerio de Educación y Ciencia (BES-2005-9031).

Reprint requests should be sent to Johan Larsson, Computational Neuroscience Group, Departament de Tecnologies de la Informació i les Comunicacions, Universitat Pompeu Fabra, Passeig de Circumval·lació, 8, 08003, Barcelona, Spain, or via e-mail: johan.petter.larsson@gmail.com.

Notes

1. To simplify the reading, only the phonetic symbols of the critical phonemes (primarily /e/ /ɛ/) will be used when transcribing the pronunciation of words.
2. For a full description, see Sebastián-Gallés et al. (2005).
3. The use of A' statistics is due to the fact that this statistic was already employed in the Sebastián-Gallés et al. (2005) study. We have run parallel analyses with the d' statistic and the same pattern of results was obtained.

REFERENCES

Abeles, M. (1991). *Corticonics*. New York: Cambridge University Press.

Bosch, L., Costa, A., & Sebastián-Gallés, N. (2000). First and second language vowel perception in early bilinguals. *European Journal of Cognitive Psychology, 12,* 189–222.

Braitenberg, V., & Schütz, A. (1991). *Anatomy of the cortex*. Berlin: Springer Verlag.

Brunel, N., & Wang, X. (2001). Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *Journal of Computational Neuroscience, 11,* 63–85.

Clopper, C. G., & Pisoni, D. B. (2005). Perception of dialect variation. In D. B. Pisoni & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 313–337). Oxford, UK: Blackwell Publishing.

Connine, C. M. (2004). It's not what you hear but how often you hear it: On the neglected role of phonological variant frequency in auditory word recognition. *Psychonomic Bulletin & Review, 11,* 1084–1089.

Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics, 34,* 269–284.

Deco, G., & Rolls, E. T. (2004). Synaptic and spiking dynamics underlying reward reversal in the orbitofrontal cortex. *Cerebral Cortex, 15,* 15–30.

Deco, G., & Rolls, E. T. (2005a). Attention, short term memory, and action selection: A unifying theory. *Progress in Neurobiology, 76,* 236–256.

Deco, G., & Rolls, E. T. (2005b). Neurodynamics of biased competition and cooperation for attention: A model with spiking neurons. *Journal of Neurophysiology, 94,* 295–313.

Deco, G., & Rolls, E. T. (2006). Decision-making and Weber's law: A neurophysiological model. *European Journal of Neuroscience, 24,* 901–916.

Drew, P. J., & Abbott, L. F. (2003). Model of song selectivity and sequence generation in area hvc of the songbird. *Journal of Neurophysiology, 89,* 2697–2706.

Eisner, F., & McQueen, J. M. (2006). Stability over time in lexically-guided auditory perceptual learning. *Conference paper, Spring meeting of the Experimental Psychology Society, Birmingham, UK, April 10–12*.

Hebb, D. (1949). *The organization of behavior—A neurophysiological theory*. New York: Wiley.

Hestrin, S., Sah, P., & Nicoll, R. (1990). Mechanisms generating the time course of dual component excitatory synaptic currents recorded in hippocampal slices. *Neuron, 5,* 247–253.

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science, 279,* 1213–1216.

Jahr, C., & Stevens, C. (1990). Voltage dependence of NMDA-activated macroscopic conductances predicted by single-channel kinetics. *Journal of Neuroscience, 10,* 3178–3182.

Koch, K. W., & Fuster, J. M. (1989). Unit activity in monkey parietal cortex related to haptic perception and temporary memory. *Experimental Brain Research, 76,* 292–306.

Kowalski, N., Depireux, D. A., & Shamma, S. A. (1996). Analysis of dynamic spectra in ferret primary auditory cortex: I. Characteristics of single-unit responses to moving ripple spectra. *Journal of Neurophysiology, 76,* 3503–3523.

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology, 51,* 141–178.

Machens, C. K., Romo, R., & Brody, C. D. (2005). Flexible control of mutual inhibition: A neural model of two-interval discrimination. *Science, 307,* 1121–1124.
Macmillan, N., & Creelman, C. (1991). *Detection theory: A user's guide*. Cambridge: Cambridge University Press.
McClelland, J. L., & Elman, J. L. (1986). The trace model of speech perception. *Cognitive Psychology, 18,* 1–86.
McCormick, D., Connors, B., Lighthall, J., & Prince, D. A. (1985). Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons in the neocortex. *Journal of Neurophysiology, 54,* 782–806.
Miller, P., & Wang, X.-J. (2002). Inhibitory control by an integral feedback signal in prefrontal cortex: A model of discrimination between sequential stimuli. *Proceedings of the National Academy of Sciences, U.S.A., 36,* 955–968.
Norris, D. G., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioural and Brain Sciences, 23,* 299–370.
Norris, D. G., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47,* 204–238.
Obleser, J., Lahiri, A., & Eulitz, C. (2004). Magnetic brain response mirrors extraction of phonological features from spoken vowels. *Journal of Cognitive Neuroscience, 16,* 31–39.
O'Mard, L. P., & Meddis, R. (1997). Computer exploration of the auditory system with LUTEAR. *British Journal of Audiology, 31,* 125–126.
Pallier, C., Bosch, L., & Sebastián-Gallés, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition, 64,* B9–B17.
Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based vs. abstract lexical entries. *Psychological Science, 12,* 445–449.
Rolls, E. T., & Deco, G. (2002). *Computational neuroscience of vision*. Oxford, UK: Oxford University Press.
Rolls, E. T., & Treves, A. (1998). *Neural networks and brain function*. Oxford, UK: Oxford University Press.
Salin, P., & Prince, D. A. (1996). Spontaneous GABA-A receptor mediated inhibitory currents in adult rat somatosensory cortex. *Journal of Neurophysiology, 75,* 1573–1588.
Sebastián-Gallés, N., Echeverría, S., & Bosch, L. (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language, 52,* 240–255.
Sebastián-Gallés, N., Rodríguez-Fornells, A., Diego-Balaguer, R., & Diaz, B. (2006). First- and second-language phonological representations in the mental lexicon. *Journal of Cognitive Neuroscience, 18,* 1277–1291.
Sebastián-Gallés, N., & Soto-Faraco, S. (1999). On-line processing of native and non-native phonemic contrasts in early bilinguals. *Cognition, 72,* 112–123.
Sebastián-Gallés, N., Vera-Constán, F., Larsson, J. P., Costa, A., & Deco, G. (submitted). Lexical plasticity in early bilinguals does not alter phoneme categories: II. Experimental evidence.
Spruston, N., Jonas, P., & Sakmann, B. (1995). Dendritic glutamate receptor channel in rat hippocampal ca3 and ca1 pyramidal neurons. *Journal of Physiology, 482,* 325–352.
Tuckwell, H. (1988). *Introduction to theoretical neurobiology*. Cambridge: Cambridge University Press.
Wang, X.-J. (1999). Synaptic basis of cortical persistent activity: The importance of NMDA receptors to working memory. *Journal of Neuroscience, 19,* 9587–9603.
Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron, 36,* 955–968.
Wilson, F. A., Scalaidhe, S. P., & Goldman-Rakic, P. S. (1994). Functional synergism between putative gamma-aminobutyrate-containing neurons and pyramidal neurons in prefrontal cortex. *Proceedings of the National Academy of Sciences, U.S.A., 91,* 4009–4013.
Wong, K.-F., & Wang, X.-J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *Journal of Neuroscience, 26,* 1314–1328.
Xiang, Z., Huguenard, J. R., & Prince, D. A. (1998). GABA-A receptor mediated currents in interneurons and pyramidal cells of rat visual cortex. *Journal of Physiology, 506,* 715–730.