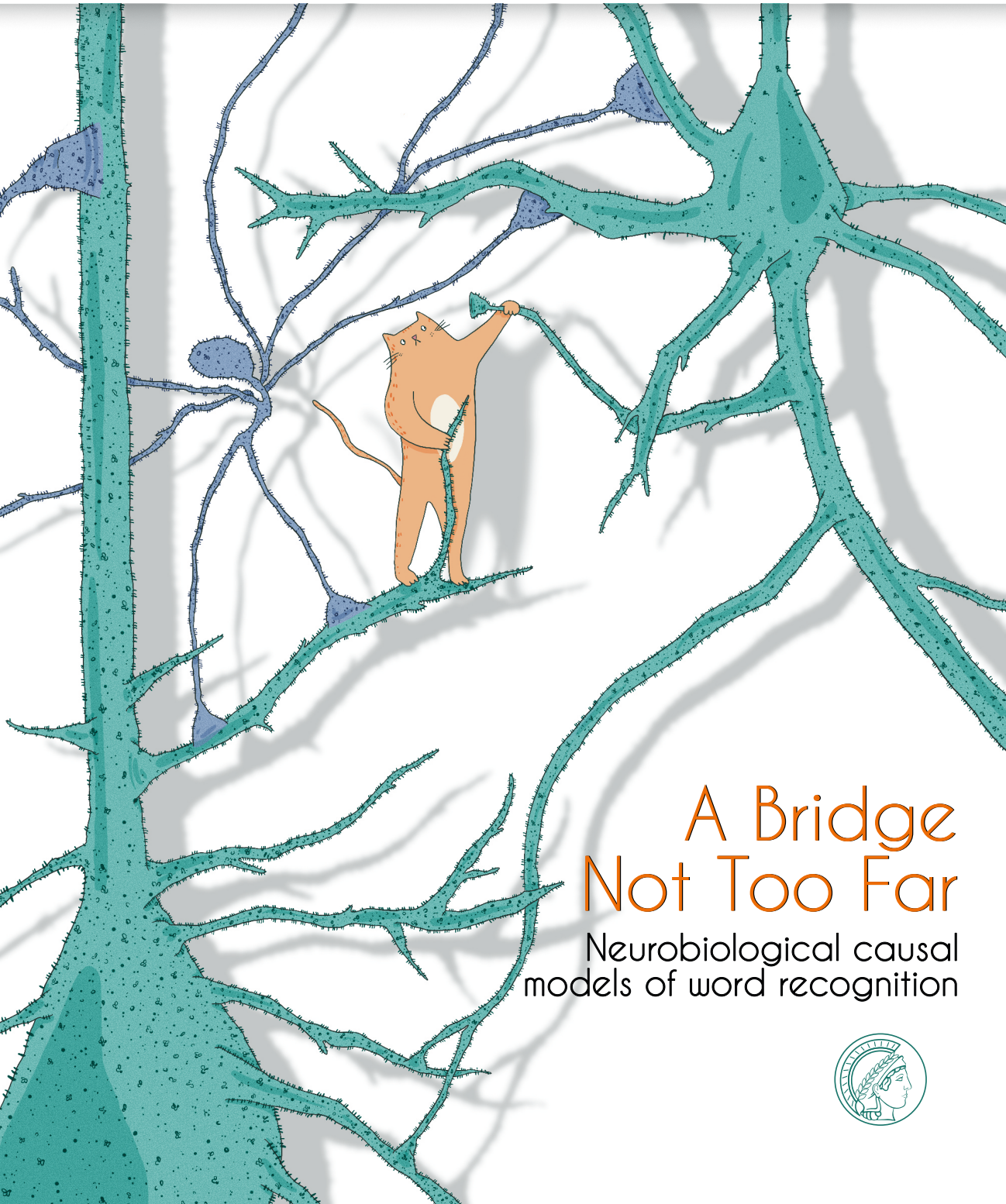
MAX PLANCK INSTITUTE
FOR PSYCHOLINGUISTICS



# A Bridge Not Too Far

## Neurobiological causal models of word recognition

# A Bridge *Not* Too Far:

# Neurobiological causal models of word recognition

**Funding body**
This research was funded by the Netherlands Organisation for Scientific Research (NWO) under Gravitation Grant 024.001.006 to the Language in Interaction Consortium.

**International Max Planck Research School (IMPRS) for Language Sciences**
The educational component of the doctoral training was provided by the International Max Planck Research School (IMPRS) for Language Sciences. The graduate school is a joint initiative between the Max Planck Institute for Psycholinguistics and two partner institutes at Radboud University – the Centre for Language Studies, and the Donders Institute for Brain, Cognition and Behaviour. The IMPRS curriculum, which is funded by the Max Planck Society for the Advancement of Science, ensures that each member receives interdisciplinary training in the language sciences and develops a well-rounded skill set in preparation for fulfilling careers in academia and beyond. More information can be found at www.mpi.nl/imprs

**The MPI series in Psycholinguistics**
Initiated in 1997, the MPI series in Psycholinguistics contains doctoral theses produced at the Max Planck Institute for Psycholinguistics. Since 2013, it includes theses produced by members of the IMPRS for Language Sciences. The current listing is available at www.mpi.nl/mpi-series


ISBN: 978-94-92910-61-5
Cover design by Margherita Ferracuti
Printed and bound by Ipskamp Drukkers, Enschede

 The research reported in this thesis was conducted at the Max Planck Institute for Psycholinguistics, in Nijmegen, the Netherlands

# A Bridge *Not* Too Far:

# Neurobiological causal models of word recognition

Proefschrift ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de rector magnificus prof. dr. J.M. Sanders,

volgens besluit van het college voor promoties

in het openbaar te verdedigen op

vrijdag 12 juli 2024

om 12.30 uur precies

door

Alessio Quaresima

geboren op 26 januari 1992

te Rome (Italië)

**Promotor:**

Prof. dr. Peter Hagoort

**Copromotoren:**

Dr. Hartmut Fitz

Dr. Karl Magnus Petersson (Max Planck Instituut voor Psycholinguïstiek)

**Manuscriptcommissie:**

Prof. dr. Marcel van Gerven

Dr. Laura Gwilliams (Stanford University, Verenigde Staten)

Prof. dr. Panayiota Poirazi (IMBB-FORTH, Griekenland)

*For Martina, Luca, and Elisa*

# Contents

# 1 | General introduction

## 1.1 Recognizing a word

Imagine yourself entering a supermarket, busy recollecting all the items on your grocery list. Suddenly, your attention is caught by a sound. You hear a fragment of a sentence from the store's speakers. You hear the word *HIMALAYA*.
Before you understand it is about a trendy type of table salt, you can't help but think of a faraway mountain peak covered with snow.
How is it possible that a word and its meaning can be retrieved from a brief sequence of sounds? The present manuscript investigates the neural mechanisms that make possible the association of sequences of sounds to words, in a biologically constrained, computational model of the human brain.


Figure 1.1: **Acoustic signals and semantic content of the word HIMALAYA** (**A**) Waveform graphs of different pronunciations of the word HIMALAYA, from three English speakers. The horizontal axis represents time, and the vertical axis represents the amplitude of the air pressure. (**B**) One of the many pictures of the Himalayan mountain chain.

When a word is uttered, the sound generated by the vocal apparatus of the speaker reaches the ears of the listeners. There, the cochlea transforms the acoustic signal into an electrochemical one. Following the projections of the auditory neurons in the brain stem, it traverses the central nervous system and reaches the temporal lobe, where the primary auditory cortex is located. Once in the forebrain, this piece of information generates a cascade of neuronal activ-

ity that ultimately produces, to the attentive listener, the recollection of a word. This entire process occurs in a few hundreds of milliseconds.

An example of a speech fragment is offered in Fig.1.1*A*. The waveforms portrayed show three different acoustic realizations of the English word HIMALAYA – the reader should intend the graphs as the sounds they represent, acoustic signals rather than visual ones. The speech fragments show two properties. First, even though they present peaks and valleys, they have no iconic relationship with the mountain chain located north of the Indian subcontinent (Fig.1.1*B*), and second, they differ from each other. Remarkably, in the face of arbitrary sound-meaning relationships and intrinsic variability across speakers, any person who previously encountered the English word HIMALAYA will, likely, recognize each of the three acoustic signals presented.

Indeed, the sound waves in Fig.1.1*A* conceal a third property, which is somehow obvious to the attentive listener. The three speech signals are composed of a sequence of sounds that most English listeners will recognize in the same discrete set of phonemes, /hɪməlaɪə/[1]. Independently of its specific acoustic realization, the word HIMALAYA is recognized whenever the listener *perceives* that sequence of phonemes. Or, in more general terms, words are retrieved when there is a sufficient match between the phonological information in the acoustic signal and the word's phonological form in the mental lexicon of the listener.

Spoken word recognition depends on the human capacity of *mapping an uninterrupted stream of phonemes onto a series of word forms stored in memory* (Magnuson & Crinnion, 2022, p. 462). Humans rely on this faculty for comprehending speech and communicating. When it is impaired, it carries dire consequences for the speaker (Mirman & Britt, 2014). Even though spoken word recognition (SWR) is central in human language, and much work has been done to characterize it (Magnuson & Crinnion, 2022; McQueen, 2007; Vitevitch, Siew, & Castro, 2018), it is not yet clear which are its neuronal underpinnings. It is not understood how lexical retrieval is *implemented* in the human brain.

The present work attempts to address this open issue and to contribute with theoretical insights toward a neurobiological explanation of SWR capacity. The study is carried out by simulating a network of biophysical neuron models. The network aims to reproduce the activation of neuronal populations representing word-form memories and phoneme-like units. By mapping a set of abstract linguistic representations to the dynamics of the modeled neurons, I propose

[1]The International Phonetic Alphabet (IPA) transcription was obtained from `http://tom.brondsted.dk/text2phoneme/`

a linking hypothesis, a bridge, between the computational primitives of word recognition and the neurobiological principles that govern brain processes.

The Introduction chapter addresses some necessary preliminaries to the investigation carried out in the experimental chapters.
Section 1.2 sets the boundary conditions for a mechanistic explanation of SWR and frames it as a computational problem. On this base, I present the requirements that a model of word recognition must satisfy to be considered biologically explanatory. The section outlines the dynamical system framework as a working hypothesis for a model that describes linguistic capacities based on neural mechanisms. Section 1.3 summarizes linguistic and neurobiological evidence on SWR. The section describes the fundamental psycholinguistic *explananda* that the model must address. Finally, Section 1.4 offers a summary of the neuroscientific evidence that must be included in the model to constrain its implementation. The biological processes described in this section are the *explanans* of the word recognition capacity.

## 1.2 Computational models of spoken word recognition

### 1.2.1 Language and the computational brain

Word recognition is a linguistic capacity that is at play within the human meaning-making cognitive system (Hagoort, 2020). This capacity is acquired during development and is mastered only if the individual has sufficient cognitive abilities and is part of a language community. Although spoken communication relies on other organs, such as the vocal system and the ears, clinical evidence in subjects with full body paralysis indicates that a functioning brain is a necessary and sufficient condition to recognize spoken words (Fedorenko & Thompson-Schill, 2014; Hagoort, 2014; Metzger et al., 2023; Rowley, Rogish, Alexander, & Riggs, 2017). Thus, the core operations supporting the recollection of a word memory occur according to the principles that govern the brain's neurobiology (Kandel, Schwartz, Jessell, Siegelbaum, & Hudspeth, 2012; Luo, 2015; Sterling & Laughlin, 2015) A neurobiological explanation of the word recognition capacity must account how this capacity is realized in the physiology of the human brain. This is not an easy task because it requires understanding two extremely complex phenomena: the human language and the brain.

Traditionally, cognitive sciences look at these two phenomena together through the lens of computation[2](CCTM Rescorla, 2020). Word recognition can be described as a linguistic computational operation, and the brain is the machine that operates this transformation. However, as we show in the following, not all computational descriptions of language capacities are equivalent. The introduction of computing machines in the early half of the last century fostered the idea that computers can be used as metaphors for the human brain (Von Neumann, 1958). In accordance, psychologists and linguists dedicated a great deal of effort to deriving computational explanations of the language function. Word recognition has been one of the first to be addressed with a computational model. The Logogen model by Morton (1969) could compute quantitative predictions of lexical decisions of human subjects during single-word recognition. Remarkably, the model was based on a threshold mechanism that vaguely resembled the logical calculus of nervous activity described thirty years earlier by McCulloch and Pitts (1943).

Over the following fifty years, several computational models were devised to reproduce the input-output relationship observed in experimental data. New models were often introduced to explain aspects of human behavior that were not matched by the earlier models or to propose simpler algorithmic solutions (Magnuson & Crinnion, 2022; Weber & Scharenborg, 2012). The underlying hypothesis of this research agenda is that an increased match between the model output and human behavior (among which neural markers) indicates that the brain computations are similar to those achieved in the model (Caucheteux & King, 2022). In this sense, they model the functions, and algorithms, of the human brain.

Notably, the pursuit of computational models was successful. In a computational sense, the problem of word form recognition is reasonably solved. For example, Bayesian models of cognition account for human behavior in extracting linguistic categories from variable speech sounds and retrieving the correct word (Kleinschmidt & Jaeger, 2015; Norris & McQueen, 2008). However, these computational models do not *explain* how the word recognition capacity relates to the human wetware, the brain circuitry (Poeppel & Idsardi, 2022). Indeed, the multiplicity of algorithmic solutions in psycholinguistics spoken word recognition models reveals that cognitive computational models are unconstrained. There are many, perhaps infinite, ways to achieve word recognition. This wealth

[2]Computation is the action of a physical system transforming inputs into outputs, following a set of rules.

of solutions becomes a problem if one aims to explain the *one* which is implemented in the human brain.

### 1.2.2 Cognitive models are not causal models of the brain...

The explanatory limits of cognitive-computational models become evident if we observe the brain as a complex information-processing system. In his influential work (Marr, 2010, I edition 1982), David Marr postulates that the processing of information in the brain can be analyzed at three levels: computational, algorithmic, and implementation. The first concerns the computational problem the system is solving (i.e., retrieving the word form that is associated with the acoustic input). The second level relates to the algorithm that the system uses to solve that problem (e.g., searching in a vocabulary? or creating a shortlist of matching items?); and the third to how that algorithm is implemented in the brain's wetware.

Cognitive computational models of word recognition address the first two levels. They carry out the expected linguistic computation with a fine-grained description of the expected input-output relationships (e.g., frequency effects in SWR, Brysbaert, Mandera, & Keuleers, 2018), and they propose algorithms to achieve them. However, these models compute on symbolic abstractions, whose relationship to the implementational substrate of the brain is opaque (Fitz, Hagoort, & Petersson, 2024). In cognitive models, it is difficult to establish a link between the model representations and the neurobiological parts and processes that govern brain activity. This has important consequences for the explanatory status of these models. According to Kaplan and Craver (2011)

> [...] the line that demarcates explanations from merely empirically adequate models seems to correspond to whether the model describes the relevant causal structures that produce, underlie, or maintain the explanandum phenomenon.

And further, for models to be explanatory

> (a) the variables in the model correspond to components, activities, properties, and organizational features of the target mechanism that produces, maintains, or underlies the phenomenon, and (b) the (perhaps mathematical) dependencies posited among these variables in the model correspond to the (perhaps quantifiable) causal relations among the components of the target mechanism.

The authors refer to these principles as the model-to-mechanism-mapping (3M) requirements. The concept of mapping is pivotal here and explicitly refers to the implementational level of description. If a cognitive model does not specify the neurobiological variables that correspond to its algorithmic components, it is non-explanatory with respect to the causal generators of behavior (Fitz et al., 2024). Similarly, Poeppel and Embick (2005) argue that existing neurolinguistic approaches fail to establish explicit links between linguistic categories and brain biology because of two intrinsic problems of the neuro-linguistics research program, the ontological incommensurability and granularity mismatch problems; together the *Mapping Problem* (MP). The first refers to the fact that linguistics and neurobiological ontologies rely on different foundational entities; the latter is that these ontologies have spatial and temporal scales that are not matched (Fig.1.2).

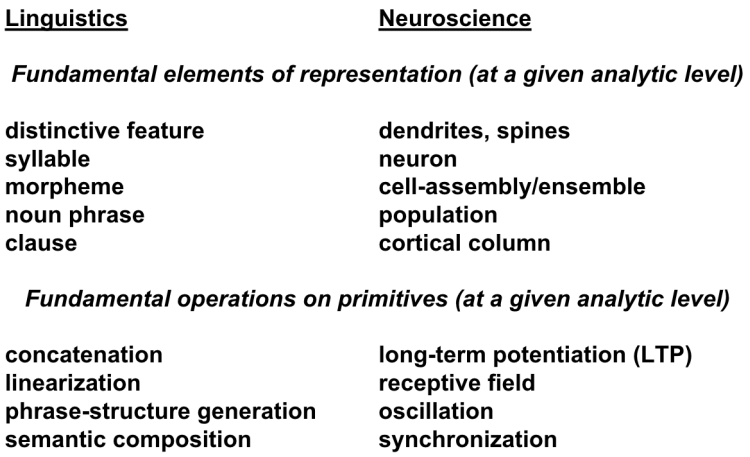



Figure 1.2: **Fundamental ontological units of linguistics and neuroscience** The lists provide a canonical inventory of neurobiological and linguistic phenomena. Each domain is governed by principled relationships between the items, denoted as vertical connections. Contrarily, the interconnections across disciplines are arbitrary (horizontal connections). Figure from Poeppel and Embick (2005).

Cognitive computational models of human and sWR the speech units and words are represented as abstract categories with a transition probability, or in some cases, as nodes of a network (Grossberg, 2003; Hannagan, Magnuson, & Grainger, 2013; Marslen-Wilson, 1987; McClelland & Elman, 1986; Norris & McQueen, 2008). The nodes and the probabilities are underspecified biological entities. Thus it is not possible to trace these units in the brain activity (MP), nor it is not possible to derive any causal dynamics to govern their interaction (3M).

### 1.2.3 ... but dynamical systems models are

An alternative strategy to establish a mapping between linguistic and neuroscientific categories is to address Marr's computational level starting from a biologically constrained level of implementation. This can be achieved if the neural and linguistic processes are described in a dynamical systems framework. Dynamical systems express the evolution of a set of variables through differential equations and are extensively used in neuroscience (Gerstner, Kistler, Naud, & Paninski, 2014). For example, the dynamical systems are used to model the membrane potential of neurons, which evolve in physical time and have physical units of measurement. Thus, if the entities of linguistics could be reformulated in the dynamical systems framework, one could aim to map language to the ontology of brain parts with causal explanatory power.

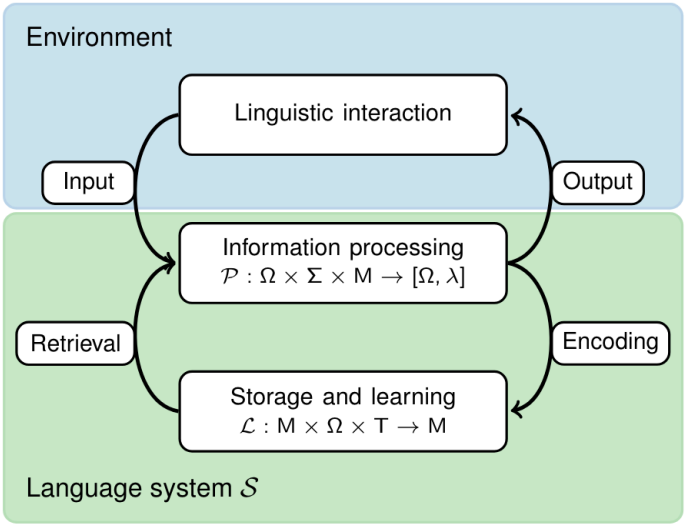



Figure 1.3: **Schematic of an adaptive information processing system for language**
The processing dynamics $P$, driven by input ($\Sigma$), state ($\Omega$), and model parameters ($M$), produces its own internal state and possibly a language output ($\lambda$). The processing is coupled with learning dynamics $L$, it enables information encoding and retrieval across timescales. On short timescales, $L$ acts as an active processing memory. Figure from Fitz et al. (2024).

One crucial step is to formulate a theory of linguistic processing that happens in physical time, as the temporal evolution of linguistic variables. This approach is called causal modeling of language, and it has recently been proposed by (Fitz et al., 2024) and investigated in a biological model of sentence processing by Uhlmann (2020). Language processing is expressed as an adaptive dynam-

ical system $S$, which couples a neural information processing device $P$ and an adaptive mechanism for storage and learning $L$ (Fig.1.3). The operations in $P$ are driven by linguistic inputs ($i_t \in \Sigma$), the internal state of the system ($s_t \in \Omega$), and the linguistic knowledge both acquired with previous experience and encoded in the human genome ($m_t \in M$). These three elements can be combined in a differential equation $P : ds = P(s, i, m)dt$ that yields a new internal state ($s' \in \Omega$) and, optionally, an output ($\lambda$). The adaptive mechanism $L$ follows takes a similar shape but operates on longer timescales. Its arguments are the internal states ($s_t$), the linguistic knowledge ($m_t$), and the developmental trajectory of the individual over time ($T_t$). The adaptive mechanism updates the speaker's linguistic knowledge which is used recursively to parse new stimuli.

In addition, the states and dynamics of a dynamical system can also be interpreted as the stages of a computation. The equations $L$ and $P$ are the infinitesimal version of the transition table of a Turing machine (Fitz et al., 2024). They describe the input-output transformations, the state updates occurring during the computation, and its memory structure. In an abstract sense, memory is an adaptive change in the system's state that encodes information about past events (Chaudhuri & Fiete, 2016). Computing devices often distinguish between a fast and volatile short-term memory (STM), and a more persistent long-term memory (LTM). In computing machines, memories are organized in data structures, which are necessary to store and re-access the memory; the data structures in which LTM and STM are organized in neurobiology are not expected to be the same. In dynamical systems, the organization in data structures is reflected in the hierarchy of dependencies and timescales among the coupled variables in the formalism. All the variables in the system evolve according to one or multiple timescales on the basis of which long and short-term memory can be distinguished. In word recognition, acoustic inputs ($i_t$) force the internal states along a trajectory $\mathscr{P}$ that is governed by processing memory of the phonological evidence that has accumulated ($s_t$), and the long-term memory of abstract linguistic categories previously acquired ($m_t$).

### 1.2.4 A biologically constrained model of word recognition

The dynamical system view ties together brain operations, linguistic processing, and computing machines. In theory, it provides a solution to the ontological incommensurability problem. However, the problem of granularity mismatch remains open. Because of their different scales and dynamics of interaction, it is not trivial to determine which linguistic and neuroscientific elements can be

mapped (Fig.1.4). Moreover, it is not clear how to evaluate the adequacy of this mapping. To solve this conundrum, I adopt three design choices common in the physical sciences, which determine the model's structure and scope and will help define the biophysical model.

- First, the model aims to elucidate the computations supporting the capacity of word recognition. It should only include those biological and linguistic elements that are necessary to it.
- Second, the model's scope is limited to a range of timescales. Only neurobiological and linguistic entities that evolve within the timescales of word recognition enter the model.
- Third, the model is a computational simulation. The simulation integrates biophysical equations whose descriptive accuracy has to be independently verified. Their evolution through time defines an input-output transformation, the linguistic computation.

The remainder of this introduction attempts to flesh out these design requirements, drawing on experimental work in psycholinguistics and neuroscience.

## 1.3 Representations and computations in spoken word recognition

### 1.3.1 From sounds to abstract representations

At the computational level of analysis, word recognition is often divided into two stages of processing, pre-lexical and lexical (McQueen, 2007; Scharenborg, Norris, Bosch, & McQueen, 2005; Vitevitch et al., 2018). The pre-lexical stage categorizes the input and transforms the sensory experience of the acoustic signal into a discrete perceptual experience of the linguistic sign (Liberman, Harris, Hoffman, & Griffith, 1957; Warren, 1970). The second computational stage concerns word retrieval from the mental lexicon. At this stage, word forms are accessed based on the pre-lexical and contextual evidence, if available (McQueen, 2005). The word form memory that optimally fits the phonological and contextual constraints is selected, and the corresponding lexical item is retrieved. (Hagoort, 2019; McQueen, 2005; Scharenborg et al., 2005; Vitevitch et al., 2018)

The pre-lexical stage infers abstract representations, such as phonemes, despite the idiosyncrasies of speakers and variable acoustic contexts. This challenge is known as the 'lack of invariance' problem (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). A shared view in the psychology of language is that speakers solve the problem through an active process of speech normalization. They infer phonological categories from the contingent sound, access word memories based on these experience-independent categories (Eisner & McQueen, 2018; Nusbaum & Magnuson, 1997), and parse the speech sounds into a hierarchy of linguistic representations (Jackendoff, 2007; McQueen, Cutler, & Norris, 2006). However, the form of these abstractions is still debated, and it is not yet clear whether the categorical representations that linguists postulate have one-to-one correlates in the neural processes that support speech comprehension.

Because of the uncertain nature of pre-lexical forms, a more coarse-grained classification divides them into segmental and supra-segmental categories. The segmental features are phoneme-sized and the suprasegmental include syllables, prosodic words, lexical stress patterns, and intonational phrases, which are necessary to distinguish among words in certain languages (McQueen, 2005). Their neural correlates are organized in an ascending hierarchy of spectro-temporal complexity, in which both linguistic and non-linguistic categories can be traced (e.g., temporal landmarks and speaker identity, Berezutskaya, Freudenburg, Güçlü, van Gerven, & Ramsey, 2017; Evans & Davis, 2015; Formisano, De Martino, Bonte, & Goebel, 2008; Hullett, Hamilton, Mesgarani, Schreiner, & Chang, 2016). Although the neurobiological evidence is inconclusive on which are the *right* pre-lexical representations that a word recognition model should consider, the spatial and temporal scales at which segmental and suprasegmental processes take place are clearer (Formisano, 2019; Sjerps & Chang, 2019; Yi, Leonard, & Chang, 2019).

Both speech and other sounds activate the primary auditory cortex (PAC, located on Herschl's Gyrus in humans), but only sounds whose spectro-temporal structure is sufficiently complex (such as phonetic features) induce activation of the Superior Temporal Gyrus (STG). The HG and the STG are anatomically adjacent, and the neural signals propagate within 50 ms from the PAC to the STG. When the speech chunks are sufficiently long (on the order of 1 s, thus containing suprasegmental information) speech-selective brain activity also invades the Superior Temporal Sulcus in a spatiotemporal pattern that moves ventromedial from the posterior STG (120 ms) to reach the mid-STS within 250 ms.

In this time period, the ventromedial gradient of activity moves towards the anterior STG and Middle-Temporal Gyrus (MTG), where the neural correlates of word form memories activate (Armeni, Willems, van den Bosch, & Schoffelen, 2019; Cibelli, Leonard, Johnson, & Chang, 2015; Ojemann, 2013; Ojemann, Schoenfield-McNeill, & Corina, 2009).

### 1.3.2 Unification of phoneme sequences into words

The retrieval of word forms from abstract representations is achieved in the lexical stage of word recognition. While the pre-lexical stage comprises a complex taxonomy of representations, the lexical stage is dominated by a single one, i.e., word form (McQueen, 2007). The lexical stage is divided into three phases that unfold in a cascaded, incremental stream. Words are pre-activated in parallel (lexical access), and when sufficient evidence is available one of the lexical candidates is selected and integrated into the sentence meaning (lexical selection, Marslen-Wilson & Zwitserlood, 1989; McQueen, 2007). Unless contextual information already excludes some of the candidates, lexical selection usually occurs at the word's *uniqueness point*, that is, the first speech unit that allows one to distinguish between the given word and other competitors in the lexicon.. Eventually, the lexical item is integrated into the sentence context (Hagoort, Hald, Bastiaansen, & Petersson, 2004).

Word form memories are accessed based on the correspondence between the pre-lexical evidence and the phonological prototype of the word form in the lexicon. In most modern languages, the ratio between the items contained in the lexicon and the number of segmental and suprasegmental speech units in the language is strikingly high. One immediate consequence of the wealth of word forms is that multiple words share the same pre-lexical features and can be distinguished only based on their sequential order. Some words are even contained in other words. For example, the lexicon of an average English speaker contains on the order of 50 thousand words, composed of less than 50 speech sounds, among which PEST (pɛst) and PETS (pɛts), and BONE (bəʊn) and TROMBONE (trombəʊn). Thus, activation of word form representations requires the composition of pre-lexical features which must be sensitive to serial order[3], and integrates information at multiple timescales (e.g., segmental and suprasegmental).

[3]The composition of pre-lexical elementary linguistic units can also be viewed as a form of Unification over the pre-lexical memories activated by the speech input (Memory, Unification, Control model, Hagoort, 2013)

Crucially, composition is achieved on short-lived, transient information that requires a processing memory.

The fact that the same phonemes presented in different order must be mapped to different words is sometimes referred to as the Temporal Order Problem (TOP) (Magnuson, Mirman, & Myers, 2013). Computational models of word recognition must be able to solve it and retrieve words with full phonological overlap. The most common algorithmic solution consists of adding an extra feature to the pre-lexical unit that specifies the position in the word, for the onset time of the phoneme. We refer to these units as time-coded or position-dependent pre-lexical representations. In connectionist models (e.g., (TRACE McClelland & Elman, 1986) and (TISK Hannagan et al., 2013)), these units, be they acoustic phones, allophones, or phonemes, activate supra-segmental nodes that use the temporal features of the pre-lexical representation to select the correct word form.

Crucially, the assumption of position-dependent representation does not reflect experimental evidence on the neural correlates of speech integration. For example, Gwilliams, King, Marantz, and Poeppel (2022) shows that phonetic content is encoded similarly at different serial positions within a word, indicating that position-specific (or context-dependent) features may not be part of the neurocomputational machinery for word recognition. A possible way out from the impasse is proposed by Yi et al. (2019). The authors suggest that the acoustic and temporal landmark features resolved in the STG are combined by a *specialized temporal circuit*, which time-stamps the phonemic features and makes them position-dependent units. However, the model is only a sketch and it is not demonstrated that this computation can be carried out in neurobiology. In any case, the synthesis of position-dependent features by itself does not explain how these features are maintained in processing memory, nor how their composition is achieved. Fig.1.4*B*, borrowed from Gwilliams, Linzen, Poeppel, and Marantz (2018), highlights that these are missing pieces of the puzzle in our understanding of human word recognition.

### 1.3.3 What must be explained?

I conclude this section with a brief recapitulation of the linguistic elements and processes that must be included in the model to be explanatory of the word recognition capacity.

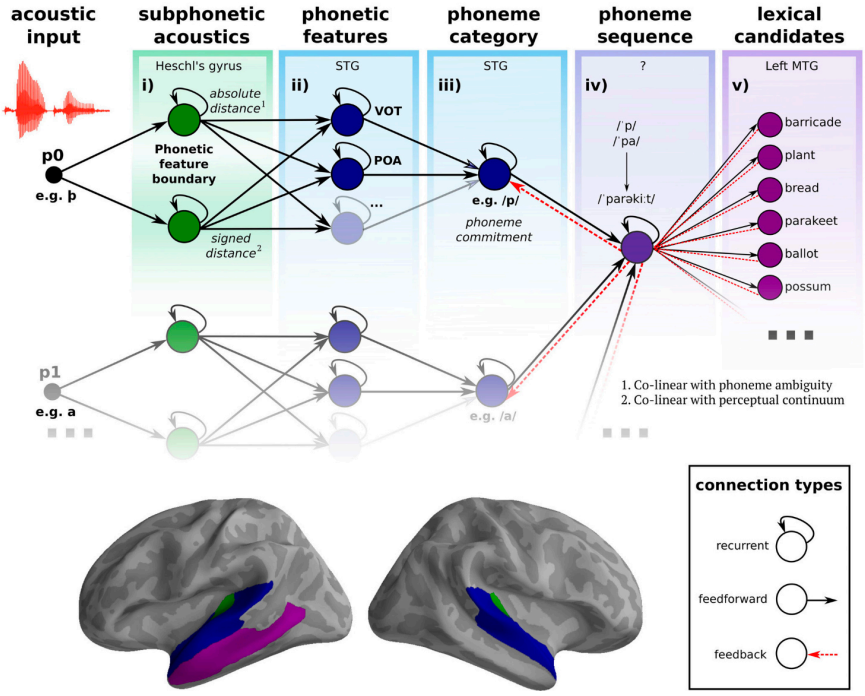



Figure 1.4: **Algorithms for combining phonological information into word forms**
The top panel shows the integration of a hierarchy of acoustic and linguistic features in a word recognition model by Gwilliams et al. (2018). The acoustic input is parsed into linguistic units of increasing granularity, up to the phoneme level. Phonemes are then composed into sequences in *iv*. It is not known in which cortical region this operation occurs (bottom panels), nor which neural mechanisms support it.

- Linguistic units activate representations distributed across the entire STG, suggesting that multiple and overlapping neural populations are co-activated for each linguistic unit (Mesgarani, David, Fritz, & Shamma, 2014; Yi et al., 2019).

- The timescales of these representations are on the order of 20 ms to 150 ms for the pre-lexical units (Gwilliams et al., 2018) and 200 ms to 500 ms for lexical units (Tucker et al., 2019).

- The computation that supports the retrieval of word memories from pre-lexical features, e.g., phonemes, is a many-to-many mapping with parallel and fast access to several word forms (Allopenna, Magnuson, & Tanenhaus, 1998; Marslen-Wilson & Welsh, 1978).

- The selection of word forms is mediated by competition between lexical neighbors (Luce & Pisoni, 1998; McQueen, Norris, & Cutler, 1994).

- Access and selection rely on fleeting sensory information that is maintained in processing memory. Because of phonological overlap, this memory must also encode the order in which the pre-lexical representations were activated (Magnuson et al., 2013; Yi et al., 2019).

## 1.4 Brain anatomical elements and computations

The brain's anatomy and physiology are the second term of the mapping hypothesis I intend to formulate. In accordance with the *Neuron doctrine* (Cajal, 1954), neurons are the fundamental computational units of the brain, and neural networks are the engine of information processing. Following this assumption, I will introduce the elements of processing in neuronal systems and brain anatomy. The interested reader can find an exhaustive description in Kandel et al. (2012), Gerstner et al. (2014), and Braitenberg and Schüz (2013).

### 1.4.1 Elements of neuronal systems

Neurons are large cells with a central body, the **soma**, and hundreds of extensions that branch from the soma to connect with other cells. One of these branches is the axon, which transmits electrical signals integrated within the soma to the neighboring cells. The other branches are called **dendrites** and, normally, receive inputs from the axons of other cells. The connections between axon-terminals and dendrites are called **synapse**. The synapse is the site where the pre-synaptic neuron makes contact with the post-synaptic cell.

When a cell is sufficiently stimulated, it emits an **action potential** and its axon terminals release a bulk of **neurotransmitters** on the post-synaptic cells. The neurotransmitters bind to a group of proteins sitting on the post-synaptic cell, the **synaptic receptors**, and favor the influx of ions in the membrane. The amount of neurotransmitters and receptors available in the synapse determines the **strength** of the pre-post synaptic interaction. This is commonly referred to as synaptic strength. Depending on their electric charge, the ions will increase or decrease the **membrane potential**. When the membrane potential of the soma reaches a certain threshold, it will lead to an action potential so that the process starts again.

Because the action potentials are all-or-none events, they carry binary information about the state of the neuron. However, post-synaptic cells also have access to the **firing rate** at which spikes are emitted. The firing rate is one of the

main indicators of the state of a cell or a population of cells in the brain network. Beyond offering a read-out of the neuronal states, the firing rate plays a role in **synaptic plasticity**, that is the process governing the increase and decrease in synaptic strength between two cells. The main form of plasticity described in neural networks is associative, or Hebbian (Hebb, 1949). Cells that are co-activated, have a large firing rate, tend to strengthen their reciprocal synapses and form **engrams**. The group of cells that is tightly connected after associative plasticity is named a **cell assembly**.

Brain networks contain several types of neurons, that divide by their electro-physiological properties and the types of connections they make with other cells. The main distinguo is between **excitatory** and **inhibitory** neurons. The former creates synapses that, upon the arrival of an action potential, **depolarize** the post-synaptic cell. Excitatory transmission is mediated by **glutamate**, and thus the excitatory cells are also called glutamatergic. In contrast, inhibitory cells release the GABA neurotransmitters, which bind to receptors that favor the influx of negative ions. They **hyper-polarize** the cell and prevent it from firing an action potential.

The processes hitherto described are relatively generic, and they apply to the nervous system of most of the animal kingdom. On the other hand, language, and spoken word recognition, is an exquisitely human capacity. To understand the differences between humans and non-human animals we have to take a step backward and look at the brain anatomy.

### 1.4.2 Neurobiological underpinnings of word recognition

The human brain comprises a network of 90 billion neurons tied together by a hundred trillion synapses. (Braitenberg & Schüz, 1998a). The brain structure that the most distinguishes humans from non-human animals is the neocortex (Molnár & Pollen, 2014). The neocortex is also considered the substrate of linguistic memories (Formisano et al., 2008; Sjerps & Chang, 2019; Yi et al., 2019). Here, cells are ontogenetically organized in micro-columns of hundreds of cells (Mountcastle, 1957, 1997), which emerge during the formation of the cortical sheet, a thin layer of 2 mm to 4 mm which contains most of the cortical cell bodies (Adesnik & Naka, 2018; Leuze et al., 2014; Senzai, Fernandez-Ruiz, & Buzsáki, 2019). Within the cortical sheet, neurons are divided into three main layers: granular, infragranular, and supragranular. Layer structure varies across the entire cortex, and the forebrain regions can be characterized by the type of cell body and connective tissue they contain, i.e., cyto- and myelo-architecture,

and the receptor types these cells express (Palomero-Gallagher & Zilles, 2017; Zilles & Amunts, 2009).

A valuable approach to determine which of these components should be considered for a biological model of word recognition is to look at the spatial and temporal scale of the cognitive problem and then parcellate the anatomical hierarchy into its characteristic scales ( Fig.1.5). Brain timescales range from tens of microseconds, at which molecular signaling occurs, to tens of seconds, at which the brain acts as a single complex system, to the lifetime of linguistic, episodic, and semantic memories. Notice that processes that live on long timescales also exist at shorter ones and the reverse; however, with certain precautions, the evolutions of dynamical systems at different timescales can be considered disentangled. In mathematical terms, this is achieved by separating equations and presenting slower or faster processes as effective parameters rather than variables (Sec. 4.6, Gerstner et al., 2014). The result is that the systems are simplified and can be considered, to an extent, in isolation.

The schematic in Fig.1.5 highlights six anatomical units that live on the hundreds of millisecond timescales; synapses, dendrites, neurons, cortical layers, cortical columns, and local circuits. All these elements are contained within a brain region, such as the temporal lobe or perisylvian region, and span a spatial scale of a few centimeters. In addition, we have to consider the two other physiological processes that enable neurons to communicate with each other and learn from experience, action potentials and synaptic plasticity.

### 1.4.3 Models of neuronal networks

The separation of the brain's spatial and temporal scales indicates that a neurobiological model of word recognition must account for the activity of neurons in the Superior and Middle Temporal Gyri. However, we do not have sufficient specifications of the cell types and connectivity patterns in this area of the human brain to constrain our network model. Similarly, it is beyond our capacity to simulate a replica of the actual system. Instead, I implement a *generic* network model which aims to account for the fundamental capacity of word recognition listed in Section 1.3.3. As I will demonstrate throughout the thesis, such a simplified model already provides precious insights into which physiological processes may support lexical access and selection.

The core of the network is the neuron model. Neuron models account for the dynamics of the membrane potential(Fig.1.6*A*, Hodgkin & Huxley, 1952; Koch, 1998). A realistic neuron model typically has several variables that describe the

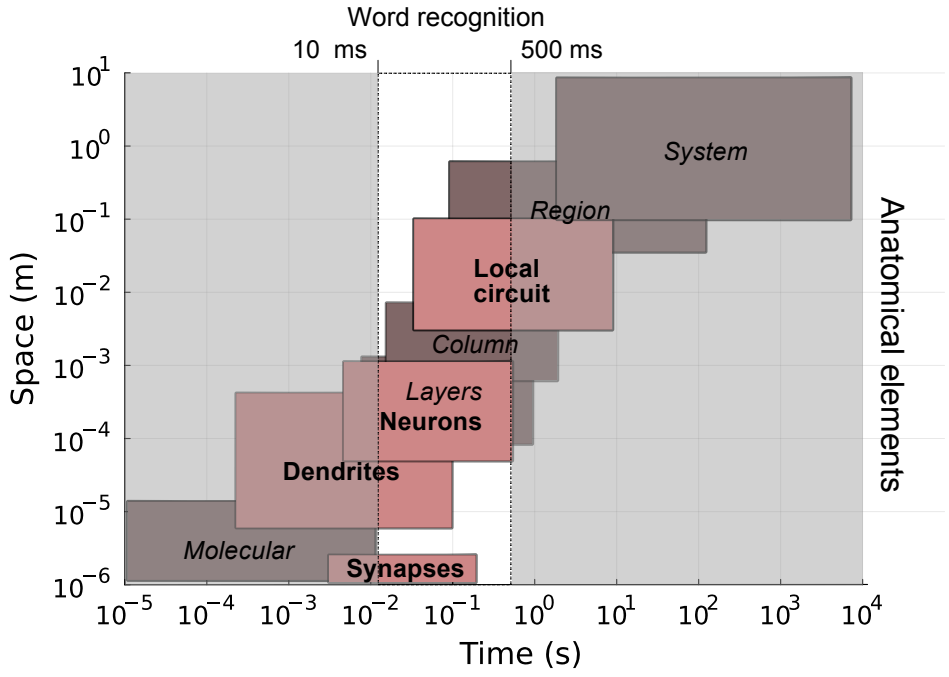



Figure 1.5: **Elements of brain neurobiology** The organization of the nervous system exhibits temporal and spatial scales. There is a significant overlap in scales, as dendrites, cells, and columns operate on similar temporal and spatial levels. Because word recognition occurs within 10 ms to 500 ms from the acoustic stimulus, we select only those elements that live in this temporal range. The elements in red shade, with boldfaced font, are the anatomical elements included in the model described in the following chapters. Image adapted from Lytton et al. (2017)

membrane potential in different regions of the cell. However, they are computationally hard to integrate, which puts limits on their use in network models. On the other end of the spectrum, reduced neuron models describe a cell with only a few differential equations that focus on the evolution of the membrane potential at the axon hillock (Brette & Gerstner, 2005). These models are called point neurons and constitute the common implementation of single cells in biological neural networks.

Despite their mathematical convenience, point neurons neglect the role of dendrites in neuronal integration and only model the spike generation mechanism at the soma. According to recent experimental and computational research, this view is excessively simplistic and may lead to overlooking the computations carried out within the cell (Fig.1.6*B*, Larkum, 2022). For example, it has been shown that dendrites foster the interaction among synapses sitting on the same branch and exhibit non-linear responses to pre-synaptic activity if their poten-

tial is sufficiently depolarized (Fişek & Häusser, 2020; Larkum, 2022; Payeur, Béïque, & Naud, 2019; J. Schiller, Major, Koester, & Schiller, 2000; Wilmes & Clopath, 2023).

The network model also defines the biophysical equations for the synapses that connect the neurons. Network models of information processing implement two types of receptors, glutamatergic and GABAergic. The former occurs when the pre-synaptic cell is excitatory, the latter when it is inhibitory. Within each class, there are fast (AMPA, $GABA_A$) and slow receptors (NMDA, $GABA_B$) (Fig.1.3*C, D*, Roth & van Rossum, 2009) Concerning the plasticity, the models follow the associative plasticity scheme described earlier. Synapse strengthens when both cells are active. A more refined form of plasticity accounts for the causal role of the pre-synaptic cell in leading the post-synaptic to fire, and include the time of the spike. This is called spike-time-dependent plasticity (STDP, Clopath, Büsing, Vasilaki, & Gerstner, 2010).

Eventually, one has to indicate the types and number of neurons and the density of neural connections Fig.1.3*E*. Because of the computational constraints discussed, network models are normally much smaller than actual brain networks. However, if the network is above a certain scale (i.e., $10^3$—$10^4$) the model can still account for the emergence of network dynamics observed in neuronal circuits, for example, excitatory/inhibitory balance (Brunel, 2000; Hiratani & Fukai, 2017; Litwin-Kumar & Doiron, 2014; Renart et al., 2010; Zajzon, Duarte, Mahmoudian, Morrison, & Duarte, 2019).

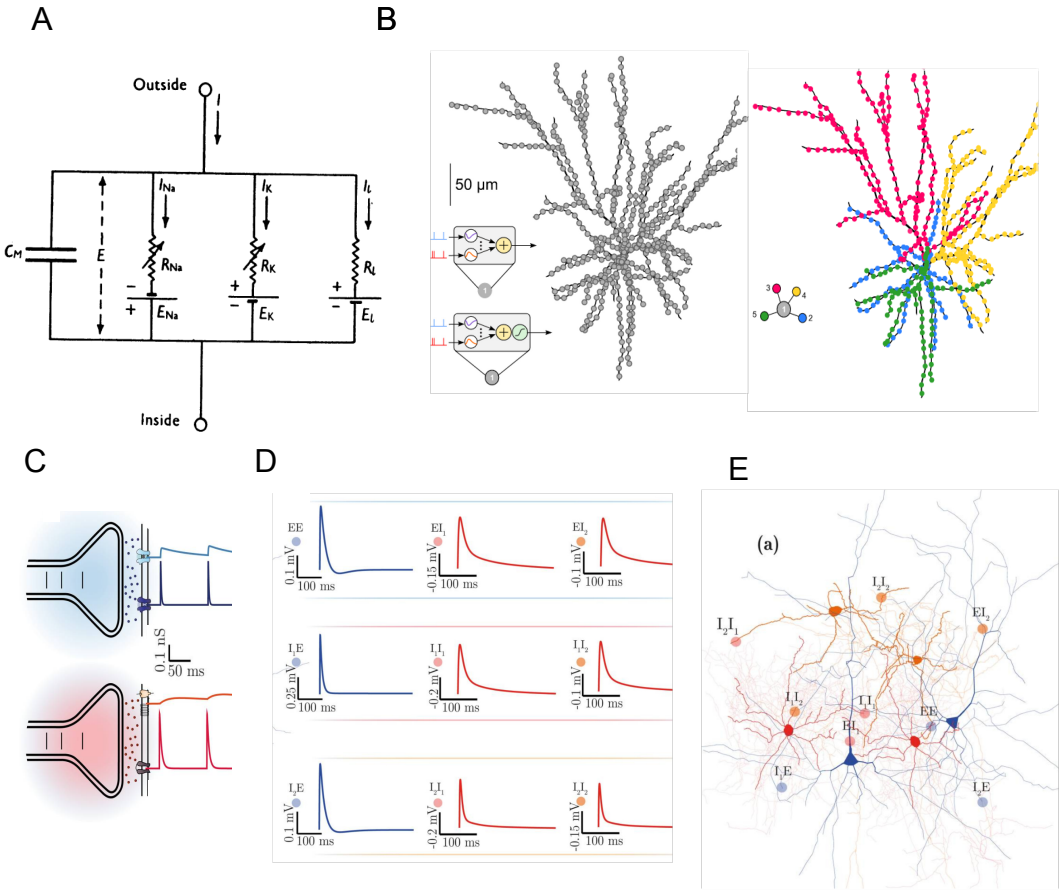



Figure 1.6: **Models of the anatomical elements supporting word recognition** (**A**) Model of action potential initiation by Hodgkin and Huxley (1952). The model describes the neuronal membrane as an electrical circuit, it introduces a system of differential equations to describe the evolution of the ion channels' opening. (**B**) Excerpt from the study by Ujfalussy et al. (2018), the panel shows two possible reductions of a complex, realistic dendritic tree. The model on the left is a point neuron reduction, and the model on the right is a four-compartment reduction. The study demonstrates that few compartments are sufficient to explain most of the computations taking place in the realistic model. (**C**) Network models implement four main receptor types. The curves in dark and light blue represent the fast and slow excitatory receptors (AMPARs and NMDARs). The receptors in dark and light red represent the inhibitory ones ($GABA_A$ and $GABA_B$ ). (**D**) The timescales of these receptors change depending on the neuron types connected by the synapse. The illustration provides the timescale of the AMPA and $GABA_A$ receptors (same colors in *C*) when it applies to excitatory (*E*) or inhibitory ($I_1$, $I_2$) cells. (**E**) An example of a small circuit with one type of excitatory neuron (blue) and two inhibitory neurons (red, orange), the neurons are connected by synapses with type-specific properties. Panels *C, D,* and *E* are adapted from Duarte and Morrison (2019).

### 1.4.4 Biologically constrained models of associative memories

The models presented describe the dynamics of brain networks and are useful for formulating hypotheses and computational experiments concerning the neural substrate of cognitive operations. For example, they have been used to prove that biological networks with Hebbian plasticity support the formation of associative memories in the form of cell assemblies (Litwin-Kumar & Doiron, 2014; Zenke, Agnes, & Gerstner, 2015). This result is remarkable because it bridges a long-standing theory of animal – and human – cognition with the activity of single cells in a brain-like network (Amit, 1995; Fuster, 1997; Ojemann & Schoenfield-McNeill, 1998).

Importantly, cell assemblies have also been indicated as the possible substrate of word memories (Pulvermüller, 1999). Garagnani, Wennekers, and Pulvermüller (2009) and Tomasello, Garagnani, Wennekers, and Pulvermüller (2018) have developed biologically constrained networks that implement lexical memories through cell assemblies. The models account for the organization of lexical memories across brain areas and can give mechanistic insights concerning the role of motor and sensory areas in cognition.

However, these models of lexical memories, akin to those of associative memories, do not consider the interaction of different cell assemblies: each memory is stored independently of the others. This is a general issue in biological models of memory, they struggle to implement any cognitive operation that goes beyond simple associations. Gallistel (2021) presents a severe critique of the problem. In the author's view, biological models that express memories through the synaptic junction fail to explain how living beings store relational memories. In the case of word memories, the relationship is the order of phonemes. This view is shared in psycholinguistics, Poeppel and Idsardi (2022) argued that there are no implementation proposals for storing word memories and, consequently, our understanding of the neural organization of word memories *[...] is somewhere between unsatisfactory and incoherent*.

The limits of present models of associative memories –namely, the lack of mechanisms to encode relationships and solve the Temporal Order Problem– is the starting point of the present thesis. The present study introduces dendritic structure in a biological model of associative memories. The remainder of this thesis will show that this is sufficient to foster the formation of network memories that are sensitive to the order of the input presented, thus encoding order relationships.

## 1.5 Overview of the thesis

The present thesis is structured into a General Introduction, four experimental chapters, and a General Discussion. Each of the experimental chapters is intended as a journal article. One of them has been published, one has been submitted and the remaining two are in preparation. The content of the experimental chapters is the following.

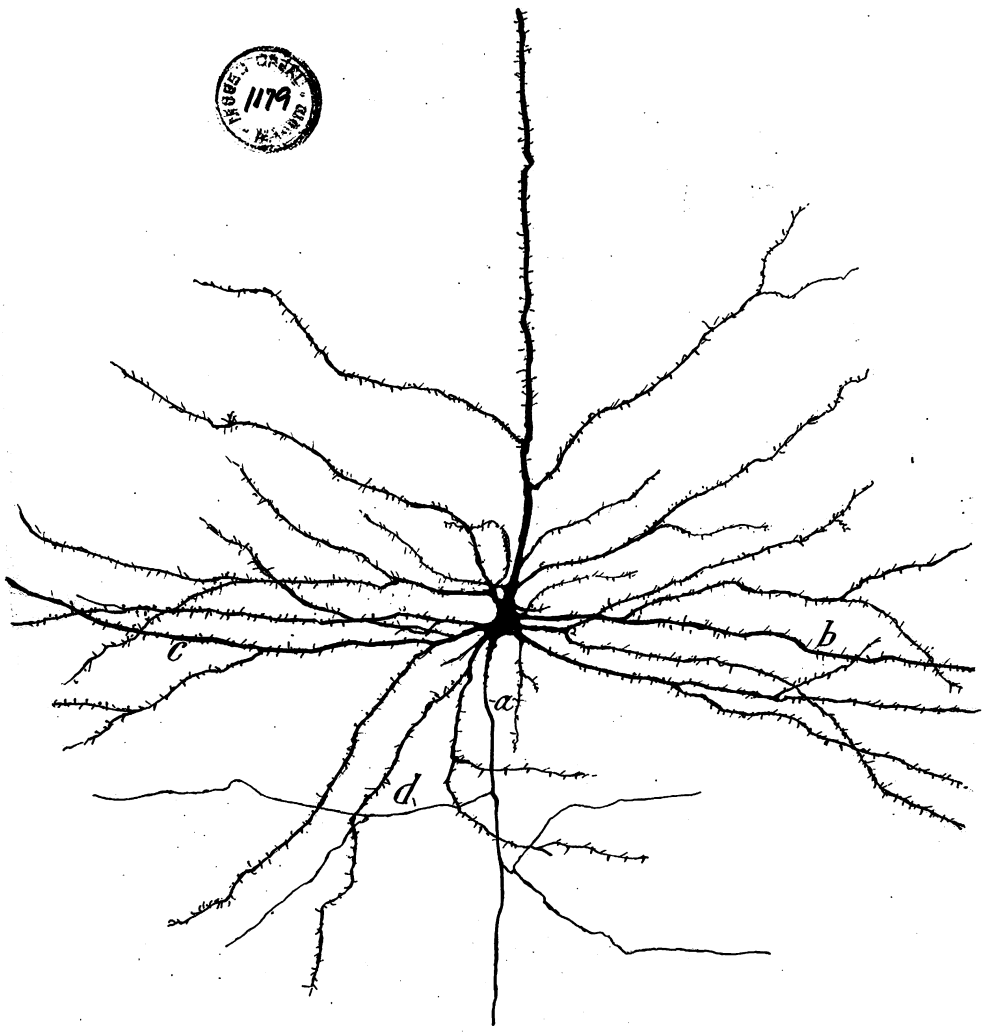
**Chapter 2** investigates dendritic integration in a novel reduced model of the pyramidal cell, the Tripod neuron. The model is endowed with two segregated dendritic compartments and NMDA and $GABA_B$ receptors on each with parameters from human cortical cells. The combination of segregation and long synaptic timescales allows the model to reproduce computations that are inaccessible to point-neuron models, such as coincidence detection, on-path inhibition, non-symmetric logical operations, and temporal integration. Temporal integration is supported by the slow decay of dendritic membrane potential, which I refer to as dendritic memory.

The first chapter individuates computational differences in the model but does not assure that these hold when the Tripod is tested in more naturalistic conditions. This analysis is achieved in **Chapter 3**. Here, the model was studied in response to a balanced stream of excitatory and inhibitory inputs. The study shows that the dendritic integration also offers the substrate for the high-conductance state and the UpDown cortical dynamics.

The Tripod model is investigated in the context of a network of neurons in **Chapter 4**. Here I study a network of Tripod neurons and inhibitory cells. The network implements excitatory and inhibitory STDP. In the chapter, I analyze the capacity of the network model to form hetero-associative memories and recognize sequences. The memories represented pre-lexical and lexical units in the form of distributed cell assemblies. I individuate in dendritic memory and in the specific network structure that emerges from the plasticity protocol the mechanism that supports the recollection of sequential memories.

Finally, in **Chapter 5**, the network model is compared to computational and behavioral results in the psychology of language. The study indicates that the model is accurate in predicting classical results in word recognition, such as the incremental and cascade nature of the process, the competition among lexical neighbors, and the robustness to variability in the inputs. In addition, the results indicate that word recognition can be achieved in biologically plausible networks without position-specific representations and open a novel perspective on the computational requirement of pre-lexical representations.

# 2 | The Tripod neuron, a minimal structural reduction of the dendritic tree



## Abstract

Neuron models with explicit dendritic dynamics have shed light on mechanisms for coincidence detection, pathway selection, and temporal filtering. However, it is still unclear which morphological and physiological features are required to capture these phenomena. In this work, we introduce the Tripod neuron model and propose a minimal structural reduction of the dendritic tree that is able to reproduce these dendritic computations. The Tripod is a three-compartment model consisting of two segregated passive dendrites and a somatic compartment modeled as an adaptive, exponential integrate-and-fire neuron. It incorporates dendritic geometry, membrane physiology, and receptor dynamics as measured in human pyramidal cells. We characterize the response of the Tripod to glutamatergic and GABAergic inputs and identify parameters that support supra-linear integration, coincidence-detection, and pathway-specific gating through shunting inhibition. Following NMDA spikes, the Tripod neuron generates plateau potentials whose duration depends on the dendritic length and the strength of synaptic input. When fitted with distal compartments, the Tripod neuron encodes previous activity into a dendritic depolarized state. This dendritic memory allows the neuron to perform temporal binding and we show that the neuron solves transition and sequence detection tasks on which a single-compartment model fails. Thus, the Tripod neuron can account for dendritic computations previously explained only with more detailed neuron models or neural networks. Due to its simplicity, the Tripod model can be used efficiently in simulations of larger cortical circuits.

## 2.1 Introduction

Biological neurons integrate complex afferent inputs within a dendritic structure which accounts for most of the spatial extent of a neuron. The dendritic arborization hosts a significant part of excitatory and inhibitory synapses and processes the input signals before the resulting signal reaches the cell body and in particular the axon hillock. In the dynamical systems theory of neural information processing, neurons function as non-linear, non-stationary (and stochastic) operators, and the dendrites determine important aspects of the neurons' transfer characteristics (Gidon et al., 2020; Larkum, Wu, Duverdin, & Gidon, 2022; Payeur, Guerguiev, Zenke, Richards, & Naud, 2021; Poirazi & Papoutsi, 2020; Stuart & Spruston, 2015).

Neuron models that explicitly consider the dynamics of the dendritic tree are typically referred to as multi-compartment models. These models capture the spatio-temporal dendritic dynamics by introducing additional state variables and differential equations that describe the dynamics of the dendritic membrane potential (Koch, 1999). Depending on the implemented dendritic architecture, membrane dynamics, and receptors/ion-channel repertoire, high-resolution multi-compartmental models can reproduce the membrane physiology in detail (Branco, Clark, & Häusser, 2010; Ujfalussy et al., 2018; Winnubst & Lohmann, 2012). Simulations with neuron models including dendrites shed light on important problems of brain functions, including unsupervised learning (Bono & Clopath, 2017; Payeur et al., 2021), signal filtering (G. R. Yang, Murray, & Wang, 2016), temporal discrimination (Branco et al., 2010), coincidence detection (Mel, 1992; Poirazi, Brannon, & Mel, 2003), structured sequence processing (Ahmad & Hawkins, 2016; Haga & Fukai, 2018), and the creation and maintenance of associative memories (Kastellakis, Silva, & Poirazi, 2016). This body of evidence suggests that dendritic processing is fundamental to nervous system computation. However, the computational cost of simulating detailed multi-compartment models impedes their use in large networks. Thus, most studies that analyze processing properties in large networks do not explicitly consider dendritic structure but often use simpler point-neuron models instead. These studies regard neural computation as the outcome of the particular network structure used, disregarding the complexity of cell-internal processes (Bastos et al., 2012; Duarte & Morrison, 2019; Haeusler, Schuch, & Maass, 2009; Potjans & Diesmann, 2014).

The present work introduces a computationally efficient, three-compartment model that includes relevant dendritic degrees of freedom and remains simple

enough to be used in larger network simulations. This model, which we call the Tripod neuron, is derived from previous theoretical and experimental work, and three main ingredients define its dynamics. First, the Tripod has two dendritic compartments. This is the minimum number of dendritic compartments, in addition to the somatic compartment, which allows a branching dendritic tree. Several studies have shown that relatively few dendritic degrees of freedom are sufficient to reproduce the non-linear integration effects of apical dendrites in pyramidal cells (Larkum, 2013; Poirazi et al., 2003). Accordingly, an extensive comparison of the number of dendritic compartments to mimic *in-vivo* dynamics indicates that two compartments are sufficient to explain most of the observed variability in the somatic membrane potential (Ujfalussy et al., 2018; Wybo et al., 2021-01-26, 2021), and models with more than two dendritic compartments show modest qualitative differences (Ahmad & Hawkins, 2016; Bono & Clopath, 2017; Kastellakis et al., 2016). Secondly, the internal dynamics of the Tripod neuron are consistent with observed neurophysiology. The dendritic structure consists of two isolated compartments connected to the somatic compartment. Each compartment integrates fast and slow excitatory and inhibitory inputs locally through conductance-based synapses, and we show that a simple circuit approximation (Koch, 1999) suggests that a single degree of freedom, the electrotonic distance from the soma, determines an integration time-scale of the dendrites and analytically defines two types of compartments, here called short and long dendrites. Finally, we investigated slow voltage-dependent NMDA receptors that mimic an important property of dendritic computation. When the post-synaptic potential exceeds a certain threshold, the NMDA receptors open to $Ca^{2+}$ ions and boost post-synaptic membrane depolarization, generating a so-called NMDA spike, or plateau potential (Antic, Zhou, Moore, Short, & Ikonomu, 2010; Mel, 1992; Tabone & Ramaswami, 2012). This non-linear phenomenon, along with self-regenerative events such as back-propagating spikes (Rapp, Yarom, & Segev, 1996) in proximal dendrites, enrich the computational toolkit of the dendrites and determine the most interesting properties of the present model. The slow voltage decay of the dendritic potential provides a short-term dendritic memory which is not accounted for by other adaptation mechanisms in single-compartment models, for example (Brette & Gerstner, 2005; Fitz et al., 2020). This aspect of our work complements previous studies of NMDARs in models with a small number of compartments (Bono & Clopath, 2017; Mel, 1992; G. R. Yang et al., 2016), and provides a basis for further explorations of the role of NMDA spikes in neuronal working mem-

ory (Fitz et al., 2020; Wang, 1999), and temporal binding (Augusto & Gambino, 2019; Baggio & Hagoort, 2011).

## 2.2 Methods

### 2.2.1 The Tripod neuron model

The *Tripod* neuron is composed of three separate computational elements or compartments. It has an axosomatic compartment, representing the soma and perisomatic locations, and two electrotonically segregated dendritic compartments coupled to the soma in a Y-shape (Fig.2.1).

***Axosomatic compartment.*** The soma was modeled as an adaptive exponential integrate-and-fire (AdEx) neuron (Brette & Gerstner, 2005). It is a two-dimensional neuron that models the dynamics of the somatic membrane potential $V^s$ and an adaptive current $w$:

$$\begin{aligned} C_m^s \frac{dV^s}{dt} &= -g_m^s \left[ (V^s - V_r) + \Delta_T \exp \frac{V^s - V_T}{\Delta_T} \right] + \\ &\quad - \sum_k g_k(t)(V^s - E_k) - w + I^{\mathrm{d}} \end{aligned} \tag{2.1}$$

$$\tau_w \frac{dw}{dt} = -w + a(V^s - V_r) \tag{2.2}$$

The leak conductance $g_m^s$ defines the permeability of the somatic membrane, $C_m^s$ its capacitance and $g_k$ the set of variable synaptic conductances (Fig.2.1*B*). The synaptic conductances and reversal potentials $E_k$ are further described in the section *Synaptic dynamics* below. We use the superscript $s$ throughout to refer to variables and parameters of the somatic compartment, whereas the superscript $d$ refers to dendritic compartments. The first equation of the AdEx neuron aims to reproduce the sub-threshold and spike-onset dynamics of pyramidal cells. For a membrane potential $V^s$ below the rheobase threshold $V_T$, the neuron behaves as a leaky integrator of the currents from the dendritic compartments $I_d$ and the somatic leakage conductances $g_k(V^s - E_k)$. For larger depolarizing events, the membrane potential exceeds the rheobase threshold $V^s > V_T$ and activates the exponential non-linearity, mimicking a spike-generation mechanism. The slope of the exponential growth is governed by $\Delta_T$. The spike events occur at times $t^f$ when $V^s$ exceeds the spiking threshold $u_{th}$. Afterward, the membrane potential is reset to $V_r$ and the adaptation current $w$ is increased by a constant value $b$.

The adaptation currently accounts for several physiological processes and decreases the excitability of the neuron after it has spiked. All parameters of the somatic compartment were fixed and set to the values used in Brette and Gerstner (2005), except for the somatic leak conductance which was set to 40 nS in agreement with the multi-compartment model of Bono and Clopath (2017), see Table 2.1. The reset potential of the AdEx model has been set to $u_r = -70.6\,\text{mV}$ as in (Brette & Gerstner, 2005) rather than to $u_r = -55\,\text{mV}$ (Bono & Clopath, 2017; Duarte & Morrison, 2019) so that the bursting behavior in the Tripod will depended only on the dendritic dynamics.

***Dendritic compartments.*** Dendritic compartments were approximated as conductive cylinders whose voltage was governed by a passive membrane-patch equation similar to the soma but lacking mechanisms for spike generation and intrinsic adaptation:

$$C_m^d \frac{dV^d}{dt} = -g_m^d(V^d - V_r) - \sum_k g_k(t)(V^d - E_k) - I_\text{d} \tag{2.3}$$

$$I_d = g_{ax}^d(V^d - V^s) \tag{2.4}$$

The current $I_d$ was computed as the potential difference between the dendritic and somatic compartment, multiplied by the axial conductance $g_{ax}^d$ (Fig.2.1*C*). Current flow was positive from the dendrites to the soma, $I_d > 0$, except when the somatic potential $V_s$ exceeded the firing threshold and the neuron emitted a spike. Consistent with Bono and Clopath (2017), we captured the backpropagation of somatic action-potentials by clamping $V_s(t^f)$ to 20 mV for 1 ms. The effect of the back-propagating action potential is illustrated in Fig.2.2**D**.

***Dendritic geometry.*** The capacitance $C_m^d$, leak conductance $g_m^d$, and axial conductance $g_{ax}^d$ of the dendritic compartments depended both on the geometry and the membrane properties. The macroscopic parameters $C_m^d$, $g_{ax}^d$ and $g_m^d$ can be

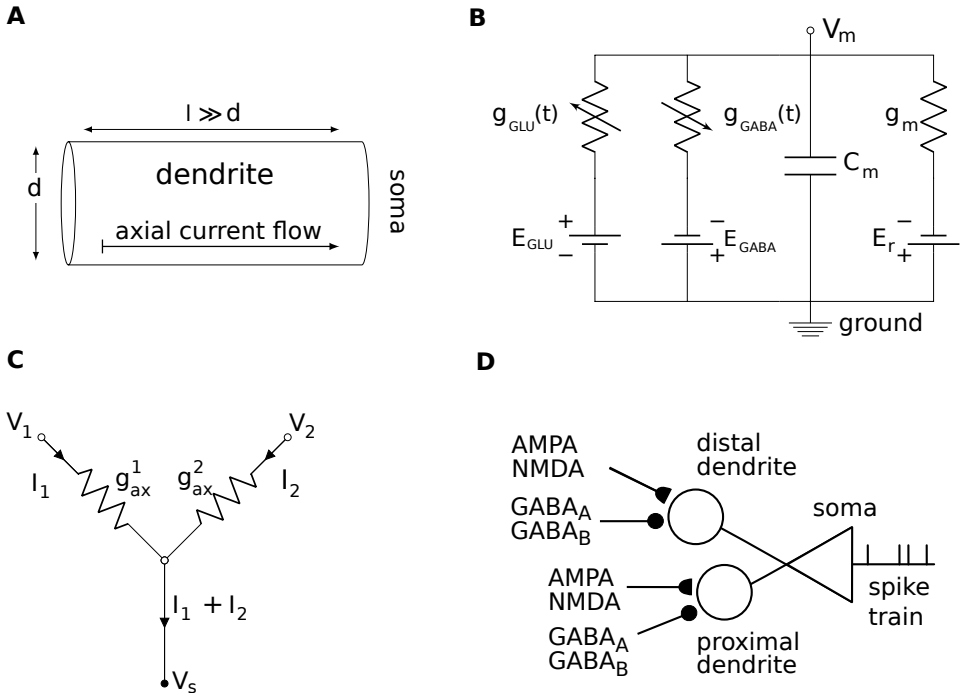



Figure 2.1: **Schematic of the Tripod neuron**
**(A)** Dendritic compartments were modeled as cylindrical segments of a cable with length $l$ and diameter $d$. Their electrical properties were set by the membrane patch equations (Eqs. 2.5, 2.6, 2.7) and membrane-specific parameters (Table 2.2). When dendrites had a larger potential than the soma, current flowed along the dendritic axis towards the soma. **(B)** Circuit diagram of a dendritic membrane patch with time-varying conductances across the membrane. Conductances were regulated by glutamatergic receptors $g_{\text{GluRs}}$ or GABAergic receptors $g_{\text{GABA}}$ with reversal potentials $E_{\text{GluRs}}$ and $E_{\text{GABA}}$, respectively (Table 2.4). The membrane reversal potential $E_r$ coupled in series with the leak conductance $g_m$ and the membrane acted as a capacitance $C_m$ with respect to the extracellular space (ground). The membrane potential $V_m$ was determined by the currents flowing to the dendritic compartment. **(C)** The dendritic potentials $V_1$ and $V_2$ were coupled to the somatic membrane $V_s$ through the axial conductances $g^1_{ax}$ and $g^2_{ax}$. The resulting current $I_1 + I_2$ flowed dromically from the dendrites to the soma. **(D)** The Tripod neuron with two dendrites and a somatic compartment. Each dendrite received synaptic input mediated by four types of receptors, AMPA, NMDA, $\text{GABA}_\text{A}$ and $\text{GABA}_\text{B}$ . Distal dendritic compartments were modeled using a smaller axial conductance compared to proximal ones. The spike-generating soma is represented as a triangle.

computed from the relative densities $c_m, r_{ax} and r_m$ via the standard cable-theory (Koch, 1999):

$$C_m = \pi c_m l d \tag{2.5}$$

$$g_m^d = \pi \frac{ld}{r_m} \tag{2.6}$$

$$g_{ax}^d = \frac{\pi}{4} \frac{d^2}{r_{ax} l} \tag{2.7}$$

where $l$ and $d$ refer to the length and diameter of the dendritic cylinder (Fig. 2.1*A*), respectively. The microscopic parameters $c_m$ and $r_m$ reflect the transmembrane capacitance and resistance per unit of surface area and $r_{ax}$ the axial resistance per units of volume that a dendritic current experiences in the direction of the axosomatic compartment. The integration timescale $\tau_d$ of a dendritic compartment is given by the effective timescale of the corresponding RC circuit:

$$\tau_d \sim \frac{C_m^d}{g_{ax}^d + g_m^d} \tag{2.8}$$

***Synaptic dynamics.*** For synaptic transmission, we considered the principal receptors concerning excitation and inhibition, including two glutamatergic receptors with fast (AMPA) and slow (NMDA) dynamics, and two GABAergic receptors with short ($GABA_A$) and long ($GABA_B$) timescales. Each receptor was modeled as a conductance with double-exponential kinetics (Roth & van Rossum, 2009):

$$g_k(t) = \bar{g}_k^{\text{syn}} \mathcal{N}_k \left( \exp\left(-\tfrac{t-t_0}{\tau_k^r}\right) - \exp\left(-\tfrac{t-t_0}{\tau_k^d}\right) \right) \tag{2.9}$$

with $k \in \{\text{AMPA}, \text{NMDA}, \text{GABA}_A, \text{GABA}_B\}$ indicating that each receptor has specific parameters. The equation describes the rise and decay of the receptor conductance $g_k$. The timescale of rise and decay is given by $\tau_r$ and $\tau_d$ while the amplitude of the curve is defined by the maximal conductance parameter $g_{syn}$. To ensure that the amplitude equals $\bar{g}_k^{\text{syn}}$, the conductance was scaled by the fixed normalization factor $\mathcal{N}_k$. This normalization factor is computed, for each receptor type, as

$$\mathcal{N}_k = \left( -e^{-t^{peak}/\tau^r} + e^{-t^{peak}/\tau_d^r} \right)^{-1} \tag{2.10}$$

$$t_k^{peak} = \frac{\tau^d \tau^r}{\tau^d - \tau^r} \ln \frac{\tau^d}{\tau^r} \tag{2.11}$$

The ratio between the maximal conductance of the NMDA and the AMPA receptor is defined as the NMDA-to-AMPA ratio (NAR). The conductance gating of the NMDAR depends on the intra-cellular depolarization which is captured by a multiplicative voltage-gating mechanism:

$$\begin{aligned} g_{\mathrm{NMDA}} &= \bar{g}^{\mathrm{syn}}_{\mathrm{NMDA}} G(v) \\ G(v) &= \left(1 + \frac{C}{3.57\,\mu\mathrm{mol/L}} \cdot e^{-\gamma v}\right)^{-1} \end{aligned} \tag{2.12}$$

where $\gamma$ regulates the steepness of the voltage-dependence. The extracellular concentration $C$ of magnesium ions $Mg^{2+}$ was fixed at $1\,\mu$mol/L. These equations and parameters were obtained from Jahr and Stevens (1990). The rise and decay timescales of the NMDAR, the NAR, and $\gamma$ assume different values in mouse (Duarte & Morrison, 2019) and human neurophysiology (Eyal et al., 2018). All compartments were endowed with excitatory and inhibitory synapses but differed in relative receptor composition and the corresponding parameters. Following previous experimental findings (Petralia, Yokotani, & Wenthold, 1994; Schulz, Knoflach, Hernandez, & Bischofberger, 2018) and modeling work (Pongracz, Poolos, Kocsis, & Shepherd, 1992), NMDARs were located only on the dendritic compartments. However, this was inconsequential in the Tripod model because the voltage threshold for NMDAR activation was larger than the somatic firing threshold, thus resulting in no contribution of NMDA channels to the somatic synaptic current. During stimulation of glutamatergic receptors, both NMDARs and AMPARs are activated. Even though the NMDAR voltage-dependent component in Eq.2.12 is continuous, its non-linear rise allows us to define a soft threshold at approximately $-40\,$mV. This value is referred to as the NMDA spike threshold throughout the manuscript. We chose $-40\,$mV because for more hyperpolarized membrane potentials (below the threshold) the NMDAR conductance is less than one-third of its AMPAR counterpart and does not trigger NMDA-spikes, as shown in Fig.2.2*C*. To parameterize the inhibitory responses, we fit the inhibitory post-synaptic potentials (IPSPs) obtained from guinea-pig hippocampus (Miles, Tóth, Gulyás, Hájos, & Freund, 1996), which characterize the dendritic versus somatic inhibition on pyramidal cells and can be considered as an effective parametrization of the differences between perisomatic and dendritic inhibition. The timescales obtained from data entail that inhibitory inputs on dendritic compartments have a slower time course, whereas somatic inhibitory inputs have a larger amplitude and faster rise and decay, and

suggest that somatic GABAergic transmission is mediated primarily by $GABA_A$ receptors (Miles et al., 1996).

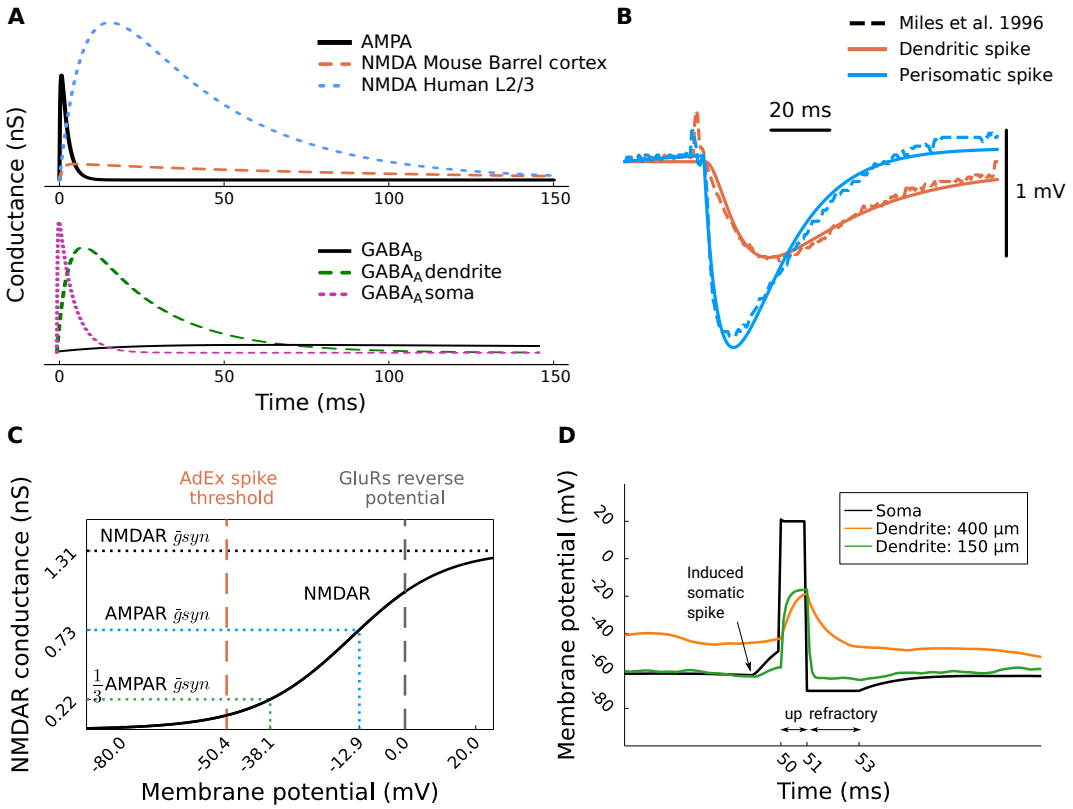



Figure 2.2: **Synaptic kinetics and conductance, and backpropagating action potential**
**(A)** The dynamics of glutamatergic (upper panel) and GABAergic (lower panel) synapses for the parameters reported in Table2.4. **(B)** Fit of $GABA_A$ timescale and maximal conductance for somatic and dendritic synapses, original data (dashed line) from Miles et al. (1996). **(C)** NMDAR conductance as a function of the compartment membrane potential. Horizontal dotted lines express the voltage-independent conductance of AMPARs and the maximal NMDARs conductance. **(D)** Back-propagating action potential in the dendrites. The backpropagation is purely due to the high membrane potential of the somatic compartment during the spike duration (1 ms). After reset the membrane potential of the soma is held fixed at the reset potential ($u_r$) for the entire duration of the refractory period (2 ms).

***Fit of inhibitory synapses.*** The fit was achieved by reproducing the somatic IPSPs reported in Miles et al. (1996). The Tripod neuron was held at resting potential and the inhibitory reversal potential was further lowered of −30 mV, similar to the experimental procedure used to record the data. The fit was performed on the minimal IPSPs, which correspond to the smallest quanta of PSP

that a single inhibitory synapse could elicit in the soma. Considering that the inhibitory neurons stimulated in the physiological experiment had more than a single synaptic contact with the pyramidal cell, we compared the fit to the stimulation of 5 simulated synapses.

### 2.2.2 Numerical simulation

Numerical integration used an improved forward Euler method (Heun's method (Ascher & Petzold, 1998)) with explicit integration and a step-size of 0.1 ms. Dendritic currents were computed from the potential difference between two coupled compartments. Because of the short integration step, the order of integration of dendrites and axosomatic compartments was unimportant. We computed the axial currents first, then the dendritic and somatic voltage changes. Note that the time-step of the explicit integration scheme used is less than half of the fastest timescale in the system and that the time scales in the model are within two orders of magnitude of each other and the explicit Table2.3; therefore, the integration scheme does not incur in numerical instability or stiffness issues at double precision computation that can emerge in the integration of cable equations in fine-grained spatial discretization models (Hines & Carnevale, 2001). Simulations were performed in Julia using custom code which can be obtained on ModelDBLINK and at `https://github.com/aquaresima/tripod_neuron`.

## 2.3 Results

Physiological parameters for pyramidal cells are difficult to reconcile across datasets because there exists significant morpho-physiological variation in the mammalian neocortex, both across species and across regions and laminae. The functional consequences of this variation can be difficult to assess. In this section, we show that some of this variability—in particular in the membrane timescale and differences in excitability between human and mouse pyramidal cells—can be explained by explicitly incorporating simple dendritic geometry and membrane physiology. We report important differences in the neuron model behavior when varying the dendritic morphology, the capacitive properties of the cell membrane, and the dendritic NMDA-to-AMPA ratio (NAR).

### 2.3.1 Geometry and physiology of dendritic compartments

#### Dendritic geometry determines activation boundaries

Excitatory synaptic input to the dendritic tree results in a forward, dromic flow of depolarizing current. This current depends on the potential difference between the perisomatic region and the location of synaptic contact, with an upper bound set by the maximum depolarization that the dendritic compartments can reach. Given the axial resistance and membrane leakage, the geometry of the dendritic branch determines whether dendritic activity can elicit somatic spikes or not. Here, we determine these *activation boundaries* as a function of dendritic length, diameter, and membrane physiology in mouse versus human pyramidal cells. Assuming that a dendrite of the Tripod is fully depolarized after a synaptic event, its capacity to generate somatic spikes is determined by the ratio between the axial conductance $g^d_{ax}$ and the membrane leakage $g^s_m$ at the soma. For integrate-and-fire neurons, the dendrite can generate a spike when the following equation is satisfied (the full derivation is given in Appendix A):

$$\frac{E_r - V_T}{V_T} =: \beta < \frac{g^d_{ax}}{g^s_m} \tag{2.13}$$

where $E_r$ is the resting membrane potential, $V_T$ is the spike threshold, and $g^s_m$ is the leak conductance of the soma. These parameters depend on the somatic compartment. In our model, $\beta$ is constant and the only variable in Eq.2.13 is the axial conductance $g^d_{ax}$ which is determined by Eq.2.7 through the cable diameter $d$, its length $l$, and the specific axial resistance $r_{ax}$ defined by the membrane physiology (Rall, 2011). Hence, Eq.2.14 defines a geometrical region where a dendrite can generate a spike. Following similar reasoning, we identify a second geometric region where full depolarization of a single dendrite is insufficient to elicit a somatic spike, but the simultaneous activation of two dendrites can:

$$\frac{\beta}{2} < \frac{g^d_{ax}}{g^s_m} < \beta \tag{2.14}$$

The two regions identified by Eq.2.13 and Eq.2.14 are shown in blue in Fig.2.3*A* and are referred to as *spiking* regions.

To test the sensitivity of the Tripod neuron to biophysical constraints, we compared two sets of membrane parameters corresponding to mouse (Dasika, White, & Colburn, 2007; Koch, 1999) and human (Eyal et al., 2016) layer 2/3 pyramidal cells. The axial conductance was the same across datasets, but the membrane

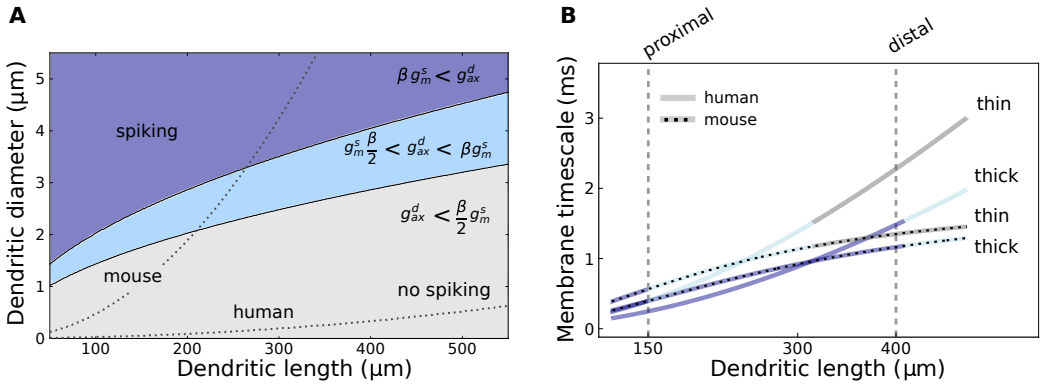



Figure 2.3: **The functional contribution of dendrites to the somatic response depends on dendritic geometry**
**(A)** Phase diagram for the axial conductance $g_{ax}^d$ as a function of dendritic diameter and length. Solid black lines show the boundaries imposed by the inequalities of Equation 2.14. They separate configurations where dendritic depolarization alone cannot elicit somatic spikes (grey region), only co-active dendritic compartments elicit somatic spikes (light blue), or depolarization of a single dendrite can elicit somatic spikes (dark blue). The geometrical regions for spike-onset onset are computed assuming the compartments clamped at $E_{\text{GluRs}}$, as described in AppendixA; because the specific axial conductance is similar for human and mouse cells, there are no species-specific differences in the geometries that lead to somatic spikes. Dotted lines mark the boundaries above which $g_{ax}^d > g_m^d$ for mouse and human pyramidal cells, respectively; the divergence between the two species is due to the larger membrane resistance of human cells with respect to mouse's cells, cfr. Table2.2. **(B)** Effective membrane timescale $\tau_m^d$ as a function of the dendritic length when the diameter is fixed at 2 µm (thin) or 4 µm (thick). Colors correspond to panel **A** and indicate the distinct functional regions of dendritic geometry. Thick dendrites influence somatic spiking more than thin ones, regardless of length. Mouse membrane timescale (dotted) converges with length while human timescales (solid) continue to increase. Throughout this work, we will use the labels *proximal* and *distal* to refer to dendrites 150 µm and 400 µm long.

conductance and capacitance differed (Table2.3). To illustrate this difference, Fig.2.3*A* shows the boundaries of effective dendrites in the Tripod neuron as a function of cable geometry. These boundaries correspond to the regions below which the membrane leakage is larger than the axial conductance, i.e., $g^d_{ax} < g^d_m$. Consequently, dendritic currents fail to reach the soma in this case and the dendrite is rendered ineffective. Within our model constraints, the dendrites of human pyramidal cells can be substantially longer than those in mice and still be functionally effective, an observation that is consistent with recent empirical evidence (Fişek & Häusser, 2020). The functional role of dendrites is also dependent on the diameter of the dendritic compartment. Thin dendrites (2 µm) have low axial conductance and their contribution to the somatic voltage is small, i.e., thin dendrites are in the no-spiking region for most of their lengths. Thick dendrites (4 µm), on the other hand, place the neuron in the spiking regime for all the lengths considered.

**Human physiology supports longer dendrites**

The effective membrane timescale characterizes the dynamics of the dendritic compartments. When the dendrite is depolarized and the soma is at the resting potential, the timescale $\tau^d_m$ for the dendritic membrane to return to the resting potential depends on the physiological parameters. It is modulated by the dendritic length and diameter, as defined in Eq.2.8. In the condition of effective dendritic transmission ($g^d_{ax} > g_m$), the current flowing out from the dendrites enters the somatic compartment, and the dendritic timescales together with the somatic membrane timescale fully determine the integration timescale in the Tripod model. Fig.2.3*B* shows the integration timescale $\tau^d_m$ for all the considered dendritic lengths, two diameters (thin 2 µm, thick 4 µm), and the physiological parameters for human and mouse (solid and dashed lines, respectively). The membrane potential in longer dendrites decays slower because the axial conductance decreases and the capacitance increases with dendritic length. For a fixed diameter, doubling the dendritic length doubles the membrane timescale. Thin dendrites have a longer timescale because of the reduced membrane leakage and axial conductance.

Overall, the differences in membrane physiology and dendritic geometry constrain the membrane's effective conductance and time constant and, consequently, the temporal integration properties of the neuron, leading to functionally relevant effects. Longer dendritic cables lead to sustained dendritic potentials, which affect the kinetics of somatic depolarization. This effect is particularly noticeable

for human physiology, suggesting that human pyramidal cells can sustain longer functioning dendrites and that length modulates neuronal responsiveness significantly. Since the functional contribution of thin dendrites is limited, we focus on thick dendrites with a diameter equal to 4 µm, consistent with previous studies (Bono & Clopath, 2017; Dasika et al., 2007; G. R. Yang et al., 2016). In the remaining work, we will study dendritic lengths in the spiking region of the phase space, and this corresponds to dendrites in the range of 100 µm to 500 µm (blue and light blue regions in Fig.2.3). For simplicity, we selected two lengths, 150 µm and 400 µm in the two spiking regions that satisfy Eq.2.13 and Eq.2.14. Following Antic et al. (2010) and Kamondi, Acsády, and Buzsáki (1998), we call a dendrite of 150 µm in length *proximal*, because it is capable of eliciting somatic spikes. A longer dendrite of 400 µm is referred to as *distal* and it can elicit somatic spikes only if co-activated with another dendrite. The proximal and distal dendrites described in the following sections are considered as roughly corresponding to the basal or apical oblique regions of pyramidal cells, respectively.

### 2.3.2 Synaptic integration with segregated dendrites

The previous section investigated how dendritic geometry and membrane physiology determine temporal integration in the Tripod neuron. We now turn to the characteristics of synaptic transmission and how the existence of segregated dendritic compartments affects neuronal responses in the model. The synaptic models used are biophysically motivated and account for relevant physiological observations.

Due to their voltage-gated component, the dynamics of NMDA receptors (NMDRs) mediates the generation of sustained plateau potentials (Major, Polsky, Denk, Schiller, & Tank, 2008) and supports coincidence detection (Rackham, Tsaneva-Atanasova, Ganesh, & Mellor, 2010; Tabone & Ramaswami, 2012). It affects the integration of excitatory input in the dendrites and the soma and plays a key role in shaping dendritic processing, synaptic plasticity, and the global input-output behavior of neurons (Doron, Chindemi, Muller, Markram, & Segev, 2017; Smith, Smith, Branco, & Häusser, 2013). Furthermore, NMDAR expression is denser in distal regions along the dendrites (Larkum, 2013; J. Schiller et al., 2000) and this suggests that there is an important relationship between the geometry and the activation of the voltage-gated receptors.

We first investigated the influence of dendritic NMDARs on somatic depolarization and the magnitude of excitatory post-synaptic potentials (EPSPs). As explained in the Methods, we included NMDAR parametrizations corresponding

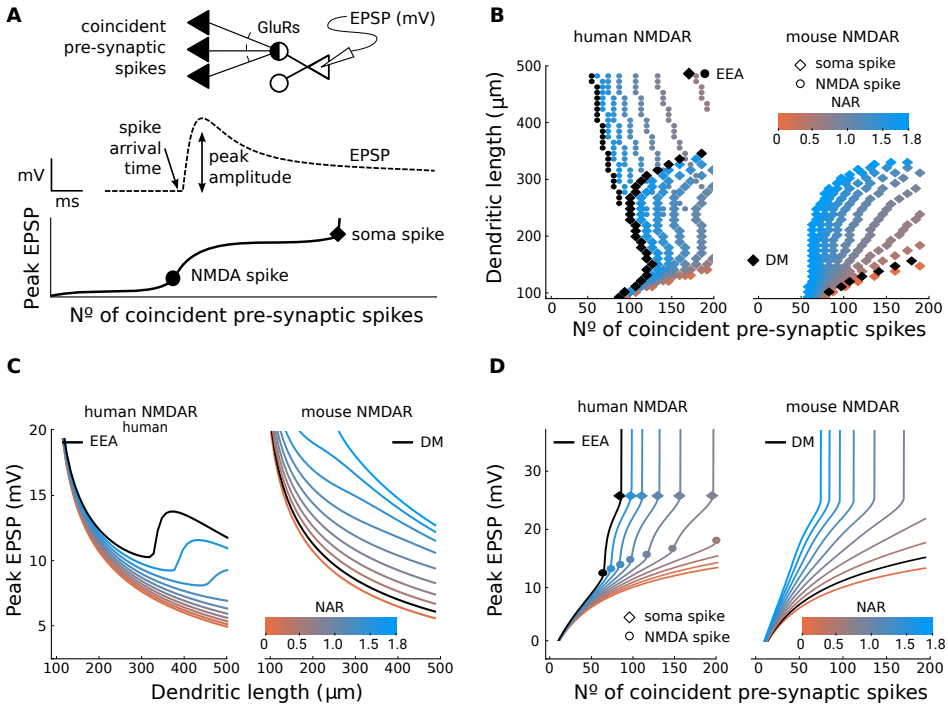



Figure 2.4: **Human-like synapses induce NMDA-related supra-linearity in EPSP peak amplitude**
**(A)** Schematic of the experimental setup. Multiple presynaptic spikes arrive concurrently at a segregated dendritic compartment with glutamatergic receptors (GluRs), and the resulting excitatory post-synaptic potential (EPSP) is measured at the soma (top). The peak amplitude of the EPSP is calculated as the difference between the membrane potential prior to stimulation and the peak membrane potential after stimulation (middle). Increasing the number of coincident presynaptic spikes results in larger peak amplitudes and causes NMDA spikes or somatic spikes (bottom). For an unbiased comparison of NMDARs between mouse and human parameters, the following simulations are based on human-like membrane parameters; when tested for mouse-like membrane, the EPSP response is weaker and sub-linear. **(B)** Tripod spike responses for *human* (left) and *mouse* NMDAR timescales and voltage gating slope (right). Each data point represents the minimal number of coincident presynaptic spikes necessary to elicit a somatic spike (diamond) or an NMDA spike (circle) for a given dendritic length (y-axis) and a specific ratio of NMDA-to-AMPA receptors (NAR, color gradient). Note that NMDA spikes are absent for mouse synaptic physiology. Black markers show the spike responses for the combination of dendritic timescale and NAR described in Eyal et al. (2016) (labeled EEA) or Duarte and Morrison (2019) (labeled DM). **(C)** Peak amplitude of the EPSP as a function of dendritic length when the number of coincident presynaptic spikes is fixed at 60. *Human*-like synaptic parameters result in an upswing of the peak EPSP relative to the increasing dendritic length, which is weaker or absent for *mouse* parameters. **(D)** Peak amplitude of the EPSP as a function of the number of coincident presynaptic spikes when the dendritic length is fixed at 300 µm. While somatic spikes occur for both human and mouse NMDARs, only *human*-like synaptic parameters cause the supra-linearity in peak EPSP that is indicative of NMDA spikes (circles).

to mouse (Avermann, Tomm, Mateo, Gerstner, & Petersen, 2012; Duarte & Morrison, 2019) and humans (Eyal et al., 2018). Compared to mouse ones, human NMDARs have shorter decay times, a larger NAR, and a steeper voltage dependence $\gamma$ in the gating mechanism. In contrast, timescales and synaptic strength of AMPARs are approximately the same for the two species.

The experimental protocol used to test the effect of varying the NMDAR characteristics is shown in Fig.2.4*A*. One of the segregated dendrites is stimulated with simultaneous spikes from excitatory presynaptic neurons, and the resulting EPSP is measured at the soma. In the synaptic model we used, coincident spikes corresponded to a single synaptic event whose efficacy was given by the peak conductance $g_{syn}$ multiplied by the number of input spikes. The peak EPSP is identified as the difference in membrane potential between the moment of spike arrival and the maximal potential reached after the spike. The peak EPSP increases with the number of co-active presynaptic neurons and converges towards a maximum value determined by the axial conductance of the targeted dendritic compartment.

Segregated dendrites with NMDARs generate a supra-linear response in the somatic EPSP which is triggered when the dendritic membrane potential reaches the threshold of the voltage-gated NMDARs. To track the onset of this supralinearity, we computed the second derivative of the EPSP peak amplitude as a function of the coincident presynaptic spikes and determined its maximum. The onset is shown in Fig.2.4*B* as a function of dendritic length and the number of coincident spikes. We distinguish between somatic spikes (peak amplitude of EPSP $\geq 30mV$, diamond markers) and the NMDAR-related supra-linearity (circles). Because the opening of NMDARs causes an all-or-none event similar to the action potential, we also refer to the NMDAR supra-linearity as an NMDA spike. When glutamatergic synapses were parameterized according to human pyramidal cells (Eyal et al., 2018, Table2.4), the NMDA-related non-linearity occurred alongside somatic spikes. When parameterized with a lower NAR, faster rise, and slower decay, corresponding to mouse synaptic physiology (Duarte & Morrison, 2019), the EPSP supra-linearity was absent, regardless of the number of synaptic inputs (Fig.2.4*B*).

The onset of NMDA spikes also depended on dendritic length. Fig.2.4*C* shows a vertical section of Fig.2.4*B* where the number of coincident spikes is fixed at 60 and dendritic length is varied between 100 µm and 500 µm. For mouse-like NMDARs, with fast rise and slow decay timescales, the peak EPSP decreased monotonically with the length of the dendrites. For human-like NMDARs, on the

other hand, dendritic stimulation resulted in an increase of the peak EPSP amplitude for dendrites longer than 300 µm when the NAR was high. This indicates that the slow rise and fast decay timescales of human NMDARs and their higher voltage sensitivity were crucial in generating NMDA spikes. Fig.2.4*D* is a horizontal section of Fig.2.4*B* with dendritic length fixed at 300 µm. Somatic spikes occurred for both human and mouse NMDARs, but only human-like synaptic parameters caused the supra-linearity in peak EPSP that corresponds to an NMDA spike.

To summarize, the results suggest that a large NAR was not sufficient to elicit NMDA spikes in mouse-like NMDARs, regardless of dendritic length and the number of coincident presynaptic spikes. Increasing the NAR (Fig.2.4*D*) raised the slope of the somatic response, but missed the supra-linear component, which indicates that the supra-linear integration depends on the NMDAR steepness ($\gamma$) and timescales, which also differ between humans and mice (Duarte & Morrison, 2019; Eyal et al., 2018). For human-like NMDARs, the occurrence of NMDA spikes was mainly dependent on the NAR and dendritic length. The length of the dendritic compartment is a crucial variable for the rise of NMDA spikes; for the opening of voltage-gated ligands of NMDARs, the membrane potential has to be sufficiently depolarized (beyond $\sim$—40 mV). Such depolarization can happen only if the compartment is sufficiently electrically segregated from the soma and the other compartments, otherwise, the membrane potential will leak towards the soma through axial currents. The dependence on the dendritic length of NMDARs' non-linearity confirms the importance of implementing voltage-dependent receptors in neuronal models with segregated dendrites

### 2.3.3 Computation with minimal dendritic structure

The above results indicate that segregated compartments are necessary for the generation of NMDA spikes. However, models with a single dendritic compartment, usually referred to as ball-and-stick models, might not be sufficient to express important dendritic computations. For instance, several dendritic phenomena depend on the interaction among synapses and therefore on their spatial arrangement on the dendrites (London & Häusser, 2005; Payeur et al., 2019), and a cascade of synapses activated from distal to proximal sites elicits a stronger response than the reverse protocol (Branco & Häusser, 2010). Hence, the question is how many compartments are needed to express these computations? We argue that a minimal model requires two dendritic compartments because it can express a minimal form of dendritic branching and it captures dendritic compu-

tations where the location of synaptic input matters. In a Y-branched dendritic tree, synaptic inputs can target the *same* or *different* dendritic branches, and the synaptic location becomes an important spatial variable of neuronal integration. This argument is in agreement with several *in-vitro* and *in-vivo* studies which have shown that two compartments are already sufficient to reproduce most of the observed processing complexity (Ujfalussy et al., 2018; Wybo et al., 2021-01-26, 2021). In the next sections, we consider the Tripod neuron in three dendritic configurations, two symmetric (distal–distal and proximal–proximal) and one asymmetric (distal-proximal). We show that in the Tripod neuron the somatic response depends on the spatial location of the inputs and that two Y-branched dendrites are sufficient to express coincidence detection (Mel, 1992), inhibition-driven pathway selection (G. R. Yang et al., 2016) logical operations (Cazé, Humphries, & Gutkin, 2013). In addition, we introduce the concept of *dendritic memory* which is the neuron's capacity to track previous activity in the voltage plateaus of distal dendrites. We show that dendritic memory can be utilized to integrate sequences of spatially distributed information and detect variations in the input stream.

**Coincidence detection**

The conductance-based mechanism that transforms presynaptic events into currents and membrane depolarization determines the EPSP response to glutamatergic inputs that occur close in time. When two excitatory synapses fire together on the same dendritic branch, the combined effect can differ from two synapses firing on separate branches. For AMPA synapses, whose receptors are not voltage-dependent, synaptic inputs across spatially segregated dendrites are known to increase the somatic EPSP response, while clustered excitation on the same dendritic branch results in weaker EPSPs (Dasika et al., 2007; Li et al., 2019). The difference between clustered and spread inputs is caused by the interaction of conductance-based synapses with the compartment voltage (Koch, 1999). An increase in synaptic conductance produces weaker depolarizing currents if the compartment is already depolarized than if the compartment is close to the resting potential. A formal derivation of this interaction is provided in Appendix B. However, as demonstrated by Mel (1992), the expression of dendritic NMDARs can yield the opposite effect. For these receptors, clustered excitation can result in larger somatic EPSPs than spread excitation, which can be interpreted as a dendritic mechanism for coincidence detection.

To test whether the Tripod neuron can reproduce these clustering effects, we compared the EPSPs generated at the soma in two conditions, clustered and spread synaptic input, and tested how the spatial distribution of the input affected somatic EPSP responses. The Tripod model is investigated here with two symmetric dendritic compartments, labeled A and B. We used dendritic lengths of 150 µm and 400 µm which are representative of compartments with weak and strong segregation from the soma. These two configurations are referred to as *proximal–proximal* and *distal–distal* configurations.

We measured the difference ΔEPSP between the somatic EPSPs resulting from excitatory input that was spread over the two compartments ($EPSP_{AB}$) or clustered on one compartment ($EPSP_{AA}$) as shown in Fig.2.5*A*. Negative values for ΔEPSP indicate that the global synaptic current was reduced for clustered input relative to spread input, whereas positive values indicate that the somatic peak depolarization was stronger for clustered input relative to spread input. ΔEPSP was measured for 200 simulations, with a random number of co-active synapses drawn uniformly from the interval $[1, 50]$ for each branch A and B in order to simulate different input intensities. The results are shown in Fig.2.5*A* where the x-axis shows the total number of co-active synapses on the two branches. There was no difference between proximal dendrites that expressed NMDARs or AMPARs only. In both cases, input spread across dendritic branches generated a larger somatic EPSP than clustered input, and this was also the case for distal dendrites with AMPARs only. However, for distal dendrites that also expressed NMDARs, clustered input caused a larger EPSP when the total synaptic input was strong, as indicated by the positive ΔEPSP (orange data points) in Fig.2.5*A* (bottom left). Thus, the Tripod neuron reproduces the AMPA spread effect and the NMDA clustering effect described in the literature (Dasika et al., 2007; Mel, 1992).

To disentangle the effects of physiology and geometry, we attempted to estimate the non-linearity of the EPSP response based on the second-order model proposed by Li et al. (2019). The original model introduced a di-synaptic matrix $\alpha_{ij}$ that determines the difference in synaptic current with respect to two synapses firing independently. The values of $\alpha_{ij}$ depend on the efficacy and the location of the synapses that are active simultaneously. They are small for synapses on different branches and negative for synapses on the same branch. To demonstrate that this second-order model is not sufficient to explain the synaptic interaction in the presence of voltage-dependent receptors in segregated den-

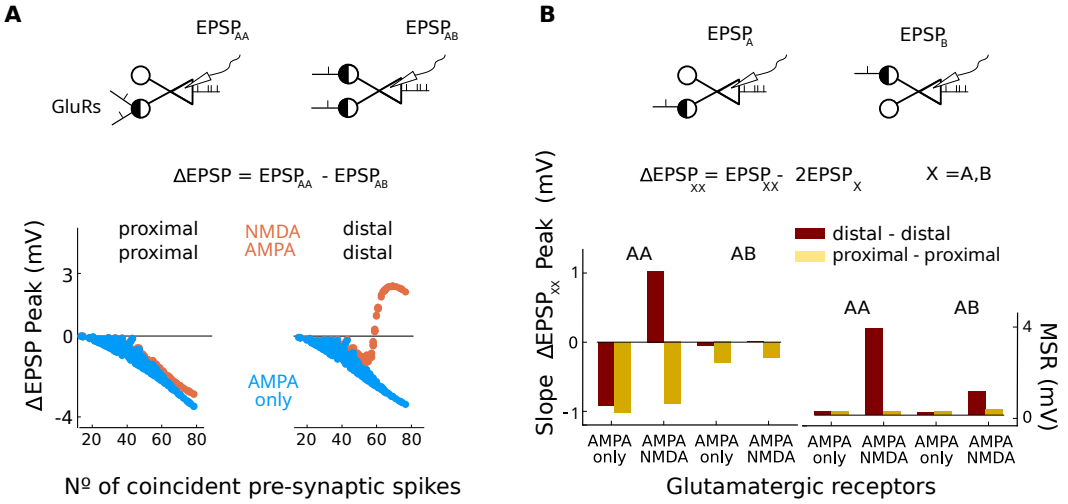



Figure 2.5: **NMDA receptors enhance somatic response in clustered condition**

**(A)** Excitatory input was applied on one dendritic branch only (clustered, AA), or on both dendritic branches (spread, AB), and the elicited EPSPs were measured at the soma. The difference between EPSPs in the two conditions is denoted ΔEPSP (top panel). The two dendritic branches had the same length and were either distal or proximal. Synapses on the two branches expressed NMDA and AMPA receptors (orange), or AMPARs only (blue). The bottom panel shows the peak ΔEPSP as a function of the number of coincident input spikes in the four conditions. For proximal–proximal dendrites, spread input resulted in stronger EPSPs for both AMPARs only and combined AMPARs/NMDARs. For distal–distal dendrites, the expression of NMDARs produced stronger responses in the clustered condition which showed a supra-linear response when the total synaptic input was sufficiently strong to activate the NMDARs (>60 co-active synapses). **(B)** The magnitude of synaptic interaction was obtained by comparing di-synaptic conditions ($XX$ = AA, AB) to input spikes on single compartments ($X$ = A, B). The top panel shows the stimulation protocol used to compute $EPSP_A$ and $EPSP_B$. $\Delta EPSP_{AA}$ and $\Delta EPSP_{AB}$ summarize the interaction for the clustered and spread conditions. The ΔEPSPs in conditions AA and AB are fit with linear regression over the global synaptic inputs and the lower panels show the slope and the mean squared residuals (MSR) of the linear fit. Di-synaptic interaction reduced somatic depolarization (negative slope of $\Delta EPSP_{AA,\ AB}$) for all input conditions, receptor types, and Tripod configurations except for clustered inputs AA on distal–distal compartments with NMDARs (third column). This configuration generated high MSRs, indicating that the interaction could not be expressed with linear di-synaptic interactions. For all conditions, the fit was computed by drawing 200 co-active synapses in the range (1,35).

dritic compartments, we stimulated the Tripod with clustered and distributed inputs and subtracted the EPSP of independent synaptic events on each branch.

$$\Delta EPSP_{XX} = EPSP_{XX} - 2EPSP_X \tag{2.15}$$

where the $X$ subscript refers to the branches A or B. Note that $\text{EPSP}_\text{A}$ is the same as $\text{EPSP}_\text{B}$ because dendrites were symmetric. $\Delta\text{EPSP}_\text{AX}$ was computed for different numbers of co-active synapses between 1 and 50, as before. The simulation was run for 8 conditions, i.e., with and without NMDARs, with two distinct geometries, and in both the distributed AB and clustered AA configurations. Following the original model, we asked whether a second-order function of the synaptic input was sufficient to explain $\Delta\text{EPSP}_\text{AB}$. Hence, we fit $\Delta\text{EPSP}_\text{AB}$ via the product of the synaptic conductances $g_e^1 \cdot g_e^2$ and obtained the results in Fig.2.5*B*. The two panels show the slope of the interaction corresponding to $\alpha_{ij}$ in Eq.2.23 and the residuals of the linear fit (right).

In the absence of NMDARs, we observed a strong attenuation of somatic EPSPs and the residuals of the linear fit were small. This effect was larger when synapses clustered on the same compartment compared to the distributed condition and this was due to the segregation of voltages in the different compartments. The EPSP attenuation effect was also stronger when dendrites were shorter (proximal–proximal configuration, yellow bars in Fig.2.5*B*). It is worth noting that the residuals of the linear fit were small for most of the configurations, suggesting that the model of Li et al. (2019) was also applicable to the Tripod neuron when only AMPA receptors were present. However, in agreement with previous results (Mel, 1992), the expression of dendritic NMDARs yielded different functional behavior and resulted in the amplification of somatic EPSPs in the clustered condition AA. This effect was dependent on dendritic length. The di-synaptic interaction still resulted in EPSP attenuation (negative) in the proximal–proximal configuration due to the reduced NMDAR contribution for proximal dendrites. For longer dendritic branches (distal–distal configurations in Fig.2.5*A*), when excitatory inputs were clustered on the same compartment, the interaction initially reduced somatic EPSP amplitudes. As the number of co-active synapses increased to around 60, however, the EPSP began to increase in a non-linear fashion. Thus, segregated dendritic compartments with voltage-dependent NMDA receptors introduce synaptic interactions that go beyond the second-order model of Li et al. (2019). These interactions cause larger EPSPs when synaptic inputs are clustered, in agreement with previous simulations

(Mel, 1992; Ujfalussy & Makara, 2019), and the magnitude of this clustering effect is strongly mediated by dendritic length.

### On-path shunting inhibition

Depending on the location of synaptic contact, inhibitory GABAergic inputs, whose ionotropic receptors have an equilibrium potential close to the resting potential, can effectively offset excitatory drive onto neighboring synapses (Koch, Poggio, & Torre, 1983). Inhibitory configurations that veto neuronal responses are referred to as *shunting* inhibition and play an important functional role. Shunting inhibition depends on the spatial distribution, the composition of inhibitory synapses, and the relative timing between excitatory and inhibitory presynaptic events. The Tripod neuron with two dendritic and one somatic compartment provides the simplest structure to study this type of inhibition.

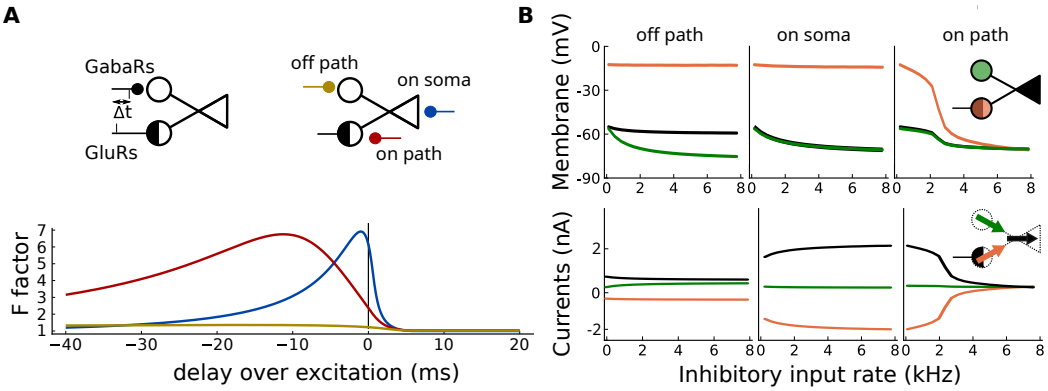



Figure 2.6: **Dendritic inhibition and shunting**
**(A)** Location and timing of inhibitory spikes determine the somatic response. The upper panel describes the experimental protocol. An inhibitory spike is delivered to the soma at an interval $\Delta t$ from the excitatory one. Then the F-factor is computed. The two panels show the EPSP attenuation for three inhibitory conditions, on-path (red), off-path (yellow), and on-soma (blue), scheme in the lower panel. In the upper panel, the dendritic $GABA_A$ receptors are parameterized with long timescales, as in Miles et al. (1996). **(B)** Average membrane potentials (upper panel) and axial currents (lower panel) for varying inhibitory input rates. The orange dendrite receives 1.5 kHz Poisson distributed excitatory input, while the neuron also receives variable inhibitory inputs at different locations (from the left: off-path, on-soma, on-path). Both dendrites are 300 µm long. Inhibition off-path has a negligible effect on the somatic membrane (black line) compared to on-path and on-soma inhibition.

We investigated different inhibitory configurations by stimulating one of the dendritic compartments with a single excitatory spike followed by an inhibitory

spike within a fixed time interval that was delivered to one of three input locations; the same dendritic compartment (*on-path*), the other dendritic compartment (*off-path*), or the soma (Fig.2.6*A*). To measure the effectiveness of inhibition, we compared the somatic EPSP in the presence or absence of GABAergic inputs. Attenuation caused by inhibition was measured as the ratio between the EPSP peaks in the two protocols (excitation versus excitation plus inhibition):

$$F = \frac{\mathrm{EPSP}_{exc}}{\mathrm{EPSP}_{exc+inh}} \tag{2.16}$$

The larger this $F$-factor, the more effective the inhibitory signal was.

Results in Fig.2.6*A* show that the impact of inhibition is determined by the relative timing of the excitatory and inhibitory inputs and it is highly location-specific. Suppose dendritic GABAergic transmission in the same compartment of excitation, *on path*. In that case, its depressing effect on the EPSP is extended in time, and it peaks when inhibitory spikes arrive around 10 ms before excitation (red line). If, on the other hand, inhibition is located on the soma, hence mediated by fast $\mathrm{GABA_A}$ receptors, then inhibition is maximally effective when inhibitory and excitatory inputs arrive simultaneously. In this condition, inhibitory spikes that arrived more than 10 ms before excitation are ineffective. When inhibition is *off-path* its effect on the somatic EPSP is negligible. Notice that the $\mathrm{GABA_B}$ receptors are active only in the dendrites, and their effect is small in the setup of Fig.2.6*A* because a single inhibitory spike is insufficient to engage these receptors.

The Tripod neuron received an excitatory Poisson input at a fixed rate of 1kHz on a single dendritic compartment and a variable rate inhibitory input on different compartments (off-path, on soma, on-path). In the absence of inhibitory input, the soma was in a depolarized state ($\langle v_m^{soma} \rangle \approx -60mV$). Fig.2.6*B* shows the mean value of the membrane potential of each compartment and the current flowing between the compartments, both averaged over a 10 s interval. When the *off-path* compartment was targeted by inhibitory inputs (leftmost panel in Fig.2.6*B*), the soma reached an equilibrium between a weak hyperpolarizing current coming from the inhibited dendrite and the depolarizing current from the excited compartment. In this condition, the soma remained depolarized, regardless of the magnitude of the inhibitory inputs. When inhibition targeted the somatic compartment (middle panels in Fig.2.6*B*), the soma received a depolarizing current from the excited dendrite and a competing hyperpolarizing current from the GABAergic synapses on the soma membrane. Because the synaptic cur-

rent depended on the somatic potential, it had a balancing effect on the compartment potential. When the inhibitory input was sufficiently strong, the soma approached the resting membrane potential. In this condition, inhibition had a divisive effect on the somatic potential. In both cases, the stimulated dendrite remained depolarized but benefitted from the NMDA boost, resulting in a large axial current. On the other hand, when the inhibition was *on-path*, that is, localized to the same compartment as excitation, inhibition pulled the dendritic potential below the NMDA threshold, and thus hyperpolarized the stimulated dendritic compartment. In this configuration, the soma remained depolarized as long as the dendritic balance of excitation and inhibition was maintained. When the inhibition overcomes excitation (around 2 kHz for this setup), the neuron was shut down, and all the compartments went to resting potential, with no axial currents flowing. Hence, somatic depolarization is more dependent on the spatial distribution of the inhibitory spikes than on the actual inhibitory input received. Furthermore, this experiment suggests that considering the somatic membrane potential alone may not be sufficient to characterize the state of the cell; in Fig.2.6*B* the membrane potential of *on soma* and *on path* conditions is similar, although the cell is in two different states and will respond differently to further stimuli. For example, for a fixed inhibitory input, increased excitation on the stimulated dendrite will only depolarize the soma if inhibition is *on path*, while it will be less effective in the *on soma* condition

**Logical operators**

Logical operators define a natural class of computations. Single-compartment neurons, which integrate inputs with a monotonic transfer function, can perform linearly separable computations but fail on non-separable ones. In contrast, theoretical and experimental work has shown that active dendrites can solve non-separable problems (Cazé et al., 2013; Gidon et al., 2020). If we consider the dendrites as independent input pathways and treat the Tripod as a binary logical gate, then the previous experiments on coincidence detection have already demonstrated that the Tripod can perform non-separable computations, matching the theoretical results in (Cazé et al., 2013).

Another possibility is to consider the neuron's dynamics explicitly. In this configuration, the input is drawn from a set of binary stimuli, e.g., A = 0, B = 1, and mapped to the input spike rates on the respective compartment, e.g., 0 = E/I balanced inactive state, 1 = E/I balanced active state (further details in Appendix C). The cell's response also has to be represented over time and cal-

culated, for example, on the output firing rate. Under this encoding, both a single-compartment neuron and the Tripod model can reproduce the truth table of multiplication (AND, true for inputs (1, 1)) and summation (OR, true for inputs (0, 1), (1, 0), (1, 1)). However, there are no mechanisms that enable single neurons to implement operators such as exclusive OR (XOR, true for (1, 0) and (0, 1) but false for (1, 1) and (0, 0)) or material implication (MI, true for (0, 1) but false for (1, 0)). Unfortunately, the same holds for dendrites with NMDA spikes; if one active dendrite is sufficient to trigger somatic spikes, two active dendrites can only increase the somatic firing rate, making it impossible to solve the XOR problem. These limitations are due to the coding scheme for the output. To avoid this, we investigated if the neuron could make the computation separable for an external linear readout. Therefore, we analyzed the sub-threshold dynamics of the somatic membrane potential (van den Broek et al., 2017) to evaluate the neural computations.

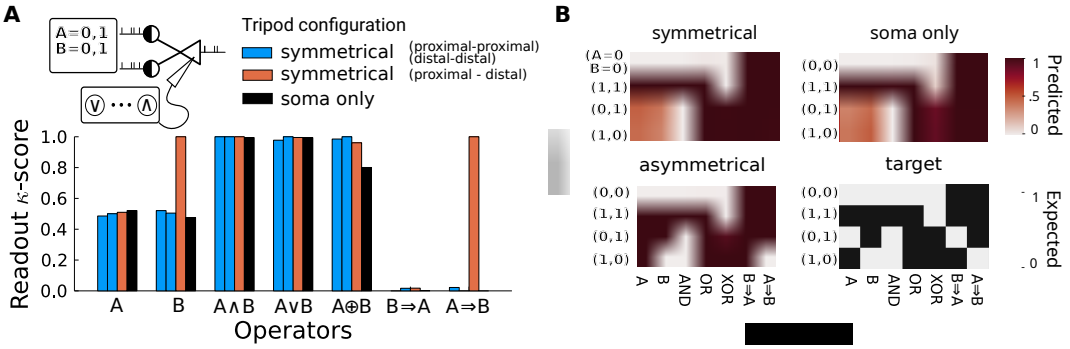



Figure 2.7: **Asymmetric dendrites enhance the separability of logical operations**
**(A)** Cohen's kappa-score accuracy of linear readout classifiers on logical operators for symmetric, asymmetric and soma-only models. The dendritic configurations are proximal–proximal and distal–distal (blue), proximal–distal (orange) and soma-only (black). **(B)** Shade of red indicates the average predicted truth value for each input condition (y-axis), operator (x-axis), and dendritic configuration (top and left panels). The black and white table (bottom-right) indicates the expected truth values. E.g., the AND operator for symmetric dendrites shows dark red (true) for condition A = 1, B = 1, and white for all the remaining conditions, corresponding to the target truth-values.

For this purpose, we stimulated the dendrites with a random sequence of four possible input configurations: (A = 0, B = 0), (A = 1, B = 0), (A = 0, B = 1) and (A = 1, B = 1). A set of seven external logistic regression readouts were used to map the neurons' somatic dynamics to the truth table of seven different operators ($Id_A$, $Id_B$, A ∨ B, A ∧ B, A ⊕ B, A ⇒ B, B ⇒ A) As mentioned above, the symbols A and B refer to the stimulated dendritic compartment and each

input is presented for a period of 200ms. The membrane dynamics was readout during the last 50 ms of the stimulus presentation; the readout had access to 5 points for the membrane potential and 5 points for the adaptive current, each spaced by 10 ms. After training, we injected a random sequence of inputs (A, B, AB, or none) and tested if the trained readout could use the information in the membrane of the soma to reproduce the correct truth table. We examined four different geometries, two symmetric one with proximal–proximal (150 µm) or distal–distal (400 µm) dendrites, one with asymmetric structure (400 µm-150 µm) and a single-compartment model. When a dendritic pathway was inactive (e.g., A = 0), the respective dendrite received a 3 kHz train of excitatory Poisson spikes, and a balanced inhibitory input. For the baseline condition (soma-only), the spikes were injected into the somatic compartment via two independent synapses, as above, the excitatory input rate was doubled for the active input condition.

After testing all the models, we measured the Coheh's kappa-score of the readout on each operator, see Fig.2.7*A*; we chose this metric to account for asymmetries in the classes' statistics, e.g., A ⇒ B has three True and one False. Symmetric configurations performed better on symmetric operators (blue bars in AND, OR and XOR operators). Conversely, asymmetric operators (red bar in $Id_B$, A ⇒ B) are best recognized with asymmetric dendrites. In the distal-proximal configuration, the activity in each dendrite is different, and input to the short dendrites is easily distinguished. The soma-only configuration scores lower than each Tripod configuration. To elucidate the computations performed, we analyzed the predicted truth value for each operator and condition (Fig.2.7*B*). As expected, the symmetric configuration (proximal–proximal, distal–distal, and soma-only) makes the same prediction concerning inputs (1, 0) and (0, 1); for asymmetric operators, this is also the case, because the readout cannot distinguish which input-pathway is activated. This is not the case for the proximal–distal condition, and the input (0, 1) is treated differently from (1, 0). In almost all conditions the Tripod neuron performed better than the single-compartment model, indicating that the inclusion of the dendritic structure was beneficial. These results show that the membrane dynamics of asymmetric Tripod models depend on the input pathway, and the neuron can act as an asymmetric logical operator.

**Dendritic memory**

When excitatory synaptic input is sufficiently strong to drive the postsynaptic voltage above the NMDA gating threshold, the ionic current flowing through the

NMDAR keeps the dendritic compartment depolarized and generates a temporally extended plateau potential (Fig.2.8*A*).

The time course of the plateau potential depends on the number of coincident presynaptic spikes, even though the dendritic potential reaches the NMDAR reversal potential (Fig.2.8*A* top-right panel). To quantify the duration of the voltage plateau, we set an arbitrary threshold at −60 mV and monitored how long the somatic membrane potential remained above this value (Fig.2.8*B*). In the presence of NMDARs with *human* timescales, NAR, and $\gamma$, long distal dendrites reached a voltage plateau whose duration increased with the number of coincident inputs and could last up to 100 ms. When dendritic length was short enough to trigger somatic spikes (Eq.2.13) the duration of the plateau potential was limited by the somatic after-spike reset potential. Because of the large conductance between proximal and somatic compartments the brief duration (1 ms) of the hyperpolarized reset potential is sufficient to prevent the continuation of the plateau-potential by pulling the dendritic potential below the NMDAR threshold. Conversely, this is not the case for distal dendrites that can sustain the plateau potential during somatic firing. Further details on dendritic membrane dynamics during and after somatic spikes are discussed in Appendix C. When the NAR was set to mouse synapses (0.25), the dendritic and somatic potentials showed a weaker, sub-linear dependence on the number of presynaptic inputs. The depolarization caused by 50 synapses was similar in extent to the depolarization caused by four times as many co-active synapses (Fig.2.8*A*, left panel). The reason why the EPSP response saturates is because the incoming synaptic current depends on the difference between the membrane potential and the synaptic reversal potential. For the remainder of this article, the dendritic parameters were set to correspond to human physiology.

We investigated whether the plateau potential generated by NMDA spikes in distal dendrites could be used as a short-term processing memory. We tested this by encoding a memory trace into distal dendrites through synaptic activity. The spike rate of the encoding signal was the critical variable and corresponded to the number of co-active synapses in the previous experiment. Then, we attempted to retrieve this memory by injecting a 1 kHz spike train on the proximal dendrite after an interval of time $\Delta t$ (illustrations in Fig.2.8*C*). The retrieval cue was weak and without previous distal inputs, the soma fired the first spike on average 50 ms after the onset of the proximal input. Note that the proximal input lasted longer than the 50 ms considered for retrieval. Thus, we considered retrieval of an encoded memory to be successful if the first somatic spike occurred earlier

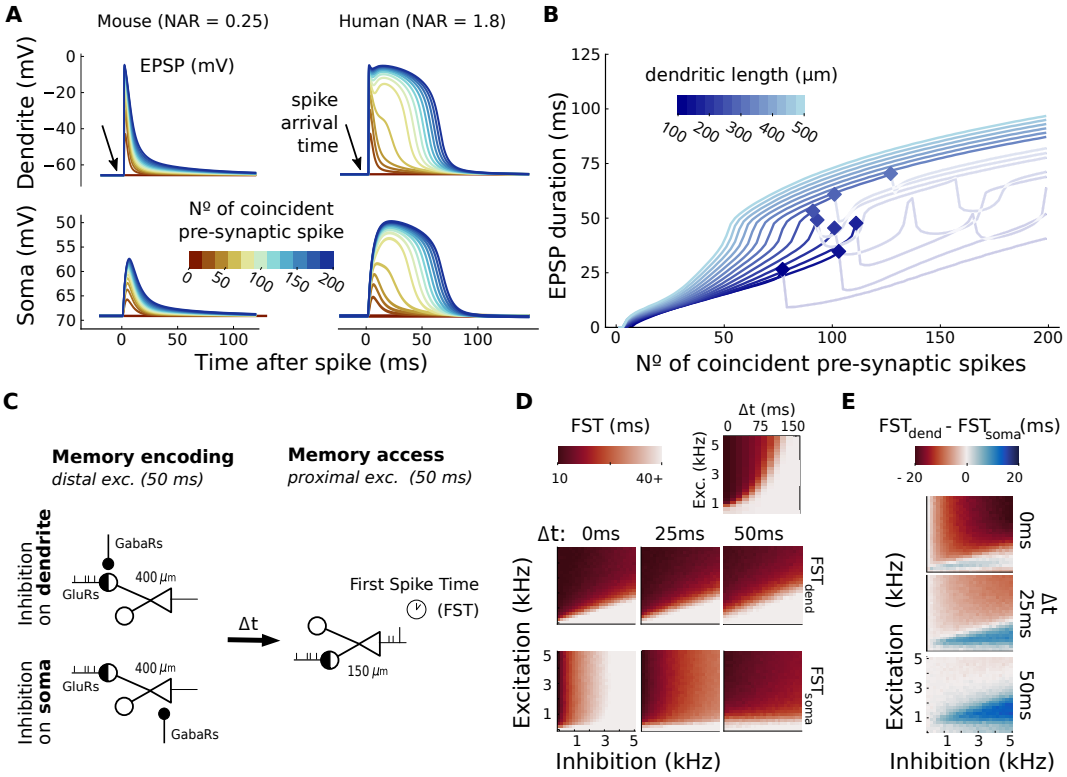



Figure 2.8: **Plateau potentials due to NMDA spikes support dendritic memory**
**(A)** Magnitude and kinetics of spike-induced EPSPs in the dendrites (upper panels) and soma (lower). Dendritic synapses are endowed with mouse (left panels) or human (right) NMDARs; AMPARs are identical for both. Input spikes arrive on a distal dendrite (400 µm), color codes for the number of coincident input spikes. For mouse-like synapses, an increase in the number of inputs did not lead to longer dendritic depolarization. Dendrites with human NMDARs show extended depolarization when the input triggers NMDA spikes. **(B)** Duration of sustained somatic depolarization (EPSP curve is above -60 mV) for simulations with *human*-like NMDARs. Color codes for dendritic length. Long dendrites result in a somatic depolarization that lasts for +100 ms, referred to as *plateau potential*. For long dendrites, the duration of the plateau potential increases monotonically with the number of simultaneous synaptic inputs. When the targeted dendrite is short enough to cause somatic spikes (diamond markers), the relation between the total presynaptic input and the duration of the depolarized state is interrupted because the somatic after-spike depolarization forces the dendrite below the activation threshold of the NMDARs. Somatic spikes do not affect the plateau potential in long dendrites because of the low axial conductance. **(C)** Input configuration for memory encoding. Memories are encoded via excitation of the distal dendrite. After an interval $\Delta t$ without input, the proximal compartment is activated and the average first-spike-time (FST) is measured. *Figure caption continues on the next page.*

Figure 2.8: **(D)** Mean FST for varying excitation strength and $\Delta t$ (upper-right). Shorter FSTs (e.g., dark red, 10 ms) indicate successful memory retrieval. Lower panels show FSTs for dendritic (upper) or somatic (lower) inhibition, measured while varying the strength of excitation and inhibition during the encoding phase. Retrieval is attempted after three intervals $\Delta t \in \{0, 25, 50\}$ ms. **(E)** Comparison of memory traces in two inhibitory configurations. Colors code for the difference between FSTs in the somatic versus dendritic inhibition condition. For short $\Delta t$, inhibition on dendrites elicits faster somatic spikes (shorter FST). For longer $\Delta t$, inhibition on dendrites is more detrimental to retrieval than inhibition on soma.

than 40 ms after the retrieval cue was injected. This measure of retrieval was called the first-spike-time (FST) and averaged over 300 independent trials in the experiment. The somatic compartment was also exposed to noisy excitatory inputs that caused random spikes during the stimulation protocol. This was not necessary for encoding and retrieval but was intended to test the robustness of plateau potentials in the presence of somatic spikes. The top-right panel in Fig.2.8*D* shows that memories encoded into long dendrites could be retrieved within about a hundred milliseconds, which was approximately the duration of the plateau potential. The lifetime of memory traces increased with the input rate that was used to encode these memories (*y*-axes). However, higher input rates during encoding did not correspond to shorter FSTs.

So far, only glutamatergic synapses have been considered. We further investigated dendritic memory in the presence of inhibition by activating GABAergic synapses during and right after the encoding phase. Inhibition was present in both somatic and distal compartments. We tested memory retrieval by monitoring the FST at three different times, separated by 25 ms each, after the encoding phase. Fig.2.8*D* shows the effect of inhibition on the distal dendrite and on the soma. When excitatory inputs on the distal dendrite were matched by dendritic inhibition, retrieval depended on the ratio between excitation and inhibition, as demonstrated by the linear separation between successful and failed retrieval. The retrieval protocol cannot distinguish between the exact amount of inhibition received during the encoding phase when memory was successfully encoded; the upper panels in Fig.2.8*D* show nearly identical success rates in memory access for the three delay intervals. This changed when inhibitory synapses fired on the soma; at first, memories were not retrievable but they became accessible when inhibitory activity ceased. Within 50 ms, there was virtually no trace of the somatic inhibition. In this condition, the magnitude of the inhibitory input modulated the retrieval success rate in a graded manner.

The difference between the two inhibitory input pathways is shown in Fig.2.8*E*. Immediately after distal activity ($\Delta t$ =0 ms), inhibition on the soma prevented spiking and memory retrieval (FST with dendritic inhibition was smaller than FST with somatic inhibition, dark red). After 50 ms, the relation between somatic and dendritic inhibition reversed and memories that were encoded during somatic inhibition were now accessible. Dendritic inhibition limited the life-span of the encoded memories and the ratio between excitation and inhibition during the encoding phase determined retrieval success. This shows that the minimal dendritic tree of the Tripod model maintained short-lived memories. Retrieval of these memories depended on the location, the input strength, and the relative timing of their encoding.

**Transition detection and sequence recognition**

Dendritic memory endows the Tripod model with two segregated memory slots, which can potentially be used to combine or discriminate incoming information over time. Here we tested whether this memory mechanism could be used to solve spatio-temporal tasks.

Excitatory and inhibitory Poissonian inputs were injected into the neuron at a constant rate. The active dendrite was set in the E/I balanced active state, the other dendrite in the inactive state (further details in Appendix C). The input targeted dendrite A or dendrite B and it was switched from one compartment to the other regularly, with frequencies in the range of 1 Hz to 100 Hz. A schematic of the input protocol is shown in Fig.2.9*A*. We first measured dendritic and somatic potentials during a sequence of switches at 4 Hz. The membrane dynamics of the three compartments are shown in Fig.2.9*B*, for models with symmetric dendrites (distal–distal) asymmetric ones (distal–proximal). The sequence of excitatory and inhibitory input spikes was the same for the two models. After a switch, the potential of distal dendrites decayed slowly while the potential at the newly stimulated dendrite started to rise. As a consequence, the depolarizing axial currents towards the soma reached their maximum right after the switch. To measure the effect of the increased axial currents we computed the average somatic potential for 300 trials with similar input statistics (Fig.2.9*B*). The somatic response to a switch differed between the two dendritic configurations. For distal–distal dendrites the response was maximal right after the switch and it was the same for the two dendrites. For distal–proximal dendrites, the somatic response was stronger during stimulation of the proximal dendrite than the distal dendrite, resulting in somatic bursts.

To further explore the Tripod's response to spatio-temporal sequences, we tested four dendritic configurations, distal–distal, proximal–proximal, distal-proximal and soma-only. For a fair comparison with the soma-only model, the switching was achieved by implementing two independent synapses that were targeted by one of the two input streams. This corresponds to a model with zero-length dendrites. We repeated the previous experiment with two input signals, one with regular switching times as above, and one where switching times were drawn from an exponential distribution with a rate equal to the switching frequency. We recorded somatic firing and averaged the output spikes over 300 trials with identical statistical realizations of the input spike train. Therefore, the reported firing rates indicate the average instantaneous somatic firing. The firing rate in response to the two input signals is shown in the left panels of Fig.2.9*C* and *D*. Somatic firing on single trials was not synchronized with the switch times (black dots), but across multiple trials it was. To quantify the dependence of output firing on the input switch, we convolved the switch times with an alpha-function (with rise and decay timescales of 10 ms) and then computed the correlation between the average firing rate and the distribution of switch times (referred to a signal/spikes correlation in the top right panels of Fig.2.9*C* and Fig.2.9*D*). Overall, the Tripod responses were correlated to the spatial switches in the input stream, for both regular and irregular switching times. As expected, the distal–distal model showed the strongest correlation with the input switch, and the peak firing rate of a Tripod with asymmetric dendrites was less synchronized. For regular switch intervals, the Tripod model lost track of signals oscillating faster than 30 Hz, although the response to non-regular signals stayed synchronized for higher frequencies. For both input conditions, the soma-only model showed zero correlation with the switching times. As suggested by the delays between the switch times and the maximal somatic response in Fig.2.9*B*, we hypothesized that the correlation might be higher at different time points. Hence, we measured the correlation backward and forward in time with delays in the range of $-200$ ms to 200 ms. The correlation with the signal was maximal when the firing response was correlated backward in time as the somatic response lagged behind the input signal. The optimal delay depended on the model, and it was shorter for shorter dendrites. The bottom panels of (Fig.2.9*C* and *D* bottom panels) show the correlation for different delays and an input signal with 6 Hz switch frequency. The delay with the highest correlation was the same across all the switch frequencies tested (data not shown), indicating that

the optimal delay depended only on the time span necessary to depolarize the dendritic compartment, which in turn depended on dendritic length.

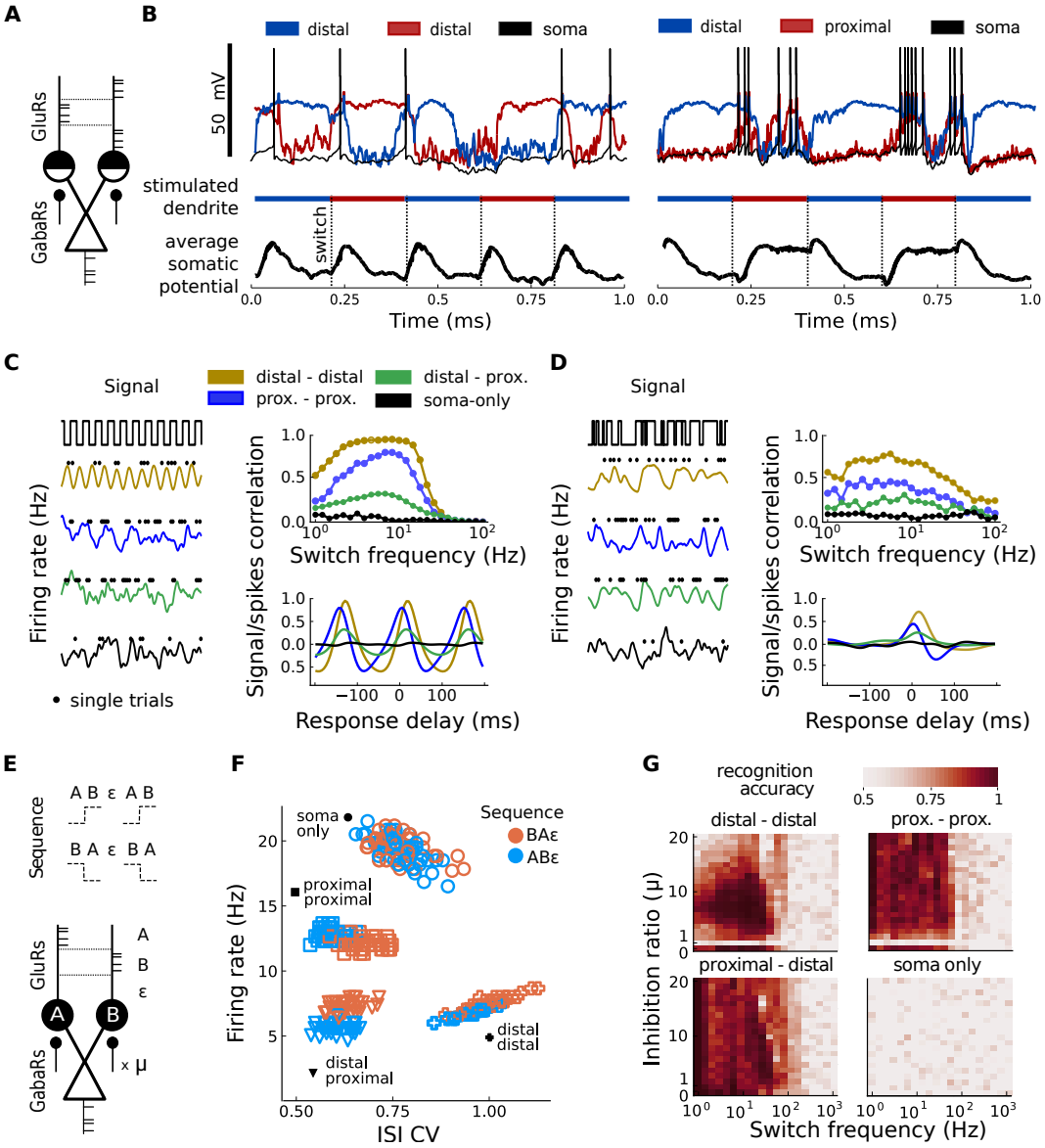



Figure 2.9: **Sensitivity to serial order**
**(A)** Excitatory and inhibitory inputs are delivered to the neuron by switching between the two dendrites periodically after a fixed interval. **(B)** Distal–distal and distal–proximal Tripod neurons receive the input described in (A). Each dendrite depolarizes during its stimulation interval. For distal dendrites, decay to rest is slow and the depolarized state overlaps in time with the rise in the potential of the other compartment. This overlap of the two depolarized dendritic states maximally depolarizes the soma, as shown in the average membrane potential of the somatic compartment (lower panels). For asymmetric dendrites, somatic depolarization is strong when the proximal compartment is stimulated. Input to the two dendrites switches at 4 Hz and the average over 300 trials is shown.
*Figure caption continues on the next page.*

Figure 2.9: **(C)** Left panel shows the average firing rate in response to a signal switching between dendrites at (6 Hz) for three Tripod configurations (colors) and a single-compartment model (black) that implements switching on two independent synaptic conductances. Black dots show spike times on one of the 300 trials used to compute the firing rate (solid lines). The top-right panel shows the correlation between firing rate and switches in the input signal as a function of switch frequency. To compute the correlation, the switch times were convolved with an alpha function. The bottom-right panel shows the correlation when the firing response is shifted in time (backward or forward) for inputs with 6 Hz switch frequency. The correlation is maximal after a delay for all the models because the soma lags behind the dendritic depolarization. Values shown in the top panel correspond to the maximal correlation obtained across all the response delays that were tested. Negative delays are due to the convolution function that maps spikes to rates. Peaks at ±150 ms are due to the oscillatory nature of the input. **(D)** As in (C) for a signal whose switch times are drawn from an exponential distribution of rate equal to the switching frequency. **(E)** Two sequences are played to the dendrites A and B of the neuron, AB$\epsilon$ or BA$\epsilon$, where $\epsilon$ is a silent pause. Dendrites receive feedback inhibition proportional to somatic activity. One of the dendrites receives $\mu$ times the feedback inhibition of the other dendrite. **(F)** With $\mu \neq 1$, the spike statistics (firing rate and $CV_{ISI}$) depend on both the sequence order (blue or orange) and the neuron's geometric properties (marker shapes). Each data point corresponds to 1 s of simulation time and the switching frequency was 6 Hz. **(G)** Sequence classification accuracy based on the somatic spike statistics in (F) as a function of inhibitory feedback ratio $\mu$ and switch frequency. Neuron configurations with dendrites outperform a soma-only model. Only the asymmetric configuration succeeded on the task when inhibitory feedback was identical on both dendrites ($\mu = 1$).

These results show that the Tripod neuron with symmetric dendrites was sensitive to transitions in the location of synaptic input. We further investigated whether the switching *direction* could be detected as well. Two input sequences were created where input was injected into dendrite A, then dendrite B, or the other way around, followed by an inputless pause $\epsilon$ (see Fig.2.9*E*), resulting in two sequences *AB*$\epsilon$ and *BA*$\epsilon$. The switched intervals were regular and we used the switching frequency to indicate the rate for rotating over the elements of the sequence (A, B,$\epsilon$). As before, the input spike trains targeting each dendrite were statistically the same. In a preliminary analysis, we measured the somatic potential during the presentation of the two sequences and verified that for symmetric models it was impossible to determine which of the two was presented. When the model had asymmetric dendrites, the order of the input on the dendrites (proximal–distal-$\epsilon$ vs. distal-proximal-$\epsilon$), changed the somatic response. To break the symmetry between the two compartments in the distal–distal and proximal–proximal configurations, we added an external inhibitory

input on both dendritic compartments. The strength of this input was proportional to the somatic activity (mimicking cortical feedback inhibition) and the input was injected by means of a conductance-based synapse, following Bono and Clopath (2017). The conductance was a double-exponential filter of the Tripod output spikes, with decay timescale of 50 ms, rise timescale of 2 ms, and peak conductance of 5 pS. The symmetry was broken by different strength feedback on the two dendrites. The B dendrite had a feedback peak-conductance that was $\mu$ times the baseline value, up to $\mu = 20$, resulting in a peak conductance of 100 pS.

We tested whether the additional inhibitory feedback would make it possible to determine which of the two sequences was presented to the Tripod, AB$\epsilon$ or BA$\epsilon$. We also compared the three dendritic configurations of the Tripod with a soma-only model. To distinguish the neural responses we calculated the average firing rate and the coefficient of variation of the inter-spike intervals ($CV_{ISI}$) that can be used to detect burstiness. These spike statistics were computed during a period of 1 s, for 100 trials, in each of the four configurations. An example of the firing rate and $CV_{ISI}$ for a switching frequency of 6 Hz and $\mu = 5$ is shown in Fig.2.9*F* for the two sequences (orange and blue). We used logistic regression to quantify whether the two sequences could be distinguished. The outcome of a grid search over $\mu$ in the range 0 to 20 and switching frequencies of 1 Hz to 1000 Hz is shown in Fig.2.9*G*. Classification accuracy was high for all dendritic configurations at switching frequencies below 100 Hz and at chance level for the soma-only neuron. These findings were robust to variations in inhibitory feedback asymmetry ( $\mu$) and switching frequency. For the distal–distal model, large dendritic feedback inhibition reduced accuracy, likely because the Tripod did not spike enough to compute reliable statistics. For both symmetric configurations, accuracy was close to chance levels when feedback was the same on both dendrites ( $\mu = 1$). The proximal–distal model could recognize the sequences also when feedback was symmetric. This shows that sequence classification can be achieved reliably when neurons are equipped with dendritic compartments, whereas a single-compartment model (in its present instantiation) fails. Consistent with the previous results on switching, sequential order in the input could be distinguished based on the somatic spike response. The experiment used a fixed input rate and a fixed number of co-active synapses for both dendrites. It is to be expected that variability in rates and the number of input channels will increase the range of dendritic input patterns that can be decoded at the soma.

In conclusion, the Tripod model shows that neurons with dendrites have computational capabilities that single-compartment models lack. Cortical neurons, which receive thousands of spikes per second, can potentially use differences in the spatial location of the input to discriminate sequential information. Dendritic integration might be able to detect this variation and transfer the result of these local computations to the soma for downstream processing.

## 2.4 Discussion

This paper has explored the computational implications of integrating dendritic compartments and voltage-gated receptors (NMDARs) into biological models of pyramidal neurons. We investigated the functional role of a simple dendritic structure in shaping the somatic response and analyzed two classes of passive dendritic compartments, proximal and distal. The present work makes three main contributions. First, we have partitioned the space of dendritic morphology, connecting the emergence and dynamics of supra-linear integration to a small number of explainable geometric and physiological parameters. Secondly, our reduced neuron model performs dendritic computations that are usually reproduced only with more complex models.And third, we have outlined how dendrites contribute to structured computation, including logic operations, frequency detection, and sequence recognition. In summary, the relatively simple Tripod neuron proposes a reduced model of dendritic structure whose functionality transcends single-compartment models. The Julia implementation of the model can be readily used in large-scale spiking neural network simulations.

In the first sections, we decomposed the model in minimal terms and investigated the contribution of various physiological and geometric factors in shaping the somatic and dendritic membrane dynamics. The comparison of human and mouse-like dendrites suggests that the former have longer integration timescales and are more excitable than their mouse counterparts, in agreement with experimental findings regarding the unique integrative properties of human dendrites (Beaulieu-Laroche et al., 2021, 2018; Fişek & Häusser, 2020). Our results confirm that human dendrites can be longer without losing the incoming current through membrane leakage; hence elongated geometries (distal thick) are possible, under our model's constraints, with human but not with mouse parameters. The maximal length obtained for the mouse is in agreement with basal and apical-oblique dendritic lengths in this species (Mohan et al., 2015). Later, we showed that independent of species-specific physiology, there is a geometric con-

straint that distinguishes between dendrites with a strong agency on the soma (100 µm to 300 µm) and those with a slow and indirect action on it (300 µm to 500 µm). The theoretical distinction between distal and proximal dendrites in terms of the maximal elicited depolarization of the soma is consistent with previous experimental and computational work (Bono & Clopath, 2017; Eyal et al., 2018; Kamondi et al., 1998; Major et al., 2008) and it refers to the electronic distance between the dendritic compartment from the soma. Overall, dendritic lengths in the range of 100 µm to 500 µm correspond to dendrites in the basal and apical-oblique region of human pyramidal cells (Spruston, 2008). Passive dendrites and cable transmission are insufficient in modeling longer dendrites (e.g., apical-tuft of layer 2/3 and 5), suggesting that active, self-regenerative mechanisms such as calcium spikes (Larkum, 2013; Larkum, Waters, Sakmann, & Helmchen, 2007) are required to transmit signals from distant dendritic input locations to the axon hillock. We associate the Tripod model to pyramidal cells rather than other types of cortical neurons for two main reasons. First, the physiological parameters adopted for both human and mouse cells and for both the membrane properties (Dasika et al., 2007; Eyal et al., 2016; Koch, 1999) and the NMDAR kinetics (Duarte & Morrison, 2019; Eyal et al., 2018) are obtained from electro-physiological studies on cortical pyramidal cells; While the interactions between dendritic integration and NMDAR non-linearity reported in the present paper could be valid for non-pyramidal cells, the different properties of NMDARs in spiny and non-spiny cells (Augustinaite, Kuhn, Helm, & Heggelund, 2014; Booker & Wyllie, 2021; Fleidervish, Binshtok, & Gutnick, 1998) may require ad-hoc model adjustments. Second, the dendritic lengths considered in the present work exceed those of other non-pyramidal cortical cells, such as layer IV spiny stellate cells (Meyer, González-Hernández, & Ferres-Torres, 1989) and aspiny cells (Maxwell, Belle, Cheunsuang, Stewart, & Morris, 2007).

We also presented a detailed analysis of the somatic excitatory post-synaptic potentials (EPSPs) when inputs are received on distal and proximal dendrites and investigated synaptic efficacy and timescales with parameters obtained from human (Eyal et al., 2018) and mouse (Avermann et al., 2012; Duarte & Morrison, 2019) *in-vitro* experiments. Our results suggest that human-like voltage-dependent receptors (NMDARs) on distal dendrites affect dendritic integration. If dendritic compartments are sufficiently segregated electrically (distal), then co-activation of neighboring synapses produces NMDA spikes and, consequently, EPSPs with a supra-linear dependence on the number of synaptic inputs. These results are in agreement with *in-vitro* empirical findings (Bono & Clopath, 2017;

Branco & Häusser, 2011; Eyal et al., 2018; Kumar et al., 2018; Polsky, Mel, & Schiller, 2004). Both electrophysiology and detailed computational models have shown that dendritic NMDA spikes can also be triggered in proximal synapses (Major et al., 2008; Mel, 1992). NMDA spikes in proximal dendrites result in larger somatic depolarization than distal ones. A few proximal NMDA spikes can drive the neuron to spike, while several distal NMDA spikes are required. Since the Tripod has only two compartments, the axial conductance to both proximal and distal synapses has to be larger than in multi-compartment models with several dendritic branches to impact the somatic membrane potential. With the present parameters, the axial conductance between the proximal and somatic compartment is large enough to trigger somatic spikes with a single depolarized proximal compartment. Therefore, our model accounts only for NMDA-induced plateau in distal dendrites because the proximal compartment can never reach the NMDA voltage-gating non-linearity without triggering a bursty response in the soma; however, this does not result in a loss of generality for our model because the amplitude of somatic depolarization remains graded with respect to the dendritic length and it is weaker for longer dendrites, as measured experimentally *in-vitro* (Major et al., 2008). Because there are only two dendritic branches in the Tripod model, we have to interpret the axial currents, the NMDA spikes, and the plateaus of the dendritic Tripod's compartments as an effective model of simultaneous depolarization in several dendritic branches of a pyramidal cell; crucially, recent evidence *in-vivo* has shown that the depolarization of a single hemi-tree of a pyramidal apical tuft, in contrast to both hemi-trees, have consequences in the behavioral scale Otor et al. (2022).

In the current literature, there is considerable variability in the parameters used to replicate NMDA spikes, in particular in the choice of the NAR, which specifies the relative difference between the peak conductances of NMDA and AMPA receptors. For example, the NAR was set to 0.25 in Duarte and Morrison (2019), 1 in Bono and Clopath (2017), 1.2 in Ujfalussy and Makara (2019), 2 (Jadi, Polsky, Schiller, & Mel, 2012), and 9 in Mel (1992). Empirical evidence, obtained mostly through indirect measurements, does report a similar level of variability. For example, NAR was found to be $\approx 0.25$ and constant throughout the dendritic tree for mice hippocampal pyramidal neurons (Strube, Gackière, Saliba, Tell, & Kessler, 2017), roughly constant across different areas of the mouse neocortex (Myme, Sugino, Turrigiano, & Nelson, 2003), and NAR was $\approx 1.8$ for human neocortical L2/3 pyramidal cells (Eyal et al., 2018). In this same spirit, we can interpret the discrepancy between the absence of NMDA spikes in mouse-like

Tripod models and the experimental evidence of NMDA-related dendritic non-linearity in mice neurons Antic et al. (2010); Larkum et al. (2022); J. Schiller et al. (2000). Rather than postulating qualitative differences between mouse and human cells, we take it as an indication of minimal requirements for the emergence of NMDA spikes in terms of timescales and steepness of NMDARs. In this respect, the variability in NMDA timescales of reduced models used in previous experiments dwarfed the difference in the NAR (50 ms (Bono & Clopath, 2017) 18.8 ms Jadi et al. (2012) 100 ms (Duarte & Morrison, 2019)). Our model identifies minimal geometric and NMDAR conditions for the occurrence of NMDA spikes and emphasizes that merely implementing NMDA receptors is not sufficient for their emergence.

Previous computational models have found that somatic EPSPs are enhanced when inputs target different, independent dendrites (Dasika et al., 2007; Li et al., 2019), in an apparent conflict with experimental and computational evidence on synaptic clustering (Bono & Clopath, 2017; Kastellakis et al., 2016; Winnubst, Cheyne, Niculescu, & Lohmann, 2015). This raises the question of whether reduced models with passive dendritic compartments are sufficiently expressive to capture dendritic integration. Our results suggest that coincidence detection can be observed under certain conditions related to the location of synaptic input and synaptic physiology. The term coincidence-detection is used to characterize several dendritic phenomena (Spruston, 2008), e.g., the generation of a spike, or an activity burst, following simultaneous excitatory inputs. In particular, it is used for both the somatic depolarization resulting from simultaneous spikes on two segregated dendrites (Dasika et al., 2007), and for the non-linear response resulting from co-activation of neighboring synapses (Mel, 1992; Ujfalussy & Makara, 2019). Our model can express both forms of dendritic coincidence detection in terms of a single variable, i.e., dendritic length. The fundamental role of dendritic length has been discussed in Jadi, Behabadi, Poleg-Polsky, Schiller, and Mel (2014) and was included in their two-layer network model of dendritic integration. However, the model only accounted for neuronal firing rates and did not model sub-threshold membrane dynamics. In addition, we explored the differences between inhibitory input onto the somatic and dendritic compartments. We associated dendritic inhibition with the activity of somatostatin interneurons (SST), and somatic inhibition to parvalbumin interneurons (PV) (Huang & Paul, 2019; Tremblay, Lee, & Rudy, 2016). From our fit on guinea pig pyramidal neurons (Miles et al., 1996), the $GABA_A$receptors on the dendritic membrane had longer timescales and their maximal conductance was smaller

than their somatic counterparts. We tested the differences between these two types of inhibition by comparing their efficacy in attenuating the somatic EPSP and showed that inhibition on the soma was effective in preventing spiking activity for a short period of time. In contrast, dendritic inhibition could silence the neuron for longer durations when applied on the same dendritic branch as excitation, but its maximal effect on the soma was limited and delayed, consistent with the current understanding of somatic and dendritic inhibition. The fast-spiking PV interneurons acting on the soma are associated with feed-forward, time-precise inhibition, while the slower action of SST cells regulates the dendritic potential via feedback inhibition (Kee, Sanda, Gupta, Stopfer, & Bazhenov, 2015; Tepper, Wilson, & Koós, 2008; Tremblay et al., 2016). In computational terms, localized inhibition allows for external gating of the dendritic stimulus by selecting which dendritic pathway is allowed to integrate the signal, and to communicate with the soma. Pathway selection has been proposed as a cortical mechanism for flexible routing of sensory stimuli (G. R. Yang et al., 2016; Zajzon et al., 2019) and, more recently, it has been demonstrated that networks that leverage dendritic gating support efficient, durable, and fast learning (Sezener et al., 2021).

The Tripod succeeds in expressing coincidence-detection and pathway-selection because of two fundamental properties of its reduced dendritic tree: non-linear integration and electronic segregation of dendritic compartments. Our principled dendritic reduction aligns well with results from data-driven reductions that have been used to distill dendritic computations in the simplest architecture that could explain the data (Beniaguev, Segev, & London, 2021; Ujfalussy et al., 2018; Wybo et al., 2021-01-26, 2021); in particular with the work by Ujfalussy et al. (2018) which shows how two compartments with non-linear integration and different timescales are sufficient to predict with high accuracy neural response under *in-vivo* stimuli conditions. However, dendritic simplification comes at a cost, synapses have no spatial resolution in the dendritic compartments but are all lumped together. Conversely, real dendrites are spatially extended and host spines, receptors, and ionic channels throughout the entirety of the dendritic cable. The interaction between synapses is determined by their relative distance and their spatial organization governs homeostatic mechanisms and heterosynaptic plasticity (Kirchner & Gjorgjieva, 2021; Oh, Parajuli, & Zito, 2015; Triesch, Vo, & Hafner, 2018; Wu, Hengen, Turrigiano, & Gjorgjieva, 2020). The continuous spatial distribution along the dendritic cable also has important implications for signal integration: single-branch synaptic activation that fol-

lows the dromic direction - from the tip towards the soma - results in stronger somatic depolarization than activation in antidromic directions (Branco et al., 2010). Such distinctions are impossible under the constraints of our model, as we neglect spatial interactions along elongated dendrites comprising multiple compartments. In addition, considering only two compartments limits the computations available to each Tripod model to the dendritic configuration instantiated in the model, e.g. symmetrical or asymmetrical. In the brain, each cell has hundreds of dendritic branches with a broad distribution of lengths, spatial arrangements and membrane physiology. Overall, the Tripod has to be considered as a compromise between accurate modeling of dendritic processes and implementing them in large-scale cortical circuits. As such, it provides a step forward from point-neuron models.

Dendritic NMDA spikes cause a long-lasting depolarization in the somatic compartment of the Tripod neuron. The duration of the depolarized state depends on dendritic length and the strength of synaptic events and it could last on the order of 100 ms, in agreement with experimental results (Branco & Häusser, 2011; Major et al., 2008; Milojkovic, Radojicic, & Antic, 2005; J. Schiller et al., 2000). This dendritic "UP-state" is governed by a self-regenerative process triggered by co-active synapses and has a timescale that is two to three times longer than the membrane's. This allows the UP-state to encode information about recent activity and the maintenance of this information can support an activity-silent processing memory at the neuronal level (Fitz et al., 2020; Stokes, 2015). Dendritic memory is similar to priming in the sense that the neuron responds faster and more strongly to a retrieval cue when the encoding signal occurs close in time. In contrast to short-lived synaptic memory (Mongillo, Barak, & Tsodyks, 2008), dendritic memory is more effective when the retrieval cue follows a different synaptic pathway than information encoding. The plateau potentials that support dendritic memory have been considered a candidate mechanism for linking neuronal to behavioral timescales (Augusto & Gambino, 2019; Bittner et al., 2015; Bittner, Milstein, Grienberger, Romani, & Magee, 2017). Dendritic memory can bind information over time, and our results suggest that it can play a role in temporal processing that is beyond single-compartment models.

Since the introduction of the NEURON simulator in 1989 (Hines, 1989), the tools for modeling dendrites have come a long way (Poirazi & Papoutsi, 2020) and the introduction of dendritic integration in cortical circuits is becoming increasingly accessible to computational research. Examples of these advances are Dendrify (Pagkalos, Chavlis, & Poirazi, 2022) and NESTML Plotnikov et al.

(2016), which allow for simulating neurons with dendrites in cortical circuits in Brian2 and NEST. The Tripod model can be easily replicated within these frameworks. In addition, recent technical advances in neuromorphic computing have successfully implemented passive dendritic compartmentalization in hardware (Kaiser et al., 2022; S. Yang et al., 2021), boosting the applicability of dendritic computation in machine-learning contexts (Guerguiev, Lillicrap, & Richards, 2017; Sezener et al., 2021). The work presented here can guide this line of implementational research as it provides a simple, scalable model that captures important computational primitives at the single neuron level beyond the point neuron.

# Tables

Table 2.1: Parameters for the axosomatic compartment of the Tripod neuron. Values corresponds to those proposed in Brette and Gerstner (2005), except for the somatic leak conductance which is set to 40 nS, as in Bono and Clopath (2017).

| Symbol | Description | Value | Unit |
|---|---|---|---|
| $g_L$ | Membrane leak conductance | 40 | nS |
| $C_m$ | Membrane capacitance | 281 | pF |
| $V_r$ | Resting membrane potential | -70.6 | mV |
| $V_T$ | Threshold potential | -50.4 | mV |
| $u_{th}$ | Spike onset threshold | 0 | mV |
| $u_r$ | Reset potential | -70.6 | mV |
| $\Delta_T$ | Slope factor | 2 | mV |
| $\tau_w$ | Spike-triggered adaptation time scale | 144 | ms |
| $a$ | Subthreshold adaptation conductance | 4 | nS |
| $b$ | Spike-triggered adaptation increment | 80.5 | pA |
| $t_{up}$ | Spike width (soma clamped at 20 mV) | 1 | ms |
| $t_{ref}$ | Refractory period | 2 | ms |

Table 2.2: Dendritic physiology parameterized for human and mouse, following Koch (1999) and Eyal et al. (2016).

| Symbol | Description | Human | Mouse | Unit |
|---|---|---|---|---|
| $r_m$ | Membrane resistance | 39 | 1.7 | $k\Omega\,cm^2$ |
| $r_{ax}$ | Intracellular resistance | 200 | 200 | $\Omega\,cm$ |
| $c_m$ | Membrane capacitance | 0.5 | 1 | $\mu F/cm^2$ |
| $V_r$ | Resting potential | -70.6 | -70.6 | mV |

Table 2.3: Parameters for dendritic compartments computed from the physiological specifics in Table2.2.

| Symbol | Description | Human | | Mouse | | Unit |
|---|---|---|---|---|---|---|
| | | distal | proximal | distal | proximal | |
| $l$ | Dendritic length | 400 | 150 | 400 | 150 | µm |
| $d$ | Dendritic diameter | 4 | 4 | 4 | 4 | µm |
| $g_m$ | Leak conductance | 1.29 | 0.32 | 29.57 | 7.39 | nS |
| $\tau_d$ | Membrane timescale | 1.48 | 0.22 | 1.11 | 0.36 | ms |
| $g_{ax}$ | Axial conductance | 15.71 | 62.83 | 15.71 | 62.83 | nS |
| $C_m$ | Membrane capacitance | 25.13 | 6.28 | 50.27 | 12.57 | pF |

Table 2.4: Parameters for mouse (Duarte & Morrison, 2019) and human (Eyal et al., 2018) excitatory synapses. Inhibitory synapse parameters derived from Miles et al. (1996).

| Symbol | Description | Human | | Mouse | | Unit |
|---|---|---|---|---|---|---|
| | *Excitatory* | AMPA | NMDA | AMPA | NMDA | |
| $E_r$ | Reversal potential | 0 | 0 | 0 | 0 | mV |
| $\tau_r$ | Rise time constant | 0.26 | 8 | 0.26 | 1 | ms |
| $\tau_d$ | Decay time constant | 2 | 35 | 2 | 100 | ms |
| $\bar{g}_{syn}$ | Peak conductance | 0.73 | 1.31 | 0.73 | 0.159 | nS |
| $\gamma$ | Voltage-gating slope | — | 0.075 | — | 0.062 | $mV^{-1}$ |

| Symbol | Description | Soma | Dendrites | | Unit |
|---|---|---|---|---|---|
| | *Inhibitory* | $GABA_A$ | $GABA_A$ | $GABA_B$ | |
| $E_r$ | Reversal potential | $V_r$ | $V_r$ | -90 | mV |
| $\tau_r$ | Rise time constant | 0.5 | 4.8 | 30 | ms |
| $\tau_d$ | Decay time constant | 15 | 29 | 400 | ms |
| $g_{syn}$ | Peak conductance | 0.38 | 0.27 | 0.006 | nS |

## 2.5 Appendix

### Appendix A: Minimal axial conductance

In order to simplify the analytical treatment, we consider the fixed point of the LIF equation, removing the exponential and the spike non-linearity. Because the slope of the AdEx nullcline is monotonous after $V_s > V_T$, there is no qualitative difference in the presence of stationary input. Additionally, we consider a neuron driven solely by excitatory inputs. With two dendrites ($i = 1, 2$), the reduced tripod circuit is described by the following system of equations:

$$C_s \frac{dV_s}{dt} = -g_m^s(V_s - E_r) + I_d \tag{2.17}$$

$$C_d^i \frac{dV_d^i}{dt} = -g_m^i(V_d - E_r) - I_d^i - g_e^i(V_d^i - E_{\text{GluRs}}) \tag{2.18}$$

$$I_d^i = -g_{ax}^i(V_s - V_d^i) \tag{2.19}$$

The system can be solved algebrically and, for $E_{\text{GluRs}} = 0$, results in:

$$V_s = E_r\left(1 - \frac{G^2 g_e^1 g_{ax}^1 + G^1 g_e^2 g_{ax}^2}{\sum_i^{1,2}(g_{ax}^i G^{i+1})(G^i - g_{ax}^i) + G^1 G^2 g_m^s}\right)$$

where $G^i = g_e^i + g_m^i + g_{ax}^i$. If the conductance between one of the dendrites and the soma is zero (neuron with single dendrite), the equation reduces to:

$$V_s = E_r\left(1 - \frac{g_e^d g_{ax}^d}{g_e^d(g_{ax}^d + g_m^s) + g_m^s(g_m^d + 2g_{ax}^d)}\right)$$

In the limit of very large excitatory conductances ($g_e \gg g_m + g_{ax}$), the neuron is a simple voltage divider and the somatic potential is given by:

$$V_s = E_r\left(1 - \frac{g_{ax}^d}{g_{ax}^d + g_m^s}\right) \tag{2.20}$$

This situation corresponds to a neuron with a single dendrite and maximally excited in the $d$-th dendritic compartment. Hence, the condition for the neuron to reach the spike threshold is:

$$g_{ax}^d > \frac{E_r - V_T}{V_T} g_m^s = \beta g_m^s \tag{2.21}$$

where $V_T$ is the firing threshold of the somatic compartment. Eq.2.13 defines the minimal condition for the dendritic compartment to elicit somatic spikes. When the axial conductance $g_{ax} > \beta g_m^s$, a full depolarization of the dendrites suffices to generate spikes in the soma.

Within the constraints of the Tripod model, some relevant parameters are fixed by the axo-somatic model used, namely the somatic leak conductance $g_m^s$, the resting membrane potential $E_r$, and the spike threshold $V_T$, which are all defined by the AdEx model Brette and Gerstner (2005). The remaining parameter for the axial conductance $g_{ax}^d$ is determined entirely by the cable geometry (diameter

$d$ and length $l$) (Rall, 2011) along with the dendritic membrane physiology as expressed in Eqs. 2.5 2.6, 2.7. Once the physiological details are defined (Dasika et al., 2007; Eyal et al., 2016), we can distinguish between geometries that elicit spikes and geometries that do not.

## Appendix B: Excitatory synaptic interactions in the passive cable

Dasika et al. (2007) shows that a model neuron with stationary conductance depolarizes more when the inputs are distributed than when synapses are localized on a single branch. This can be demonstrated by determining the equivalence between a circuit with two active synapses on different branches ($G_1$ and $G_2$) and a circuit with one single active conductance $G_s$ ($G_1 = G_s$ and $G_2 = 0$). The following equivalence holds:

$$G_s = \frac{G_1 + G_2 + 2G_1G_2\frac{1}{g_{ax}^d \cdot g_m}}{1 - G_1G_2\left(\frac{1}{g_{ax}^d \cdot g_m^d}\right)} \tag{2.22}$$

where $g_{ax}$ and $g_m$ are the axial and the leak conductances of the passive membrane patch (the dendritic compartments), respectively. The equation shows that, in the presence of segregated dendritic compartments ($g_{ax} < \infty$), $G_s$ is always greater than $G_1+G_2$. The interaction has been further simplified in (Li et al., 2019). The authors reduce the interaction between synapses in a second-order approximation where the total current, in case of simultaneous firing synapses, is given by:

$$\begin{aligned} I_{syn} &= \sum g_e^i(E_e - v) + \Delta_I \\ \Delta_I &= \sum_i \sum_j \alpha_{i,j} g_e^i g_e^j (E_e - v) \end{aligned} \tag{2.23}$$

where the $E_e$ states the receptor reverse potential. The interaction has been approximated to a binary function where the second-order synaptic contribution $\alpha_{i,j}$ is almost zero for synapses on different branches and negative for synapses on the same branch.

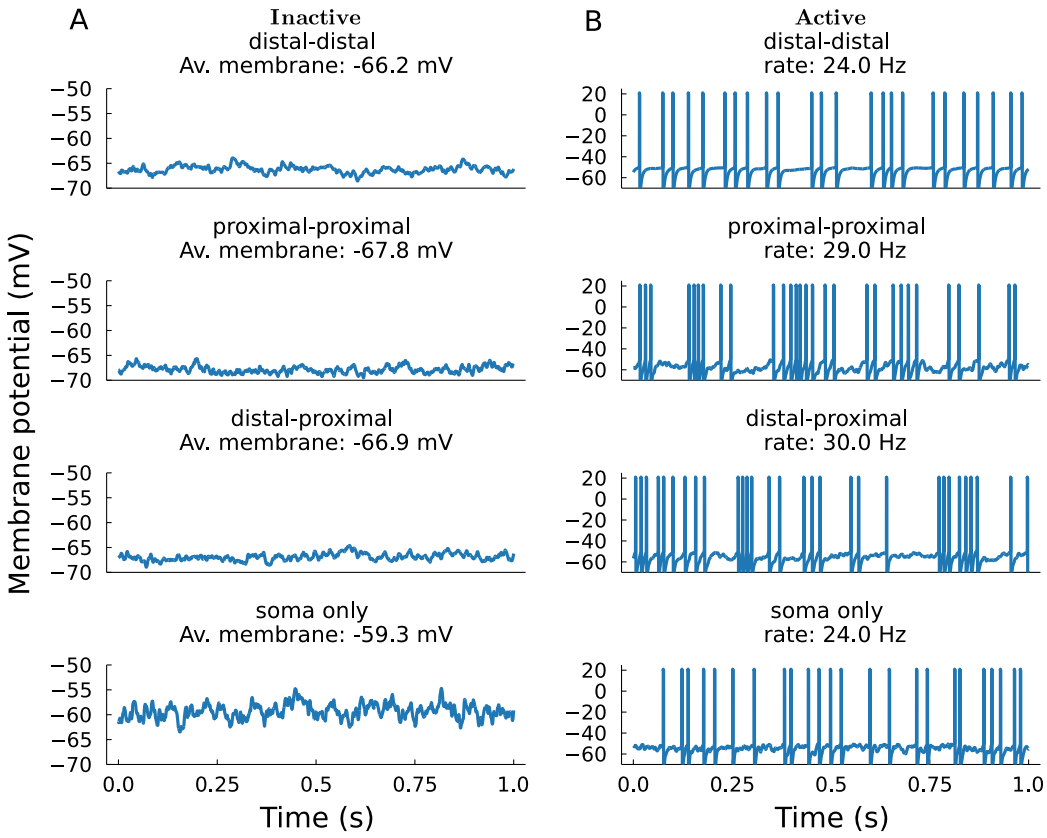
## Appendix C: Membrane dynamics across experiments

### Balanced inputs condition

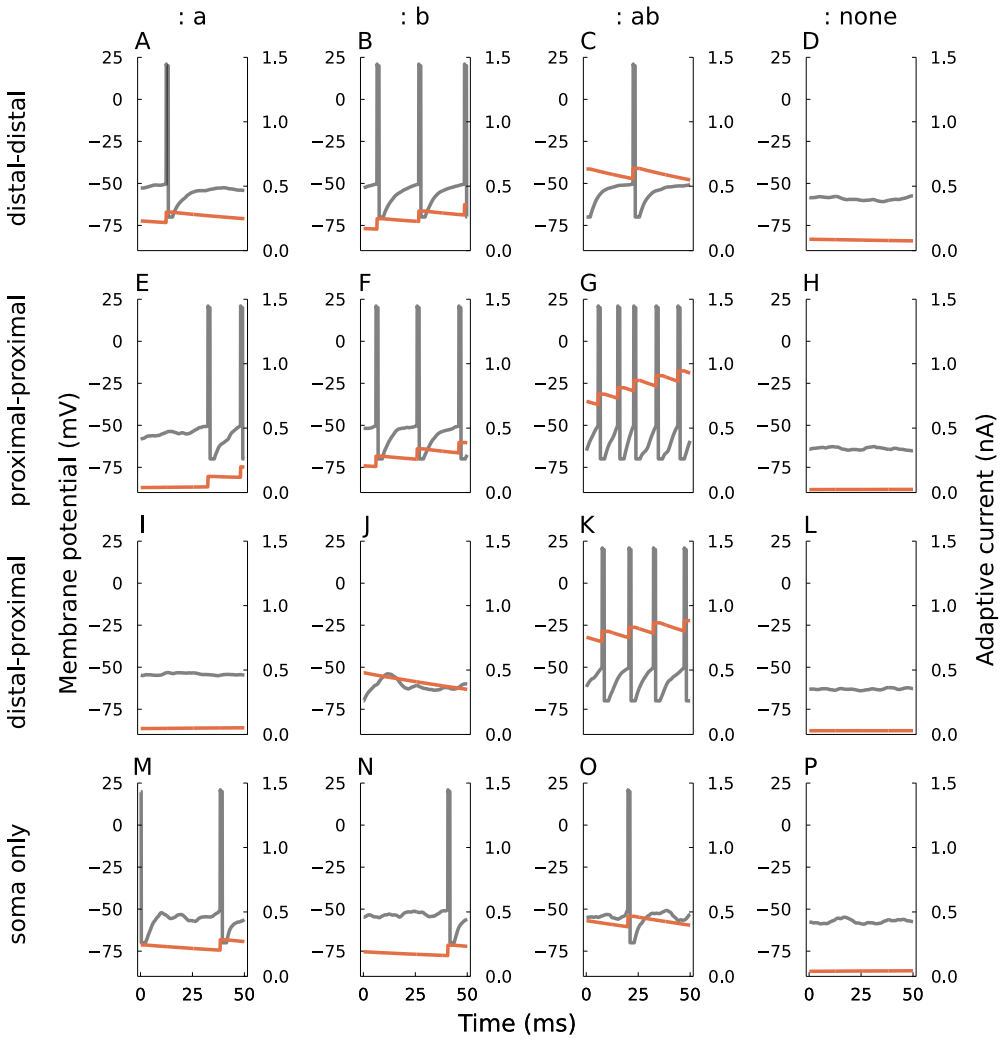
To study the model in naturalistic conditions we stimulated the Tripod with excitatory and inhibitory spike trains. We defined a balanced condition such that the somatic compartment is depolarized and both glutamatergic and gabaergic conductances are large; this input configuration ensures that the dendritic computations investigated are not artifacts of the unrealistic setup. The balance is obtained by fixing the excitatory firing rate to 3 kHz and varying the corresponding inhibitory rates. This procedure results in inhibitory firing rates of 3 kHz for distal dendrites (400 μm); 4.8 kHz for proximal dendrites (150 μm); and 1 kHz for the soma only model. With these inputs, the neuron (almost) never fires and the somatic compartment rests around −67 mV for the three dendritic configurations, distal-distal, distal-proximal, and proximal-proximal. Following the protocols presented in the Results section, each dendrite was activated by doubling the excitatory input; when this happens, the dendrite depolarizes and causes the neuron to fire. When both dendrites are activated the neuron's firing rate is approximately 30 Hz, with little variations between different dendritic configurations. In experiments Fig.2.7, Fig.2.8, Fig.2.9 additional excitatory noise was injected in the somatic compartment to ensure firing activity when one single dendrite was activated.

The *soma only* balanced configuration was also defined on a similar basis, although the soma compartment needs to be more depolarized −60 mV to initiate spikes when one of the input pathways is activated. Fig.2.1 illustrates these effects and shows the three Tripod configurations and the soma-only condition in the inactive (A) and active (B) states.

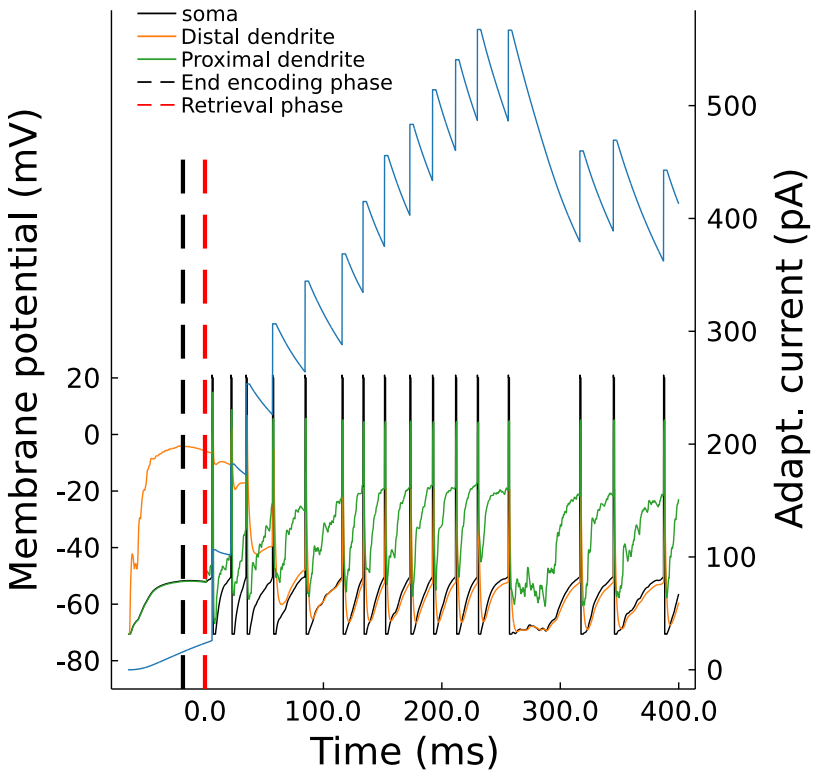
An important outcome of the balanced configuration is to avoid artifacts of the AdEx model, as discussed in Górski, Depannemaecker, and Destexhe (2021). When the AdEx is strongly excited, for example with strong GluRs stimulation or injected currents, the neuron starts firing and the adaptive current rapidly rises. If the stimulation terminates abruptly, the adaptive current pulls down the membrane voltage, generating unnatural hyper-polarization. In Fig.2.2, Fig.2.3, Fig.2.4 we show that, due to our balance condition, these artifacts are not observed in the Tripod model and realistic membrane dynamics can be observed.



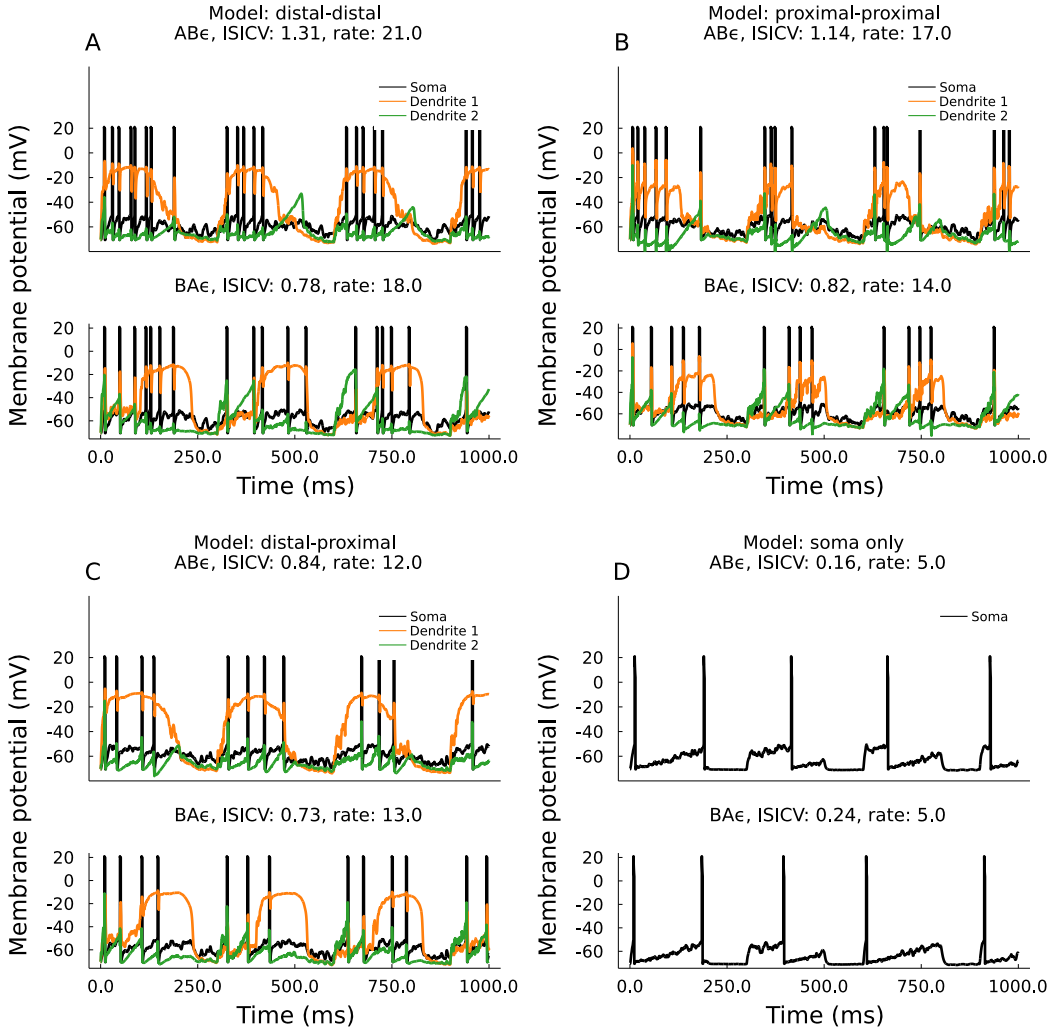

SI Figure 2.1: **Membrane dynamics of Tripod models**
(**A**) Model activity in the inactive condition, with 3 kHz excitatory inputs and dendritic-length-specific values for inhibition (see main text). (**B**) Membrane dynamics in the active mode with doubled excitatory input rates.




SI Figure 2.2: **Tripod dynamics in the Logical operators task**
**(A)-(P)** Somatic membrane potential (gray) and adaptive current (red) for the 4 dendritic configurations (vertical arrangement) and the 4 input configurations (horizontal arrangement).

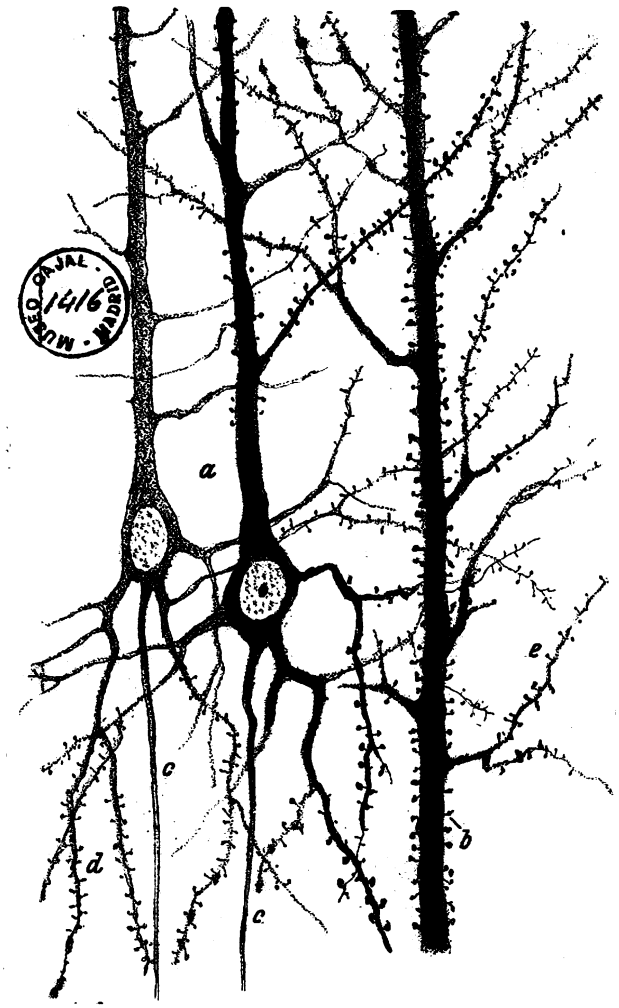



SI Figure 2.3: **Tripod dynamics in the Dendritic memory operators task** The figure depicts somatic membrane voltage and adaptive current (black and blue) as well as dendritic voltages (red and green) for the distal-proximal dendritic configuration. Vertical dashed lines in red and black indicate the end of the encoding phase and the onset of the retrieval phase, see main text for clarifications.




SI Figure 2.4: **Tripod dynamics in the Sequence recognition task**
Somatic membrane potential (black) and dendritic membrane potentials (red and green) for the four Tripod configurations. The two panels in **(A)** - **(D)** show the dynamics of the Tripod in the AB$\epsilon$ and the BA$\epsilon$ sequence, see main text. To facilitate the comparison, these simulations were run with frozen noise input. As such, the difference between the upper and lower panels is only the strength of the excitatory inputs which is doubled in the stimulated compartment: input A in the AB sequence (top panels) and input B in the BA sequence (bottom panels).

# 3 | Dendritic non-linearities enable up-down states in single neurons



## Abstract

Under intense synaptic bombardment, cortical neurons tend to operate in a high-conductance state (HCS) where the cell membrane hovers close to the firing threshold and is, therefore, more sensitive to input fluctuations. Typically associated with active processing, these states are transient and their characteristics depend on input density, neuronal and synaptic physiology, as well as computational demands. Under certain operating regimes, cortical neurons show persistent transitions between the HCS (also referred to as an Up-state) and a hyperpolarized Down-state. This type of bistability has been ascribed to cellular and synaptic anatomy and physiology. Whereas the dependence of Up-state on input fluctuations and the role of NMDA receptors has been well-characterized, less is known about the detailed mechanisms that govern the Up-Down state (UDS) transitions and whether this bistability is primarily due to the input or to neuronal and synaptic properties.

The present work aims to fill this gap by investigating the emergence of UDS in a minimal, three-compartment neuron model. This neuron has two segregated passive dendrites and an axosomatic compartment and incorporates dendritic geometry and physiology from human cortical pyramidal cells. Dendrites and soma express conductance-based synapses with glutamatergic and GABAergic receptors. We drive the dendritic compartments with fluctuating, balanced excitatory/inhibitory Poisson inputs (E/I balance). We show that only neurons expressing NMDARs support distinct UDS transitions. Although balanced inputs themselves do not trigger the UDS, the bistable dynamics emerge when input fluctuations increase within regimes observed in cortical activity ($CV_{ISI} < 1.2$). Neurons expressing NMDARs display sharper bistability than models without NMDARs. Voltage-gated receptors in dendrites make it easier to pick up input correlations, whereas a point-neuron model or passive dendrites would require large fluctuations to capture the same effects.

This work connects local, compartmentalized E/I balance in dendrites with two important cortical properties: HCS and UDS transitions. It also clarifies the role of NMDARs in governing the response to fluctuations in the input.

## 3.1 Introduction

Word processing happens at multiple time scales because the relevant pieces of pre-lexical information are delivered with different characteristic time constants. This also requires multiple matching regimes of integration in relation to the input rate of the relevant linguistic building blocks. As we have seen in the previous chapter, the Tripod neuron provides the right spatial and temporal organization to support the required computational machinery to recognize sequences of inputs at different timescales. In this chapter, we verify that this computational machinery generates output consistent with the dynamics observed in vivo recordings of single cells' membrane potential. We study the dynamics of the Tripod model when it undergoes high-rate pre-synaptic firing, as observed in cortical networks. The analysis shows that the segregated dendrites and voltage-dependent NMDARs, which play a major role in dendritic computation, are also necessary to reproduce the transition between the up and down states of the cell membrane and favor the sparse and irregular firing of the Tripod when in the balanced state.

### The balanced state in cortical networks

In cortical networks, neurons are subjected to a balanced barrage of spikes from glutamatergic and GABAergic pre-synaptic neurons (Shu, Hasenstaub, & McCormick, 2003); thus, their membranes are traversed by excitatory and inhibitory (E/I) synaptic currents correlated in time and intensity (Okun & Lampl, 2008). Due to the continuous synaptic bombardment, the synaptic conductance increases by a factor of 3 relative to the membrane conductance in the cell's resting state, resulting in the so-called high-conductance state (HCS) (Destexhe, Rudolph, & Paré, 2003). In the HCS, the neuronal membrane potential is strongly depolarized and stays close to the spiking threshold; somatic firing then becomes sensitive to small variations (fluctuations) in the pre-synaptic activity (Baker, Zhu, & Rosenbaum, 2020; Destexhe et al., 2003; Ebsch & Rosenbaum, 2018; Zhou & Yu, 2018). The E/I balance condition is met throughout the cell, and the dendritic arborization accommodates for locally fine-tuned clusters of both glutamatergic and GABAergic synapses (local E/I balance, Gjorgjieva, Drion, & Marder, 2016; Iascone et al., 2020; Liu, 2004). As a result, the cells are in a critical state that allows for dynamic compartmentalization of computation with sub-regions of the dendritic tree driving the generated somatic spikes (Otor et al., 2022).

The HCS is measured regularly *in-vivo* and can be considered as the operating point of cortical neurons in awake animals (Destexhe, Hughes, Rudolph, & Crunelli, 2007). However, the recordings reveal that the HCS is only one aspect of the complex cortical dynamics. The somatic membrane of cortical neurons is typically bi-stable, and its membrane potential fluctuates around two metastable values known as Up-Down States (UDS). The up-state has the properties of the HCS. The down-state corresponds to a silent neuron, with membrane potential fluctuating close to the resting state (Cowan & Wilson, 1994; Wilson & Kawaguchi, 1996). The transitions between up- and down-states have been explained as the switching from active to inactive processing in local cortical circuits, driven by the collective dynamics of the network (Jercog et al., 2017; Luczak, Bartho, Marguet, Buzsaki, & Harris, 2007). Analytical insights and network simulations have shown that two conditions are necessary for the emergence of the up-down dynamics in spiking networks: balanced excitation and inhibition as well as strong excitatory recurrent connections (Compte, Sanchez-Vives, McCormick, & Wang, 2003; Cossart, Aronov, & Yuste, 2003; Droste & Lindner, 2017; Tukker, Beed, Schmitz, Larkum, & Sachdev, 2020). The network models that are used to reproduce the UDS are normally hyperpolarized until external noise, intrinsic network fluctuations, or special pacemaker cells trigger the onset of the up-state. Once stimulated, the recurrent excitatory connections maintain the up-state because of their strong recurrence until adaptation or homeostatic inhibition takes over and restores the down-state. Crucially, the conditions required for such bi-stable networks are relatively narrow, and the models need to be fine-tuned to avoid pathological behaviors and elicit the intended up-down dynamics (Brunel, 2000; Maksimov, Diesmann, & van Albada, 2018).

In contrast to network-driven bi-stability, experimental studies have suggested a physiological explanation for the UDS which emphasizes the relationship between the up states and the onset of plateau potentials in the dendrites (Antic et al., 2010; Milojkovic et al., 2005; Oikonomou, Singh, Sterjanaj, & Antic, 2014). This view suggests that, whereas the up-state is triggered by the network activity, its bi-stability is primarily a physiological phenomenon shaped by voltage-gated receptors and other regenerative dendritic events (Antic et al., 2010; Larkum, 2013) or rectifying ionic currents (Loewenstein et al., 2005; Sanders, Berends, Major, Goldman, & Lisman, 2013). However, it remains unclear whether NMDARs have a role in driving the up-down cortical dynamics, especially in light of recent findings showing intact UDS despite NMDAR blocking (Palmer et al.,

2014; Smith et al., 2013). Theoretical explanations also fall short in this regard and the mechanisms for intra-cellular bi-stability remain uncertain. On the one hand, the UDS seems to emerge from the dendritic balance between several, short-lived excitatory and inhibitory currents (AMPA/GABAa) distributed across the entire cell (Larkum, 2022); on the other hand, computational insights suggest that the dendritic architecture and the physiology of the NMDARs are essential for picking up correlations in the input and inducing the up-state (Benucci, Verschure, & König, 2004; Papoutsi, Sidiropoulou, & Poirazi, 2014).

Given the previous work, three questions remain open. First, is the local dendritic balance sufficient to cause the onset of UDS, independently of dendritic NMDA spikes? Second, how does the NMDARs' non-linearity interact with the local E/I balanced state to trigger the onset of the up-states? and third, what is the dendritic resolution necessary to reproduce these effects in reduced computational models? Here we investigate the relationship between E/I network balance, high conductance states, and up-down states, and we propose a simple, integrated description of how these phenomena relate to each other: up-down states result from fluctuations around the E/I balanced state and are amplified by dendritic non-linearities. In order to investigate these issues in this study, we leveraged a recently published model of dynamic dendritic integration, the Tripod neuron (Quaresima et al., 2022). We exposed this neuron to barrages of presynaptic spikes at increasing rates. In the balanced condition, we trigger up-down transitions by modulating the coefficient of variation of the inter-spike-interval ($CV_{ISI}$) of the excitatory synaptic inputs. We demonstrate that local balance and small fluctuations in the inputs are sufficient to explain the rise of strongly bi-stable dynamics. The UDS in the Tripod neuron exhibits similar characteristics to those observed in pyramidal cells *in-vivo*.

## 3.2 Methods

### 3.2.1 The Tripod neuron model

We investigated up-down dynamics in the *Tripod* neuron which is composed of three computational elements, or compartments (Quaresima et al., 2022). The neuron has an axosomatic compartment, representing the soma and perisomatic dendritic segments, which is modeled as an adaptive exponential integrate-and-fire neuron (Brette & Gerstner, 2005). In addition, it has two electrotonically segregated passive dendritic compartments that are coupled to the soma in a Y-

shape (Fig.3.1*A*) with membrane timescales determined by the dendritic length (100 µm to 500 µm). A characterization of the model can be found in the previous Chapter of this thesis, and we do not repeat the details of the Tripod implementation here. In this study, we used human parameters for the dendritic and synaptic physiology, which in the previous study proved to have greater computational capacity. We focus on the synaptic mechanism of the NMDA receptor (NMDAR), which is crucial for the present work.

***Synaptic dynamics.*** The Tripod neuron implements two excitatory glutamatergic receptors with fast (AMPA) and slow (NMDA) dynamics, and two inhibitory GABAergic receptors with short ($GABA_A$) and long ($GABA_B$) timescales. The receptors dynamics were modeled as conductances with double-exponential kinetics (Roth & van Rossum, 2009):

$$g_k(t) = \bar{g}_k^{\text{syn}} \mathcal{N}_k \left( \exp\left(-\frac{t-t_0}{\tau_k^r}\right) - \exp\left(-\frac{t-t_0}{\tau_k^d}\right) \right) \tag{3.1}$$

where $t_0$ is the pre-synaptic spike time, $k \in \{\text{AMPA}, \text{NMDA}, \text{GABA}_A, \text{GABA}_B\}$ and $\tau_k^r$ and $\tau_k^d$ are the specific rise and decay time constants of each receptor (see Table3.3).

The conductance gating of the NMDA receptor was dependent on the intracellular depolarization which was captured by a multiplicative voltage-gating mechanism:

$$\begin{aligned} g_{\text{NMDA}} &= \bar{g}_{\text{NMDA}}^{\text{syn}} G(V) \\ G(V) &= \left(1 + \frac{C}{3.57\,\text{µmol/L}} \cdot e^{-\gamma V}\right)^{-1} \end{aligned} \tag{3.2}$$

where $V$ is the membrane voltage of the dendritic compartment and $\gamma$ controls the steepness of the voltage dependence. The extracellular concentration $C$ of magnesium ions $Mg^{2+}$ was fixed at 1 µmol/L. The equations and parameters for the NMDA receptor are based on Jahr and Stevens (1990).

To investigate the role of NMDA in regulating the onset and duration of the up states, we considered two synaptic configurations, one with and another one without NMDA receptors. In the condition without NMDA receptors (AMPA-only), the AMPA conductance was scaled such that its maximal conductance matched the sum of AMPA and NMDA conductances in the condition with NMDA receptors. In this way, we could focus our investigation on two characteristics of the NMDA receptor activation: its voltage dependence and its extended timescale.

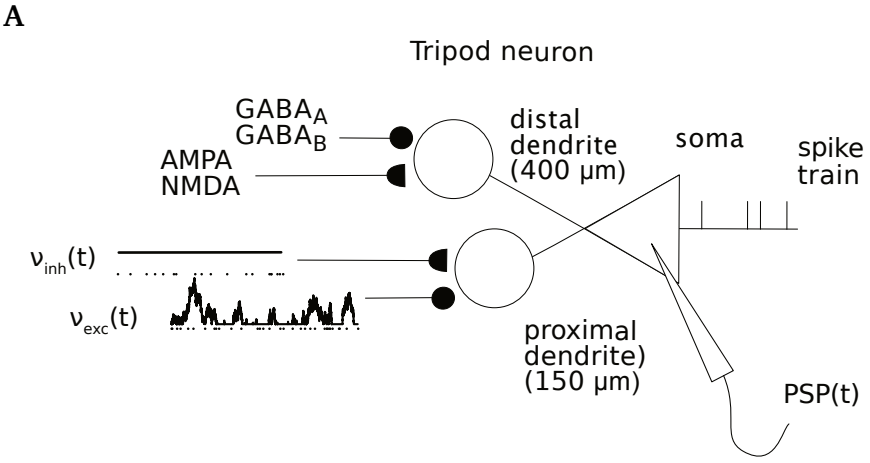



Figure 3.1: **Schematic of the Tripod neuron and inputs**
The Tripod neuron has two dendritic and one axosomatic compartment. Proximal dendrites are 150 µm long, distal dendrites are 400 µm long. Each dendritic compartment hosts glutamatergic (AMPA, NMDA) and GABAergic ($GABA_A$, $GABA_B$) receptors. Input to the neuron is drawn from a Poisson distribution with rates $\nu_{exc}(t)$ and $\nu_{inh}(t)$. The excitatory input rate was varied between 0.5 kHz to 80 kHz and the inhibitory was adjusted to the $\nu_{ext}(t)$ times a factor ($k_{E/I}$) to maintain the balance condition.

Model parameters for the somatic and the dendritic compartments are shown in Table3.1 and Table3.2, respectively. They are the same as in the original Tripod model (Quaresima et al. (2022) and Chapter 2 of this thesis), except for the reset potential $u_r$ which was set to −55 mV, following Bono and Clopath (2017) and Duarte and Morrison (2019). Similarly, all the parameters for the synapse types used in the Tripod neuron are described in Table3.3.

## 3.2.2 Excitatory and inhibitory inputs

Cortical neurons are exposed to continuous, intense synaptic bombardment from local and long-range afferents. During awake activity, the E/I input streams are balanced in cortical networks and the post-synaptic neurons remain depolarized and fire at sparse, low rates. We replicated these physiological conditions in the Tripod neuron by stimulating the model with Poisson-distributed excitatory and inhibitory (E/I) spikes at a high rate. The excitatory synaptic strength is maintained fixed throughout the study ($J_{exc} = \bar{g}_{syn}^{\text{GluRs}}$, Table3.3) while the excitatory input rate was varied in the range of $\nu_{exc} = 0.5$ kHz to 80 kHz. This stimulation protocol results in synaptic conductances larger than the membrane leak conductance (HCS). However, if the excitatory stimuli overcome the inhibitory ones, the neuron is fully depolarized and enters a non-physiological

firing regime; similarly, it remains hyperpolarized in the opposite case. These two input configurations are discussed in Appendix A. Thus, to reproduce the physiological condition of cortical cells we adjust the inhibitory firing rate such that the two input streams are balanced.

**Balance via modulation of the inhibitory rate**

In line with experimental evidence on the HCS and previous computational studies (Kuhn, Aertsen, & Rotter, 2004; Wilson, 2008), we define the balance excitatory/inhibitory (E/I) condition via the average membrane potential of the somatic compartment. The E/I balance is achieved when the somatic membrane potential hovers right below the threshold potential for spike generation. We chose this $V_s$ =−55 mV because it corresponds to the average potential of the up-state measured *in-vivo* (Wilson & Kawaguchi, 1996) and was also used in previous modeling studies of the balanced state (Kuhn et al., 2004). To reach this average potential, we simulate a Tripod neuron in the free-membrane condition, i.e., without the spike generation mechanism in the soma, and compute the necessary inhibitory rate via gradient-free optimization. We ran a grid search for all the excitatory rates (0.5 kHz to 80 kHz) and the inhibitory rate ratios ($k_{E/I}$= 0 to 2, with resolution 0.01). We chose the value $\bar{k}_{E/I}$such that the average of the membrane potential on 10 simulations was closer to the sub-threshold membrane potential.

The illustration in Fig.3.2*A* briefly recapitulates the method; the inhibitory rate (marked red) is used to attain the balanced condition. We compute the E/I ratio for Tripod models with NMDA and AMPA-only receptors and for the soma-only model. The $k_{E/I}$ resulting from the optimization algorithms are portrayed in Fig.3.2*B*. The upper panels show the relative strength of the excitatory to inhibitory rate necessary to reach the balance ($\bar{k}_{E/I}$). The minimum rate for the balance condition to be achieved, at least for some dendritic length, depends on the presence or not of NMDARs. It is approximately 0.7 kHz for the standard Tripod model, and 1.5 kHz for AMPARs-only models; these values are indicated on the panels with a dashed black line. For the soma-only model, the minimum input is 2 kHz circa. The color shades in the panels indicate the dendritic length of the balanced model. Crucially, the balance is geometry-dependent in both the NMDARs and AMPARs conditions; longer dendrites require a higher E/I ratio than shorter dendrites, corresponding to higher values of the $\bar{k}_{E/I}$. Further insights into the difference between the geometry dependence of the $\bar{k}_{E/I}$ are presented in Fig.3.1

### Input fluctuations

The high conductance state (HCS) reached via balanced E/I synaptic inputs is part of a richer set of cortical states. These comprise the up-state (characterized by the HCS) as well as a hyperpolarized, inhibition-dominated, up-state. It has been proposed that up-down transitions originate from fluctuations in the balance of E/I inputs (Benucci et al., 2004; Jercog et al., 2017; Papoutsi et al., 2014); thus, we study which neural mechanism can amplify the E/I fluctuations and give rise to the bi-stable dynamics. For this purpose we introduce controlled fluctuations in the input firing rates, following the method described in Vogels, Sprekeler, Zenke, Clopath, and Gerstner (2011), and modulated the instantaneous rate ($\nu_t^{input}$) of an inhomogeneous Poisson process. The rate $\nu_t^{input}$ is used to draw the number of excitatory spikes on the Tripod's synapses.

The stochastic process used, detailed in the Supplementary Methods section, allows us to generate a spike train with a defined rate and inter-spike-interval coefficient of variation ($CV_{ISI}$). It depends on three parameters, the correlation timescale (fixed at $\tau$ =50 ms, (Vogels et al., 2011)), the baseline firing rate $\nu_0$, and the fluctuation size $\beta$. The latter parameter maps linearly to the $CV_{ISI}$ of the input. Some realizations of these spike trains, for fixed firing rate ($\nu_0$ =1 kHz) and $CV_{ISI}$ in the range 1 to 1.5 are shown in Fig.3.2*B*.

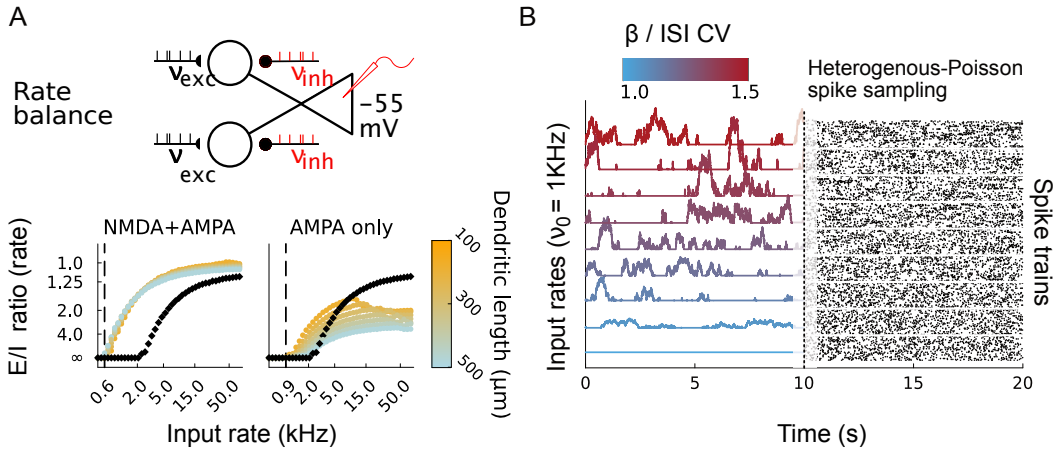



Figure 3.2: **Excitatory inhibitory balance rates and fluctuations in the input rates**

**(A)** The upper panel illustrates a scheme of the protocol used to reach the E/I balance. The rate protocol modulates the inhibitory firing rate (red) until the membrane potential stabilizes around −55 mV. The balance E/I ratio ($\bar{k}_{E/I}$) reached (lower panels) depends on the dendritic lengths and input rates. The panels report the E/I ratios reached via the rate protocol for models with and without NMDARs (left and right panels) and the soma-only model (black dots). The vertical dashed lines indicate the minimum input for which the balance condition is met by at least one dendritic length. The color codes for dendritic lengths as illustrated in the legend. In general, for increasing excitatory inputs, the E/I ratio decreases (that is, less excitation is required to meet the balance condition). Similarly, the E/I ratio increases with dendritic length (excitation is more effective for shorter dendrites). (**B**) Profile of the input spike rates and spike trains used in the simulations. The panel's left side illustrates nine input rate samples, with increasing $CV_{ISI}$ (fluctuation size, $\beta$) (blue to red). The right side shows fifty spike train samples for each rate, drawn from non-homogeneous Poisson distribution. With increasing $\beta$, the signal fluctuations increase in size (left), and the spikes are less regular (right). The signal is scaled, so the average rate remains constant ($\nu_0 = 1\,\text{kHz}$).

## 3.3 Results

### 3.3.1 Local E/I balance in the Tripod neuron

#### Local E/I balance depends on dendritic geometry and physiology

To understand the response of the Tripod neuron to barrages of pre-synaptic excitatory and inhibitory spikes, we exposed the dendrites to intense synaptic bombardment and measured somatic activity. Preliminary analyses of the neuron's response to pre-synaptic E/I firing indicate that dendrites govern the neuronal transfer function. In Appendix A, we show that the somatic firing depends jointly on the input rate and the dendritic geometry and configuration. However, for input rates beyond 5 kHz to 10 kHz, the model's firing is not in the biological regime for most of the dendritic configurations. We know that the inconsistency is due to an imbalance of excitatory and inhibitory afferents in the dendritic compartments. Indeed, the E/I balance is a fundamental characteristic of synaptic distribution on dendrites; it is locally fine-tuned and depends on the distance from the soma (Iascone et al., 2020; Liu, 2004). Thus we set the model in the biological firing regime by balancing E/I inputs, which is individuated by imposing the physiological properties of the high-conductance state in the somatic compartment (Destexhe et al., 2003). The procedure used to reach the balance state is described in the Methods.

We disentangle the effect of segregated dendritic compartments and NMDARs-based non-linear integration by comparing Tripod neurons with and without voltage-gated receptors. First, we computed the average membrane potential for balanced inputs in models with NMDARs and AMPARs, with increased granularity at low firing rates (inputs in the range 0.3 kHz to 5 kHz). The membrane potential of dendritic and somatic compartments is illustrated in Fig.3.3*A*; the panels portray symmetric models with increasing dendritic length. The upper panels show the membrane potential reached by the dendrites in the balanced condition, the lower panels show the membrane potential of the somatic compartment. The color shade indicates the input rate with which the model was stimulated. The dendritic potentials will eventually accumulate on a straight line for high input rates (blue shades), indicating that the balance condition has been reached. The dashed line indicates the minimum rate for which the balance condition is met, for at least one dendritic length.

In both synaptic conditions, the dendritic membrane potential increases linearly with the dendritic length; the increase indexes the segregation and is due

to the larger axial impedance of longer dendritic branches. For dendrites with NMDARs, the dendritic potentials split into two dense groups below and above the dashed line. This threshold corresponds to the minimum dendritic input necessary to activate the voltage gates of the NMDARs ($\sim$−50 mV). Above it, the dendrites converge to the balance condition; below it, the neuron remains hyperpolarized. In contrast, for AMPARs-only models, the membrane potential grows slowly and steadily; every logarithmic increase in rate corresponds to a linear increase in dendritic potential. The differences between these two different trends are also visible in the somatic membrane potential, portrayed in the lower panels. First, for NMDAR models, there is a narrow interval of input rates for which the balance condition is reached by some but not all models. For AMPAR models, the interval spans 1 kHz. Second, for NMDAR models, the first dendritic configurations that reach the balance are the ones with long dendrites, advantaged by the electrical segregation and the NMDAR boost. The opposite is the case for models without voltage-gated NMDA receptors. The rapid convergence of models with NMDARs implies a mechanism for compensating electrical distance in dendrites; because of the NMDARs boost, all dendritic configurations reach the balance condition for the same input rate. This is not the case for dendrites without NMDARs.

Under the E/I balance condition, the dendritic compartments transmit an amount of current to the soma that is specified by the Tripod neuron physiology (including somatic membrane permeability and its capacitance). To verify the balance being independent of the dendritic arrangement, we measured the somatic membrane potential for models with asymmetrical dendrites across all permutations of `d1,d2` = 100 µm to 500 µm. The local balance is sufficient to achieve the target membrane potential in all configurations.

**Membrane potential and firing activity in the balanced state**

When in the balance state, the neurons are in the so-called fluctuation-driven regime, or high conductance state (HCS), the somatic activity depends on the fluctuations in the frequency of excitatory and inhibitory spike trains rather than on the average amount of pre-synaptic activity. This condition is considered the operative state of cortical neurons and has implications for the cell and the network level computations (Baker et al., 2020; Destexhe et al., 2003; Higley & Contreras, 2006; Maksimov et al., 2018). We therefore investigated the consequences of the dendritic E/I balance on somatic activity and analyzed the statistics of the output spike train in this condition.

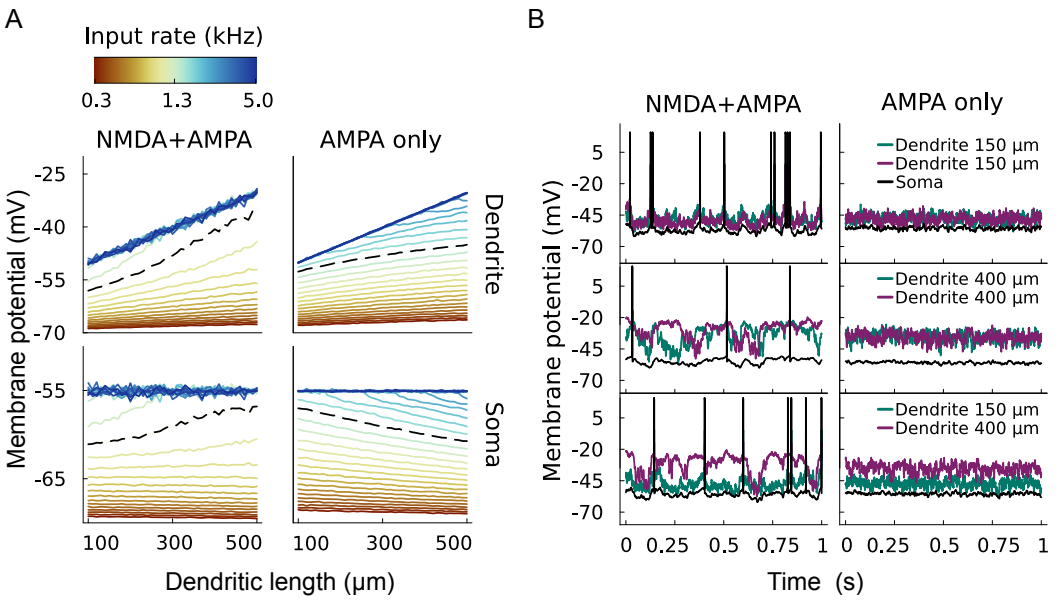



Figure 3.3: **The Tripod in the high conductance state**
**(A)** Membrane potential of dendrites (upper panels) and soma (lower panels) in models with and without NMDARs (left and right respectively) with balanced E/I inputs. The panels illustrate the average membrane potential for the symmetrical models' compartment, for all the dendritic lengths (x-axis), and input rates in the range 0.3 kHz to 5 kHz (color code). As before, the dashed lines evidence the minimal input necessary to meet the balance condition. The upper panels show that the balance dendritic potential increases with the dendritic length. Models with NMDARs have a steep convergence; for most rates, all or no dendritic configurations reach the balance condition. For AMPARs-only models, the convergence is slow, and there is a range of inputs for which the condition is met in short dendrites but is not in longer dendrites. **(B)** Membrane dynamics in dendrites and soma, for NMDARs (left) and AMPARs-only (right) models for 1 s time interval. Three dendritic configurations are explored, proximal-proximal (150-150 μm), distal-distal (400-400 μm), and proximal-distal (150-400 μm). The same frozen spike trains are used to stimulate the compartments in all six simulations. Dendrites with NMDARs exhibit larger fluctuations than AMPARs only and drive the soma to fire.

To start with, we explored the activity of Tripod neurons with and without NMDARs and in different dendritic configurations. In Fig.3.3*B*, we show the membrane potential of soma and dendrites for three dendritic configurations, distal-distal, proximal-proximal, and distal-proximal with input $\mathscr{I}^{rate}_{balance} = (5\,\text{kHz}; \bar{k}^{rate}_{E/I})$. The labels *distal* and *proximal* are borrowed from the original Tripod paper and correspond to dendritic compartment lengths of 400 μm (distal) and 150 μm (proximal) (Quaresima et al., 2022). For simplicity of comparison, we used a frozen noise input and applied the same excitatory and inhibitory spike trains to

the six simulations. From visual inspection of 1 s of simulation, it emerges that Tripod neurons with NMDARs, in the balance condition, have an active soma and produce spikes at a low firing rate. Conversely, the Tripod models with AMPARs-only synapses do not emit somatic spikes. Despite the average potential being the same in both conditions, both for the soma and the dendrites (dark blue in Fig.3.3*A*), the dendrites endowed with NMDARs present larger fluctuations than the dendrites of AMPARs-only models. In the NMDAR condition, dendritic depolarization is boosted by the activation of the NMDAR voltage-dependent non-linearity. Fluctuations are smooth and last in the order of 100 ms. Both in long and short dendritic configurations, the soma responds with firing when two depolarized fluctuations overlap. In contrast, membrane fluctuations in AMPARs-only models are small and high frequency; thus, the low-pass filter integration in the somatic compartments results in a flat membrane that fails to reach the threshold for spike onset.

### 3.3.2 Up- and down- states with NMDARs

#### NMDARs boost membrane fluctuations in the balance state and facilitate firing

For an exhaustive insight into the Tripod's activity under balance conditions, we measured the statistics of the somatic firing (mean rate and $CV_{ISI}$), the somatic membrane potential standard deviation and the cross-correlations between compartments' membrane potential, for all input rates and dendritic configurations. The scheme in Fig.3.4*A* illustrates these measures and refers to the panel where they are displayed. For visualization purposes we divided the dendritic configurations space (1600 dendritic configurations) into three subsets, grouping dendrites in the range 100 µm to 300 µm (proximal symmetrical), 300 µm to 500 µm (distal symmetrical), 300 µm to 500 µm×100 µm to 300 µm (asymmetrical, twice).

The average over the four configuration subsets is portrayed in Fig.3.4*B1* for models with and without NMDARs, and soma-only models. The first panel shows that, in the balance conditions, only the models with NMDARs present somatic activity. Crucially the output rate is non-monotonic; the firing rate reaches a peak around 3 kHz to 4 kHz, and then it decays to zero (the rebound at the end of the rate range is due to numerical instabilities, see the Supplementary Methods). As expected, the distal dendrite group has lower firing rates, while the short proximal dendrites fire more. The asymmetrical group presents intermediate

values, which correspond to the average over the four sets. The $CV_{ISI}$ remains stable around 1 for input rates in the range 1 kHz to 10 kHz, and shows a slight increase for higher rates. In this case, the distal dendrites show lower $CV_{ISI}$, indicating more regular fire. For the AMPARs-only models, there are almost no spikes in any dendritic conditions, consequently, the $CV_{ISI}$ statistics are poor. The dynamics of the somatic potential is captured in Fig.3.4*B3*, which illustrates the standard deviation of the membrane potential in the free-membrane condition (without firing threshold), being the average −55 mV by definition. Consistently with the results hitherto presented, the fluctuations are much larger (by a factor of 2-3) in the NMDARs models than in the AMPARs. However, the differences are less pronounced than in the firing rates, which suggests that the equilibrium condition amplifies small somatic fluctuations into sizeable firing rates.

The measures for the soma-only model are in significant agreement with the results reported in (Kuhn et al., 2004), confirming the procedure used to obtain the E/I balance points. Unlike the model studied by Kuhn and colleagues, the fluctuations of our soma-only model are insufficient to elicit somatic spikes. The difference is due to the smaller maximal conductance ($g_{syn}^{\text{AMPA}}$) of our synaptic model. For a proper comparison, we have to factor in the synaptic efficacy in the spike train statistics; assuming that for a unitary conductance, the input stream has $CV_{ISI}$ equal 1, then the spike train with the same rate has $CV_{ISI}$ equal to $g_{syn}^{max}$, for a synapse with such maximal conductance. The maximal synaptic conductance used in the soma-only model by Kuhn et al. (2004) is 10 times larger (Table3.3) than in our model. An exhaustive comparison of soma-only models with different parameters (Table3.4 is presented in Fig.3.1.

In the case of the NMDAR models, the dendritic configuration appears to determine somatic activity in interaction with the input rate. In panel Fig.3.4*C1* we observe that the firing response depends on the dendritic lengths and the rate of the pre-synaptic spike train. For low input rates, the response is stable around 2 Hz and homogeneous for all dendritic lengths. When the input rate increases, models with long dendrites progressively reduce their firing rate till they fail to reach the spike threshold. In contrast, models with shorter dendrites increase the firing rate up to 10 Hz reaching a maximum before 10 kHz. At approximately 20 kHz, the firing rate drops to zero for all dendritic lengths. Interestingly, the models with the highest firing rates are not those with the shortest dendrites, but those with approximately 200 μm-long configurations. Similarly, the configurations with peak response change over the input rate; for low rates, the most robust response comes from neurons with long, distal, dendrites; for larger rates,

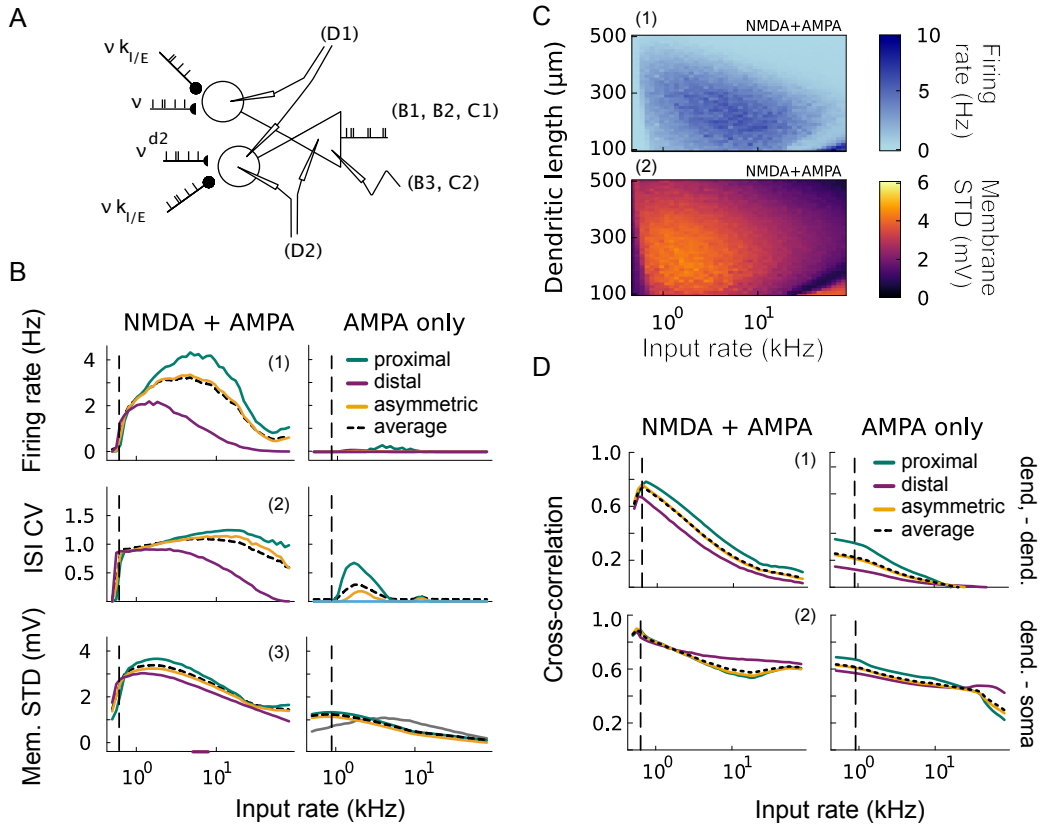



Figure 3.4: **The Tripod in the high conductance state**
(**A**) Schematics of the measures performed on the Tripod neuron; firing rate, $CV_{ISI}$, membrane standard deviation (STD) and cross-correlation between compartments. **(B)** Statistics of somatic firing and membrane potential averaged over four regions of the dendritic lengths (100 µm to 300 µm (proximal), 300 µm to 500 µm (distal), 300 µm to 500 µm × 100 µm to 300 µm), for both NMDARs, AMPARs-only and soma-only models. From top to bottom, the panels depict the firing rate, interspike interval coefficient of variation, and standard deviation of the somatic compartment membrane potential. The panels show that only NMDAR models are actively firing, consistently with large fluctuations in the somatic membrane. Both the average firing rate and the membrane fluctuations have a peak at intermediate frequencies and then decay to zero. (**C**) Firing rate and standard membrane deviation for NMDARs models, for symmetrical models with all dendritic lengths. The panels show that the length of the model's compartments has an effect on the somatic activity **(D)** Cross-correlation between the two dendritic branches and between the dendrites and the soma. The measure compares symmetric and asymmetrical models, with ranges as in *A*. In all conditions, the cross-correlation decreases with the increase of the input rate but dendrites with voltage-gated receptors are twice more correlated than in the AMPARs condition.

it shifts towards shorter configurations. This dependency on the dendritic length of the neuron's response offers a mechanism for network selectivity in neurons

with different integration timescales that dominate the cortical activity, depending on the network's activity. The bottom panel of Fig.3.4*C* illustrates the membrane's standard deviation, which assumes a similar pattern to the firing rate. However, the standard deviation is globally shifted towards lower input rates (as it is also visible in Fig.3.4*B3*) and it is more stable and homogeneous across the dendritic lengths. The input rates around 1 kHz are the most interesting; models that receive this input frequency present low to no firing rate but exhibit a sizeable somatic membrane variation. In the next section, we show that this discrepancy is because the neuron enters into a bi-stable dynamics resembling the up-down dynamics of cortical cells.

Using the same grouping technique we measured the cross-correlation between the membrane potential of the dendritic (Fig.3.4*D1*) compartments. The analysis reveals a stark difference in the internal dynamics between the NMDARs and AMPARs conditions. In both synaptic configurations, the cross-correlation between the two compartments decreases with the rate. However, in the presence of NMDA receptors, the cross-correlation between dendrites is three times larger than in the AMPARs only. The high correlation indicates that the entire neuron depolarizes and hyper-polarizes globally, compensating for the noise in the spike trains arriving at each dendrite. The dendrites communicate among them through the axial resistance and the somatic compartment, which constitutes a low-pass filter. The Tripod amplifies low-frequency fluctuations and cap high-frequency ones. With the increase of the firing rate, the fluctuations become faster, and the somatic filtering decorates the dendritic compartments. Nonetheless, the soma remains in between dendritic voltage fluctuations and keeps high cross-correlations with both dendrites (Fig.3.4*D2*). The high cross-correlation between soma and dendrites suggests the presence of meta-stable states in the membrane dynamics in which soma and dendrites drift, guided by external fluctuations. We test this hypothesis in the next section and explore the impact of increasing fluctuations at the dendritic sites. With increasing the input $CV_{ISI}$, the correlations remain high in the NMDARs models (data not shown), and the dendrites favor the onset of meta-stable up-down states.

### Bistable somatic dynamics: Up- and down-states require non-linear dendritic integration

When the E/I ratio is balanced, the Tripod neuron has the property of the HCS and becomes sensitive to fluctuations in the barrages of pre-synaptic spikes. However, *in-vivo* network activity usually has more variability than the homoge-

nous Poisson input spike trains studied in previous sections (Maksimov et al., 2018; Paré, Shink, Gaudreau, Destexhe, & Lang, 1998). The network fluctuations to which a cortical cell is exposed have been connected to the onset of somatic up and down states (Jercog et al., 2017; Papoutsi et al., 2014; Wilson & Kawaguchi, 1996). We test these hypotheses on the Tripod neuron and, in agreement with previous work, show that increasing the input $CV_{ISI}$ leads to a bi-stable somatic membrane in Tripod and soma-only neurons. However, bi-stability emerges for smaller fluctuation sizes in Tripod neurons with NMDARs than in the two control conditions (AMPARs-only and soma-only), indicating a significant role of dendritic non-linearity in picking up over the afferent network's state. In order to test the consequences of fluctuations in the input spike trains, we devised a stochastic process to modulate the amplitude of the input rate without affecting its average. Thus, we produce spike trains that correspond to locally inhomogeneous Poisson distributions, with the same average rate; the stochastic process is discussed in detail in the Methods. The fluctuation sizes studied ($\beta$ parameter) correspond to an $CV_{ISI}$ of the input spike train in the range 1 to 2, consistent with the upper range of activity measured *in-vivo* (Holt, Softky, Koch, & Douglas, 1996).

First, we visually inspected the membrane dynamics for Tripod neurons with and without NMDARs. Fig.3.5 shows the membrane potential of one the dendrites (distal, 400 µm) and the soma, and the all-point-histogram of the somatic potential, for a balanced input spike train with increasing fluctuations ($\beta$ =1 to 1.15). In the example, the input range was set to 1.6 kHz, which corresponds to the input range with the largest standard deviation for distal dendrites. The results evidence three clear aspects of the neuron's dynamics. First, the dendritic membrane of models with NMDARs (Fig.3.5*A*) have excursions of more than 30 mV, whereas dendrites without voltage-gated receptors are less sensitive to input fluctuations (Fig.3.5*B*). Second, somatic firing follows the oscillations of the dendritic membrane; the neuron exhibits a burst of spikes when the membrane is most depolarized. Third, with increasing size of input fluctuations (increasing $\beta$), the somatic compartment of models with NMDARs enters into a bi-stable dynamic, revealed by the APH. Contrarily, without NMDARs, the soma has one single peak around −60 mV.

We used a Gaussian-kernel-estimate (GKE) approach to quantify the somatic bi-stability, it allowed us to count the peaks and estimate the strength of the APH bi-stability. The GKE is parameterized by the size of the window used for the Gaussian kernel (details in the Supplementary Methods). With increasing

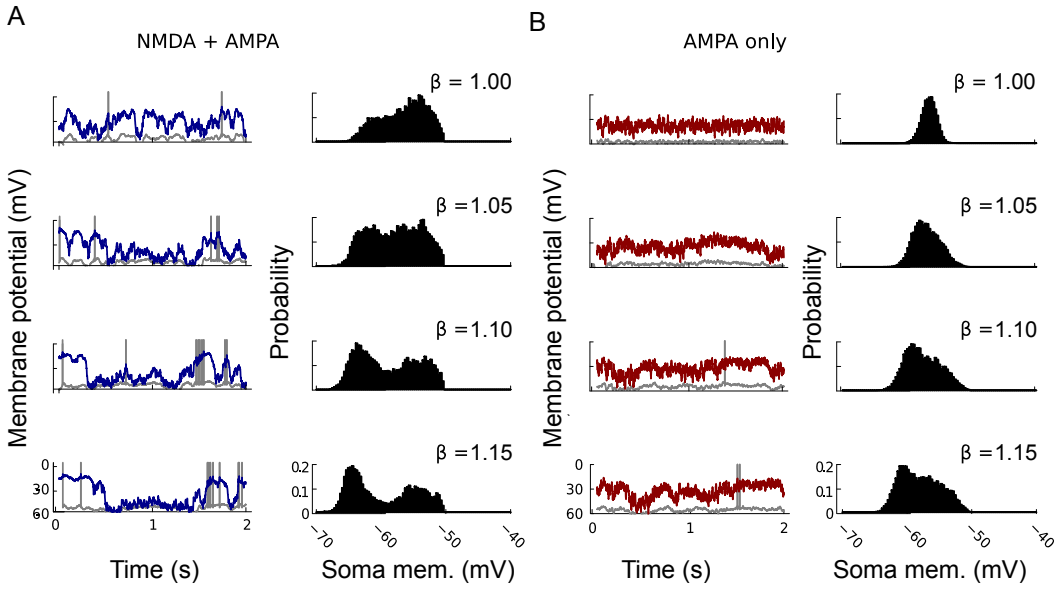



Figure 3.5: **Input fluctuations lead to bi-stable membrane potential in NMDARs models**
**(A)** Membrane potential and all-point-histogram (APH) of a distal-distal model endowed with NMDARs and exposed to intense synaptic bombardment with an average input frequency of 1.6 kHz. The model has balanced E/I inputs. Both excitatory and inhibitory input spike trains have increasing rate fluctuations with $\beta$ ($CV_{ISI}$) in the range 1 ( to 1.25 (panels below, from top to bottom). The dendritic membrane responds with large fluctuations visible in the somatic activity (left) and the increasing bimodality in the APH distribution (right). **(B)** Equivalent configuration and inputs than in *(A)* but with an AMPARs-only model; without NMDARs receptor, the somatic model is less sensitive to fluctuations in the inputs. The external inputs are frozen and the same in all panels to facilitate comparisons.

window size, the number of peaks will monotonically decrease. An example of the GKE for increasing window size is portrayed in Fig.3.6*A*. Our bi-stability measure estimates the maximal window size that retains two peaks in the GKE of the APH. There is one caveat for our measure, which corresponds to APH with more than two peaks; for example, for high frequencies (beyond 30 kHz), the APH will sometimes show three peaks (as in the APH of the distal-distal model in Fig.3.6*A*). The appearance of three peaks at an increasing rate is due to the decorrelation between the two dendrites, which will, independently, be in a depolarized (up) or hyperpolarized (down) state. When permuting over the dendritic states, there are four cases (three in symmetrical dendrites), up-up, down-down, up-down, and down-up. Because the dendritic decorrelation is higher for longer dendrites and higher rates, the three peaks structure will occur

more often in those cases. To avoid counting the triple peaks in the bimodal index, we count the APH with more than two as non-bimodal.

We investigated the emergence of bimodality in three dendritic models representing the possible dendritic configurations: distal-distal, proximal-proximal, and distal-proximal. As a control, we also measured the bimodality index in models without NMDARs and in the soma-only model. We tested the models for all input rates in the range 0.5 kHz to 80 kHz and for $\beta$ ($CV_{ISI}$) in the range 1 to 2. The results of the measures are reported in Fig.3.6*B*. Tripod neurons with NMDARs have a high bimodal index when the input's $CV_{ISI}$ is larger than 1.2 and for inputs within the range 0.5 kHz to 5 kHz. The bimodal membrane distribution is robust across symmetrical and asymmetrical dendritic configurations. The bimodality index slowly fades for higher inputs because the pre-synaptic bombardment forces the neuron dendrites into decorrelated states and flattens the APH. In these conditions, bimodality is recovered only for large fluctuations in the inputs, $CV_{ISI} > \approx 1.5$. Tripod neurons with AMPARs-only synapses (right panels in Fig.3.6*B*) require massive fluctuations to enter the bimodal dynamics, and the bimodality index remains low. The soma-only model, surprisingly, has intermediate properties between the AMPARs and the NMDARs models. Its bimodality index is high in a short interval of the input rate (4 kHz to 8 kHz) and for fluctuations beyond $\beta = 1.4$, in both the NMDARs and AMPARs condition. To further explain the relation between bimodality and fluctuations in the inputs, ($\beta$) we examined the distal-proximal model over inputs generated varying both the fluctuation timescale ($\tau_r$) and size ($\beta$), as expressed in Eq.3.4. The results are illustrated in Fig.3.2. The best predictor of the bimodality index was neither the $\beta$ nor the $\tau_r$; rather, bimodality was higher when the input $CV_{ISI}$ was higher.

The measure of the bimodality index for Tripod neurons with NMDARs ensures that the model will enter the up-down dynamics for $CV_{ISI} > 1.2$ This result is compatible with the range of $CV_{ISI}$ observed in the cortex (Maksimov et al., 2018; Ostojic, 2011; Shinomoto, Sakai, & Funahashi, 1999). On the other hand, the measure in Fig.3.6*B* does not clarify the role of the excitation/inhibition balance in eliciting the up-down states. To test this, we set the $\beta = 1.2$ and measured the bimodality while varying the E/I ratio ($k_{E/I}$), the results are illustrated in Fig.3.6*C*. We varied the $k_{E/I}$within 5 % to 200 % of the $\bar{k}_{E/I}$such to explore nearly all possible input configurations for the excitatory and inhibitory rates; the measure was carried on the four models previously indicated, two symmetrical (distal and proximal), one asymmetrical, and the soma only. Consistently with the methods used to set the balance, we maintained the excitatory rate fixed and varied the in-

hibitory rate, $\mathscr{I}^{rate}_{balance} = (\nu_{exc} = 0.5\,\text{kHz to } 80\,\text{kHz}; \nu_{inh} = \alpha\nu_{exc} \cdot \bar{k}^{rate}_{E/I}; \beta = 1.2)$, with $\alpha = 5\,\%$ to $200\,\%$. The results show that the balance is not a necessary outcome for input rates below 1 kHz but becomes critical for larger ones. The UDSs occur only in a narrow band of possible E/I ratios and are centered around the $\bar{k}_{E/I}(alpha = 100\%)$. We also computed the average membrane potential of the soma, expressed in Fig.3.6*C* by the red-blue color code. For the Tripod neuron, the soma will persist in the up (red) or down (blue) state when the balance is increased (+) or diminished (-). For soma-only models, the fluctuation size chosen as input was not sufficient to elicit bimodal states; however, the measure indicates that when the excitatory input is sufficiently strong (2 kHz), the model with $\bar{k}_{E/I}$will stay stable around $-55$ mV, as expected from the balance condition imposed. Interestingly, variations in the $\bar{k}_{E/I}$have less effect on the average somatic potential in the soma-only rather than in the Tripod neuron.

Finally, we report the firing rate in all the eight models measured before for the entire input range and $\beta = 1$ to 2 The firing rate of Tripod neurons with NMDARs, AMPARs only, and the soma-only models are portrayed in Fig.3.6*D*. In the absence of fluctuations (lower extreme of the plot, $CV_{ISI}$ equal 0), the same output rates curves illustrated in Fig.3.4*Ciii* are recovered; the Tripod neurons fire at approximately 5 Hz, with a peak a low frequency and a slow decay for higher input rates. In the distal-distal model, the fluctuation size does not affect this trend, the firing rate increases at low input rates, reaches a peak, and decreases. Conversely, for models with NMDARs receptor and at least one proximal dendrite (distal-proximal and proximal-proximal configurations), the firing rate increases abruptly for $\beta > 1.1$ The response differs structurally from the firing rate measured in the absence of input fluctuations (Fig.3.4*Ciii*). First, it does not depend on the input rate; the response is stable at 15 Hz for the distal-proximal model and around 8 Hz for the proximal-proximal model. Second, in the presence of fluctuations, the firing rate does depend on the dendritic configuration, and the asymmetrical (distal-proximal) model has a firing rate about 90 % larger than in the symmetrical case (proximal-proximal). This difference between the symmetrical and asymmetrical model was absent at $CV_{ISI}$ equal 1, indicating that asymmetries increase the model's sensitivity to fluctuations. In contrast, the soma-only and the AMPARs-only models have maintained the classical input rate dependence; they show a peak at intermediate frequencies. The AMPARs models have a second peak around 10 kHz, which may be due to the interactions with the inversion of $k_{E/I}$reported in Fig.3.2*B*. In both cases, the firing rate is larger than zero only when fluctuations are beyond $\beta = 1.2$.

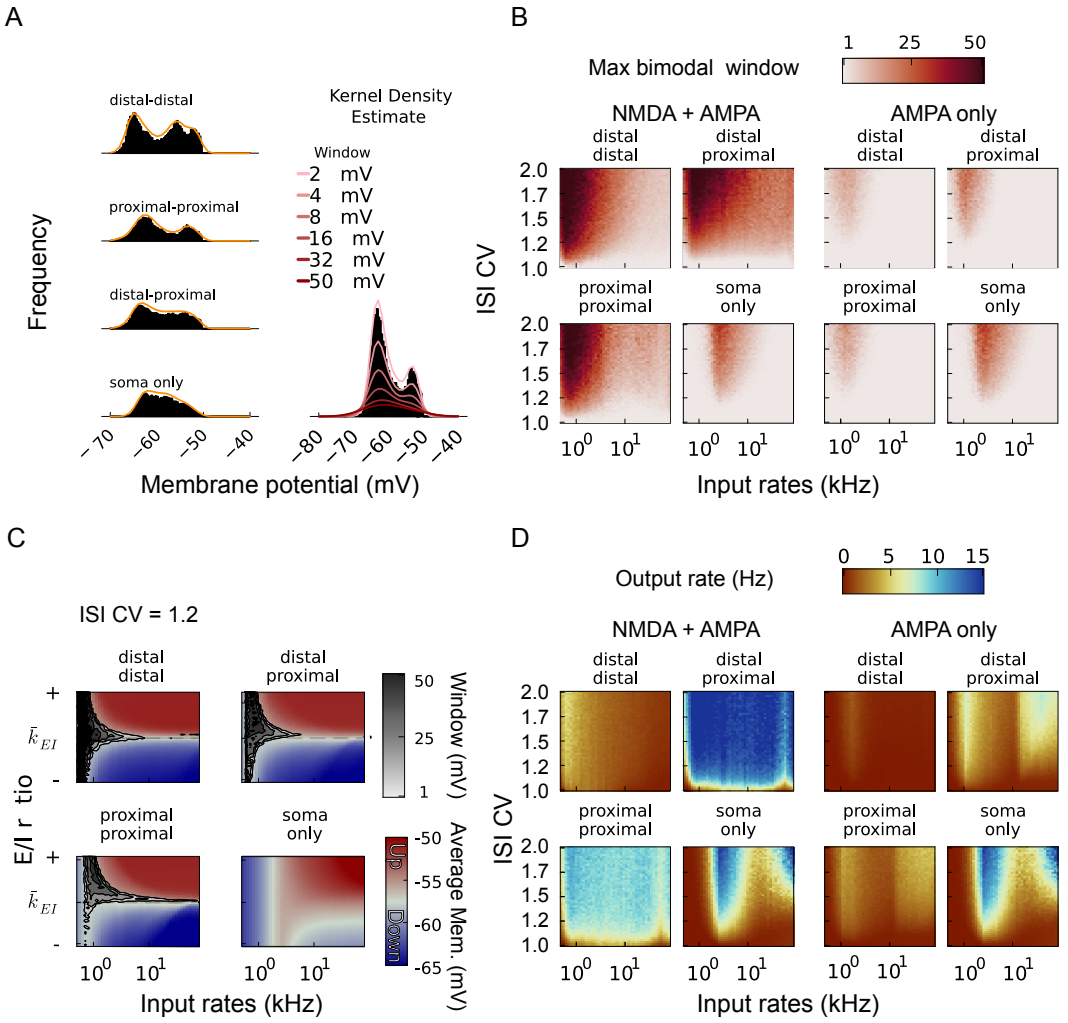



Figure 3.6: **UDS depends on input fluctuations but requires E/I balance and non-linear dendritic integration to express**
**(A)** (Left) All-point histogram (APH) for three NMDARs Tripod neurons and soma-only model. The orange line illustrates the Gaussian Kernel Density Estimate (GKDE) for a window of 2 mV. The input rate is 1.6 kHz. (Right) APH of a distal-proximal model with the same input rate, the color-shaded lines show the GKDE for increasing window lengths (panel's legend). For large windows, the bimodality (two peaks in the distribution) of GKDE reduces monotonically. The minim window with bimodal distribution indexes the bimodality of the APH. **(B)** Bimodality index for symmetrical and asymmetrical Tripod neurons, with and without NMDARs, and soma-only models. The bimodality is tested against increasing fluctuation size ($CV_{ISI}$) in the input (y-axis, 1 to 2)) and for all the input rates (x-axis 0.5 kHz to 80 kHz). Only Tripod neurons with NMDARs show bimodality in the APH. *Figure caption continues on the next page.*

Figure 3.6: **(C)** Bimodality index and average membrane potential against variations in the $k_{E/I}$, for symmetrical and asymmetrical models, and $\beta$ ($CV_{ISI}$) equal 1.2 kHz. For increasing input frequency, the range of $k_{E/I}$for which the APH is bimodal shrinks around the balance $\bar{k}_{E/I}$. The membrane potential indicates that exceeding the $\bar{k}_{E/I}$turns in permanent hyperpolarized (down, blue) or depolarized (up, red) states. **(D)** Firing rate for the eight models analyzed. The presence of fluctuations in the NMDARs Tripod neurons causes higher firing rates. The asymmetrical models benefit from fluctuations more than the symmetrical models.

We have shown that including dendrites and NDMA receptors in single neuron models significantly impacts the somatic dynamics, both in the absence of input fluctuations and when exposed to inputs with varying rates. Critically, the Tripod neuron shows strong bi-stability when the fluctuations are in the regime of those measured *in-vivo*, ($CV_{ISI}$ approximately 1.2). In the next section, we analyze the temporal traits of the emerging UDS.

### 3.3.3 Statistics of up-down states in the Tripod neuron

Following the analysis performed in (Benucci et al., 2004) we measure the statistics of state duration, firing rate, and membrane potential of the UDS in the Tripod neuron. For simplicity, we focus on Tripod neurons with distal-proximal compartments, input rate of 1.6 kHz, and fluctuations corresponding to $CV_{ISI} = 1.2$ We chose this setup because the geometry reflects the heterogeneity of real dendritic harbors, and the rate and $CV_{ISI}$ are consistent with common cortical conditions. The results, however, generalize to other models and fluctuation sizes which are not illustrated here.

First, we inspect visually the somatic membrane fluctuations in the Tripod neuron with and without NMDARs, and in the soma-only model (Fig.3.7*A*). The three models were stimulated with the same input spike train to facilitate the comparison. In agreement with the analysis of the previous sections, the Tripod with NMDARs shows two distinctive meta-stable states in the membrane potential; the same oscillations in the potential are visible in the other two models, but the variations are less sharp and pronounced. To quantify these states we used the method proposed in (Anderson, Lampl, Gillespie, & Ferster, 2000) and defined two thresholds for the up-down states. When the somatic membrane crosses the UP threshold we count a somatic upstate, it lasts until the potential crosses the DOWN threshold; the thresholds are set, respectively, to three fourth (0.75) and one-fourth (0.25) of the range spanned by the membrane potential

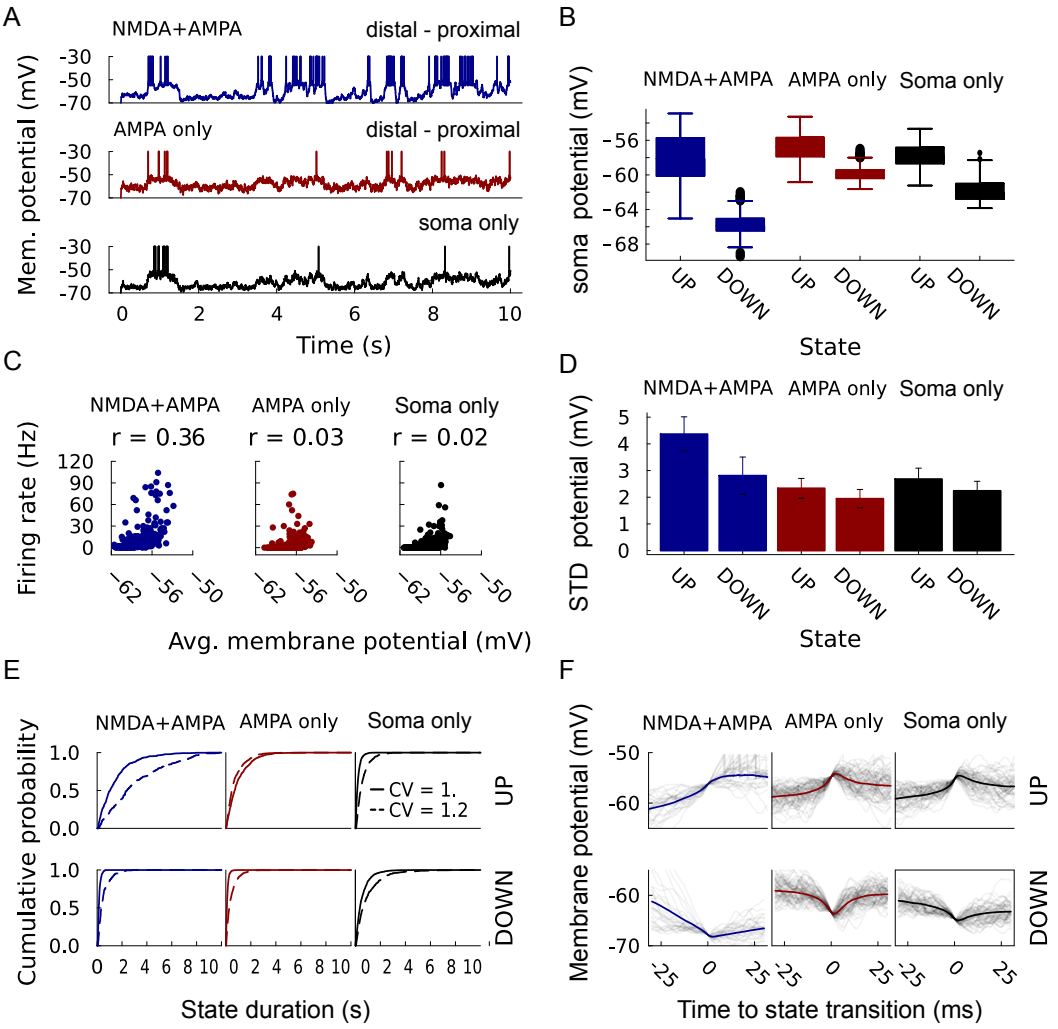



Figure 3.7: **Statistics of up-down states in the Tripod neuron**
**(A)** Membrane potential dynamics under fluctuating synaptic input. The presynaptic spike train is the same for the three models presented, Tripod with NMDARs (blue), without NMDARs (red) and the soma-only model (black). The model with NMDA fires more and has more pronounced UDS. To quantify the membrane dynamics, the potential has been clustered into an up and down region using two thresholds (Anderson, Lampl, Gillespie, & Ferster, 2000). (**B**) Membrane potential of the up and down states for the three models. The down-state is about −65 mV for the Tripod with NMDARs, but more depolarized for the two other models. Thus, the difference underlying the bimodality is more pronounced in models with NMDARs. *Figure caption continues on the next page.*

Figure 3.7: (**C**) For each up-state, we compared the average membrane potential and the instantaneous firing rate; models with NMDARs have a high correlation between the potential and the firing rate ($\rho$ =0.36), which is not the case for the other models. (**D**) Standard deviation of the membrane potentials clustered in the UDS. (**E**) Cumulative distribution of the state duration for two conditions of the input $CV_{ISI}$ (1, solid line; 1.2 dashed line); the two conditions correspond to awake and anesthetized animal and the Tripod NMDAR model matches qualitatively the physiological data. (**F**) Average dynamics of the membrane potential around the UDS transition. The Tripod NMDARs transition lasts on the order of 25 ms and is pronounced and stable immediately after. Conversely, the UDS transitions are faster and less pronounced in the tripod without NMDARs and in the soma-only model; the membrane potential moves in an unimodal distribution and the UDS are its tails.

during the simulation. Once the UDS intervals are defined, we compute the average membrane potential in the state and compare it across the three models (Fig.3.7*B*). All models have significantly different membrane potential averages in the UDS, ensuring that the threshold technique used to individuate them is effective. Nonetheless, the difference between up and down states is more pronounced in the Tripod with NMDARs than in the two control conditions and only presence of voltage-gated receptors the down-state is strongly hyper-polarized (−65 mV). Furthermore, the up states of NMDAR models share several characteristics with those measured *in vivo* by Anderson, Lampl, Reichova, Carandini, and Ferster (2000) or reproduced with detailed neuron models (Benucci et al., 2004). In particular, we report a significant correlation between the instantaneous firing rate and the membrane potential in the up-state (Fig.3.7*C*); and the larger variability in the membrane potential in the up-state than in the down-state (Fig.3.7*D*). To test the functional properties of the up-down dynamics emerging in the Tripod, we compared the cumulative probability of the states' duration with and without correlations in the input streams ($CV_{ISI} = 1$ and 1.2). The analysis aims to reproduce the variations measured *in vivo* when the animals were exposed to visual stimuli or rather deprived of it (Anderson, Lampl, Gillespie, & Ferster, 2000). The effect of the manipulations, shown in Fig.3.7*E* is in large agreement with the experimental evidence; an increase of the correlations in the input stream causes the probability distribution to shift forward (left-top panel) but has little effect on the distribution of down states (left-bottom panel). While the Tripod model with NMDARs qualitatively reproduces *in vivo* measures, it is not the case for Tripod models with AMPARs only or soma-only neurons (Fig.3.7*C, D, E*). The up-down dynamics emerging in the Tripod model show similarities with the up-down measured in the cortex also from a temporal per-

spective. By averaging the somatic membrane potential we measured the time course of the transition between the UDS. It results that the onset of up states is faster and has a sigmoidal shape, in contrast, the transition to the down-state is slower and has a linear trend (Fig.3.7*F*), these differences are consistent with the temporal characteristics of the transition between UDS reported in Wilson and Kawaguchi (1996).

## 3.4 Discussion

We investigated the response of a minimal dendritic model (the Tripod neuron, (Quaresima et al., 2022)) to high-rate E/I balanced inputs and showed the emergence of distinctive cortical traits, such as the high conductance state and the up-down states. The Tripod neuron has two segregated passive compartments that accommodate NMDA receptors. From systematic comparisons with simpler models (no NMDARs and no dendrites), we conclude that NMDARs and segregated compartments are sufficient for detecting fluctuations in the synaptic inputs, thus reproducing functional aspects of cortical cells. The model indicates that the UDS are the natural consequence of fluctuations of the balanced E/I activity on the dendritic arborization. The UDS of the Tripod model reflects the properties of UDS in cortical cells, such as the correlation between the membrane potential and the firing rate, the bi-modal distribution of the sub-threshold states, and the dynamics of the up-down transition.

Concerning the contribution of the dendrites to the HCS, our study indicates that fine-tuned dendritic balance shapes the membrane potential of the neuron and drives somatic depolarization. For unbalanced (E/I) inputs, the neuron increases its firing rate exponentially, consistently with the traditional results on leaky-integrate-and-fire (LIF) frequency-current curves (Nordlie, Tetzlaff, & Einevoll, 2010), although not with physiological recordings (Anderson, Lampl, Gillespie, & Ferster, 2000). Computational and experimental results converge on the fact that cortical neurons are exposed to balanced E/I pre-synaptic firing (Destexhe, Rudolph, Fellous, & Sejnowski, 2001; Higley & Contreras, 2006; Kuhn et al., 2004; Shu et al., 2003) and cortical neurons leverage different mechanisms to reach the E/I balanced state (Turrigiano, 2011; Vogels et al., 2013). We investigated two E/I balance mechanisms and showed that, even if leveraging different degrees of freedom (i.e., membrane potential and firing rate), they reach similar excitatory/inhibitory equilibrium points. Crucially, we found that the organization of excitation and inhibition depends on the dendritic geometry,

with longer dendrites requiring more excitation than shorter dendrites to achieve somatic balance; our results are in considerable agreement with whole-neuron recordings on the distribution of excitatory and inhibitory dendritic spines (Iascone et al., 2020) and respect the dendritic democracy hypothesis, with stronger synapses in distal regions (Häusser, 2001; Magee & Cook, 2000).

The E/I dendritic balance yields a somatic activity profile similar to the one reached by balancing somatic conductances (Kuhn et al., 2004), suggesting that theoretical insights based on the properties of balanced networks should be extended to networks with dendrites (Baker et al., 2020; Denève & Machens, 2016; Ebsch & Rosenbaum, 2018; Renart, Moreno-Bote, Wang, & Parga, 2007). However, two differences between somatic and dendritic balance must be considered for future computational work. First, passive segregated dendrites have a low-pass filter effect on the signal; consequently, fast fluctuations are attenuated in models with AMPARs. The attenuation applies to models with linear or sub-linear integration in the segregated compartment and is consistent with previous work that explored coincident detection in the dendrites (Dasika et al., 2007). Second, dendritic non-linearity, such as NMDAR, recovers the sensitivity to fast fluctuations and introduces an interplay between the pre-synaptic rate and synaptic efficacy and the dendritic location of the synapses (i.e., the dendritic length in our model).

Our study suggests the untested hypothesis that input correlations at different timescales are best expressed by synaptic clusters located at different dendritic lengths. In particular, short dendrites should pick fast fluctuations, and long dendrites should pick fluctuations on longer timescales. The geometry-rate interaction is the natural consequence of the timescale introduced by the segregated dendritic compartment (Quaresima et al., 2022). Such selectivity of the dendritic branch is in line with recent physiological studies that show complex arrangements of the pre-synaptic axons on the dendrites of sub-cortical and cortical cells (Callan, Heß, Felmy, & Leibold, 2021; Lafourcade et al., 2022). An additional macroscopic consequence of dendritic non-linearity is the high correlation between dendritic compartments. Because input correlations are null in our setup, it suggests that membrane potential information propagates more efficiently in non-linear dendrites and accounts for the correlation between local and global events observed in single branches of whole-cell imaging *in-vivo* (Otor et al., 2022; Palmer et al., 2014; Russell et al., 2022). The linear drop of the dendrite-dendrite correlation suggests that the cell can leverage multiple regimes of integration depending on the input rate.

The Tripod neuron model explains how cellular bi-stability can emerge by the synaptic bombardment that produces the HCS. By varying the second moment of the input spikes (CV ISI), the model switches from the depolarized state to the stereotypical up-down dynamics. We investigated the constraints on the local activity and showed that models with NMDARs naturally enter the UDS for fluctuations within physiological boundaries. Our results are in agreement with the early hypothesis that the onset of NMDA spikes and the consequent dendritic up-state drive the UDS, as proposed in Antic et al. (2010); Milojkovic et al. (2005). Furthermore, there is a striking similarity between the profile of the dendritic up-states in the Tripod model and the membrane potential of dendritic branches measured *in vivo* in a subclass of L5 cells (Dembrow & Spain, 2022). In the experiment, however, the inhibitory synapses were blocked, which limits the range of comparison to Dembrow and Spain (2022), to the lower range of the Tripod input rates. On the other hand, the dependence of the UDS on the NMDARs' non-linearity is not in line with other experimental works that leveraged intracellular MK-801 to block NMDARs in single cells (Chen, Rochefort, Sakmann, & Konnerth, 2013; Palmer et al., 2014; Smith et al., 2013). These studies have shown the emergence of UDS independently of the voltage-dependent receptor activity. Such evidence posits a different hypothesis on the relation between NMDA spikes and somatic UDS (Larkum, 2022), the sub-threshold UDS are governed by the barrage of synaptic inputs activating the AMPA receptors; whereas the NMDA spikes are additional sparks, signaling that spatially-clustered synapses were jointly activated.

While our model elicits somatic firing only in the presence of NMDARs - matching the Larkum (2022) hypothesis, it cannot retrieve the bimodal membrane potential for models without NMDARs when the fluctuation size is small; for AMPARs models, the bimodal state requires $CV_{ISI} > 1.5$. We advance three, possibly concurrent, interpretations for the differences between our results and the latter experimental evidence. First, it may be that the blockage of NMDARs does not prevent other non-linear mechanisms, as evidenced in the entorhinal cortex by Digby, Bravo, Paulsen, and Magloire (2017). In this case, our model has to be interpreted as expressing the general function of a dendritic non-linearity. Second, the explanation may rely on the role of inhibition in driving UDS. In our experiments we remain close to the E/I balanced conditions, both in the up and down states, however, it is possible that the up-state corresponds to the balanced state while the down-state is dominated by strong inhibition (inhibition-stabilized, Jercog et al. (2017)), or even the absence of excitatory sensory stimuli (Dem-

brow & Spain, 2022). In this second case, the model with AMPARs-only should receive inputs with leveled-up inhibition, which may result in UDS. Third, we ought to consider the possibility that the UDS emerges from the complex dendritic morphology of real cells. The early computational findings by Benucci et al. (2004) suggest that the temporal filtering of many dendritic branches with heterogeneous timescales supports the soma's up-state with broad and continuous depolarizing currents. Our models do not account for the sub-threshold summation of currents from multiple dendritic branches; rather, we show that the extremely reduced dendritic morphologies are sufficient if the dendritic compartments meet the local balance conditions and express non-linear voltage-gated receptors.

Previous hypotheses on the rise and maintenance of up-states suggested that cells use inward-rectifying potassium channels (Cunningham et al., 2006; Sanders et al., 2013), or similar ionic mechanisms (Loewenstein et al., 2005) to maintain the membrane depolarization following a strong current pulse. However, it is unclear how these fine-tuned systems may maintain stability in the face of noisy network dynamics. In contrast, we provided evidence that reduced dendritic models can account for the UDS robustly via dendritic non-linearities. The present results have broad implications for computational studies because they indicate that dendrites relax the conditions for neuronal bi-stability in the presence of synaptic noise. The implication of dendrites on the UDS was already investigated in previous computational studies by Papoutsi et al. (2014) and Benucci et al. (2004). Our conclusions are in agreement with these previous findings, although our model meets the E/I dendritic balance explicitly, resulting in firing statistics and membrane dynamics remarkably similar to cortical cells. In addition, the Tripod can be efficiently implemented in medium to large networks. Thus, it allows future studies to investigate the interaction between local and network-level bi-stability.

## 3.5 Supplementary Methods

### Input spike trains and notation

Excitatory and inhibitory input spikes are drawn randomly from Poisson distributions. The excitatory input rate is the same for both dendrites and varies within the range 0.5 kHz to 80 kHz. Conversely, the synaptic efficacy for the excitatory inputs is fixed and corresponds to the peak amplitude of the glutamatergic receptors $\bar{g}_{syn}^{\text{GluRs}}$ (Table3.3). Similarly, for the inhibitory inputs, the inhibitory synaptic efficacy is fixed by the receptors' peak amplitude and their firing rate varies. A complete description of the set of parameters that characterize the model inputs can be expressed by the tuple $\mathscr{I} = (\nu_{exc}, \nu_{inh}, \beta, \bar{g}_{syn}^{\text{GluRs}}, \bar{g}_{syn}^{\text{GABA}}) = (\nu_{exc}, \nu_{inh}, \beta)$, where the last term accounts only for independent variables.

We modulate the rate of the inhibitory inputs while controlling for the balance condition. In principle, there can be multiple values of the inhibitory rate that satisfy the balance condition. However, our results show that for a given excitatory rate there is only one inhibitory rate that achieves the E/I balance condition. Therefore, we express the inhibitory rate as a fraction $k_{E/I}$ of the excitatory rate. Therefore, in the absence of fluctuations ($\text{CV}_{\text{ISI}}$ equal to 1), the input to the models are fully charecterized by

$$\mathscr{I}_{balance} = (\nu_{exc}; \quad \nu_{inh} = k_{E/I}^{-1} \cdot \nu_{exc}; \quad \beta = 0) \tag{3.3}$$

Which can be compactly expressed with $\mathscr{I}_{balance} = (\nu_{exc}; k_{E/I})$. Note that we will be referring to the E/I ratio as the factor $k_{E/I}$, whose inverse modulates the inhibitory input rate and synaptic strength in Eq.3.3. We chose to use the inverse of $k_{E/I}$ instead of $k_{I/E} = k_{E/I}^{-1}$ for consistency with previous studies (Kuhn et al., 2004).

### Fluctuations in the pre-synaptic firing rate

For excitatory spike trains, the instantaneous input rate is computed by multiplying the input rate ($\nu_{exc}^0$) with a non-negative, time-dependent signal $\bar{r}_t$ which is defined as follows:

$$\begin{aligned} \xi &:= \text{rand}[-0.5, 0.5](uniform) \\ z_{t+1} &:= \xi - (\xi - z_t)e^{-\frac{dt}{\tau}} \\ r_t &:= 1 + \tilde{\beta} \max(0, z_t) \end{aligned} \tag{3.4}$$

The variable $z_t$ is a discrete Ornstein-Uhlenbeck process and $\tau$ is its autocorrelation time; it is fixed at 50 ms as in Vogels et al. (2011), however there are no qualitative differences for values within a reasonable, physiological range (10 ms to 100 ms). Finally, $\tilde{\beta}$ is a free parameter and determines the fluctuation size of the rectified process $r_t$. at 50 ms Vogels et al. (2011), We fix the signal average to 1 by normalizing it with its average over the simulation interval $\bar{r}_t = \frac{r_t}{\langle r \rangle_T}$. Thus, the instantaneous input rate is given by

$$\nu^{exc}_{(}t) = \nu_0 \frac{r_t}{\langle r \rangle_T} = \nu_0 \bar{r}_t \tag{3.5}$$

The average input rate $\langle \nu^{exc}(t) \rangle$ is constant for each input spike train and equal to $\nu_0$.

The parameter $\tilde{\beta}$ can be transformed with a simple lineartransformation; we define $\beta \doteq 10^{-3}\tilde{\beta} + 1$, or, equivalently:

$$\tilde{\beta} = 10^3 (\beta - 1) \tag{3.6}$$

With the substitution in Eq.3.6 the $CV_{ISI}$ of the spike trains computed from the heterogeneous Poisson distribution (Fig.3.2), matches the values of the parameter $\beta$. Thus, the $CV_{ISI}$ of the spike train generated has the same range of the $\beta$ parameter, namely, 1 to 2 and we can use the $CV_{ISI}$ as a proxy for the fluctuation size of $\nu_t^{input}$. Further comparisons between the $CV_{ISI}$ and the parameter $\beta$ are illustrated in Fig.3.2. The panels also show that the membrane bimodality depends on the input statistics only through the $CV_{ISI}$ and not the parameters $\tau$ or $\beta$.

### Quantifying up-down states

***All points histogram.*** The all-time-points histogram (APH) is a measure of the average time the membrane potential falls within a certain voltage range. It was first used in Wilson and Kawaguchi (1996) to characterize *in vivo* intracellular recordings from cortical and striatal spiny neurons. In the presence of up-down states, the APH shows a bimodal distribution of somatic membrane potentials. Values most frequently assumed by the membrane are either near the firing threshold (up state) or close to the resting potential (down state). We compute the APH for simulations of 10 s if not otherwise specified. The APH has a resolution in voltage of 1 mV.

***Bimodality of APH.*** We measured the presence of a bimodal distribution through Gaussian kernel density estimate (GKDE). Following the method in Silverman (1981), we computed the GKDE for the APH as

$$GKDE\,[\mathrm{APH}, h](x) = \sum_{i=-90\,\mathrm{mV}}^{-40\,\mathrm{mV}} \frac{\mathrm{APH}_x \left(\exp -((x-i)/h)^2\right)}{h} \tag{3.7}$$

where $\mathrm{APH}_x$ is the magnitude of the histogram bin corresponding to the voltage $x$. The function $GKDE\,[\mathrm{APH}, h](x)$ is a smooth curve that depends on the particular realization of the APH and the window size $h$ used for the Gaussian convolution. It was demonstrated that, for a function obtained from Eq.3.7, the number of local maxima of the curve $GKDE(x)$ is a monotonically decreasing function of the window size ($h$) (Silverman, 1981). Thus, to estimate the bimodality of an APH, we compute the $GKDE$ for windows in the range 1 mV to 50 mV and count the number of maxima. Because the number of maxima is a monotonically decreasing function, we can use the minimal window size ($\bar{h}$) for which there is a unique maximum in the GKDE (unimodal distribution) as an index of the APH bimodality. The larger the minimal window $\bar{h}$, the stronger the APH bimodality, indicating the presence of up-down states. Notice that because of intrinsic randomness in the APH, a local maximum is considered such only if it is larger than 30% of the global maximum.

### Numerical simulation

The differential equations of the model were integrated using the forward Euler method (Heun's method (Ascher & Petzold, 1998)) with explicit integration and a step-size of 0.1 ms. The implementation followed the original Tripod study. In the previous study, the low-rate pre-synaptic spikes did not result in numerical stiffness. However, at high input rates (beyond 50 kHz), the synaptic conductance becomes large, and even small changes in the membrane potential lead to large incoming synaptic currents. In this condition, the membrane potential varies largely in one single time step (tens of milli-volt in 0.1 ms), while the coupled equations of the synaptic conductance changes over the time scale of tens or hundreds ms, causing a stiffness problem. To solve this issue, we limited the derivative of the membrane potential by capping to 1 nA the maximal current flowing through each compartment ionic channel. Simulations were performed in Julia using the same code for Quaresima et al. (2022), except for the capped dendritic currents. The code can be obtained on GitHub at `https://github.com/aquaresima/tripod_neuron`.

Table 3.1: Parameters for the axosomatic compartment of the Tripod neuron (Quaresima et al., 2022) which is modeled as an adaptive exponential neuron (Brette & Gerstner, 2005).

| Symbol | Description | Value | Unit |
|---|---|---|---|
| $g_L$ | Membrane leak conductance | 40 | nS |
| $C_m$ | Membrane capacitance | 281 | pF |
| $V_r$ | Resting membrane potential | -70.6 | mV |
| $V_T$ | Threshold potential | -50.4 | mV |
| $u_{th}$ | Spike onset threshold | 0 | mV |
| $u_r$ | Reset potential | -55 | mV |
| $\Delta_T$ | Slope factor | 2 | mV |
| $\tau_w$ | Spike-triggered adaptation time scale | 144 | ms |
| $a$ | Sub-threshold adaptation conductance | 4 | nS |
| $b$ | Spike-triggered adaptation increment | 80.5 | pA |
| $t_{up}$ | Spike width (soma clamped at 20 mV) | 1 | ms |
| $t_{ref}$ | Absolute refractory period | 2 | ms |

Table 3.2: Parameters for proximal and distal dendritic compartments in the Tripod neuron model.

| Symbol | Description | Distal | Proximal | Unit |
|---|---|---|---|---|
| $l$ | Dendritic length | 400 | 150 | µm |
| $d$ | Dendritic diameter | 4 | 4 | µm |
| $g_m$ | Leak conductance | 1.29 | 0.32 | nS |
| $\tau_d$ | Membrane time constant | 1.48 | 0.22 | ms |
| $g_{ax}$ | Axial conductance | 15.71 | 62.83 | nS |
| $C_m$ | Membrane capacitance | 25.13 | 6.28 | pF |

Table 3.3: Synaptic parameters of the Tripod neuron used in the simulations in the Results section.

| | $E_r$(mV) | $\tau_r$(ms) | $\tau_d$(ms) | $\bar{g}^{\text{syn}}$(nS) | $\gamma\left(\frac{1}{\text{mV}}\right)$ |
|---|---|---|---|---|---|
| **Excitation** | | | | | |
| ***Soma*** | | | | | |
| AMPA | 0 | 0.26 | 2 | 0.73 | – |
| ***Dendrite with NMDA + AMPA*** | | | | | |
| AMPA | 0 | 0.26 | 2 | 0.73 | – |
| NMDA | 0 | 8 | 35 | 1.31 | 0.075 |
| ***Dendrite with AMPA-only*** | | | | | |
| AMPA | 0 | 0.26 | 2 | 2.04 | – |
| **Inhibition** | | | | | |
| ***Soma*** | | | | | |
| $GABA_A$ | −75 | 0.1 | 15 | 0.38 | – |
| ***Dendrite*** | | | | | |
| $GABA_A$ | −75 | 4.8 | 29 | 0.27 | – |
| $GABA_B$ | −90 | 30 | 400 | −0.006 | – |

## 3.6 Appendix

### Appendix A: Dendritic integration under synaptic bombardment

To understand the response of the Tripod to barrages of pre-synaptic excitatory and inhibitory spikes, we expose the dendrites to intense synaptic bombardment and measure somatic activity. For this experiment, we stimulate the neuron with stationary excitatory and inhibitory Poisson input at fixed rates, $\mathscr{I} = (\nu_{ext}; k_{E/I} = 1.25)$, while systematically varying the length of each compartment independently (Fig.3.1*A*). The synaptic strengths correspond to the maximal synaptic conductance for excitatory and inhibitory synapses. We measure the rates and the coefficient of variation of the inter-spike-interval ($CV_{ISI}$) for all dendritic lengths ($x$ and $y$ axes in Fig.3.1*A* and *B*, 100 µm to 500 µm) and increasing input rates $\nu_{exc}$ (8 panels, 1.5 kHz to 50 kHz). The output rates and the $CV_{ISI}$ of the models are illustrated in Fig.3.1 *A* and *B*, respectively. For low input rates ($\nu_{exc}$ <1.5 kHz), the neuron is not sufficiently excited to produce spikes, regardless of the dendritic length. Increasing the input rate of both excitatory and inhibitory neurons ($\nu_{exc}$ from 1.5 kHz to 15 kHz), the Tripod activates and responds with increasing spike rate to the stimulus. Beyond 15 kHz, the firing response stabilizes in the interval 10 Hz to 50 Hz for most dendritic configurations, with the exception of the Tripod models with the shortest compartments (bottom-left corners) and the longest ones (top-right corners) which tend to 70 Hz and 0 Hz respectively. Non-symmetric configurations (top-left and bottom-right corners) have intermediate rates, suggesting that each dendritic length contributes with the same depolarizing current to the soma as it does in its symmetric counterparts. With regards to the firing regularity (Fig.3.1*B*), sparse firing with biological $CV_{ISI}$ (0.5 to 1.5) is observed only in the intermediate input rates; for high input rates the model's response tends to be the most regular, with $CV_{ISI}$ approaching zero. Dendritic asymmetries have an impact on the regularity of somatic spiking, too. In the intermediate ranges, dendritic asymmetries lead to a higher coefficient of variation of the inter-spike interval ($CV_{ISI}$). The impact of asymmetry is visible through the darker off-diagonal shades in the central panels of Fig.3.1*B*.

In the experiment illustrated in Fig.3.1*A* and Fig.3.1*B*, the E/I ratio ($k_{E/I}$) is arbitrarily fixed to 1.25 to ensure a substantial firing from the Tripod. Similar results were obtained by varying the value within a finite interval. The panel in Fig.3.1*C* shows the firing rate for the symmetrical model averaged across all

dendritic lengths and for $k_{E/I}$ in the range 0.5 to 5 The results are highly consistent and there is no qualitative difference between simulations with different $k_{E/I}$. The balance ratio only determines the minimal excitatory rate necessary to trigger somatic firing (it shifts the response pattern on the x-axis) and limits the firing response (lowering the plateau from 60 Hz to 30 Hz on the y-axis); crucially the steepness of the response function remains the same for different balance ratios. In the case of small values of the E/I balance ratio (i.e., $k_E/I < 1$), the inhibition dominates, and the cell has little to no activity (lighter blue lines).

It is worth looking deeper into the response of the Tripod models at low input regimes, illustrated in Fig.3.1*D*. For inputs in the range (1.5 kHz to 5 kHz), the models with higher rates are not those with the shortest dendrites but those with intermediate lengths. This trend is clearly shown in Fig.3.1*D*, which indicates the dendritic length with the strongest response (highest output rate) for each input rate, and for each E/I ratio. Long dendrites benefit from the electrical segregation from the soma; in these input conditions, the NMDAR's non-linearity has a pronounced effect. On the other hand, short dendrites do not reach the necessary potential to activate the NMDARs and remain hyperpolarized. The same trend also emerges in the balance conditions discussed below. Around 5 Hz, all dendrites receive enough inputs to activate the NMDARs voltage-gating fully; in these conditions, the high axial conductance of short dendrites favors the Tripod with proximal configurations.

## Appendix B: Balance with inhibitory synaptic plasticity

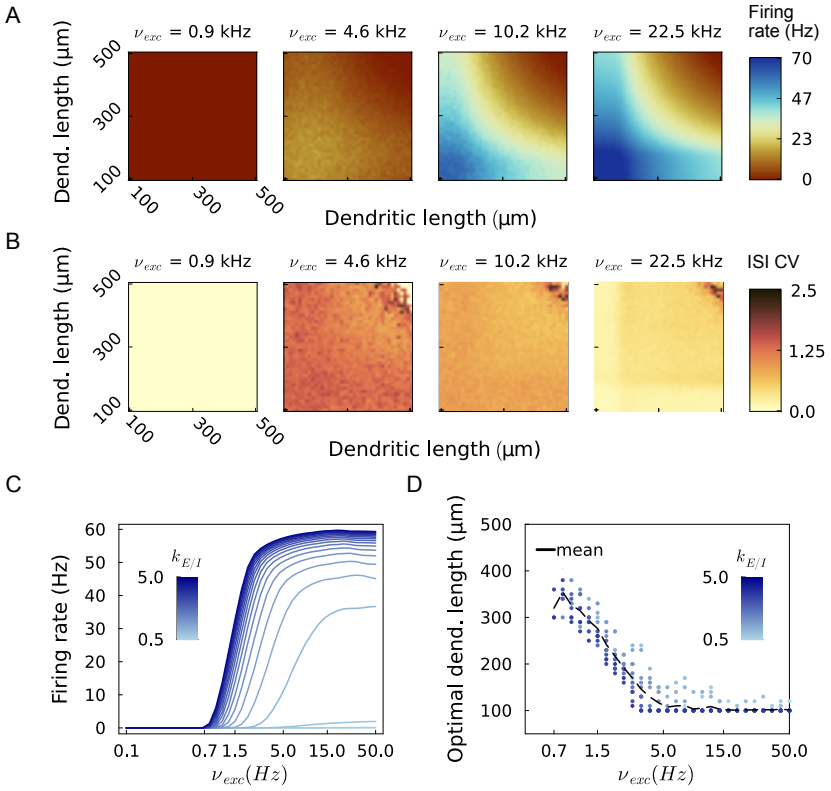
To verify the robustness of our results on the balanced condition, we computed the balanced state using a different operative definition of the balanced state (the neuron firing at 10 Hz). For this purpose, it was not possible to use the free membrane potential and we adjusted the inhibitory stream via inhibitory spike-timing-dependent plasticity (Vogels et al., 2011). We impose the E/I balance such that the Tripod fires sparsely (coefficient of variation of inter-spike-interval (CV ISI)$\approx$ 1 and firing rate within 5 Hz to 10 Hz), under continuous and supra-threshold excitatory stimuli. These requirements match with computational and experimental evidence (Destexhe et al., 2003; Higley & Contreras, 2006; Maksimov et al., 2018; Vogels et al., 2011).

We set the excitatory and inhibitory rates equal and an inhibitory spike-time-dependent plasticity (iSTDP) learning rule to compute the appropriate inhibitory synaptic efficacy ($g_{syn}^{\text{GABA}}$) (illustration in Fig.3.2*A*). The iSTDP rule from Vogels et al. (2011) finds the inhibitory synaptic efficacy such that the Tripod firing rate $\nu_{out}$ converges to a target value $\alpha$, which is set to 5 Hz in this study. For each dendritic compartment, the synaptic efficacy is updated whenever a pre-synaptic inhibitory spike arrives on the Tripod ($X^I = \delta(t_I - t)$) or the Tripod fires ($X^{out} = \delta(t_{out} - t)$):

$$J_{inh} \leftarrow J_{inh} + \eta \left( X^{out}(\nu_{out} - \alpha) + X^I \nu_{inh} \right) \tag{3.8}$$

This rule acts as a homeostatic feedback; it increases the inhibitory synaptic strength when the Tripod is firing more frequently than the target rate $\alpha = 5$ Hz and decreases it when it is firing at a lower rate. The $\bar{k}_{E/I}$reached with the iSTDP protocol are portrayed in Fig.3.2*B*.

On the theoretical ground, it is hard to equate the balance reach via synaptic efficacy and rate in the Tripod model. Because the NMDARs' voltage-gated component prevents linearizing the neuron's membrane potential and applies Campbell's theorem, as proposed in Kuhn et al. (2004). However, we could evaluate the correlation between the $\bar{k}_{E/I}$reached with the two protocols. For the three models (NMDARs, AMPARs, soma-only), the rate and iSTDP protocols converge on highly similar E/I ratios ($\rho_{\text{AMPA}}$ =0.95, $\rho_{\text{NMDA}} = 0.94$, $\rho_{\text{soma}} =$ 0.99). The relationship between rate and iSTDP protocols is portrayed in the scatter plots of Fig.3.2*C*). Each point indicates the $\bar{k}_{E/I}$achieved with rate balance (y-axis) and the rate reached with iSTDP balance (x-axis), color-coded for all dendritic lengths and inputs. The $\bar{k}_{E/I}$values for the rate-balanced condition



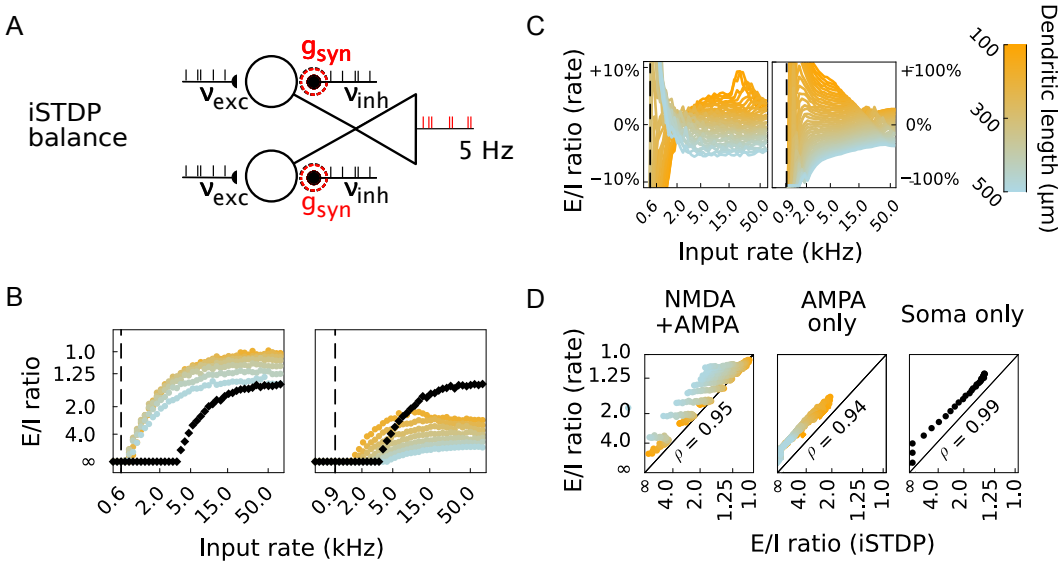

Appendix Figure 3.1: **Spiking activity with unbalanced inputs**
**(A)** Firing rate for all geometric configurations in the range 100 µm to 500 µm (x and y axes) under increasing synaptic bombardment (panels 1-4). The $k_{E/I}$ ratio is fixed to 1.25. The excitatory input rate is varied between 0.1 kHz to 50 kHz, of which four intermediate input frequencies are shown. The firing rate of the Tripod depends strongly on the dendritic lengths, and the difference in response between models with long and short dendrites increases with the input rates. **(B)** Same inputs configuration as in *A*, the four panels indicate the inter-spike-interval coefficient of variation ($CV_{ISI}$) of the model under synaptic bombardment. The $CV_{ISI}$ is larger for asymmetrical models (near each panel's left and bottom axes) and also depends on the external input rate. The models simulated in this experiment have biologically plausible responses only in the intermediate input ranges (panels 2 and 3). **(C)** Firing rate averaged across all dendritic lengths as a function of the input rate. The panel compares it over several excitation/inhibition ratio values $K_{E/I}$ (color shades). The firing response has a consistent pattern against $k_{E/I}$; the neuron is unresponsive up to a threshold, then rapidly converges to a firing rate plateau that remains unchanged for high input rates. The excitation/inhibition ratio sets the firing threshold (shifts on the x-axis) and the magnitude of the firing plateau (shifts on the y-axis); however, the response curve remains the same. **(D)** Dendritic length of the model (symmetrical only) with maximal firing response plotted against the input rate. Low frequencies show a more robust spike response in models with long dendrites.

Appendix Figure 3.1: **(D)** Relative variations in $\bar{k}_{E/I}$across dendritic lengths, in models with and without NMDARS. The variation is computed with respect to the average. In the presence of voltage-gated receptors the relative variations are contained within the 10%, furthermore, they invert around 10 kHz. In a model with AMPARs only, the spread of $\bar{k}_{E/I}$is ten times larger.

are smaller (shifted left w.r.t. the diagonal) than the $\bar{k}_{E/I}$values for the iSTDP protocol. Such difference is due to the larger somatic depolarization necessary to achieve the imposed firing rate than to stabilize the soma at −55 mV. The perfect match between the two conditions is reached for higher values of the target rate for the rate-balance protocol, namely $V_m = -50$ mV, which was tested for control.

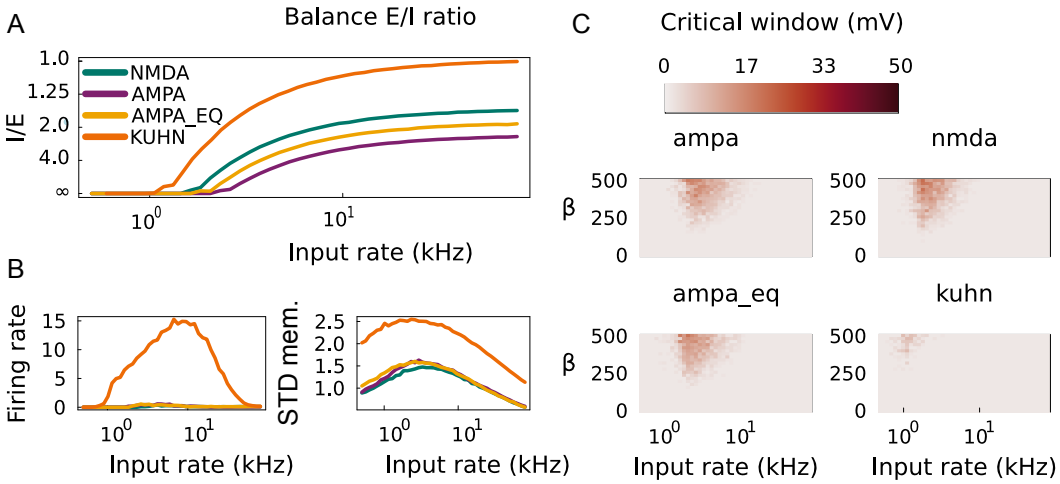
The two protocols differ concerning the dependence of $\bar{k}_{E/I}$on the dendritic length. For each input rate, the fork between the shortest and the longest dendrite is larger when the balance is reached via iSTDP than when the balance is reached with the rate condition. This is visible by confronting the panels in Fig.3.2*B* and Fig.3.2*B* and by considering the horizontal clusters of increasing dendritic lengths in the NMDARs condition of Fig.3.2*D*. We hypothesize that the dendritic length - $\bar{k}_{E/I}$interaction is due to the integration timescale of each dendrite, which depends on the length. Slower or faster integration timescales can accentuate or attenuate the shunting inhibition effect, thus affecting the $\bar{k}_{E/I}$obtained via the iSTDP protocol.

The interaction between dendritic length and dendritic-non linearity in determining the E/I balance is made clear in Fig.3.2*C*; the panel shows the relative difference between the $\bar{k}_{E/I}$obtained with the rate-modulation method, for all the dendritic lengths. The difference is computed with respect to the average $\bar{k}_{E/I}$. For NMDAR models, the equilibrium $\bar{k}_{E/I}$is contained within the 10 % of variation for all the dendritic lengths. It means that for a dendrite of 500 µm, the $\bar{k}_{E/I}$is less than 20 % larger than for a dendrite 100 µm long. When dendritic non-linearities are absent in the dendrites, the variation reaches the 100 %. Thus, the presence of NMDARs weakens the dependency on the dendritic lengths, indicating that the inhibitory neurons require less tuning for the dendritic target.




Appendix Figure 3.2: **Balanced excitation/inhibition with iSTDP**
**(A)** The iSTDP protocol acts on the synaptic conductance of the gabaergic synapses; it varies following an iSTDP learning rule until the output rate of the neuron is 5 Hz **(B)** The panels illustrate the E/I ratio obtained by modulating the inhibitory synaptic efficacy (iSTDP protocol). **(C)** The dependency on the dendritic length is larger in models with AMPA only, similar to what was observed in the $\bar{k}_{E/I}$obtained with the rate model. However, the iSTDP protocol results in stronger dependency of the $\bar{k}_{E/I}$from the dendritic length, possibly because of the effect of shunting inhibition on the spike-generation mechanism (see main text). (**D**) Correlation, model per model and rate per rate, of the iSTDP and rate-modulation $\bar{k}_{E/I}$. Although E/I with iSTDP shows higher dependence on the dendritic length, the two methods maintain extremely high correlations in the $\bar{k}_{E/I}$.

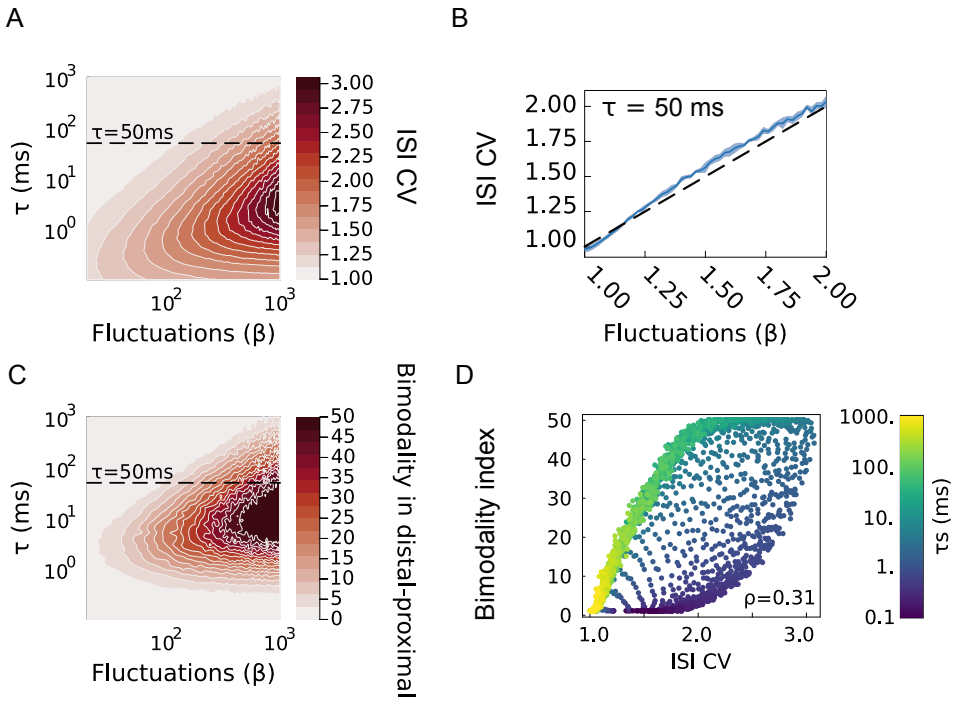
## 3.7 Supplementary Material




Supplementary Figure 3.1: **Somatic models do not express bistability, independently from their synaptic strengths and timescales**
In the three panels, we analyze E/I balance points, firing rate and membrane standard deviation, and bimodality index of four soma-only models with different synaptic efficacy and timescales (Table3.4). Two of the models were used in Quaresima et al. (2022), a third one in Duarte and Morrison (2019), and a fourth model matches the one investigated in (Kuhn et al., 2004). **(A)** $\bar{k}_{E/I}$values for the four models computed varying the inhibitory firing rate. **(B)** Firing rate and somatic membrane standard deviation. The rate and membrane fluctuations with *KUHN* synaptic parameters use a double-exponential fit of the $\alpha$-function synapses used in Kuhn et al. (2004); the output rate and membrane fluctuations in this synaptic configuration reproduce the original work. Differences with the remaining soma models are due to the smaller maximal conductance ($g_{syn}^{\text{AMPA}}$). The maximal synaptic conductance used in the soma-only model by Kuhn et al. (2004) is 10 times larger (Table3.4) than in our model. If we factor in the synaptic efficacy in the spike train statistics, the spike train has $\text{CV}_{\text{ISI}}$ equal to $g_{syn}^{max} * \text{CV}_{\text{ISI}}^0$; thus, the KHUN model is exposed to a spike train with larger $\text{CV}_{\text{ISI}}$than the other models. **(C)** Bimodality index for the four models. The soma models report bimodality indices consistent with those shown in Fig.3.6*B*. Crucially, the model with *KUHN* synapse shows no fluctuations.

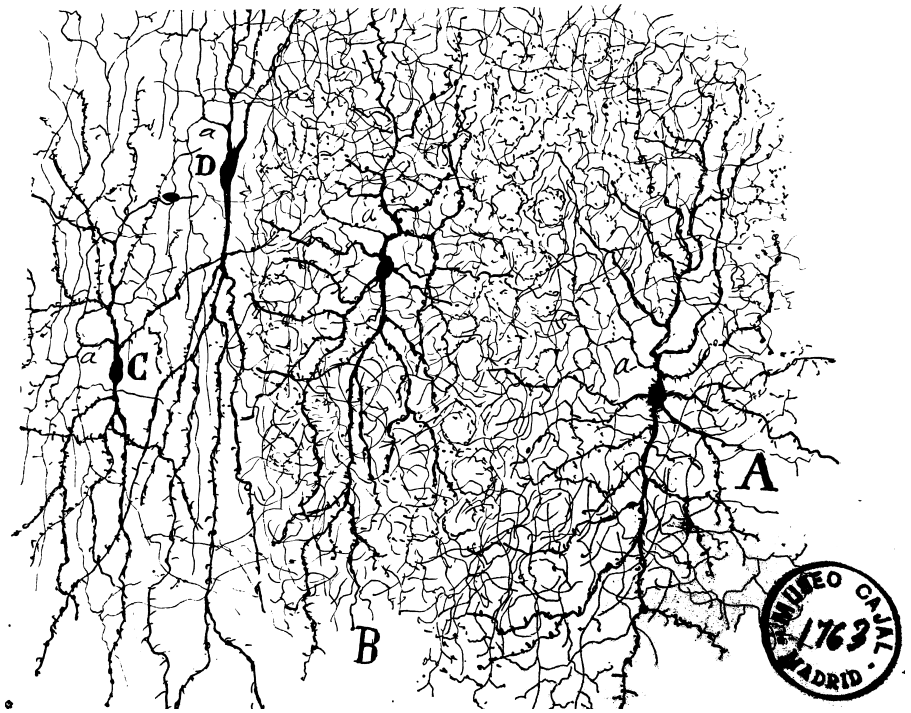
Table 3.4: Synaptic parameters of soma-only models

| Receptor | $E_r$(mV) | $\tau_{rise}$(ms) | $\tau_{decay}$(ms) | $g_{syn}$(nS) | $\gamma\left(\frac{1}{\text{mV}}\right)$ |
|---|---|---|---|---|---|
| **Tripod soma** (Quaresima et al., 2022) | | | | | |
| AMPA | 0.0 | 0.26 | 2. | 0.73 | - |
| NMDA | - | - | - | - | |
| **Tripod soma AMPA-equivalent** | | | | | |
| AMPA | 0.0 | 0.26 | 2. | 2.2 | - |
| NMDA | - | - | - | - | |
| **Soma model** (Duarte & Morrison, 2019) | | | | | |
| AMPA | 0.0 | 0.26 | 2. | 0.73 | - |
| NMDA | 0.0 | 1.00 | 100 | 0.159 | 0.62 |
| **Soma model** (Kuhn et al., 2004) *(alpha function synapse)* | | | | | |
| AMPA | 0.0 | 0.25 | 0.30 | 7. | - |
| NMDA | - | - | - | - | |




Supplementary Figure 3.2: **Tripod neuron bimodality depends on input** $CV_{ISI}$ The stochastic process used to generate the input spike train (Eq.3.4) has two free parameters: the autocorrelation timescale ($\tau$) and the fluctuation size ($\beta$). In the main text, we use $\tau$ =50 ms, consistently with previous work from Vogels et al. (2011) and show that, in this condition, the $CV_{ISI}$of the spike train is proportional to the second parameter ($\tilde{\beta}$, Eq.3.6). Crucially, the results on the onset of UDS in the Tripod neuron do not depend on the particular combination of $\tau$ and $\beta$, but only on the resulting $CV_{ISI}$. (**A**) $CV_{ISI}$of the spike train for $\tau_r$ =0.1 ms to 1000 ms and $\beta = 0$ to 1000, in log-log scale. (**B**) Linear relationship between $CV_{ISI}$ and $\beta$ for $\tau$ =50 ms. (**C**) The bimodality index measured in the Tripod neuron with distal-proximal compartment and excitatory firing rate equal 5 kHz. (**D**) The interaction between $CV_{ISI}$and bimodality index is quasi-linear for $\tau$ above 10 ms, which is a reasonable timescale for cortical processes. It has an overall Pearson correlation of 0.31. If only the values of $\tau$ larger than 10 ms are taken into account, the correlation reaches the 0.95. The results indicate that our model processes the inputs similarly, whether their fluctuations are fast and small or slow and big.

# 4 | Dendrites support formation and reactivation of sequential memories through Hebbian plasticity



## Abstract

Storage and retrieval of sequences require memory that is sensitive to the temporal order of features. For example, in human language, words that are stored in long-term memory are retrieved based on the order of phonemes. It is currently unknown whether Hebbian learning supports the formation of memories that are structured in time. We investigated whether word-like memories can emerge in a network of neurons with dendritic structures. Dendrites provide neuronal processing memory on the order of 100 ms and have been implicated in structured memory formation. We compared a network of neurons with dendrites and two networks of point neurons that have previously been shown to acquire stable long-term memories and process sequential information. The networks were equipped with voltage-based, spike-timing-dependent plasticity (STDP) and were homeostatically balanced with inhibitory STDP. In the learning phase, networks were exposed to phoneme sequences and word labels, which led to the formation of overlapping cell assemblies. In the retrieval phase, networks only received phoneme sequences as input, and we measured the firing activity of the corresponding word populations. The network with dendrites correctly reactivated the word populations with a success rate of 80%, including words composed of the same phonemes in a different order. The networks of point neurons reactivated only words that contained phonemes that were unique to these words and confused words with shared phonemes (success rate below 20%). These results suggest that the slow timescale and non-linearity of dendritic depolarization allowed neurons to establish connections between neural groups that were sensitive to serial order. Inhibitory STDP prevented the potentiation of connections between unrelated neural populations during learning. During retrieval, it maintained the dendrites hyperpolarized and limited the reactivation of incorrect cell assemblies. Thus, the addition of dendrites enables the encoding of temporal relations into associative memories.

## 4.1 Introduction

Speech perception results from the integration of continuous streams of acoustic information over time. Understanding how this capacity is grounded in the underlying neural activity remains a challenge for the brain sciences. One aspect of the phenomenon to be explained concerns how the brain accesses long-term memories on the base of environmental stimuli with temporal extent. During spoken word recognition, it corresponds to recollecting words from a lifelong learned lexicon based on the phonological information (McQueen, 2007). Although many of the neural markers of this process are known, a mechanistic theory of how word memories are learned, maintained, and accessed is lacking (Poeppel & Idsardi, 2022). One hypothesis is that word memories are stored in the long term in the mental lexicon in the form of cell assemblies that are acquired through Hebbian learning (Garagnani et al., 2009; Pulvermüller, 1999); this view is in agreement with cognitive and computational theories of cortical processing based on the synaptic junctions (Amit, 1995; Fuster, 1997; Gastaldi, Schwalger, De Falco, Quiroga, & Gerstner, 2021). However, because of the *phonological overlap* between words in the human lexica (words being composed of the same phonemes, such as /kæt/ and /tæk/), accessing the word forms entails that cortical memories are sensitive to the order of the phonemes in the stimulus. According to a long-standing critique of associative memories, these computational requirements cannot be achieved in the synaptic connections because associative memories lack the necessary expressiveness to encode relationships (Gallistel, 2021; Gallistel & King, 2011), i.e., the order among phonemes in human words. Thus, the main theoretical issue concerning the implementation of a biologically constrained word recognition model is the capacity of a spiking neural network to discriminate between sequences of similar inputs. We refer to this problem as the sequence detection problem. In the next three sections, we outline the requirements for long-term and short-term memory imposed by sequence recognition and finally present a model to solve this cognitive-computational problem.

#### Formation and retrieval of word memories

A cell assembly is a functional population of neurons that emerges from the repeated co-activation of cells (Hebb, 1949). Reciprocal firing leads to synaptic potentiation and the formation of synaptic engrams in the form of interconnected subnetworks. The engrams are carved in the neuronal tissue through

experience, following associative plasticity (Langille & Brown, 2018). Accordingly, assemblies are considered the core unit of memory (Poo et al., 2016). The strengthening of their synaptic connections, that is, the formation of an engram through long-term potentiation (LTP), is considered the causal landmark of memory formation (Dringenberg, 2020).

In the cell assembly view, the recognition of a familiar stimulus during sound perception corresponds to the reactivation of groups of auditory neurons and generates a burst of activity (Sakurai et al., 2018). Reliable, correlated cell activity is widely observed in the experimental literature on human and non-human animals (Almeida-Filho et al., 2014; Cohen et al., 2020; Hemberger, Shein-Idelson, Pammer, & Laurent, 2019), for a critical review of this phenomenon see Langille and Gallistel (2020). Compelling evidence on word recognition comes from two studies that obtained intracortical recordings in humans. Chan et al. (2014) observed bursts of neural activity in localized populations of the temporal cortex upon acoustic presentation of single words, but not when the stimulus was played backward. Secondly, Vaz, Wittig, Inati, and Zaghloul (2020) isolated neuronal bursts in the middle temporal gyrus during recall of spoken word associations. Cortical neurons reactivated in ordered sequences during the learning period and when the cue-target pair was correctly recollected, but they lacked sequential structure when the association was wrong. Taken together, these studies suggest that accessing spoken word memories entails sequential firing of cell assemblies in speech areas.

Sequential spikes within assemblies, named phase-sequences, were already indicated as one of the distinctive features of memory access by Hebb (1949). Sequential firing is pivotal for the passage of information to downstream neural populations with causal consequences on behavior (Buzsáki, 2010). Several models have shown how Hebbian plasticity leads to the generation of phase-sequences in spiking neurons; networks can be trained to repeat sequences and activate assemblies in succession (Cone & Shouval, 2021; Haga & Fukai, 2018; Maes, Barahona, & Clopath, 2020; Scott & Frank, 2023; Shouval, 2011). These models are based on two properties of cortical learning. First, memories can be instilled in the network connectivity via the repeated presentation of the stimuli (Litwin-Kumar & Doiron, 2014; Tomasello et al., 2018; Zenke et al., 2015). Second, Hebbian plasticity, in the form of spike-time-dependent plasticity (STDP), supports the formation of directed connections, and thus sequentially sensitive activity (Clopath et al., 2010; Fiete, Senn, Wang, & Hahnloser, 2010). However, the fact that engrams can be arranged to produce sequential activity does

not explain how such spike sequences can be read out from downstream neural assemblies.

The problem becomes pressing if one intends to explain word recognition based on the neural mechanism of cell assembly; the order of phonemes in phonological overlapping words is key to distinguishing between word memories. In this respect, two questions must be addressed. First, can sequential activity trigger the recollection of order-dependent memories based solely on associative synapses? Second, does Hebbian associative synaptic plasticity support the acquisition of such memories? On the one hand, detecting sequences based on synaptic connections is achievable; networks with fine-tuned asymmetries in the weights structures can re-activate assemblies based on the order of the stimuli (Sequence Detector Network, Knoblauch & Pulvermüller, 2005) and supervised plasticity rules render neurons sensitive to the order of pre-synaptic firing (Tempotron neuron, Gütig & Sompolinsky, 2006). On the other hand, these solutions have not been demonstrated to be valid in realistic conditions. For example, when the STDP rule was tested in a sequence labeling task in a biologically constrained network, plasticity among excitatory cells offered only a moderate contribution (Duarte & Morrison, 2014). Thus, it remains unclear how to induce order-sensitive memories in recurrent spiking models using Hebbian-like plasticity.

**Integration of long-term and short-term phonological memories**

One possibility is that biological networks require additional computational primitives to perform sequence detection. For example, variables to carry information forward on the time scale of the memory to be accessed (Chaudhuri & Fiete, 2016). Theoretical and experimental work indicates that short-term memory (STM) of phonological information must be held in the system by a dedicated mechanism, distinct from long-term memory (LTM) (Norris, 2017). The volatile auditory information storage must also allow for learning of the new association in the LTM. Arguably, with a compatible STM storage, the STDP rule might be sensitive to the sequential structure of the stimulus.

We hypothesize that the STM must satisfy certain requirements in order to achieve sequence labeling. First, it must encode time information, for example, in its slow decay to the resting state; second, there has to be a silent memory such that rapid compensatory mechanisms do not erase it (Zenke et al., 2015); and, third, it should interact with the long-term rule in the network to induce potentiation of the salient synapses. Of the three main STM proposals in the

computational literature, persistent activity (Papoutsi et al., 2014; Wang, 1999, 2021), synaptic short-term memory (Mongillo et al., 2008), and neuronal memory (Fitz et al., 2020; Salaj et al., 2020); the constraints described appear to be satisfied only by synaptic short-term memories. Crucially, few computational studies have tested the interaction of this type of STM with the formation of cell assemblies. A previous study by Cone and Shouval (2021) demonstrated that networks can learn sequences on the timescale of human words using synaptic decay to encode STM and eligibility traces to update long-term memories. For each neuron, the synaptic memory was shared across pre-synaptic cells, which is not realistic considering the actual geometry of synaptic arrangements. In our opinion, the slow, post-synaptic decay that was implemented is better expressed by regenerative events in segregated dendritic compartments, for example, the NMDA receptor spikes. The crucial role played by dendritic processes in detecting spatiotemporal sequences is indeed demonstrated both experimentally and computationally (Bhalla, 2017; Branco et al., 2010).

We propose that dendritic memory (Quaresima et al., 2022) can satisfy these requirements as well as support the formation and maintenance of sequence memories on the timescale of human spoken words. Dendritic memory is expressed by long-lasting (100 ms) plateau potential, elicited by NMDA receptor spikes; the dendrites undergo long depolarized states that can bind the sequences of incoming phonemes. In addition, because the dendritic compartments are segregated, the somatic firing activity is only weakly coupled to the dendritic membrane potential and supports the silent encoding of short-term memories. The hypothesis is that if STM is expressed as dendritic memory, it will induce the formation of long-term memories through vSTDP, thus supporting sequence detection in biological networks.

### The Tripod network model and the sequence recognition task

Dendritic neurons are modeled as Tripod neurons (Quaresima et al., 2022), a three-compartment neuron model with two dendritic compartments and a soma. The dendritic compartments are endowed with NMDA receptors and voltage-gated channels that allow for dendritic memory in the order of a hundred milliseconds (Fig.4.1*A*). The memory is provided by the long-lasting dendritic depolarization following an NMDA spike.

The network is composed of 2000 excitatory Tripod neurons, 175 fast-spiking interneurons, and 325 slow-spiking inhibitory neurons, modeled as point neurons ($I_1$, $I_2$). Each neuron is connected to 20% of the other neurons in the net-

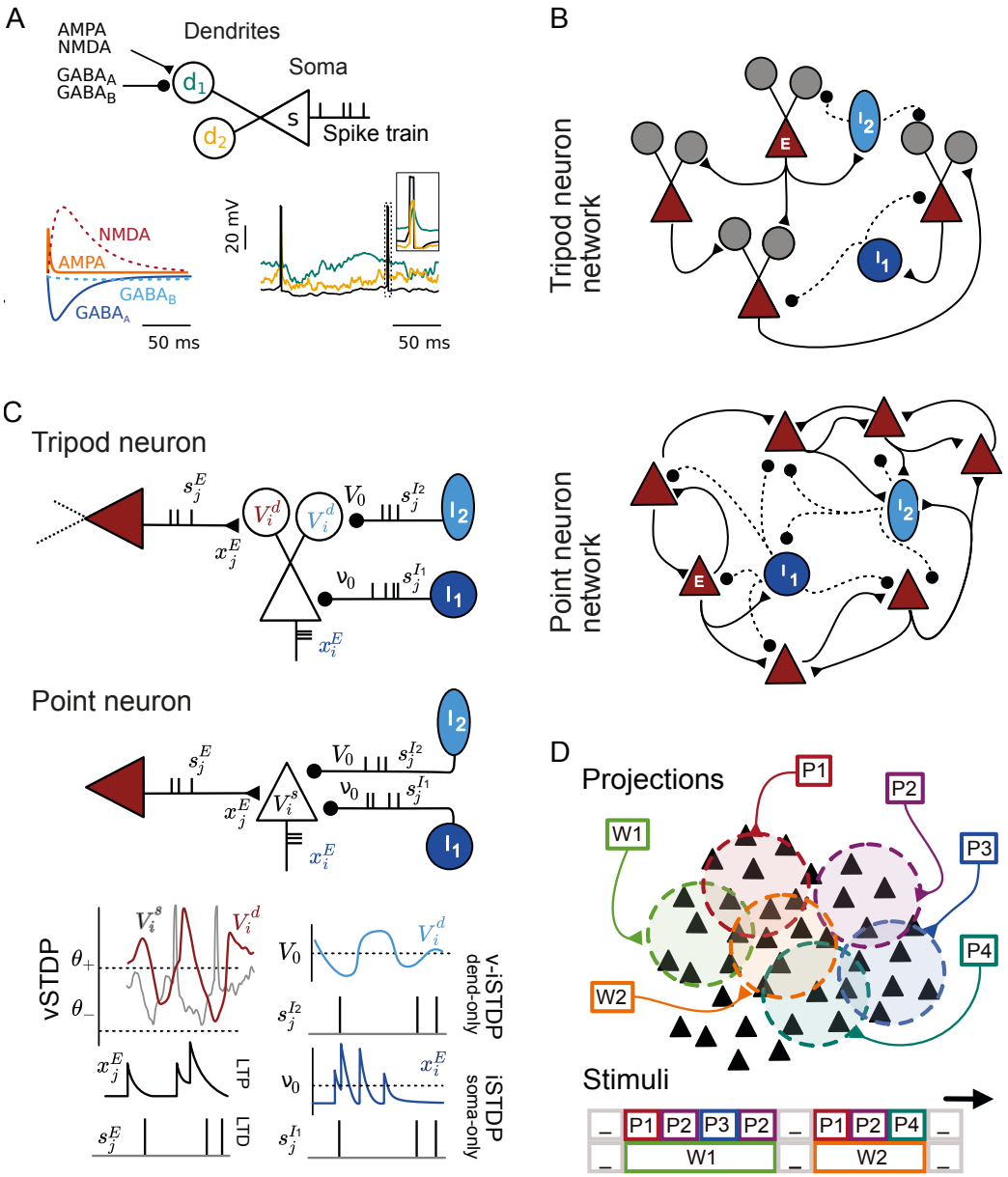



figure 4.1: **Tripod network: neuron models, connectivity, and plasticity rules**
*Figure caption continues on the next page.*

figure 4.1: **Tripod network: neuron models, connectivity, and plasticity rules.**
(**A**) The Tripod neuron has two dendritic compartments (circles) and a soma (triangle). The dendritic length determines the electrical properties of the compartment. Synapses included glutamatergic (AMPA, NMDA) and gabaergic ($GABA_A$, $GABA_B$) receptors with the illustrated timescales. The soma was modeled as an adaptive-exponential neuron and somatic spikes were backpropagates to the dendrites, as shown in the inset.
(**B**) In the Tripod network, excitatory neurons targeted the dendrites of other excitatory neurons and the soma of inhibitory neurons. Inhibitory neurons $I_1$ (fast-spiking) and $I_2$ (adaptive) targeted the soma and dendrites of Tripod neurons. They also form recurrent connections within the inhibitory populations. In the networks of point neurons, all synapses are connected to the somatic compartment.
(**C**) Schematics of excitatory and inhibitory synapses plasticity. The left side illustrates voltage-based STDP in the glutamatergic synapses between excitatory neurons ($V_i^d, V_i^s$ are the membrane potential of dendritic and somatic compartments). The plasticity rules are the same in the dendritic and somatic compartments, except for the thresholds of long-term potentiation (LTP) and depression (LTD), respectively $\theta^+$ and $\theta^-$. Right side; the inhibitory synapses onto the excitatory cells are subject to iSTDP. The equilibrium values depend on the neuron types. $I_2$cells aim to stabilize the dendritic potential at $V_0 = -70\,\text{mV}$ (v-iSTDP) and $I_1$cells regulate the firing rate of the soma, $x^E = 10\,\text{Hz}$ (iSTDP). Excitatory and inhibitory plasticity are driven by the pre-synaptic activity, in the form of filtered spike train ($x_j^E$) or sum of delta functions ($s_j^E, s_j^{I1}, s_j^{I2}$)
(**D**) External projections target a subset of excitatory neurons. The phoneme and word projections are fixed and target 5% of the excitatory population in both dendritic and point-neuron models. The phonemes are stimulated in sequence, and the corresponding word is stimulated throughout the interval.

work. The synaptic connections onto Tripod neurons are located on the dendrites (Fig.4.1*B*) and subject to voltage-dependent spike-timing-dependent plasticity (vSTDP, Bono & Clopath, 2017; Clopath et al., 2010). Excitatory plasticity supports the formation of engrams by strengthening the post-synaptic connections of neurons that fire onto depolarized dendrites. Because of the strongly connected cell assemblies, the network is prone to runaway activity. We used two additional mechanisms to prevent this from happening. First, homeostatic plasticity keeps the total incoming synaptic strength constant (multiplicative synaptic scaling, Tetzlaff, Kolodziejski, Timme, & Wörgötter, 2011). Second, the network is stabilized by means of fast compensatory mechanisms in the form of inhibitory spike-timing-dependent plasticity (iSTDP, Vogels et al., 2011). iSTDP is implemented via plastic connections between $I_1$ neurons and the soma of the excitatory cells. In addition, we introduce a voltage-dependent iSTDP rule between the $I_2$ population and the targeted dendritic compartment to reach a balanced excitatory-inhibitory synaptic input on the dendrites (Fig.4.1*C*).

We investigate word recognition in a Tripod network simulation. Word memories are intended as long-term associations between sequences of phonemes and words assemblies. The assemblies are induced on randomly selected subsets of the excitatory population, with word and phoneme assemblies loosely overlapping (Fig.4.1*D*). We use external input projections to induce the strengthening of synapses within and across the cell assemblies. The projections stimulate the dendritic compartments.

The network simulation is divided into two phases. The first phase is the associative one and we present both words and phonemes stimuli, simultaneously. On the base of the STDP rule among excitatory neurons, the co-activation of phonemes and word populations is expected to form auto-associative (recurrent engrams) and, possibly, hetero-associative (feedforward connections, between phonemes and words) memories. In the second phase, the recall phase, we turn off the plasticity mechanism and present only phoneme sequences. During recall, we measure the activity of all word populations and consider word recognition successful if the target word population is the most active one in the pool of the word populations.

We test word recognition on seven lexica that differ in the number of words they contain (8 to 17) and the amount of phonological overlap among the words. Two of them, *Identity* and *No overlap*, contain words whose phonemes are not shared across the words in the lexicon; thus, each phoneme is univocally associated with a word. Conversely, the other lexica comprise words with phonological

overlap. An example from the lexicon *Overlap* are three words *log, dog,* and *god*; the first two words share two phonemes, but the second and the third share all the phonemes, and they are distinguishable only on the base of the order of the phonemes presented. Thus, correct word selection requires phonemes-word associations sensitive to the order of the phonemes in the stimulus rather than their sole identity features. The network must create order-sensitive associations to achieve correct word recognition. We compare the Tripod network with two point-neuron models endowed with similar plasticity mechanisms and measure the differences in the recognition score, its dynamics, and the synaptic structure formed following the associative phase.

## 4.2 Results

### 4.2.1 Word recognition in the Tripod network model

We investigated the capacity of the Tripod network to form and recall time structure-dependent memories in a word recognition task. Because the dendritic neuron model has not been studied in a recurrent network before, we first set the network to a realistic operational point and verified the absence of pathological dynamics. We tuned background noise and synaptic strengths such that the network was in a regime of sparse firing. When the three synaptic learning rules apply (v-STDP, iSTDP, v-iSTDP), the network's baseline activity has a low firing rate and tends to synchronize in slow bursts of activity. Conversely, when receiving external input, the network shows a low degree of synchrony and a sparse firing rate (Appendix A). The analysis indicates that the firing rate is heterogeneous across cells, both in the associative and recall phases: neurons receiving external projections fire more than neurons that do not. In the associative phase, a strong response from both phonemes and word populations is expected because of the external stimuli on both populations. Conversely, in the recall phase, word assemblies are not stimulated, and their activation originates from the reverberation of activity in the phoneme populations.

The following sections analyze in detail the activation of word assemblies in the recall phase. First, we show that for each sequence of phonemes, the most reactivated population is the one corresponding to the word associated with the phoneme sequence (the word *doll* and the sequence *D, O, L, L*). Then, by using the firing activity of word assemblies as the index of word recognition, we estimate the network capacity to detect sequences in a set of lexical with in-

creasing phonological overlap. In the remaining sections, we investigate the mechanisms that allow the network to recall word memories. To determine the role of the dendrites in the network, we test four additional networks of point-neuron and dendritic models on the same task and evaluate their recognition capacity and dynamics. Thus, we analyze the network structure that emerges during the learning phase. We demonstrate that the phonemes-to-word connections strengthen (long-term storage) as well as that the dendritic non-linearity (short-term memory) determine the network's capacity to recognize sequences.

**Word assemblies are sensitive to phonological order**

We start the analysis of the network recognition capacity by taking a closer look at the coordinated firing of word and phonemes population during the associative and recall phases. To highlight the structure of the network activity, we ordered the network neurons by the projections they receive and plotted their spikes in the raster plots in Fig.4.2*A*. The figure shows the network activity for the eight phonemes and ten words in the lexicon *Overlap* during both phases. In the recall phase, only phonemes populations receive external inputs. The lexicon used in this experiment includes words with phonological overlap. Some words are contained within others (e.g., *poll*/*pollen*), some are anagrams of one another (e.g., *lop*/*poll*) or reversed (*dog*/*god*).

We first analyze the activity of the assemblies in the associative and recall phases. The raster plot shows that the activity of phoneme assemblies lasts longer than the stimulation intervals (grey vertical lines). Phonemes are stimulated for 50 ms only, but their firing persists for about 50 ms after the offset of the stimulus. Although the external input was the same for all phoneme populations (8 kHz for 50 ms), phoneme assemblies were not reactivated with the same strength. For example, the phonemes *D*, *G*, and *L* respond more strongly to the external inputs than the phoneme *O*. Comparing the activity across the entire simulation, it appears that the average activity of the phoneme assemblies depends on the total number of occurrences of the phoneme (App. Fig.4.1 ). The low firing of frequent phonemes is due to both neuronal adaptation, which tends to hyperpolarize the frequently stimulated populations, and the homeostatic component of the plasticity rule, which penalizes recurrent connections associated with frequent phonemes. Similar to phoneme assemblies, neurons in the word populations fire in the two phases. In the associative phase, the word populations are activated along with the phonemes, and the external stimulus dominates their time course. The beginning of the assembly firing coincides

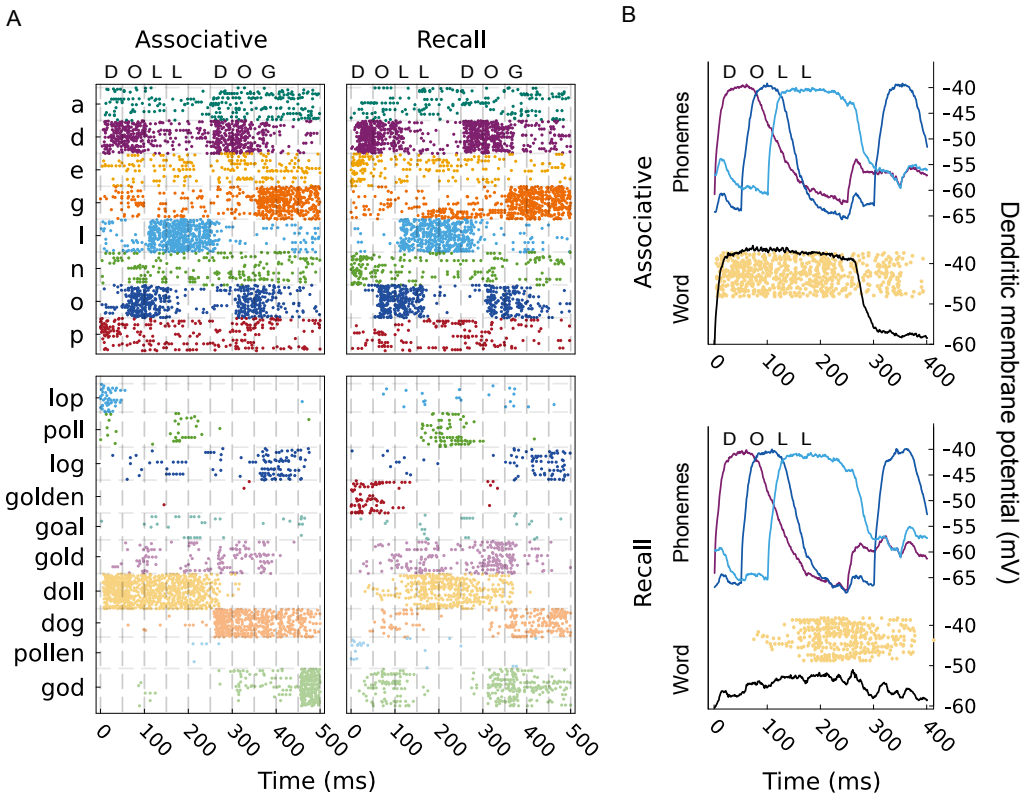



figure 4.2: **Network activity in the associative and recall phases**
(**A**) Raster plot of phoneme and word populations during the presentation of the words *doll* and *dog*. Individual phonemes are stimulated for 50 ms each with a silent interval of 50 ms between words. In the associative phase, the word assemblies are stimulated for the entire interval in which the words are activated (*doll* 200 ms, *dog* 150 ms) (left). Conversely, they are not stimulated in the recall phase but activate from the reverberation of activity related to phoneme associated populations (right). (**B**) Average dendritic membrane potential for the assemblies associated with the phonemes *D, O* and *L,* and the word *doll*; the panels show the associative (top) and recall (bottom) phases. In the associative phase, the external projections strongly depolarize the phonemes and word assemblies. In contrast, in the recall phase, only the phonemes reach the −40 mV depolarization. The membrane potential of the word assembly slowy builds up and reach the maximum depolarization at word offset. Similarly, the firing activity of the word population changes in the two phases. In the associative phase the word assembly starts firing at word onset (the yellow dots indicate the assembly's spikes), conversely, in the recall phase, the word population fires only after 50 ms to 100 ms from the onset of the phoneme sequence.

with word onset and quickly fades when the following word is presented. In the associative recall phase, the word populations do not receive input but are nonetheless activated by the activation of phoneme populations; they receive the external stimulus through the reverberation of the phoneme assemblies. All words that contain the input phonemes in some position are partially reactivated. For example, for the input sequence *D, O, L, L*, the words *dog*, *poll*, and *gold* are activated, however the population of the word *doll* fires more; presumably because it matches both the identity and the order of the phonemes (Fig.4.2*A*). This behavior indicates that the word assemblies activate in a manner that depend on the order of inputs and that the network can detect sequences.

Before proceeding with the analysis of the sequence recognition capacity, we explored the dynamics of the dendritic membrane potential of phonemes and word assemblies. Fig.4.2*B* offers insights into the prolonged activity of the assemblies. The panels illustrate the membrane potential of three phonemes (D, O, L) and the word assembly *doll* during the presentation of the corresponding phoneme sequence (*D, O, L, L*). The dendritic membrane potentials of both the phoneme- and word populations remain depolarized beyond the duration of the stimulus. Such slow decay is supported by the dendritic memory of the Tripod neuron, and it is due to the rise of NMDA spikes in segregated dendrites. In the intervals in which the phoneme population activity overlaps with the depolarized word population, the synaptic connections are potentiated by the STDP rule. Despite the average word membrane being below the LTP threshold, certain cells are sufficiently depolarized and form stable connections within and between assemblies. Thus, the large dendritic depolarization due to stimulation enables the formation of phonemes-to-words associations that are sensitive to the co-activation of phonemes.

In the recall phase, the reactivation of the target word population takes place gradually during the presentation of the phonemes. Because the NMDA spikes are all or none events in which the membrane potential reaches approximately $-20$ mV, the slow building up of dendritic potential in the lower panel of Fig.4.2*B* indicates the recruitment of more neurons in the word's assembly. The slow depolarization transforms into a burst of activity after 100 ms, when enough neurons are depolarized. The consequence of this dynamics is that the peak activity of the word population lags behind the onset of the phonological stimulus. Such delayed responses are observed across all the words inspected, and it appears to be an intrinsic property of word assembly reactivation in the dendritic model. In the following, we individuate an optimal interval in which word populations

are maximally activated and use this interval to quantify the word recognition in the model.

### Word recognition latency

The previous analysis indicates that in the Tripod network word populations are reactivated by the activity in the phoneme assemblies. In addition, the response of the target word population has a latency of approximately 100 ms. We now define a measure of word recognition that allows us to identify correct word recognition and accounts for the delay in word reactivation.

To start with, and elucidate how the measure works, we analyze the firing activity of the assembly associated with the word *doll*. The word lasts four phonemes, which corresponds to 200 ms. To inspect if the network reactivates the correct memory, we compare the assembly firing rate of the correct word population with those of the other populations during the interval in which the input is presented. For each interval in which the sequence *D, O, L, L* is presented, the corresponding word population should activate. Thus we can identify each of these intervals as trials to test word recognition. In addition, because of the word's population latency, we also test an interval shifted forwards of 100 ms. An illustration of the comparison among word population rates is presented in Fig.4.3*A*; the eight panels show the average firing rate of the target population *doll* against five competitor words (*dog, god, log, poll* and *goal*) and across the 110 trials where the sequence of phonemes (*D, O, L, L*) was presented. The activity of *doll* is shown on the vertical axis of each panel, and the horizontal axes show the activity of the word competitors. The black circles present each trial, and the red dots indicate when the average firing across the entire simulation is significant. The number annotated on each panel is the chance-corrected Cohen's ($\kappa$) word recognition score, based solely on the firing rate (Methods).

When the assembly's activity is measured in the same interval of the phonological input (0 ms to 200 ms, upper panel), the firing rate of the target is systematically less than when it is measured with a delay of 100 ms, in some cases, it is not even sufficient to distinguish between the target and the competitor (*god*, second panel). By taking the average on all words, including the remaining five, we estimate that when the word recognition measure is delayed 100 ms, the word assembly *doll* is correctly recognized in 90% of the trials, while only in the 60% when it is measured in the same interval of the phonological input. Because the delay is computed starting from the offset point, we refer to this as the offset point (OP) measure.

We can generalize the procedure described for the word *doll* and determine the optimal word recognition latency in the network for all the words. One issue is the difference in length among the lexical items; some words contain more phonemes than others. For example, the OP measure of *doll* computes the average firing over four phonemes, (200 ms), but its competitors have three, four, or six phonemes. This difference can introduce biases when comparing the firing rates of the associated assemblies. For this reason, we integrate the OP measure with a second one with a fixed length. We measure the firing rate for the duration of one phoneme (50 ms) and do so at the uniqueness point (UP) of the word. The UP is the phoneme that makes the word unique in the lexicon; it depends on the presence of other words with shared onset phonemes. As before, we shifted the interval forward or backward in time to test the timing of word reactivation. We refer to this shift as the interval delay. The UP and OP measures are illustrated in Fig.4.3*B* and described in further detail in the Methods.

For each trial, we select the word associated with the population with the highest firing rate as the retrieved lexical item. Then we use it to compute a confusion matrix. The confusion matrix indicates which word was expected (target) and which was retrieved (reactivated). Correct word recognition implies that the target and reactivated words are the same and that the matrix has most of its trials on the diagonal. An example of the evolution of the confusion matrix with the delay is portrayed in Fig.4.3*C*. The stacked heatmaps show the confusion matrix for the lexicon *Overlap* based on the uniqueness point measure. The bottom matrix corresponds to 0 ms of delay, and the pairing between target and reactivated words appears random. The reactivated words match the target more and more as the delay increases, until a maximum of 100 ms and start decreasing. This is somewhat unexpected because all words are optimally recognized with the same delay despite the lexicon containing words with 3–6 phonemes. Moreover, it is important to notice that the 100 ms interval overlaps with the presentation of the following word. From the confusion matrix, we computed the $\kappa$-score as described in the Methods, the $\kappa$-score for the *Overlap* lexicon is 0.77%

To characterize the model's capacity to recognize sequences in the presence of shared phonemes, we compared the recognition accuracy on seven different lexica that varied in size, number of shared phonemes (phonological overlap), and average word length. Some of the seven lexica have been used in previous studies on word recognition. The lexica and their statistical differences are described in the Methods. For the first two lexica (Identity and No overlap), words can be identified based on the first phoneme because it is unique to each word.

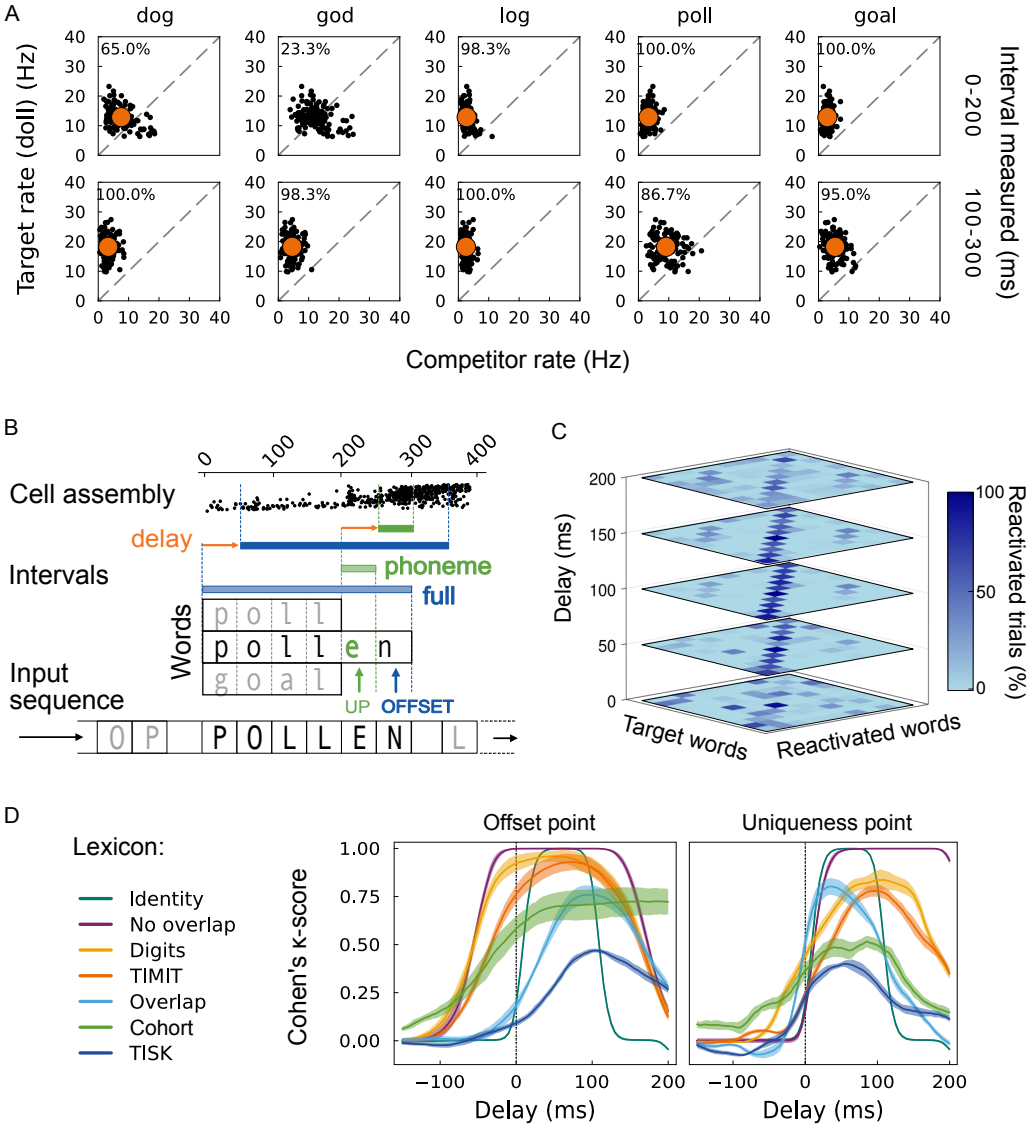



figure 4.3: **Recognition accuracy at word offset and uniqueness point**
(**A**) Firing rates of the target word population against competitor populations for the phoneme sequence */P, O, L, L/*. Rates are measured over the duration of the target word, shifted forward in time by 100 ms. Each dot represents one trial. When located above the diagonal, the target population was more active than the competitor word. Orange dots show the mean over 100 trials. *Figure caption continues on the next page.*

figure 4.3: (**B**) Word recognition is tested at the word offset and uniqueness points (blue and green arrows). Assembly activity is measured as the average firing rate in the interval between word onset and offset (full interval, blue bar) or the interval of the phoneme at the uniqueness point (single phoneme interval, green bar). The word with the highest firing rate in the measured interval is chosen as the network output. The measured interval is shifted in time (delay, orange arrow) to test the evolution of the network dynamics recognition time frame. (**C**) Network responses are summarized in the confusion matrix that indicates the percentage of trials in which a word was activated for each target word. Recognition accuracy is higher when the diagonal stands out, indicating that the correct word was reactivated for most of the trials. Off-diagonal values different from zero indicate failed recognition. The vertical axis shows the time shift at which the matrix was computed. The matrices that are shown were obtained for the *Overlap* lexicon at the offset point. (**D**) Cohen's $\kappa$-scores for the seven lexica averaged across ten samples for the offset (left panel) and uniqueness (right panel) point measures, plotted as a function of the interval delay. The network reaches maximum accuracy for delays between 50 ms to 150 ms for all lexica. The sharp increase of the $\kappa$-score in the right panel highlights the fact that word recognition cannot be achieved before the uniqueness point.

Thus, recognition does not require sequence memory, and we refer to these lexica as memoryless. The other five lexica have phonological overlap and require temporal integration of symbols in a sequence. We computed the $\kappa$-score on the seven lexica for the offset and uniqueness point measures and varying the delay between $-100$ ms to 200 ms. The accuracy measure associates every delay shift with an average recognition score. Thus, it quantifies the latency of word recognition. The recognition $\kappa$-score is plotted against the delay in Fig.4.3*D*.

In agreement with the confusion matrices, the $\kappa$-score increased when shifting the measured interval forward in time. For both the OP and UP measures (left and right panels), recognition accuracy peaked for delays in the range 50 ms to 150 ms. The offset point measure has a smoother ramp than the UP because it averages the firing rate on a longer interval. Recognition accuracy was generally higher for the OP measure than the UP measure. At the OP, the highest accuracy is above 70 % for all but the *TISK* lexicon. At the UP, the recognition is proportional to the full interval measure (OP) but slightly lower (the Pearson correlation between the two measures is 0.9). The similarity between the two measures indicates the network activity before and after the short UP interval contributes marginally to word recognition. Most importantly, the UP measure (right panel) shows that recognition is not possible ahead of the word's uniqueness point; for delays below 0 ms, most lexica had a recognition score close to zero. The UP measure indicates that before the UP, the network was in a state

of co-activation of multiple assemblies with insufficient information to discern between words (flat confusion matrix, $\kappa \approx 0$). Accuracy increases towards the UP for lexica with longer words because some of these words can be ruled out already before the UP, resulting in a higher chance of recognition (non-flat confusion matrix, $\kappa > 0$). For example, in the *Cohort* lexicon, the sequence *C, A, P, I, T* will exclude the words *capias, capillary* and *capistrate* although *T* occurs before the uniqueness point. We tested ten randomized realizations of the network, which reached similar accuracy on each lexicon. The standard deviation of the $\kappa$-score is below 5 % for all lexica and conditions, except *Cohort* in the OP measure where it reached 10 %.

**Dendrites are needed to recognize words with phonological overlap**

To understand the computational role of dendrites in the formation and reactivation of word assemblies, we compared the Tripod network with four other models, two with dendrites and two without. The networks of point neurons were chosen from the literature for their biological plausibility and relevance to the task at hand. The networks with dendrites are instead copies of the original model but stripped of the asymmetry in the dendritic compartments. By comparing the models, we aim to isolate the network mechanisms that support word recognition. We hypothesize that the dendritic memory, provided by electrical segregation and NMDA receptors in the dendritic compartments (Quaresima et al., 2022) is the computational primitive that supports the recognition of words with phonological overlap.

The first point-neuron model was the network described in Litwin-Kumar and Doiron (2014) (LKD for short). It has previously been shown to learn and maintain stable cell assemblies in the presence of ongoing plasticity and background noise, using interacting forms of excitatory and inhibitory STDP. Assemblies developed for 20 stimuli randomly injected into the network during the early phase, similar to our associative phase. Later, in the absence of inputs, the network spontaneously and robustly reactivated the assemblies, demonstrating that the network learned the memories presented in the early phase. The second model is based on the work of Duarte and Morrison (2019) (DM for short), which modeled an L2/3 circuit of mouse cortex. The original study investigated the role of various sources of heterogeneity on the network's computational capabilities, such as temporal integration and delayed decision-making. The network had no plasticity in the synaptic connections, and its recurrent weights were calibrated to achieve a balanced excitatory-inhibitory state. Because word recognition re-

quires associations between phonemes and words, we endowed the DM network with STDP on the synapses connecting excitatory neurons. Analysis showed that the DM network was unstable and prone to oscillatory dynamics. To stabilize the network, we added iSTDP and the v-iSTDP homeostatic inhibition, similar to the Tripod network but with both inhibitory plasticity mechanisms related to the soma of the excitatory neurons. Both the LKD and the DM networks implemented conductance-based synapses; only the DM network modeled NMDA receptors, which, independently from the results on dendrites, are known to play a role in working memory in that they support the onset of persistent activity in network cliques (Papoutsi et al., 2014; Wang, 1999, 2021). In addition, the two models differ in the density of their recurrent connectivity, the number of cells considered, and, most importantly, the presence of distinct classes of inhibitory neurons. The LKD model has 4000 excitatory cells and 1000 fast inhibitory neurons. Conversely, the DM has 2000 excitatory neurons, 175 fast-spiking gabaergic cells ($I_1$) and 325 slow-spiking ones ($I_2$). An illustration of the point-neuron models and their synaptic learning rules is shown in Fig.4.1*B* and *C*. Equations and parameters of the models can be found in the Methods.

Moreover, we compared the Tripod network with two other dendritic models, one with only one dendritic compartment and one model where the input and recurrent connections targeted both dendrites (symmetric network). The three models differ in their dendritic and afferent configurations but are all endowed with the dendritic memory provided by segregated compartments and NMDARs. Compared to networks of point neurons, the networks with dendrites were more robust to parameter variations. Network stability is due to the location of the recurrent connections on the dendrites; because the axial conductance limits the maximum current flowing from the dendrites to the soma the networks are less prone to epileptic firing. The network with dendrites configurations of the three models are illustrated in Fig.4.4*A* and further described in the Methods. Note that symmetry here refers to both the synaptic connectivity pattern and the dendritic lengths.

The five models were exposed to the seven lexica introduced in the previous section. For each pair of models and lexicon, we computed the recognition accuracy according to the OP and UP measures, with delays in the range of $-50$ ms to 150 ms. We introduce two novel measures to elucidate the time course of word recognition: the average recognition delay (ARD) and the optimal delay. The ARD is a linearly weighted average of the delay intervals, using the $\kappa$-score as weight (Eq.4.18, Methods); the optimal delay is instead the delay with which the

model reached the maximum score on the lexicon. The comparison between the models' accuracy, optimal, and average recognition delays in the word recognition task is shown in Fig.4.4*B*. The panels refer to the UP measure. The analysis for the OP measure is qualitatively similar ( SI Fig.4.6 ). The greyscale bars show the two lexica that do not require recognition memory, while colored bars refer to lexica that do. The $\kappa$-scores for the latter group were averaged for each model and compared pairwise between models in Fig.4.4*C*. The red color intensity codes for differences within +10 percentage points between models on the y-axis and models on the x-axis, with the significance level of the relative differences between pairs of models indicated.

The top panel shows that networks with dendrites have systematically higher recognition scores than point-neuron models in lexica with phonological overlap. The point-neuron models perform well on the memoryless lexica, but their accuracy drops when temporal integration is required. The confusion matrices (SI Fig.4.6 *A*) showed that, in the latter case, the networks reactivated a subset of the word assemblies for any input sequence, suggesting that the functional association between the phonemes and word populations was not established. The point-neuron models did not learn the sequential structure of word memories. In contrast, the three dendritic models recognized all the words, in all lexica; the errors were distributed equally among the remaining words (SI Fig.4.6 *B*).

In addition, the ARD and optimal delay measures reveal that the time course of recognition differs in the two models. Dendritic models have high recognition, on average, for delays in the range 50 ms to 100 ms from the offset point (middle panel). Conversely, the delays corresponding to the maximum score (bottom panel) span a large range and depend on the lexicon property. The ARD of the point-neuron models is negative and its optimal delay is negative or close to zero, which indicates that the correct word populations are maximally re-activated before the offset point. For the lexicon that requires memory, the delay measure of point-neuron models is not informative because their scores are close to the chance level. The analysis indicates that immediate access to word memories is successful only if the identity of the early phonemes is sufficient to disambiguate the word and access the correct word memory, as it happens for the first two lexica. Otherwise, if the early phonemes contain insufficient information, the recognition will be, at best, based on the marginal distribution of the phoneme's identity over the lexicon. The probability that the first phoneme belongs to the correct word decreases with the increase in phonological overlap. In networks with dendrites, the reactivation of word memories relies on the tem-

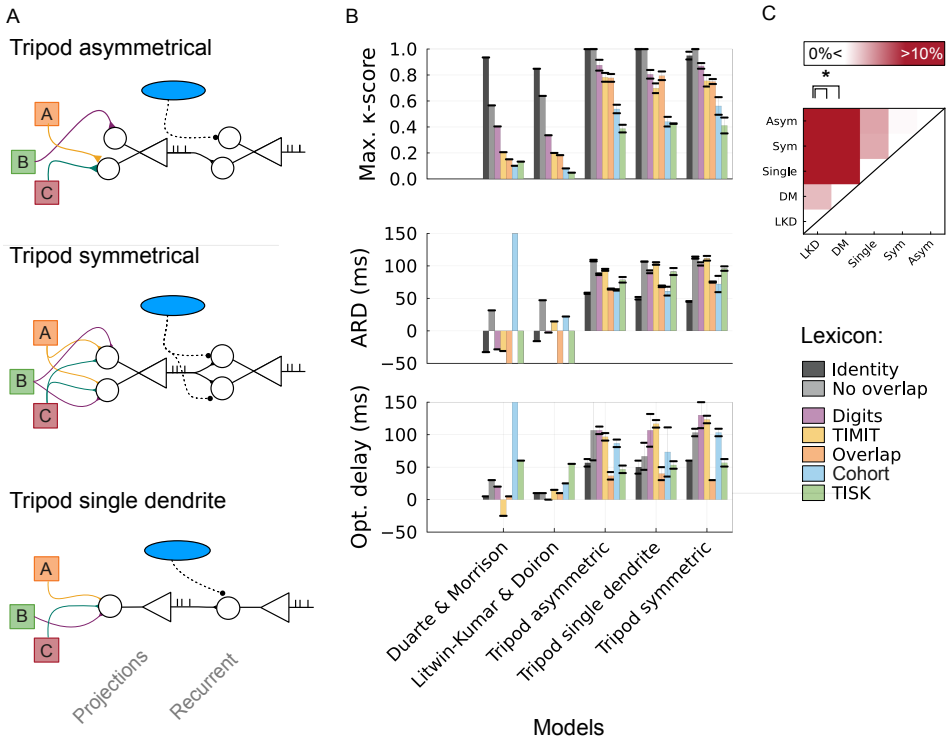



figure 4.4: **Comparison of point-neuron and networks with dendrites**
(**A**) Dendritic configurations and connectivity patterns of the Tripod-based network models. All three variants have dendrites whose length is uniformly random in the range of 150 µm to 400 µm. The asymmetric neuron has randomized pre-synaptic connections with density $\rho$ =0.2 that target only one of the two post-synaptic compartments; the length of the dendrites is also drawn independently from each other. The symmetric network receives the same pre-synaptic connections, external and recurrent, on both dendrites, with half the connection strength than the asymmetric case; synapses have the same density ($\rho$); the two dendrites have the same length in the symmetric neuron model. Finally, the single dendrite neuron has only one segregated compartment, which receives all synaptic connections.
(**B**) The three panels illustrate the maximal $\kappa$-score - across all delays, the average recognition delay (ARD), and their optimal delay (delay with maximum score) for the five network models and the seven lexica tested. Grayscale bars show the two memoryless lexica and colored bars show the five lexica that require temporal integration of phonemes for recognition. point-neuron models perform best on memoryless lexica with negative delays; it indicates that recognition is achieved before the offset point. Dendritic models outperform them on lexica with overlapping phonemes and reach high accuracy with a delay between 50 ms to 100 ms from the uniqueness point. The dendritic models' average and standard deviations refer to a sample of three independent initializations per model. (**C**) The matrix shows the difference in accuracy between a model on the y-axis and a model on the x-axis, with a resolution of five percentage points, for the five lexica that require memory for word recognition.

poral integration of the inputs, confirmed by the large values (50 ms to 150 ms) of average and optimal delays. Phoneme activity is carried forward in the dendrite's membrane potential and the word population fully activates when the last disambiguation piece of the input sequence is presented at the uniqueness point.

Concerning the role of asymmetry in the network, the present results indicate that its contribution is marginal. Averaged over three independent samples, the asymmetric Tripod model scored higher than the other two dendritic models, but the differences are less than 5 to 10% and not significant Fig.4.4*C*. The delay measures of the three dendritic models resemble each other in offset and uniqueness point measures. The models' similarities indicate that dendritic memory carries information over time independently of the connectivity and dendritic configurations tested (Fig.4.4*A*). In addition, the difference in accuracy among the three models does not depend on the lexicon; they all decrease their recognition accuracy for the lexicon with larger phonological overlap, suggesting neither asymmetry nor the number of dendrites provides additional mechanisms to cope with the increasing phonological overlap. Finally, the lack of significant differences between the two point-neuron models indicates the presence of two classes of inhibitory neurons with distinct inhibitory plasticity rules ($I_1$, $I_2$) is not the core mechanism for the word recognition task at hand.

From the comparison between the five models, we deduce that the dendritic memory endowed by the NMDARs in segregated compartments is the mechanism that allows the network to solve the word recognition task. However, because sequence detection requires the interaction of short-term and long-term memories, it is not yet clear if the point-neuron models fail to establish the hetero-associative connections or, rather, cannot perform temporal integration. In the following section, we analyze the network connectivity emerging from the associative phase in the five models and show that only networks with dendrites form strong associations between phonemes and word assemblies.

### 4.2.2 Network configuration emerging from Hebbian plasticity

#### Network with dendrites have strong connections from phonemes to words

The present section investigates the network structure that supports sequence detection. To perform the network analysis, we reduce the model's connectivity to an effective matrix that accounts only for words and phonemes assem-

blies. We define the effective matrix ($\mathscr{C}$) as the average connectivity between the cells belonging to two assemblies (Methods). The effective matrix is a non-symmetric square matrix computed from the learned connectivity matrix at the end of the associative phase. The $\mathscr{C}$ is composed of four blocks, the connections between word assemblies ($\mathscr{C}^{W\to W}$), between phoneme assemblies ($\mathscr{C}^{P\to P}$), from phonemes to words ($\mathscr{C}^{P\to W}$), and from words to phonemes ($\mathscr{C}^{W\to P}$). An example of the $\mathscr{C}$ matrix from the asymmetric Tripod model for the *Overlap* lexicon is shown in Fig.4.5*A*. The top-left quadrant shows the feedback connections from word assemblies to phoneme assemblies, and the bottom-right quadrant shows the feedforward connections from phonemes to words. The top-right and bottom-left quadrants show the connectivity within phoneme and word assemblies, with strong recurrent connections on the diagonal. For completeness, the effective connectivity matrix for the remaining models is shown in SI Fig.4.4 .

The four connection types of the effective matrix contribute differently to the word recognition task. To test their impact, we selectively removed each of the four blocks from $\mathscr{C}$. The resulting effective matrices are similar to the original Fig.4.5*A*, except for the removed block, whose elements are set to the average connection strength (SI Fig.4.5 ). Thus, we measured the recognition score and calculated the percentage of recognition loss in each of the four conditions. The recognition loss is computed as $\mathscr{L} = \frac{\kappa'-\kappa}{\kappa}$, where $\kappa'$ is the score with the modified matrix and $\kappa$ is the baseline score. The loss is computed for each of the seven lexica and then averaged. The results of the ablation experiments (Fig.4.5*B*) indicate that the phonemes to word connections ($\mathscr{C}^{P\to W}$) are pivotal for word recognition, recognition drops of the 100% when they are removed. Such connections are necessary because they relay externally-driven activity in the phoneme populations to the word assemblies. They emerge in the associative phase when phoneme and word populations are simultaneously activated. Interestingly, the phoneme-to-word connections for the *Overlap* lexicon have a similar, although weaker, structure in the DM network but are absent in the LKD network (SI Fig.4.4 ). The word-to-word connections are also important but not strictly necessary; in this case, the recognition drops of 25%. The other two blocks in Fig.4.5*A* involve synapses targeting phoneme populations. Their contribution to word recognition is marginal compared to the others, although the loss is significantly larger than zero. These connections have a smaller impact on the network dynamics because they are weaker than the external projections that stimulate the phonemes populations.

Based on the outcome of the previous experiment, we inspected the effectivity connectivity matrix of the five models (point-neuron and network with dendrites) to determine if the recognition capacity of the models was due to differences in $\mathscr{C}$. To this aim, we calculated which phoneme and word populations should be associated for word recognition to succeed. We thus define a group of *lexical* connections as the connections from any phoneme to the words that contain it and vice versa, plus all recurrent connections within phoneme and word assemblies. The three groups that compose the lexical connections are illustrated in Fig.4.5*C*. We measured the average strength of the four types of lexical connections for the Overlap lexicon for the five models tested. To facilitate model comparison, the columns were scaled between the maximum average synaptic connection of each model's $\mathscr{C}$ matrix (e.g., 26 pF in Fig.4.5*A*). The minimum corresponds to the initial synaptic strength of the excitatory to excitatory connections, which determines the synaptic weight budget for synaptic normalization. The initial weights are 10 pF for the dendritic models and vary for the two point-neuron models (DM: 0.45 pF, LKD: 2.76 pF). It was not possible to set the weights of the somatic models to be the same as the dendritic models because of instabilities in the network activity. We return to this issue in the Discussion.

Differences and similarities between the five models for the four lexical connection types are shown in the bottom panel of Fig.4.5*D*. For the dendritic models, the hetero-associative connections (phoneme-to-word) have strength comparable to auto-associative ones (words-to-words and phonemes-to-phonemes). In contrast, the point-neuron models have strong auto-associative synapses but weak connections binding phonemes to word assemblies. Interestingly, the recurrent phonemic connections are as strong as the word-to-word in the point-neuron and asymmetric models, but they are weaker for the symmetric and single dendrite models. All models have relatively weak word-to-phoneme connections. Overall, the analysis of the effective connectivity matrices indicates that the presence of dendrites promotes the strengthening of non-recurrent connections between phoneme and word assemblies. An overview of the effective connectivity matrices for the five models on all the lexica is presented in SI Fig.4.7 . Comparing the panels across different lexica indicates that the formation of hetero-associative connections is achievable in point-neuron models if the words have no phonological overlap. A comparative analysis of these four types of connections is also presented in Appendix B, which analyzes the learning dynamics

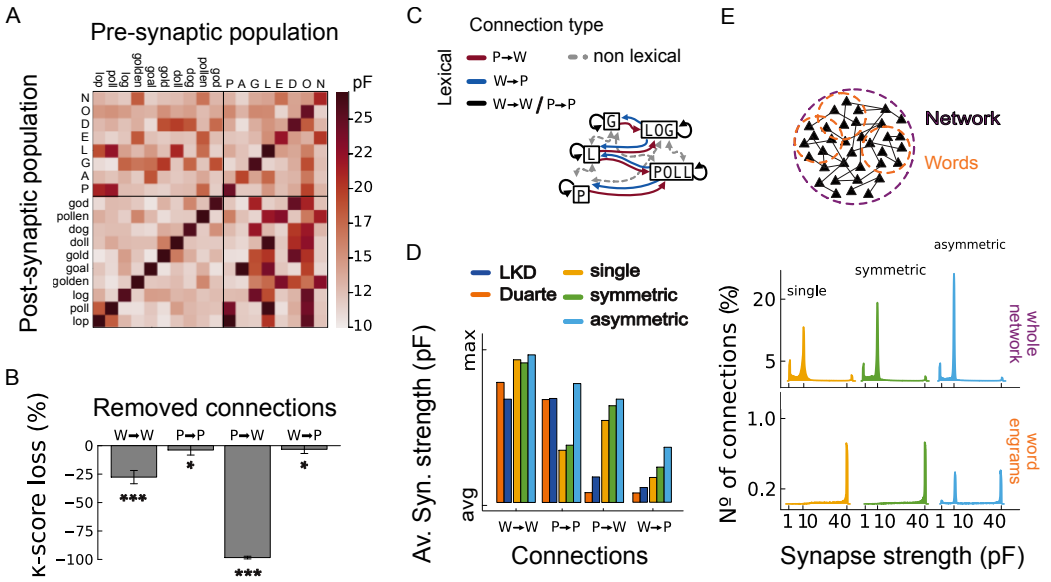



figure 4.5: **Feedforward structure in network with dendrites support word recognition**
(**A**) Effective connectivity matrix of the asymmetric network with dendrites for the *Overlap* lexicon. The heatmap shows the average synaptic strength between assemblies, measured in pF. The four quadrants show the recurrent connections within phoneme and word populations, the feedforward connections from phonemes to words, and the feedback connections from words to phonemes. (**B**) The word recognition $\kappa$-score changes when each of the four quadrants of the effective connectivity matrix is leveled to the initial synaptic weight (the resulting matrices are shown in SI Fig.4.5 ). The four bars indicate the loss in recognition in each condition. They are all significant. However, removing the phonemes to word connection causes a drop in recognition of the 100%. (**C**) The schematic at the top illustrates the four types of lexical connections and separates them from the non-lexical ones, which are not expected to contribute to word recognition. The non-lexical connections bind phonemes with words that do not contain them or those among different words. (**D**) Average synaptic strength of the four types of lexical connections for the point-neuron and dendritic models. All models have strong connections within word assemblies, but only dendritic models have strong phoneme-to-word connections. (**E**) Distribution of synaptic strength over the entire network (top) and the word recurrent connections (bottom). The top panel shows that most synapses remain close to their initial strength of 10 pF, less than 3% of the connections developed stronger synapses (40 pF). The bottom panel (zoom in the range 0 to 1% of synaptic connections) indicates that for the symmetric and single dendritic models, the word assemblies have only strong recurrent connections. In contrast, in the asymmetric network, recurrent word engrams have both weak and strong synapses within the engram.

of the five models during the associative phase. the results are fully consistent with the ablation study presented here.

Further differences between the dendritic models can be gleaned from an analysis of the synaptic weights landscape of the three networks. The histograms in Fig.4.5*E* show the percentage of connections (y-axis) against their respective synaptic strengths (x-axis) for the entire network (top) and within the recurrent word engrams (bottom, grey shades show the entire network for comparison). The top row indicates that the dendritic models reach similar network configurations through Hebbian learning. The distributions of synaptic strength each have three peaks at 1.78, 10, and 40 pF. The two extrema are the minimum and maximum synaptic strengths allowed in the plasticity rule (Methods), and the mid-value is the synaptic strength with which the model is initialized. Synaptic normalization (i.e., homeostasis) enforces that if a compartment develops a strong incoming synapse, some other synapses will get weaker; one fully potentiated synapse requires four fully depressed ones. Thus, homeostasis explains the leftmost peak as a consequence of the strengthening of lexical connections. The most pronounced peak occurs at 10 pF. These synapses remain unchanged by learning and do not contribute to any engram. The asymmetric Tripod network has roughly twice as many idle connections as the single dendrite model and 40 % more than the symmetric one. The difference is visible in the lower panel; for the single and symmetric Tripod models, most of the recurrent word connections are fully potentiated, whereas only half are for the asymmetric model. Comparing the panels in Fig.4.5*D* and *E* with the effective connectivity of the symmetric and single dendrite network SI Fig.4.6 , we infer that the asymmetric configuration recruits fewer synaptic resources for the recurrent word-to-word connections. The STDP redistributes the synaptic budget among the other connection types. The larger availability of synaptic resources explains why the asymmetric model has phoneme-to-phoneme and word-to-phoneme synapses whose strength is comparable to the word-to-word connections, which is not the case in the other two dendritic models.

### Structures in the effective connectivity matrix mediate recognition of words with phonological overlap

Because the phoneme stimuli do not encode temporal information, i.e., the position of the phoneme in the word, the network must rely on recurrent connections to process sequential order and reactivate the correct word assemblies. For example, the words *dog* and *god* share the same phonemes (lexicon *Over-*

*lap*), but their order is different. When presented with the phoneme sequence *D,O,G*, the word assemblies *dog* and *god* should assume different states such that sequential order information triggers the reactivation of *dog*. Because hetero-associative synapses are prominent in mediating the activation of word assemblies, we looked for the serial order mechanism in this set of connections. We notice that the phonemes-to-word connections show variability in their synaptic strength; the variability is visible in the rows of the lower-right block of Fig.4.5*A*. Following the theoretical results by Knoblauch and Pulvermüller (2005), we hypothesized that the different strengths encode the order of each pre-lexical unit and support sequence recognition capacity in the model. Thus, we now investigate the architecture of the synaptic weights in the phonemes-to-word connections and whether it also contributes to the word recognition capacities of the networks with dendrites.

To test the hypothesis that average synaptic weights encode the phoneme serial position, we assigned an index to each phoneme, matching its serial position in the word, and determined the average synaptic connection between the phoneme and the word assemblies. For this analysis we considered the five lexica requiring memory. Because the weights of the phoneme-to-word connections vary across words and even more across the lexica, we normalized the synaptic weights by the strongest phoneme-to-word connection in each word. Similarly, because words have different lengths, we centered all the samples to their uniqueness points. The schematic on top of Fig.4.6*A* illustrates the serial position of the phonemes relative to the UP of the word *golden* from the *Overlap* lexicon. Here, the phoneme *E* corresponds to the uniqueness point, and the other phonemes are ordered according to it. The lower panel shows the average synaptic weight of each phoneme-to-word connection (y-axis) corresponding to the phoneme position in the word (x-axis). The minimum on the y-axis corresponds to the initial synaptic strength of the connections ($W_0$).

The plot in Fig.4.6*A* shows significant differences among the three phonemes before the UP. The strength of the projections from phoneme to word increases with the serial position of the phoneme in the word: assemblies associated with early phonemes have weaker connections than those closer to the uniqueness point. In the present work, we have not investigated the dynamics leading to the formation of the weight architecture in Fig.4.6*A*. However, in the following, we show that this structure in the synaptic weights is necessary for recognizing words with phonological overlap, and it is not a spurious property of the network.

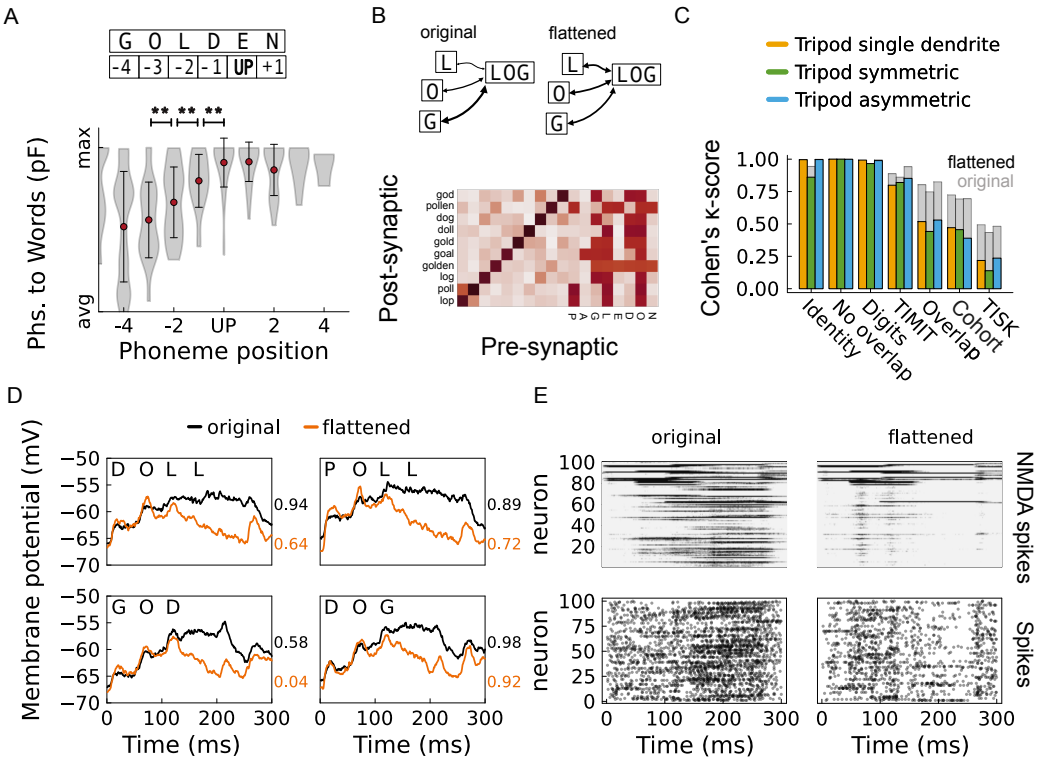



figure 4.6: **Feedforward structure in network with dendrites support word recognition**
(**A**) Strength of the phoneme-to-word connections organized by the serial order of phonemes in the word. The phoneme position is computed relative to the uniqueness point (top panel). The graph shows data pooled from all words in the five lexica that require memory. Red dots indicate averages, while the violin plots account for the entire sample. Significative differences in the average synaptic weights are evidenced between the three phonemes preceding the uniqueness point. (**B**) Lower quadrants of the connectivity matrix with flattened structure in the phoneme-to-word connections. The matrix shows the modified phonemes-to-word connections, the original being in panel Fig.4.5*A*. Flattened connections maintain the associations between phonemes and words but remove the internal structure necessary for distinguishing phonologically overlapping words. (**C**) Word recognition accuracy in networks with flattened connections, compared with the original connectivity matrix (grey). The $\kappa$-score decreases for the dendritic models when lexica have a large phonological overlap. (**D**) Average dendritic membrane potential with the original and the flattened connectivity. The membrane potential is portrayed for four distinct word assemblies. The black line indicates the normal dendritic dynamics, and the orange is the one in the flattened condition. The four plots show that in the flattened condition, the cell assemblies have a weaker sustained depolarization after word offset. (**E**) Assembly dynamics for the word *doll*, also comparing the original and the flattened network. The upper panels show the dendritic membrane potential, with black traits indicating potential above 20 mV (NMDA spikes). The lower panels show the somatic firing activity. Both the membrane potential and the firing activity are pooled over 10 samples of the sequence *D, O, L, L*. The panels show that the lack of sustained depolarization is due to fewer NMDA spikes and results in weaker somatic activity.

To test whether the increase of synaptic strength with serial order contributes to word recognition, we manipulated the network weights. We flattened the phoneme-to-words connections such that the curve in Fig.4.6*A* would appear flat (Methods). The panels in Fig.4.6*B* show a schematic of the transformation (top) and the resulting effective connectivity matrix (bottom) for the asymmetric Tripod network tested on the *Overlap* lexicon. For each post-synaptic word (y-axis), the strength of all incoming connections from the phoneme assemblies (x-axis) was identical. We then tested recognition accuracy as before and compared the two conditions (Fig.4.6*C*). We found that the structure of phoneme-to-word connections matters for the lexical items that require memory but not for the other two lexical items. Crucially, the *Digits* and *TIMIT* lexicon require memory, but they are less affected by the flattening of connections. The explanation is that if the words in the lexicon are distinguishable by the specific combination of phonemes, they do not rely on the serial order. The weight structure does matter when the phonological overlap is large, and the serial order is necessary to distinguish among words. Recognition accuracy drops for all dendritic models, losing 40 % to 60 % for the three lexical items with the highest degree of phonological overlap (i.e., *Overlap, Cohort, TISK*).

Further insights into the role of weight differences are obtained by comparing the average dendritic potential of words' assemblies in the original and flattened conditions. The four panels in Fig.4.6*D* portray four words of the *Overlap* lexicon (*doll, poll, god,* and *dog*). The black line refers to the original model, and the orange line refers to the one with flattened phonemes to word connectivity; the values indicate the recognition score of the specific word. A zoom-in in the assembly dynamics is offered for the word *doll*. Fig.4.6*E* shows the membrane potential (top) and somatic activity (bottom) for all the neurons in the assembly in both the original network and flattened conditions; the traces are obtained averaging 10 presentations of the word. In the upper panels, the color scale is chosen such that only the NMDA spikes are visible (black reveals dendritic potential larger than 20 mV). Taken together, Fig.4.6*D* and *E* indicate that tampering with the effective connectivity matrix affects the average membrane dynamics of the word's assembly. In the flattened condition, the membrane potential increases upon the presentation of early phonemes: between 0 ms to 100 ms, the assembly activity resembles the one in the original network. Then, at the presentation of the third phoneme (100 ms to 150 ms), the assembly fails to ignite the NMDA spikes that trigger the full reactivation. The result is that the membrane potential decays faster (orange traces), and the word population has sparser fir-

ing activity. Crucially, this does not entail that the word cannot be recognized at all, nor that all the words are equally affected, as evidenced by the word scores in Fig.4.6*H*. Rather, the weight structure offers an additional network mechanism that supports the distinction of words with full phonological overlap.

The present results illuminate pivotal role of dendrites in forming a structured and functional connectivity matrix. Models with dendrites support the formation of hetero-associative connections (phonemes to words), which are necessary for word recognition. The differences in the connection weights, which encode both the identity and order of the phonemes, are read by the activity of the phonemes assemblies and maintained in the dendritic memory of the word's population. We now shift the focus from network structure to the short-term memory mechanism and show that it results from the interplay of non-linear dendritic integration and dendritic inhibition.

### 4.2.3 Dendritic memory and inhibitory control

#### Dendritic non-linearity and inhibition governs temporal integration

The results have shown that networks of dendritic neurons are better at achieving word recognition than networks of point neurons. The sequence detection capacity follows from delayed temporal integration and feedforward connectivity between phonemes and word assemblies. We evidenced how dendritic memory in segregated compartments is the mechanism supporting the computation in networks with dendrites. However, to activate and maintain the dendritic memory, the cells must be at a specific operational point that allows the expression of NMDA spikes. In addition, to avoid encoding the wrong memory, the network has to control the dendritic non-linearity and suppress NMDA spikes upon spurious activity in the assemblies. The voltage dependency in the NMDARs and tight dendritic inhibition permit such a computational state. We now show that the interactions of these mechanisms are necessary for the network model to achieve word recognition.

In the Tripod network, the dendritic non-linearity is governed by the ratio of NMDA-to-AMPA peak-conductances of the glutamatergic receptors (NAR) and the decay timescale of the NMDAR receptors ($\tau_d$). Fig.4.7*A* illustrates the impact of variations in the NMDA receptor on the dendritic membrane potential of the post-synaptic cell. The panels show the excitatory post-synaptic potential (EPSP) of a 300 µm dendrite following an excitatory spike on a synapse with weight 100 nF. In the left panel, the NAR varies between 0 and 2.7; in the right

panel, the decay timescale $\tau_d$ ranges from 2 ms to 50 ms. In both cases, the non-varied parameter maintains the baseline value of the model, NAR $= 1.8$ and $\tau_d = 34$ ms. The dashed lines portray this couple of parameters. The curves in the left panel show a NAR threshold for the onset of the non-linear response, approximately at NAR $= 1.2$. Above this NAR, the dendritic membrane enters a long-lasting depolarized state called *plateau potential*. These plateau potentials can be viewed as a form of intra-cellular dendritic memory on short timescales, and the conditions that elicit such states in the Tripod neuron have been described in Quaresima et al. (2022). The right panel in Fig.4.7*A* shows that the duration of plateau potentials was proportional to the timescale of the NMDA decay, $\tau_d$.

Inhibitory control is determined by the potential $V_0$ that sets the target value for the voltage-dependent iSTDP rule on the dendrites of $I_2$neurons. To determine the role of inhibitory plasticity v-iSTDP on synaptic learning, we varied the dendritic target voltage $V_0$ in the range of $-90$ mV to $-40$ mV. The histograms in Fig.4.7*B* show the distribution of the post-synaptic weights from $I_2$ neurons onto the dendritic compartment after the association phase. For a target value of $V_0 = 70$ mV or lower, the synaptic weights accumulated at the maximum value 243 pF that we allowed for inhibitory synapses. When the target $V_0$ was set higher, inhibitory synapses were potentiated less during the association phase, and the weight distributions flattened towards smaller values.

Changes in the dendritic non-linearity and inhibitory control do not imply that the network cannot process the phonemic stimuli correctly. Indeed, because the connectivity matrix remains unchanged, it may still drive correct assembly reactivation against weaker non-linearity or higher dendritic noise. In the present study, the structure of the excitatory connections is frozen to the one obtained from the associative phase. Thus, it could be that the parameter changes do not affect the task performance but only change the network dynamics. We measured the word recognition performance on all the parameter ranges to test this hypothesis. The three panels in Fig.4.7*D* illustrate the performance loss corresponding to networks simulated with each parameter variation. The score obtained for baseline values is used as the comparison term, and the loss is the difference between the score in the test condition and the score achieved in the baseline ($\kappa$-score $\approx 0.85$). From the comparison of the three parameters swaps, it results that the inhibitory control plays the most significative role, with performance dropping of 60 % when the $V_0$ is increased to $-55$ mV. To match the

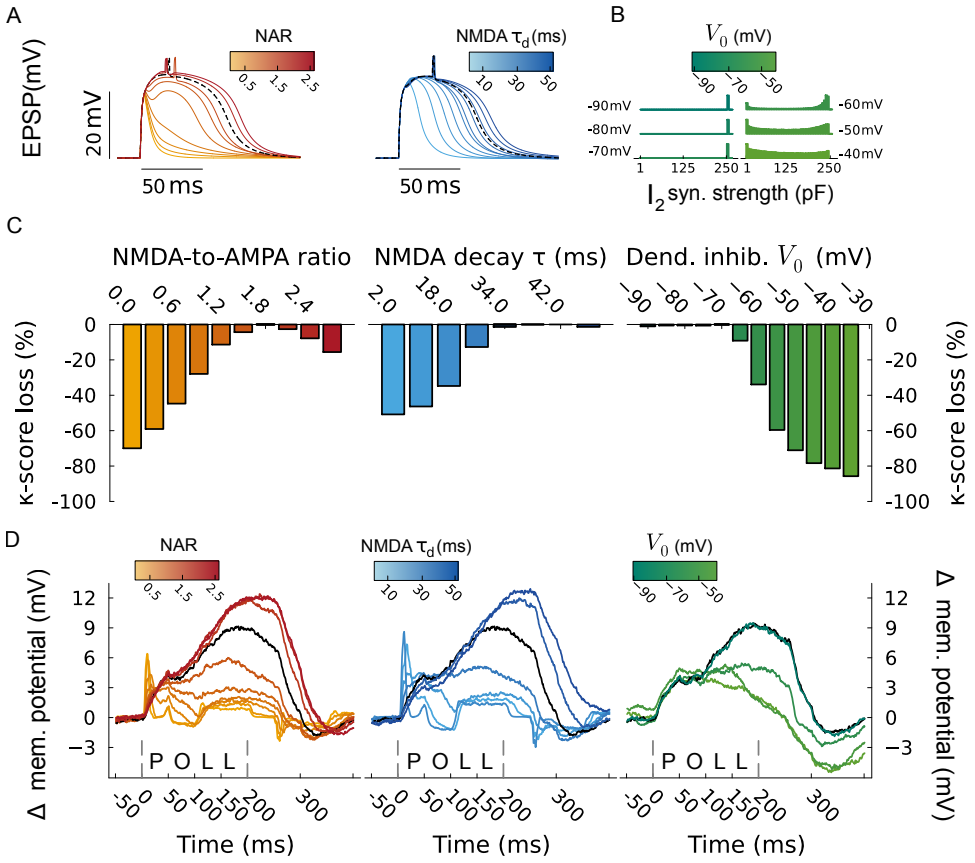



figure 4.7: **Dendritic non-linearity and inhibitory control enable word recognition**

(**A**) EPSP of the dendritic membrane in Tripod neurons after stimulation of a single synapse. Curves show the membrane potential for varying NARs and NMDA decay timescales. The dashed line indicates the model that was used in the previous section, corresponding to a Tripod neuron with human synapses, with NAR $= 1.8$ and $\tau_d = 35$ ms. (**B**) Histograms of the inhibitory synaptic strength onto dendrites, measured at the end of the recall phase, for increasing v-iSTDP target potentials $V_0$. Synapses potentiate towards their maximum value for $V_0$ lower than $-70$ mV. (**C**) Loss of $\kappa$-score for variations in the parameters NAR, $\tau_d$, and $V_0$. The bars with the darker shade correspond to the original model. Accuracy reduces up to 60 % when the dendritic non-linearity is absent (NAR $< 0.6$) or short ($\tau_d < 18$ ms). The changes are larger for increased values of the dendritic inhibition target potential ($V_0$). (**D**) Difference in membrane potentials upon presenting the phonemes */P, O, L, L*. The plots express the difference between the average dendritic potential of the assembly associated with the word *poll* and the dendritic potential of the entire network. The difference in membrane potential is measured for all the parameter variations, i.e., NAR, $\tau_d$, $V_0$. The black solid line is the same for the three panels and refers to the baseline asymmetric Tripod model.

same loss, the NAR has to diminish to 0 to 0.6 and the timescale to 2 ms to 10 ms, which corresponds to absent plateau potentials.

Variations in the NMDA non-linearity and the strength of inhibitory control can also be traced in network activity. To investigate changes in the dynamics, we compared the average dendritic membrane potential of a word population with the average in the rest of the network while presenting the associated phoneme sequence. The difference between the control and the dendritic potential of the target population (*poll*) is shown in Fig.4.7*C* for the variations in the NMDARs and dendritic inhibition parameters. The left and middle panels show similar trends in that the difference in membrane potential stays close to zero for weak and short dendritic non-linearity. The activity of the target word assembly peaks after the first phoneme (*P*) but then returns to that of the control condition. On the other hand, if the non-linearity is parameterized as in the baseline model (black line), the dendritic membrane potential increases throughout the presentation of the phoneme sequence until it reaches a peak around the word offset and. A different pattern can be observed when the target potential $V_0$ of v-iSTDP is varied. The rightmost panel in Fig.4.7*C* shows that the ERP is wider when dendritic inhibition is strong, although weaker inhibition also generates a substantial ERP relative to control. When $V_0$ is equal to $-70$ mV or lower, the difference in membrane potential is nearly identical because the $I_2$ inhibitory synapses that have developed during the association phase are similar (see Fig.4.7*B*). The changes in the network activity expressed by the difference in membrane potential (Fig.4.7*D*) match the $\kappa$-score loss of the word recognition task match (Fig.4.7*C*). This indicates that the dynamics portrayed in Fig.4.7*D* are causal for the task.

Crucially, in the latter case, the difference in membrane potential has a direction that seems counter-intuitive regarding the inhibitory function; more potent dendritic inhibition should cause weaker dendritic depolarizations. Hence, we inspected the average membrane potential for the word assembly and control condition to shed light on the ongoing dynamics. The four panels in SI Fig.4.9 show the soma and dendritic average potentials and solve the paradox of the role of v-iSTDP control. When the inhibitory target potential is larger, e.g., $V_0 = -50$ mV, the network dendrites are permanently depolarized, and the difference between the target population activity and the control condition is less significant. Thus, the dendritic inhibition acts as a signal-to-noise control mechanism that allows the assembly to reactivate only upon the correct stimulus, maintaining the cells hyperpolarized otherwise.

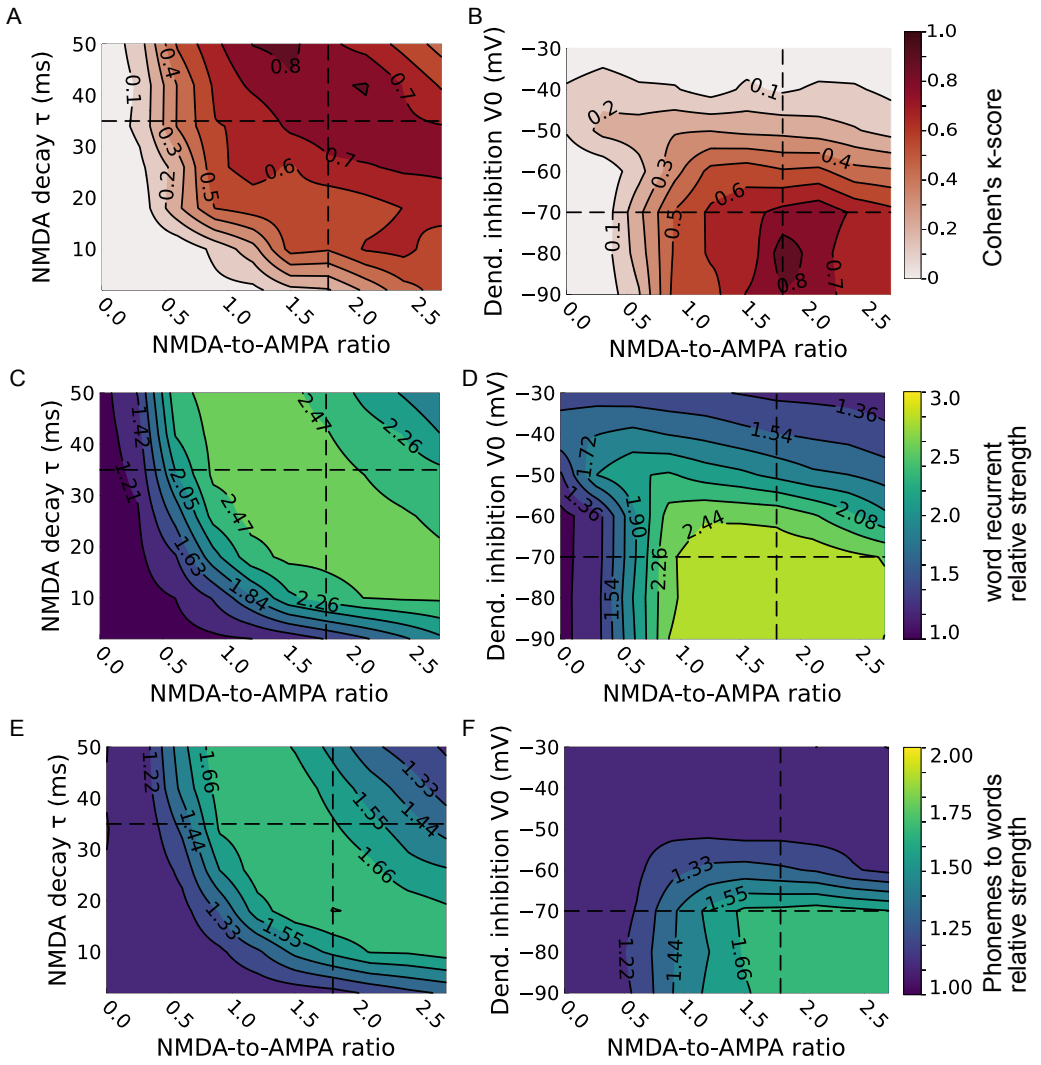



figure 4.8: **Dendritic non-linearity and inhibitory control enable the formation of phonemes to word connections**
(**A**, **B**) Cohen's $\kappa$-score of the Tripod asymmetric model for pairs of NMDA-to-AMPA ratio (NAR) and $\tau_d$ and for pairs of NARs and $V_0$. The model's parameters are varied in both the associative and recall phases. The dashed lines indicate the baseline model. The left panel indicates that the $\kappa$-score is stable if changes in the NMDAR's timescale are compensated with the NAR. Conversely, the interaction between the NAR and $V_0$ variables is weaker (right panel); the performances decay linearly on both axes. (**C**, **D**) Strength of the recurrent word connections for the same parameter swap. Connections' strength is divided by the average synaptic weight (10 pF) to evidence the formation of engrams. The contour plot of the words-to-words connection mirrors the $\kappa$-score panels, indicating that the changes in the dendritic non-linearity and inhibitory control also limit the formation of hetero-associative connections. (**E**, **F**) As in previous panels, for the phonemes to word connections. The hetero-associative connections respond similarly to the auto-associative for variations in the dendritic non-linearity (panel *E*). Conversely, weaker dendritic inhibition is more detrimental for the phonemes-to-word connections than for the recurrent ones.

### 4.2.4 Dendritic network features are necessary to develop sequence detection networks

We reasoned that the loss in recognition could be caused by testing recall with parameters different from those used in the associative phase, hence the mismatch. Thus, to exclude this possibility, we modulated dendritic non-linearity and inhibitory control during the associative phase and again tested word recognition. The panels in Fig.4.8*A, B* show the $\kappa$-score for a range of NARs, $\tau_s$, and $V_0$ and indicate that enabling STDP plasticity does not solve the issue. Recognition drops for NAR below 1.0, for timescales shorter than 10 ms, and inhibitory target potential ($V_0$) larger than $-60$ mV. In addition, there is an interaction between the NMDAR timescale and synaptic efficacy (NAR); for weaker non-linearity, longer timescales are beneficial. Rather than stabilizing, changes in the dendritic non-linearity and inhibitory control that occur during the associative phase also affect the formation of recurrent and hetero-associative connections. The average strength of the recurrent connections diminishes for most of the parameter variations and so does for the phonemes-to-word, in both cases the trends are similar to the recognition score (Fig.4.8*C, D*). These latter results indicate that both dendritic non-linearity and inhibitory control are necessary for the network with dendrites to form assemblies. In the case of the dendritic linearity, this can be explained by the high threshold for synaptic strengthening in the STDP. The threshold is at $-20$ mV and if the dendrites are not sufficiently depolarized they will not enter the potentiation range. For the inhibitory control, the mechanism is more subtle. The strong external stimuli will override any network activity. Consistently with this observation, recurrent associations form despite weaker inhibitory control (Fig.4.8*D*). Conversely, the hetero-associative connections seem to require a high signal-to-noise ratio in the assemblies, and when inhibitory control weakens, this class of connections rapidly fades (Fig.4.8*F*).

To conclude, we have shown that dendritic non-linear excitability is necessary for the network's word recognition capacity. The contribution of dendritic inhibition is also crucial because it governs dendritic depolarization and thus reduces the possibility that NMDA spikes originate from the intrinsic fluctuations in the network activity. The analysis suggests that NMDA spikes and tight inhibitory control are necessary during both the recall and associative phases. In the former, they contribute to forming the phonemes-to-word connections; in the latter, they allow for the temporal integration that is necessary for word recognition.

## 4.3 Discussion

The present work investigated the formation and reactivation of Hebbian cell assemblies by stimuli with temporal structure. To this scope, we propose a novel, biologically constrained, spiking neural network model with dendrites and plastic excitatory connections. We evaluated the model in a phonemes-word association task and compared it with control networks with and without dendrites. The networks received sparse and overlapping excitatory projections, representing words and phonemes. The task was to activate the correct word assembly following the phonemic sequence. We demonstrated that the introduction of segregated dendritic compartments endowed the network with the capability to recognize sequences of inputs with overlapping features. In dendritic models, paired stimulation and STDP organized the networks in auto-associative and hetero-associative engrams. Word memories were then correctly recollected in vocabularies with partial or complete phonological overlap (as in the words *dog* and *god*). In contrast, the point-neuron models tested failed in the word recognition task when the vocabulary contained words with shared phonemes and did not form hetero-associative connections. In the following, we first discuss the significance of the present result for theories of biological memory, then we outline the contribution of dendritic non-linearity and inhibition in memory access.

#### Recognizing word memories with phonological overlap

The perception and recognition of phonological sequences is a fundamental human cognitive capacity (Dehaene, Meyniel, Wacongne, Wang, & Pallier, 2015; Povel & Essens, 1985). The recollection of sequence memories entails that the neural processes transform temporal patterns of activity into spatially-coded and time-compressed representations of the stimulus (Bagur et al., 2022; Chan et al., 2014; Fox, Leonard, Sjerps, & Chang, 2020; Vaz et al., 2020). However, how such a computation is carried out in biological networks remains poorly understood. First, it is not clear how word memories are stored in the neural substrate through Hebbian plasticity (Poeppel & Idsardi, 2022). Second, it is unknown how the networks integrate the short-term memory of the acoustic stimuli with the long-term memory of the lexical item to be accessed (Norris, 2017).

Concerning the storage of long-term memories, the dominant hypothesis is that consolidated memories are maintained in the strong recurrent connections of cell assemblies (Amit, 1995; Fuster, 1997; Poo et al., 2016; Pulvermüller, 1999). Cell assemblies and synaptic engrams successfully account for associa-

tive memories and their principles have been implemented in several computational models. Some of these studies have shown that associative plasticity, such as spike-time dependent plasticity (STDP), supports the acquisition of long-term synaptic memories and their recollection (Garagnani et al., 2009; Litwin-Kumar & Doiron, 2014; Tomasello et al., 2018; Zenke et al., 2015). However, these forms of associative memories seems to fall short in encoding relationships (Gallistel, 2021), such as the sequential order of phonemes in word memories. Our Tripod model shows that introducing dendritic compartments resolves this conundrum and supports the formation of order-sensitive memories.

The Tripod network model acquired word memories via STDP and maintained the memories in the hetero-associative connections between phonemes and word assemblies. The network memories were sensitive to both the identity and the order of the phonemes presented in the input. Information about the order was stored in the synaptic connections, and the strength depended on the serial position of the phoneme in the word. Weights were larger for phonemes closer to the uniqueness point. A similar mechanism for sequence detection in cell assemblies was already proposed by Knoblauch and Pulvermüller (2005). In contrast, the two point-neuron models investigated did not form hetero-associative connections for vocabularies with phonological overlap, let alone the sequence detection synaptic architecture. Remarkably, the DM model from Duarte and Morrison (2019) did better than the LKD model (Litwin-Kumar & Doiron, 2014). The main difference between the two models was related to the presence of longer timescales in the somatic NMDA receptors in the DM model. Similar to the Tripod network, the presence of slow-decaying depolarizing currents contributes significantly to the formation of the phonemes-to-words connections. The fact that these connections form in the case of memory-less vocabularies rules out that the models fail because of a shortage of synaptic budget in the face of the homeostatic mechanism.

**Transitory stimuli are integrated over time through dendritic memory**

The presence of hetero-associative connections is, however, not sufficient for recollecting the word memories. The model requires that the dendritic compartments express NMDA spikes. The nonlinearity in the segregated dendrites of the Tripod neuron mediates the interaction of short (dendritic) and long (synaptic) memories. Upon the presentation of external phonemic stimuli, the synaptic variables are read and encoded into the dendritic plateau potential of single cells. The integration of successive pieces of information is expressed in the slow

build-up of the assembly dendritic membrane potential. The dendritic memory allowed for the integration of these sources of information over the word's timescale. Crucially, the synaptic and dendritic memories have different computational trade-offs in terms of stability, robustness to noise, and duration of their transients (Chaudhuri & Fiete, 2016). The dendritic memory is encoded within a few tens of milliseconds and erased within hundreds while the engrams form over tens of seconds and remain stable over time (Appendix B). While stable memories have been shown to form in point-neuron models Litwin-Kumar and Doiron (2014), the novelty of our results is that also hetero-associative connections are stable and can re-activate overlapping memories.

To understand the relevance of the present contribution we must clarify how the word recognition task presented distinguishes from other instances of sequence learning. Sequence memories can also be instantiated as a chain of neural assemblies that activate in fixed order (Almeida-Filho et al., 2014, Hebb's phase sequences). This neural phenomenon is observed during hippocampal replay and in bird song (Buzsáki, 2010) and can be reproduced in classical spiking network models (Clopath et al., 2010; Fiete et al., 2010; Gillett, Pereira, & Brunel, 2020; Gjorgjieva, Clopath, Audet, & Pfister, 2011; Haga & Fukai, 2018; Maes et al., 2020; Rajan, Harvey, & Tank, 2016; Reifenstein, Bin Khalid, & Kempter, 2021; Riquelme, Hemberger, Laurent, & Gjorgjieva, 2023). In most cases, these models leverage hetero-associative connections between the chained assemblies. Nonetheless, the sequences investigated commonly include items that do not repeat. When there is overlap among the items, the model expresses additional features that allow distinguishing the identity of the sequence presented. For example, the network in Cone and Shouval (2021) implements an external reservoir that depends on the identity of the sequence, while the network in Maes et al. (2020) leverages an external neural clock.

Two elements set our model apart from these previous studies. First, the word memory is accessed at the word's uniqueness point rather than at the end of the sequence. Early word memory access, at the uniqueness point, indicates that the memories are immediately retrieved when sufficient information is available in the input. Second, the sequence overlaps for long intervals, which impose temporal integration on the scale of a hundred milliseconds. Few studies have directly investigated whether STDP is sufficient for learning and retrieving word-like sequences. Among these, Duarte and Morrison (2014) investigated a point-neuron model with biological constraints and reported no net contribution of glutamatergic plasticity in sequence detection. Similarly, our attempts to

implement sequence detection in the two control point neuron networks were not successful.

A few more recent studies have included dendritic computations in network models with weak physiological constraints (Bouhadjar, Wouters, Diesmann, & Tetzlaff, 2022; Hawkins & Ahmad, 2016; Leugering, Nieters, & Pipa, 2023). Similarly to our model, these networks leverage plateau potentials to implement sequence detection. However, they present substantial abstractions in the dynamics of dendritic integration and synaptic plasticity. It is now well-known that dendritic processes determine the transfer function of single cells (I. S. Jones & Kording, 2021; Koch, 1998; London & Häusser, 2005; Payeur et al., 2019; Poirazi et al., 2003; Poirazi & Papoutsi, 2020; Ujfalussy et al., 2018). Their contribution is remarkable in synaptic clustering, processing memory, expanded memory storage, signal filtering and processing (Cazé & Stimberg, 2021; London & Häusser, 2005; Mel, 1992; Poirazi et al., 2003; Quaresima et al., 2022; Spruston, 2008). However, it is not yet well understood how these computational primitives interact with the network dynamics and whether explicit modeling of dendritic processes can enrich the computational capabilities of neural networks (Larkum, 2022). Thus, the present work contributes to clarifying the computational implications of considering dendrites in realistic biological networks. Our study implements a reduced three-compartment dendritic model (the Tripod neuron, Quaresima et al., 2022), and studies the effect of the dendritic computations in networks. By comparing dendritic models with one single compartment and with symmetric synaptic connections, we deduced that dendritic memory is the most important ingredient of the Tripod neuron for the task at hand. Dendritic memory is retained in the long-lasting depolarization (plateau potential) following the onset of NMDA spikes. In our network model, dendritic memory is strictly necessary for word recognition in the recall phase and contributes to the formation of phonemes-to-word connections in the associative phase. We have not tested if the same computations could be achieved leveraging other types of short-term memories, such as synaptic short-term adaptation (Cone & Shouval, 2021; Mongillo et al., 2008). Conversely, we verified that the same task cannot be achieved via neuronal memory (Fitz et al., 2020). The point neuron models were endowed with such memories in the form of membrane and spike-threshold adaptation.

Along with short-term memory, the two dendritic compartments of the Tripod neuron offer segregated pathways for signal integration. Each dendritic compartment supports a synaptic cluster; thus, it is balanced by dendritic inhibition

and can autonomously elicit an NMDA spike (Quaresima et al., 2022). This property reflects the spatial distribution of synaptic clusters along the dendrites (Kastellakis & Poirazi, 2019; Larkum, 2022; Poirazi et al., 2003). Synaptic clusters allow the neuron to take part in multiple engrams because memories are maintained and reactivated independently (Kastellakis et al., 2016; Legenstein & Maass, 2011). We tested the implications of this capacity by comparing three dendritic models, of which only one model had two segregated dendritic pathways. Although their sequence-recognition performances were not significantly different, the asymmetric model required half of the specialized connections than the other dendritic models.

A third contribution of the dendritic compartments concerns the stability of the voltage-dependent plasticity rule. The segregated dendritic compartment spans approximately 60 mV between the hyperpolarized and depolarized state, and the potentiation region is accessible only to NMDA spikes. Thus, the dendritic non-linearity ensures that only salient stimuli are encoded in the long-term memory. Our results are in agreement with previous work on STDP in dendrites. These studies show that voltage-dependent STDP in the dendrites maintains stable memories across time (Bono & Clopath, 2017; Bono, Wilmes, & Clopath, 2017; Kastellakis et al., 2016) and that decoupling dendritic potential from somatic activity mitigates the stability-plasticity dilemma (Wilmes & Clopath, 2023).

**Dendritic inhibition controls memory encoding and access**

The strengthening of excitatory connections makes the network prone to runaway excitation, a type of winner-takes-all dynamics (Ermentrout, 1992), and it destabilizes the network dynamics, causing continuous burst activity. In our model, the problem is solved by plasticity on the synapses of fast-spiking inhibitory neurons, which target the soma of tripod neurons and act fast compensatory mechanisms (Litwin-Kumar & Doiron, 2014; Zenke & Gerstner, 2017). However, somatic inhibition does not prevent dendrites from becoming strongly depolarized. This interferes with the formation of synaptic engrams. Following the experimental evidence on the role of inhibitory control on dendrites, (Chiu et al., 2018; Herstel & Wierenga, 2021; Zucca et al., 2017), we introduced dendritic inhibition as an additional compensatory mechanism, which can selectively silence dendritic branches, prevent the overriding of synaptic memories, and overall change the neuronal transfer function. The computational advantages

of tight inhibitory balance on dendritic compartments were already reported by G. R. Yang et al. (2016) and Mikulasch, Rudelt, and Priesemann (2020).

The present work introduces voltage-dependent iSTDP in gabaergic synapses targeting dendritic compartments. The rule steers the dendritic compartment toward a hyperpolarized state. When the dendritic compartment is depolarized, the inhibitory synapses are strengthened. Our results indicate that this form of inhibitory control is necessary for sequence detection and for the formation of hetero-associative connections. However, our implementation of the voltage-based dendritic iSTDP presents some limitations. Our model requires strong dendritic inhibition, resulting in the inhibitory plasticity driving all the weights to the maximum synaptic efficacy, so there is no neuronal specificity in the inhibitory connections. We observed that lowering dendritic inhibitory control results in heterogeneous inhibitory synaptic weights but also reduces the recognition score. Future work should explore the inhibitory plasticity mechanism and individuate the correct parameters to develop neuron or assembly-specific inhibitory connections. This could be done by changing the plasticity rule with a non-spike-time-dependent one (Miehl & Gjorgjieva, 2022; Pedrosa & Clopath, 2020) or changing the connection probability to have more gabaergic synaptic contacts per dendrite.

Moreover, the lack of neuron-specificity in the inhibitory weights indicates that dendritic inhibition does not act as a competition mechanism. This is in contrast with other models of sequence processing (Cone & Shouval, 2021; Maes et al., 2020, Vlachos et al., in preparation) or word recognition (Hannagan et al., 2013; McClelland & Elman, 1986), in which inhibition mediates competition among the excitatory assemblies. Somatic inhibition could still play this role because its synaptic strength depends on the neuron's firing rate (Appendix A). However, the lack of plasticity in the glutamatergic connections of the inhibitory cells ($E \rightarrow I_1$) makes it such that inhibition acts as a uniform blanket rather than a competition mechanism. Thus, our network solves the word recognition task without leveraging lateral inhibition among lexical competitors.

The homeostatic mechanism in the vSTDP, which maintains the incoming synaptic strength fixed for each neuron, is the only mechanism that can mediate competition, and its effect is expressed in the weak reciprocal connections among words. However, homeostasis seems ineffective. The connectivity matrices in the results and the appendices indicate that the assemblies that share some phonemes have, on average, stronger excitatory interactions than those that do not. We deem that implementing an explicit competition mechanism would

increase the capacity of the network to distinguish among words with shared phonemes.

## 4.4 Methods

In this study, we examine the activity of spiking network models when exposed to external stimuli representing phonemes and words. In the following sections, we first describe the equations characterizing the neuron models and their synaptic dynamics. Then we describe the network architecture, the recurrent connectivity, and the plasticity rules for each network class. Finally, we describe the experimental protocol used to model word recognition.

### 4.4.1 Neurons and synapses

The network models investigated are composed of two types of neurons, excitatory and inhibitory. The excitatory neurons have been modeled as three-compartment neurons in the Tripod networks and single-compartment units in the point-neuron networks. The inhibitory neurons have two sub-types, both modeled as single-compartment neurons. The models and parameters used vary for each network, and the variations are constrained to ensure stability during the simulation. For the two previously published networks, namely the L2/3 cortical circuit by Duarte and Morrison (2019) and the balanced network by Litwin-Kumar and Doiron (2014), we used models and parameters identical with the original publications; except for the plasticity rules introduced in the Duarte and Morrison (2019). In the following particular attention is reserved to the network with dendrites, as it is the most complex model and constitutes the novel contribution of the present work.

#### Excitatory neurons

***Tripod neurons*** The Tripod neurons comprise three electronically segregated compartments: an axosomatic compartment and two membrane patches. A detailed description of the model is presented in Quaresima et al. (2022), which also constitutes the second chapter of the present manuscript. The somatic compartment is modeled with the adaptive exponential integrate-and-fire model (Brette & Gerstner, 2005), it is two-dimensional and comprises the dynamics of membrane potential and adaptive hyper-polarizing currents (Eqs.4.1,4.2).

$$C_m^s \frac{dV^s}{dt} = -g_m^s \left[ (V^s - V_r) + \Delta_T \exp \frac{V^s - \theta}{\Delta_T} \right] - \sum_k g_k^s(t)(V^s - E_k) - w + \sum_{d=1,2} g_{ax}^d (V^d - V^s) \quad (4.1)$$

$$\tau_w \frac{dw}{dt} = -w + a(V^s - V_r) + b \sum \delta(t - t^f) \quad (4.2)$$

$$\text{if} \quad V^s > \theta \quad \rightarrow V^s = V_r \quad ; \quad (4.3)$$

$$C_m^d \frac{dV^d}{dt} = -g_m^d (V^d - V_r) - \sum_k g_k^d(t)(V^d - E_k) - g_{ax}^d (V^d - V^s); \quad \text{d} = 1,2 \quad (4.4)$$

The leak conductance $g_m^s$ defines the electrical permeability of the somatic membrane, $C_m^s$ its capacitance, and $g_k^s$ the set of variable synaptic conductances. The membrane potential $V^s$ is reset to $V_r$ after a spike, and the adaptation current $w$ is increased by a constant value $b$ (Eq.4.2). Spikes occur at times $t^f$ when the potential $V^s$ exceeds a threshold $\theta$ (Eq.4.3). Parameters for the somatic compartment were fixed and set to the values used in Brette and Gerstner (2005) and are listed in Table4.1.

The somatic compartment is coupled to the dendrites by the term $g_{ax}^d$ (Eq.4.1), it accounts for the axial conductance of the soma and the dendrite. The electrical properties of the dendritic compartment are governed by the passive membrane-patch equation Eq.4.4 (Koch, 1998). The equation's parameters depend on the dendritic geometry combined with physiological parameters. The physiology concerns the membrane permeability, the membrane capacitance, and the axial impedance to the soma ($\bar{g}_m$ =25.6 µS cm$^2$; $\bar{C}_m$ =0.5 µF/cm$^2$; $\bar{g}_{ax}$ =0.5 S cm$^2$) (Eyal et al., 2016; Koch, 1998). In the present study, we use the parameters obtained from human pyramidal cells Eyal et al. (2016) because they show enhanced dendritic memory (Quaresima et al., 2022). The dendritic diameter is fixed at 4 mm, and dendritic electrical properties are determined only by the dendritic length $L_d$, in agreement with the equations:

$$\text{Dendrite}[L_d, d = 4\,\mu\text{m}] = \begin{cases} C_m = & \pi \bar{c}_m \cdot L_d \\ g_m^d = & \pi \bar{g}_m \cdot L_d \\ g_{ax}^d = & \frac{\pi}{4} d^2 \bar{g}_{ax} \cdot {L_d}^{-1} \end{cases}$$

In all the dendritic models tested in the present study, the dendritic compartments have heterogeneous lengths in the range 150 µm to 400 µm.

The current flow between soma and dendrites depends on $g_{ax}\,(V_d > V_s)$. It usually is positive (dromic); conversely, during somatic firing, the soma compartment's membrane potential is clamped to 20 mV for the duration of the spike (spike width, 1 ms). During the spike interval, the current flow is antidromic ($V_d < V_s$), and the current flowing to the dendrites acts as a backpropagating axon potential (the dendritic and soma potentials are illustrated in Fig.4.1*A*, the inset shows the potentials' dynamics during somatic firing).

The Tripod neurons express AMPA, NMDA, $\mathrm{GABA_A}$, and $\mathrm{GABA_B}$ receptors on the dendrites and only AMPA and $\mathrm{GABA_A}$ on the soma. The synaptic transmission is modeled with a double-exponential equation (Roth & van Rossum, 2009). The equation describes the rise and decay of the receptors' conductance $g_k$. The synaptic conductance $g_k$ is given by:

$$g_k(t) = \bar{g}_k^{\mathrm{syn}} \mathcal{N}_k \left( \exp\left( -\frac{t - t_0}{\tau_k^r} \right) - \exp\left( -\frac{t - t_0}{\tau_k^d} \right) \right) \tag{4.5}$$

$$\mathcal{N}_k = \left( -e^{-t^{peak}/\tau^r} + e^{-t^{peak}/\tau_d^r} \right)^{-1}$$

$$t_k^{peak} = \frac{\tau^d \tau^r}{\tau^d - \tau^r} \ln \frac{\tau^d}{\tau^r} \tag{4.6}$$

with $k \in \{\mathrm{AMPA}, \mathrm{NMDA}, \mathrm{GABA_A}, \mathrm{GABA_B}\}$ indicating that each receptor has specific parameters. The timescales of rise and decay are given by $\tau_r$ and $\tau_d$ while the amplitude of the curve is defined by the maximal conductance parameter $g_{syn}$. To ensure that the amplitude equals $\bar{g}_k^{\mathrm{syn}}$, the conductance was scaled by the fixed normalization factor $\mathcal{N}_k$ Eq.4.6. The ratio between the maximal conductance of NMDA to AMPA receptors (NAR) governs the non-linear response to glutamatergic stimuli, in the present work is set as in human-like cells in Eyal et al. (2018); Quaresima et al. (2022). The synaptic parameters of the Tripod neuron are listed in Table4.3.

***Point neuron models*** The excitatory cells of the point-neuron networks are also implemented with the AdEx model (Eq.4.1). For the point-neuron models, the membrane potential threshold $\theta$ changes over time and increases after each

spike. The dynamic threshold facilitates voltage-dependent synaptic plasticity during the bursting intervals, it is governed by the equation:

$$\tau_\theta \frac{d\theta}{dt} = -(\theta - \theta_0) + \theta_+ \sum \delta(t - t^f) \tag{4.7}$$

$$\tag{4.8}$$

For the network borrowed from Duarte and Morrison (2019), the soma hosts four types of receptors, as in the Tripod. Conversely, the model from Litwin-Kumar and Doiron (2014) only has AMPA and $GABA_A$ receptors. The parameters of the point-neuron AdEx models and synapses are listed in Table4.1 and Table4.3 .

**Inhibitory interneurons**

Inhibitory neurons are modeled as single, isopotential compartments with leaky integrate-and-fire dynamics. The membrane potential $V_i$ is governed by the following equations:

$$\begin{aligned} C_m \frac{dV_i}{dt} &= -g_m(V_i - V_r) - \sum_k g_k(t)(V_i - E_k) - w \\ \tau_w \frac{dw}{dt} &= a(V_i - V_r) - w \end{aligned} \tag{4.9}$$

$$\text{if } V > V_T \rightarrow V = V_r \tag{4.10}$$

Following previous modeling efforts (e.g. Duarte and Morrison (2019); Park and Geffen (2020); G. R. Yang et al. (2016)) and consistently with physiological description (Tremblay et al., 2016), we implement two types of inhibitory interneurons: fast-spiking and non-fast-spiking. The first class corresponds to a set of well-defined parvalbumin-expressing (PV) GABAergic neurons, among which basket-cells; these cells primarily target the perisomatic regions and are indicated in this work as $I_1$ population. Conversely, the second group includes parvalbumin-negative GABAergic neurons expressing somatostatin. For notational simplicity, we refer to the second group as $I_2$. Following Duarte and Morrison (2019), the fast-spiking neurons are modeled as not adaptive while the $I_2$ neurons are. The two classes differ in LIF model parameters (Table4.2), and synaptic properties (Table4.3). In our model, they also differ for the compartment of the Tripod neuron they connect to. $I_1$ neurons target the soma compartment, and their synapses express fast $GABA_A$ receptors. $I_2$ neurons, on the other

hand, target the dendrites and, due to the additional slow $GABA_B$ component of the dendritic synapses, have a longer timescale. The inhibitory neurons' synapses express AMPA, $GABA_A$ receptors, and the synaptic transmission is modeled with a double-exponential model. The parameters for the inhibitory neurons are borrowed from Duarte and Morrison (2019) and are reported in Table4.1 and Table4.4.

### 4.4.2 Network architectures

We implement three main classes of networks that vary for the number and types of neurons considered and for the connectivity between them. The novel network model is composed of Tripod neurons as excitatory cells. The network counts 2000 Tripod neurons with dendritic lengths uniformly distributed in the range150 µm to 400 µm; along with 175 fast-spiking neurons, and 325 slow-spiking neurons. The proportion of inhibitory neurons follows Duarte and Morrison (2019). The network lacks geometrical structure and is conceived as a local cortical circuit. The Tripod neurons connect reciprocally through the dendrites and to the soma of inhibitory neurons. The fast-spiking neurons target the Tripod neurons' soma, while the slow-spiking neurons target the dendrites. Inhibitory neurons also connect reciprocally. All connections have a probability of $p$ =0.2. The connections targeting the two dendrites of a Tripod neuron are drawn independently from each other. Hence some neurons will connect through one dendrite only, and others will connect through both dendrites (pre-synaptic spikes arrive separately on both compartments). The connectivity pattern described is the default network configuration and is identified as the *asymmetrical* network. As a result, the Tripod network has 10 different connection types; the Tripod-Tripod and $I_2$-Tripod are doubled because they account for the two dendrites.

The remaining two networks are composed of point-neuron models, illustrated in the right panel of Fig.4.1*B*. One of the models implements an exact copy of the network described in Litwin-Kumar and Doiron (2014) (LKD model). The network is composed of 4000 excitatory neurons and 1000 inhibitory neurons. Then neurons are sparsely connected with a probability, $p$ =0.2. The third network is composed of 2000 excitatory neurons, 200 fast-spiking neurons and 200 slow-spiking neurons. The network is modeled after the non-heterogeneous model of cortical L2/3 (Duarte & Morrison, 2019). Neurons are connected with probability and connectivity strength that depends on the type of connection, for a total of nine types of connections. The network is dubbed the DM network.

The networks' neuron types and the number of cells are listed in Table4.6. In all the models, the synaptic weights change according to the connection types and are listed in Table4.7 with the respective connection probability. An illustration of the Tripod and point neuron networks connectivity is offered in Fig.4.1*B*

### 4.4.3 Synaptic learning rules

The synapses connecting excitatory neurons (EE) and from the inhibitory to the excitatory neurons (IE) are subject to spike-time-dependent plasticity (STDP). The two plasticity rules are illustrated in Fig.4.1*C*.

***Excitatory STDP*** The glutamatergic synapses undergo a voltage-dependent STDP rule (Clopath et al., 2010). The voltage-dependent STDP (vSTDP) strengthens the connection when, following a pre-synaptic spike, the post-synaptic neuron's membrane potential is depolarized. The equations governing the vSTDP rule are:

$$\begin{aligned} J_{ij}^{EE} \leftarrow \quad J_{ij}^{EE} \quad & +dtA_{LTD} \quad s_j\left[u_i(t)-\theta^-\right]_+ \\ & +dtA_{LTP} \quad x_j\left[V_i-\theta^+\right]_+\left[v_i(t)-\theta^-\right]_+ \end{aligned} \tag{4.11}$$

$$\tau_u du_j = \quad -(u_j - V_j)dt \tag{4.12}$$

$$\tau_v dv_j = \quad -(v_j - V_j)dt \tag{4.13}$$

$$\tau_x dx_i = \quad -x_i dt + \delta(t - t_f) \tag{4.14}$$

The $J_{ij}^{EE}$ indicates the connection between the excitatory neurons, post-synaptic $i$, and pre-synaptic $j$. The variable $s_j = \sum_t \delta(t - t_j^f)$ is the spike train of the presynaptic neuron, and $x_i$ is the filtered spike train of the post-synaptic cell. In addition, the variables $V_j, u_j, v_j$ are the membrane potential, the slow, and the fast-filtered membrane potential of the targeted compartment. The two thresholds $(\theta^+, \theta^-)$ mark the region where long-term potentiation (LTP) and long-term depression (LTD) occur, coupled with the learning rates $A_{LTP}$ and $A_{LTD}$. In the point-neuron models, the vSTDP is applied to the soma $(V_i = V_i^s)$, while in the Tripod neurons, the vSTDP is applied to the dendrites $(V_i = V_i^{d=1,2})$. For each couple of connected Tripod neurons, there are two values $J_{ij}^{E^1E}$ and $J_{ij}^{E^2E}$, one for each dendrite, and they are updated independently. In addition, the vSTDP rule includes a homeostatic mechanism that maintains the sum of pre-synaptic

weights fixed; the homeostatic rule follows the prescription of synaptic scaling (Tetzlaff et al., 2011; Triesch et al., 2018; Turrigiano, 2011) and is multiplicative. Conversely, the homeostatic rule in Litwin-Kumar and Doiron (2014) is additive. The multiplicative homeostatic rule is implemented as follows:

$$J_{ij}^{EE} \leftarrow J_{ij}^{EE} \frac{\sum_k J_{ik}}{\sum_k J_{ik}^0} \tag{4.15}$$

where $J_{ij}^0$ is the initial value of the synaptic weight. Homeostasis is applied every 20 ms.

***Inhibitory STDP*** For the connections from inhibitory to excitatory neurons, the networks implement the iSTDP rule from Vogels et al. (2011). The iSTDP governs the network stability by regulating the inhibitory synaptic strength. It is a necessary ingredient for the formation of assemblies in the original LKD network because it prevents winner-takes-all dynamics. We implemented the same mechanism for the $\mathrm{I}_1$or fast-spiking interneurons in the three networks, ($E^s\mathrm{I}_1$); the iSTDP maintains the firing rate of the excitatory cells by modulating the synaptic strength of the inhibitory pathways. In addition, we modified the iSTDP rule for the slow interneuron type such that they can control the dendritic non-linearity of the Tripod's dendrites. The $\mathrm{I}_2$iSTDP targets the membrane potential of the post-synaptic neuron, rather than its firing rate. We refer to it as voltage-dependent iSTDP or v-iSTDP. It applies to synapses on the soma ($E[s]\mathrm{I}_2$) in the DM and to synapses on the dendrites ($E[d]\mathrm{I}_2$) in the Tripod network. Because there are non $\mathrm{I}_2$neurons in the LKD network, this plasticity rule does not apply there. The two iSTDP rules follow the equations:

$$\begin{aligned} &\text{if} \quad j \in \mathrm{I}_1, \quad i \in \mathrm{Exc}: \\ &\qquad J_{ij}^{EI} \leftarrow J_{ij}^{E[s]I} + \eta dt \left[ x_j s_i^E + (x_i - \alpha) s_j^I \right] \qquad (4.16) \\ &\text{if} \quad j \in \mathrm{I}_2, \quad i \in \mathrm{Exc}: \\ &\qquad J_{ij}^{EI} \leftarrow J_{ij}^{EI} + \eta dt \left[ x_j s_i^E + (v_i - V_0) s_j^I \right] \qquad (4.17) \end{aligned}$$

The $s_j^I$ is the spike train of the inhibitory neuron $j$, $x_i$ is the filtered spike trains of the excitatory neuron $i$ and $v_i$ is the dendritic or somatic potential (depending on the network to which the plasticity rule applies), respectively. The $\alpha$ is the target firing rate for the iSTDP of $\mathrm{I}_1$neurons, while the $V_0$ is the target membrane potential for the iSTDP of $\mathrm{I}_2$neurons. $\eta$ is the learning rate. The parameters of the iSTDP rules are listed in Table4.5.

### 4.4.4 Stimuli

***Phonemes and word projections*** The networks were stimulated through external projections encoding phoneme and word inputs. Depending on the lexicon used, each network has 20 to 35 distinct projections, each representing a distinct phoneme or a word. Each projection targets $\rho = 5\,\%$ of the network, creating an assembly of co-activated cells. The assemblies were approximately 100 neurons in the Tripod and DM network and 200 in the LKD network. The projections are distributed randomly to the network's neurons; therefore, the assemblies lack any geometrical structure and each cell can be contained in more than one assembly. The phoneme and word assemblies were stimulated with spike trains. All pre-synaptic projections have fixed firing rate (8 kHz) and synaptic efficacy (2.78 pF for point-neuron models, 20 pF for the Tripod model). These numbers were optimized to have a strong neuronal response but not unstabilize the network. Stimuli spike trains are drawn from a Poisson distribution at every presentation, with the rate being the same for all projections; as a consequence, the stimuli cannot be interpreted as rate- or time-coded. The only information available to the networks to discriminate between the stimuli is the identity of the activated assemblies, which constitutes a spatial code. The phoneme stimuli were presented for 50 ms each, and the word stimuli were presented for the entire duration of the stimulus. Each word was followed by a silent interval without external stimuli, the pause lasts 50 ms.

***Sequence association and recall*** The stimulation protocol consists of two phases during which the phonemes to words associations are first established and then recall is tested, both phases last 5 min. During the associative phase, we present the network with external stimuli on both phonemes and word projections. The stimuli are presented in sequence; one word is drawn randomly, and the corresponding phoneme projections are activated. The phonemes populations are activated for 50 ms each, while the word ensembles are activated for the entire duration of the word (that is 50 ms times the number of phonemes contained in the word). The number of words presented in the two phases depends on the average length and the interval duration; for the experiments discussed here, each word is presented 50 to 100 times. All words are presented approximately the same number of times; therefore, there are no frequency effects. The phase duration has been chosen to be long enough for the synaptic weights to converge, but short enough to allow for several simulations and maintain the study reproducible.

***Lexica*** The lexica were chosen from previous publications or assembled for this specific study. The seven lexica tested vary for the number of items, the average length of words, and the phonological overlap. Phonological overlap is computed as the ratio between the number of words in the lexicon and the number of phonemes: $\mathcal{O} = N_w/N_p$. The lexica and their characteristics are presented in Table4.8.

The TIMIT lexicon has been selected based on the work from Dong, Huang, and Xu (2018). The Cohort lexicon was obtained from Granger, Whitson, Larson, Lynch, and Lynch (1994). The TISK lexicon was obtained from Hannagan et al. (2013). The remaining lexica are original to this work.

### 4.4.5 Word recognition measure

***Uniqueness and offset point*** We individuate two salient time points for word recognition for every phoneme sequence presented during the recall phase. One is the offset of the word (offset point, OP), that is, the last phoneme of the sequence. The other is the word's uniqueness point (UP). The UP corresponds to the phoneme that permits distinguishing a word from any other in the lexicon (Marslen-Wilson, 1973); when the word is contained in another word (such as *gold* and *golden*), the UP is the phoneme following the last shared phoneme, i.e., the silence after the end of the word in case of *gold*. As in the previous section, we estimate word recognition for every trial by measuring the firing rate in the adjacent interval. We associate the OP with an interval lasting the entire length of the phonemes' sequence, from the word onset to the word offset (50 ms times the length of the word); this measure is the same one portrayed in Fig.4.2*D*. Conversely, the UP interval is shorter and corresponds to the duration of one single phoneme (50 ms); in this case, the interval starts and ends with the UP phoneme. An example of the uniqueness and offset points of the word *golden*, and their intervals, is illustrated in Fig.4.3*A*; the offset and UP time-points are marked on the phoneme sequence, respectively, by the blue and green arrow marks. Because words are longer than one phoneme, the interval associated with the offset is, on average, longer than the one associated with the UP.

***Recognition $\kappa$-score*** Because the reactivation of word assemblies is delayed with respect to the onset ofset of the phoneme sequence, we introduced a variable time shift in the measure of the firing rate to individuate the optimal interval of word recognition. For both the offset and UP measures, the measured intervals were probed with delays from −150 ms to 200 ms, sampling every 5 ms. The

delayed interval is also illustrated in Fig.4.3*A*; the applied time shift is indicated with the orange arrow. We computed the lexicon confusion matrix (CM) for any given delay. The confusion matrix indicates which word is reactivated when a target word is presented; the sum of each column is one and corresponds to the fraction of trials (column) assigned to each target (row). Values close to one on the CM's diagonal indicate correct recognition, while off-diagonal elements indicate mistaken phoneme sequences. The average recognition score is computed from the confusion matrix using Cohen's $\kappa$ to correct against chance.

***Average Recognition Delay*** The average recognition delay is computed for both the uniqueness and offset measures starting from the $\kappa$-score. The ARD is the linearly weighted average of the delay function, with respect to the $\kappa$-scores. That is:

$$\mathrm{ARD} = \frac{\sum_{d=d_0}^{d=d_f} \kappa(d) d}{\sum_{d=d_0}^{d=d_f} \kappa(d)} \tag{4.18}$$

***Effective connectivity matrix*** Excitatory cells have recurrent connections with 20 % of the remaining network's cells. These connections are formed between the soma and the soma in the point-neuron models and the soma and the dendritic compartment in the dendritic models. We label these connections $W^s_{ij}$ or $W^d_{ij}$, where $s$ and $d$ are the somatic or dendritic compartment of the post-synaptic neuron $i$, onto which the pre-synaptic neuron $j$ projects.

For the asymmetric model $W^1_{ij} \neq W^2_{ij}$ and the two matrices were drawn independently. For the symmetric model $W^1_{ij} = W^2_{ij}$, and for the model with a single dendrite only $W^2_{ij}$ is set to zero. For all the dendritic models, $W^s_{ij}$ is also set to zero. The opposite is true for the point-neuron models, for which only $W^s_{ij}$ is defined. We define the effective connectivity matrix ($\mathscr{C}$) as the average of the connections between neurons from the pre-synaptic assembly $A$ to the post-synaptic assembly $B$

$$\mathscr{C}_{BA} = \frac{1}{N_A N_B} \sum_{i \in A;\ j \in B} \mathscr{N}_c (W^{[1]}_{ij} + W^{[2]}_{ij} + W^{[s]}_{ij})$$

where $N_A$ ($N_B$) is the number of neurons that belong to the assembly $A$ ($B$) and the superscripts in $W^{1,2,s}$, indicates whether the post-synaptic connection targets the two dendrites or the soma. The factor $\mathscr{N}_c$ accounts for the compartments that receive inputs, it is 0.5 in the symmetric and asymmetric dendritic models and 1 otherwise. The $\mathscr{C}$ is composed of four blocks, the connections between word

assemblies ($\mathscr{C}^{W\to W}$), between phoneme assemblies ($\mathscr{C}^{P\to P}$), from phonemes to words ($\mathscr{C}^{P\to W}$), and from words to phonemes ($\mathscr{C}^{W\to P}$).

***Block-wise update of the effective matrix*** To test which connection types are necessary for word recognition, we leveled the synaptic strength of selected synapses to the initial synaptic strength value ($W_0$). The synapses are updated based on their pre and post-synaptic cell assemblies. The synaptic update is defined by:

$$\begin{aligned} &\text{for each } i \in A \text{ and } j \in B: \\ &W^d_{i,j} = \begin{cases} 0, & \text{if } W^d_{i,j} = 0 \\ W_0, & \text{otherwise} \end{cases} \end{aligned} \tag{4.19}$$

where $i, j$ are the indices of two excitatory cells, $A, B$ are the post and pre-synaptic assemblies, and the superscript $d$ indicates the dendritic compartments, which are updated independently. The update rule in Eq.4.19 is iterated among all the $A, B$ assemblies belonging to the block that has to be flattened. For example, to level the phonemes-to-word connections ($\mathscr{C}^{P\to W}$), the update rule is repeated for all the assemblies $A$ associated with words, and all the assemblies $B$ associated with phonemes.

***Flattening of lexical connections*** The strength of the connections between the phonemes and word assemblies depends on the serial position of the phoneme in the word (Results, Fig.4.6*A*). To evaluate the computational role of the differences in synaptic weights, we modified $W^d_{i,j}$ such that all connection strengths from each phoneme to the word populations are the same. The result is an effective matrix in which the connections from any phoneme to the word that contains them have the same strength. To this aim, we used a similar procedure to Eq.4.19, selecting $A, B$ only among the lexical connections. In this case, the synaptic value was not updated with the initial value but with the average of all the synapses targeting the word population $\langle W^d_{i,j} \rangle_{\mathscr{P}_A \to A}$. Namely,

$$\begin{aligned} &\text{for each} \quad A \in \mathscr{W} \text{ and } B \in \mathscr{P}_A: \\ &\text{for each} \quad i \in A \text{ and } j \in B: \\ &W^d_{i,j} = \begin{cases} 0, & \text{if } W^d_{i,j} = 0 \\ \langle W^d_{i,j} \rangle_{\mathscr{P}_A \to A}, & \text{otherwise} \end{cases} \end{aligned} \tag{4.20}$$

Table 4.1: Parameters for the axosomatic compartment (Duarte & Morrison, 2019; Litwin-Kumar & Doiron, 2014; Quaresima et al., 2022).

| Symbol | Description | Tripod | LKD | DM | Unit |
|---|---|---|---|---|---|
| $V_r$ | Resting membrane potential | -70.6 | -70. | -76.43 | mV |
| $g_L$ | Membrane leak conductance | 40 | 15. | 4.64 | nS |
| $C_m$ | Membrane capacitance | 281 | 300 | 116.5 | pF |
| $V_T$ | Threshold potential | -50. | -52. | -44.45 | mV |
| $u_{th}$ | Spike onset threshold | 0 | 0 | 0 | mV |
| $u_r$ | Reset potential | -55 | -60 | -54.18 | mV |
| $\Delta_T$ | Slope factor | 2 | 2 | 2 | mV |
| $a$ | Adaptation conductance | 4 | 4 | 4 | nS |
| $\tau_w$ | STA[1] timescale | 144 | 150 | 144 | ms |
| $b$ | STA[1] increment | 80.5 | 80.5 | 80.5 | pA |
| $t_{up}$ | Spike width (20 mV) | 1 | 1 | 1 | ms |
| $t_{ref}$ | Refractory period | 2 | 2 | 2.1 | ms |

[1] Spike-triggered adaptation

where $\mathscr{W}$ is the set of word assemblies and $\mathscr{P}_A$ is the set of phonemes contained in the word $A$. The result of the update rule in Eq.4.20 is illustrated in Fig.4.6*B*.

Table 4.2: Parameters for the axosomatic compartment of the inhibitory neurons for the three network models (Duarte & Morrison, 2019; Litwin-Kumar & Doiron, 2014). The inhibitory neurons of the Tripod network have the same parameters as those in the Duarte and Morrison model.

| Symbol | Description | $I_1$ (LKD) | $I_1$ (DM) | $I_2$ (DM) | Unit |
|---|---|---|---|---|---|
| $g_L$ | Membrane leak conductance | 15 | 9.75 | 4.61 | nS |
| $C_m$ | Membrane capacitance | 300 | 104.52 | 102.86 | pF |
| $V_r$ | Resting membrane potential | -62 | -64.33 | -61 | mV |
| $V_T$ | Threshold potential | -52 | -38.97 | -34.4 | mV |
| $u_{th}$ | Spike onset threshold | $V_T$ | $V_T$ | $V_T$ | mV |
| $u_r$ | Reset potential | -57 | -57.47 | -47.11 | mV |
| $a$ | Adaptation conductance | - | - | 144 | nS |
| $\tau_w$ | STA[1] time scale | - | - | 4 | ms |
| $b$ | STA[1] increment | - | - | 80.5 | pA |
| $t_{up}$ | Spike width (20 mV) | 1 | 1 | 1 | ms |
| $t_{ref}$ | Refractory period | 1 | 0.5 | 1.3 | ms |

[1] Spike-triggered adaptation

Table 4.3: Synaptic parameters for the excitatory and inhibitory neurons in the Tripod network.

| Receptor | $E_r$(mV) | $\tau_{rise}$(ms) | $\tau_{decay}$(ms) | $g_{syn}$(nS) | $\gamma\left(\frac{1}{\text{mV}}\right)$ |
|---|---|---|---|---|---|
| **Tripod soma** | | | | | |
| AMPA | 0.0 | 0.26 | 2. | 0.73 | - |
| $GABA_A$ | -70 | 0.1 | 15 | 0.38 | |
| **Tripod dendrites** | | | | | |
| AMPA | 0.0 | 0.26 | 2. | 0.73 | - |
| NMDA | 0.0 | 8. | 35. | 0.73 · NAR | 0.075 |
| $GABA_A$ | -70 | 4.8 | 29.0 | 0.27 | |
| $GABA_B$ | - 90 | 30 | 400 | 0.006 | |
| **$I_1$soma** | | | | | |
| AMPA | 0.0 | 0.18 | 0.7 | 1.04 | - |
| $GABA_A$ | -75.0 | 0.19 | 2.5 | 0.84 | - |
| **$I_2$soma** | | | | | |
| AMPA | 0.0 | 0.18 | 1.8 | 0.56 | - |
| $GABA_A$ | -75.0 | 0.19 | 5.0 | 0.59 | - |

Table 4.4: Synaptic parameters for the excitatory and inhibitory neurons in the point neuron models (Duarte & Morrison, 2019; Litwin-Kumar & Doiron, 2014).

| **Receptor** | $E_r$(mV) | $\tau_{rise}$(ms) | $\tau_{decay}$(ms) | $g_{syn}$(nS) | $\gamma\left(\frac{1}{\text{mV}}\right)$ |
|---|---|---|---|---|---|
| **LKD soma** | | | | | |
| AMPA | 0.0 | 1.0 | 6.0 | 1.0 | - |
| $GABA_A$ | -75.0 | 0.5 | 2.0 | 1.0 | - |
| **LKD $I_1$** | | | | | |
| AMPA | 0.0 | 1.0 | 6.0 | 1.0 | - |
| $GABA_A$ | -75.0 | 0.5 | 2.0 | 1.0 | - |
| **Duarte soma** | | | | | |
| AMPA | 0.0 | 0.25 | 2.0 | 0.73 | - |
| NMDA | 0.0 | 0.99 | 100.0 | 0.16 | - |
| $GABA_A$ | -75.0 | 0.5 | 6.0 | 0.26 | - |
| $GABA_B$ | -90.0 | 30.0 | 100.0 | 0.01 | - |
| **Duarte $I_1$** | | | | | |
| AMPA | 0.0 | 0.1 | 0.7 | 1.6 | - |
| NMDA | 0.0 | 1.0 | 100.0 | 0.0 | - |
| $GABA_A$ | -75.0 | 0.1 | 2.5 | 1.0 | - |
| $GABA_B$ | -90.0 | 25.0 | 400.0 | 0.02 | - |
| **LKD $I_2$** | | | | | |
| AMPA | 0.0 | 0.2 | 1.8 | 0.8 | - |
| NMDA | 0.0 | 1.0 | 100.0 | 0.01 | - |
| $GABA_A$ | -75.0 | 0.2 | 5.0 | 0.7 | - |
| $GABA_B$ | -90.0 | 25.0 | 500.0 | 0.02 | - |

Table 4.5: Parameters for the v-STDP, iSTDP, and v-iSTDP plasticity rules. The v-STDP parameters for the point-neuron models (LKD and DM) are obtained from Clopath et al. (2010), and identical to the original study by Litwin-Kumar and Doiron (2014); for the dendritic model is obtained from Bono and Clopath (2017). The iSTDP and v-iSTDP rules follow the parameters by Vogels et al. (2011) with variations on the learning rate $\eta$ and the target of the homeostasis for the two interneurons.

| Symbol | Description | LKD | DM | Tripod | Unit |
|---|---|---|---|---|---|
| **v-STDP** | | | | | |
| $a^-$ | LTD learning rate | 8.0E-05 | 8.0E-05 | 4.0E-05 | Hz |
| $a^+$ | LTP learning rate | 1.4E-04 | 1.4E-04 | 1.4E-04 | Hz |
| $\theta^-$ | LTD threshold | -70 | -70 | -40 | mV |
| $\theta^+$ | LTP threshold | -49 | -49 | -20 | mV |
| $\tau_u$ | slow membrane filter | 10.0 | 10.0 | 15.0 | ms |
| $\tau_v$ | fast membrane filter | 7.0 | 7.0 | 45.0 | ms |
| $\tau_x$ | spike train filter | 15.0 | 15.0 | 20.0 | ms |
| $\tau_s$ | homeostasis timescale | 20.0 | 20.0 | 20.0 | ms |
| $j^-$ | minimum synaptic weight | 2.78 | 2.78 | 2.78 | pF |
| $j^+$ | maximum synaptic weight | 21.0 | 4.0 | 41.4 | pF |
| **iSTDP** | | | | | |
| $\eta$ | learning rate | 1.0 | 0.2 | 0.2 | Hz |
| $j^+$ | maximum synaptic weight | 243.0 | 243.0 | 243.0 | pF |
| $j^-$ | minimum synaptic weight | 2.78 | 2.78 | 2.78 | pF |
| **$I_1$ neurons** | | | | | |
| $r_0$ | target firing rate | 5 | 10 | 10 | Hz |
| $\tau_y$ | rate filter timescale | 20.0 | 20.0 | 20.0 | ms |
| **$I_2$ neurons** | | | | | |
| $vd$ | target dend. potential | - | -70 | -70 | mV |
| $\tau_d$ | dend. potential filter timescale | - | 5.0 | 5.0 | ms |

Table 4.6: Number and types of neurons for each model, intensity of the background noise.

| Parameter | Description | Tripod | DM | LKD | Unit |
|---|---|---|---|---|---|
| №E | Excitatory neurons | 2000 | 2000 | 4000 | - |
| №$I_1$ | Fast inhibitory neurons | 175 | 175 | 1000 | - |
| №$I_2$ | Excitatory neurons | 325 | 325 | - | - |
| Excitatory Noise | Excitatory noise | 4 | 4 | 4 | kHz |
| SST Noise | SST noise | 0.5 | 0.5 | 0.5 | kHz |
| PV Noise | PV noise | 0.5 | 0.5 | 0.5 | kHz |

Table 4.7: Connectivity for the three network models. The connections subject to STDP and iSTDP are in bold text, the synapses are listed in the couple post-pre synaptic and the squared parentheses indicate the post-synaptic target compartment. The numbers indicate respectively the connection probability and initial synaptic strength. The synaptic strength is in parenthesis, expressed in pF.

| Synapse | LitwinKumar & Doiron | Duarte & Morrison | Tripod |
|---|---|---|---|
| **E[d]E** | - | - | 0.2 (10.78) |
| **E[s]E** | 0.2 (2.76) | 0.168 (0.45) | - |
| $I_1E$ | 0.2 (1.27) | 0.575 (1.65) | 0.2 (5.27) |
| $I_2E$ | 0.2 | 0.244 (0.638) | 0.2 (5.27) |
| **E[s]**$I_1$ | 0.2 (48.7) | 0.6 (5.148) | 0.2 (15.8) |
| **E[d]**$I_2$ | - | - | 0.2 (15.8) |
| **E**$^s I_2$ | - | 0.465 (4.85) | - |
| $I_2I_2$ | - | 0.381 (0.83) | 0.2 (16.2) |
| $I_1I_1$ | 0.2 (16.2) | 0.55 (2.22) | 0.2 (16.2) |
| $I_1I_2$ | - | 0.379 (1.47) | 0.2 (1.47) |
| $I_2I_1$ | - | 0.24 (1.4) | 0.2 (0.83) |

Table 4.8: Lexica that are used in the simulations. The table presents the words that compose each lexicon, the average phonological overlap, and the lexicon size

| Vocabulary | Words | Phon. Overlap | Lexicon size |
|---|---|---|---|
| Identity | a, b, c, d, e, f, g, h, i, j | 1 | 10 |
| No overlap | abcd, efghi, jkl, mno, pqrst, uvw, xyz | 1 | 7 |
| TIMIT | all, ask, carry, dark, don't, had, like, me, oily, rag, she, suit, that, wash, water, year, your | 2.63 | 17 |
| Digits | zero, one, two, three, four, five, six, seven, eight, nine | 3.33 | 10 |
| Overlap | pollen, gold, golden, doll, lop, god, log, poll, goal, dog | 5.0 | 10 |
| Cohort | capias, capillary, capistrate, capital, capitalise, capitalism, capitalist, capitate, capitol, capitoline, capitulate, capitulation | 7.69 | 12 |
| TISK | art, artist, arts, rats, star, start, stat, tar, tars, trist | 8.33 | 10 |

## 4.5 Appendix

### Appendix A: Network activity during the learning and recall phase.

To characterize network activity, we measured the average firing rate and the variability in the interspike intervals (ISI) (Fig.4.2*B*). We observe the firing rate of the excitatory neurons is heterogeneous and ranges from 0 Hz to 15 Hz. The rate of the excitatory population (Tripod neurons) divides into two groups. One group comprises neurons that receive external stimuli and thus belong to an assembly (*Exc. with inputs*); these cells fire at various rates in the range 5 Hz to 15 Hz. The second class is composed of neurons that do not receive external stimuli; they fire at a lower rate, in the range 0 Hz to 10 Hz. Because of the heterogeneity in the neurons' rate, we cannot use the standard coefficient of variation for the ISI measure, it would associate higher variability to neurons with higher rates (Holt et al., 1996). For an unbiased estimate of the $CV_{ISI}$ we computed the coefficient of variation based only on adjacent interspike intervals ($CV_{ISI}2$). Accordingly, the differences between excitatory neurons are less visible in the ISI variability; neurons with no inputs have $CV_{ISI}2$ in the interval 0.5 to 1.3 while the measure is slightly lower for neurons that receive external stimuli. Finally, interneurons of types $I_1$ and $I_2$ also fire at heterogeneous rates in the interval 5 Hz to 50 Hz, their coefficient of variation remains around 1, consistent with Poisson input they receive from the remainder of the network.

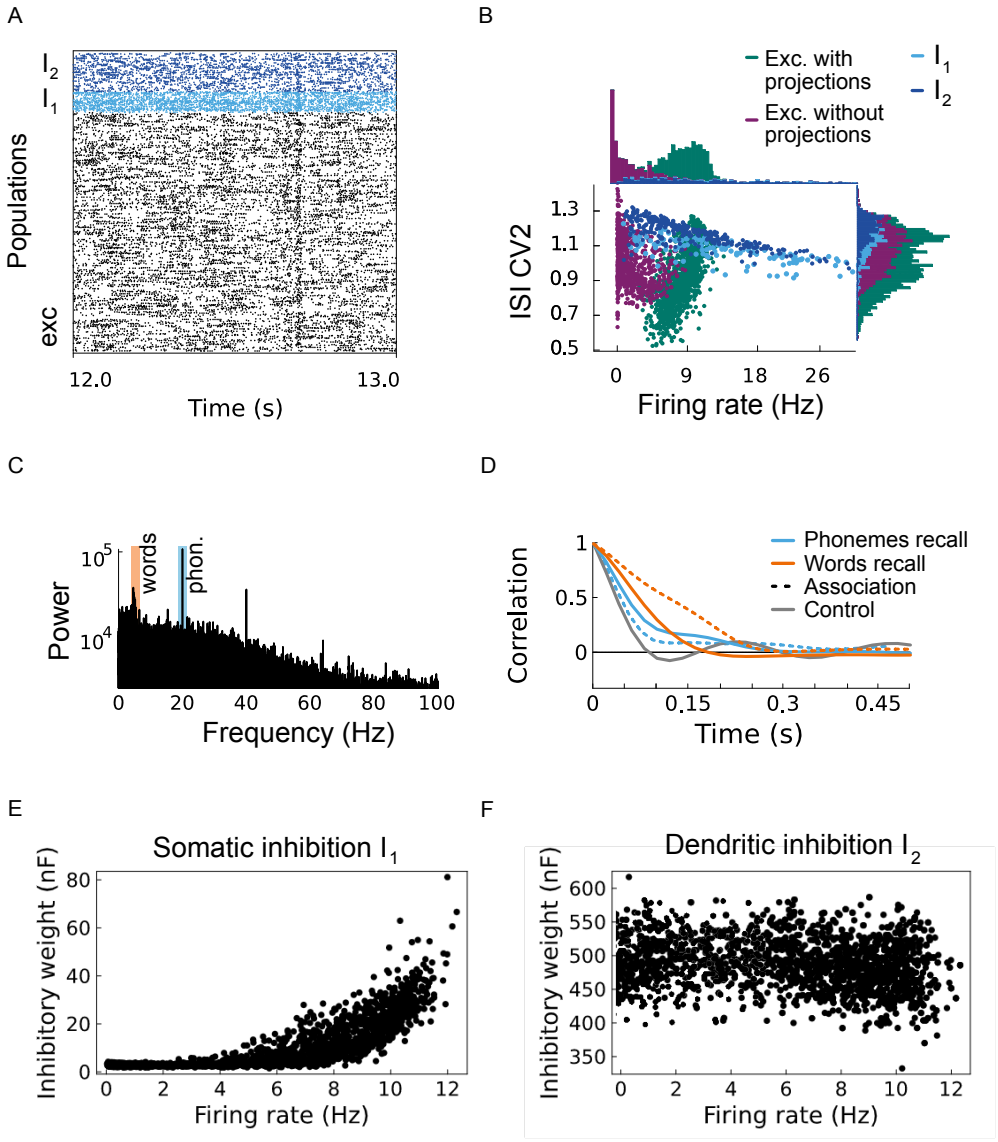
The variability observed in the rate of excitatory neurons is unexpected. Indeed, the rapid compensatory mechanism implemented in the network (iSTDP on somatic compartments) should steer the neurons to fire at 10 Hz. In point-neuron models (both 4.6 and Litwin-Kumar and Doiron (2014)), the rate is tightly centered at the homeostatic value, although some neurons receive inputs, and some do not. These initial results show that including dendrites in the neuron model greatly impacts the network dynamics and leads to heterogeneity in the firing rate in the face of counteracting homeostatic mechanisms.

We exclude the presence of pathological synchronicity, we computed the fast Fourier transform (FFT) of the network firing rate during the associative and recall phases. At this scope, the spikes of the excitatory and inhibitory populations were merged and binned in intervals of 5 ms, corresponding to Nyquist frequency of 100 Hz for the FFT. The spectrogram for the recall phase is illustrated in App. Fig.4.1 *C*. The frequency analysis reveals clear peaks at the frequency of the phonemic input 20 Hz and 40 Hz - corresponding to the fundamental fre-

quency and the first harmonic of the phonemic stimuli (50 ms). The peaked spectral reveals the resistor-capacitor (RC) circuit subserving the Tripod neuron's dendritic compartments; the dendritic compartments start charging when the phoneme projections activate and slowly decay when it is turned off. A smaller but more interesting peak is visible around 5 Hz (orange band in the panel); this frequency interval corresponds to word inputs. Because no external activity is carried out on the word projections in the recall phase, the peak in the spectrum must correspond to the reactivation of the word assemblies, indicating an internally generated response. The fact that words are interleaved by 50 ms of silence (absence of inputs) offers a possible confound, however, the peak at low frequency is present also when there is no silence among words in the recall phase.

Additional evidence that the word assemblies populations are reactivated in the recall phase is obtained by the comparative analysis of the firing rate correlations. In App. Fig.4.1 *D*, we illustrate the average autocorrelation function of network subpopulations corresponding to phonemes, words, or randomly sampled neurons (control group of 100 neurons); we mark the half-height (HH) correlation time, that is, the time for the firing rate to decorrelate of 50 %. For the phonemes populations, the differences between the associative and recall phases are expected to be minor because they receive the same external input in the two phases; the HH is 48 ms in the associative phase and (55 ms) in the recall. Despite being a marginal difference, the HH in the recall phase is systematically (10 network samples) longer than in the associative phase; we hypothesize this to be due to the fewer external inputs in the former condition, which interfere less with the reverberations of the assemblies. The word assemblies shows a radically different scenario. In the early phase, the populations are maintained in continuous firing by the external projections, lasting 150 ms to 350 ms. The assemblies have HH correlations longer than 100 ms (orange dashed lines). When the external stimulation is removed (recall phase, orange solid lines), the correlation time of the word assemblies shrinks to 75 ms. The difference in the two conditions indicates that the retrieved activity is less stable than the input activity. Nonetheless, the HH of word populations is twice the HH of the control populations (grey, 40 ms), which points to the role of the acquired word engram in fostering activity in the assemblies.

Finally, we report the strength of the inhibitory neurons with respect to the firing activity of the excitatory cells. The difference in the inhibitory plasticity rule decouples the two inhibitory populations. In the case of the $I_1$, the synaptic




Appendix Figure 4.1: **Firing activity during the associative and recall phase**
*Figure caption continues on the next page.*

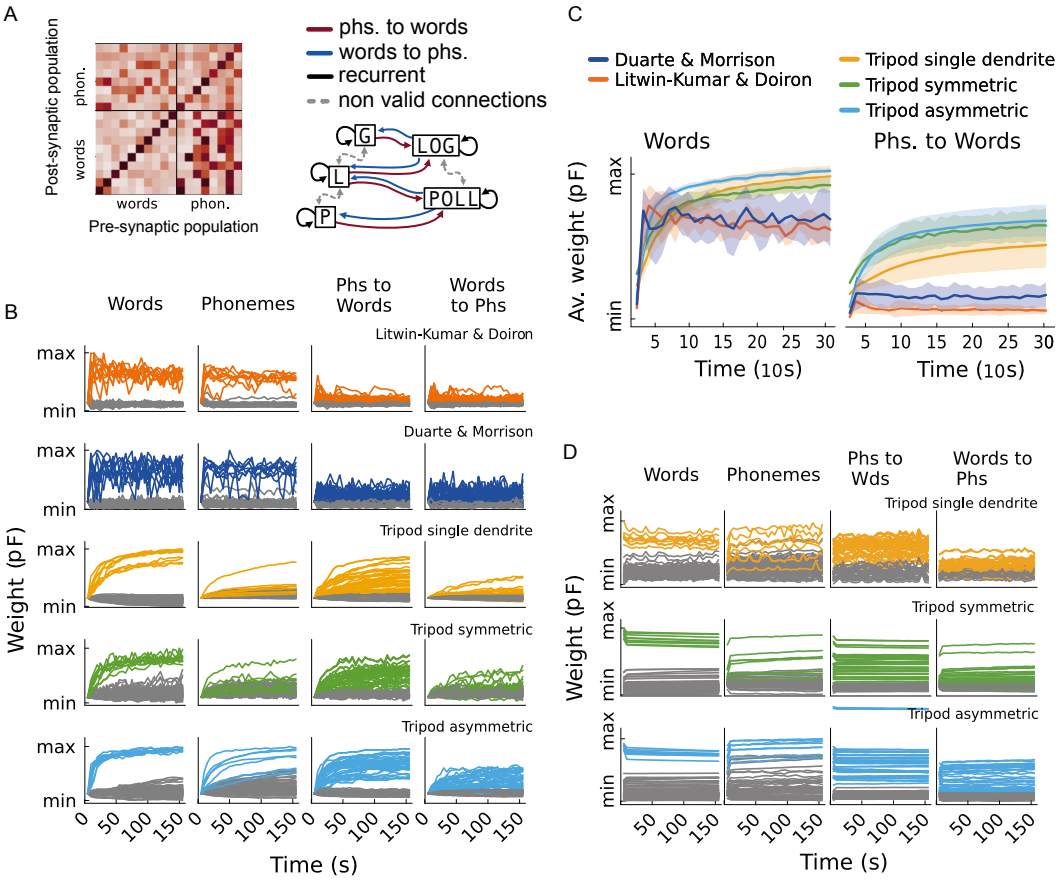
Appendix Figure 4.1: (**A**) Raster plot of the Tripod network during an interval of 1 s. Excitatory dendritic neurons are shown in black, fast-spiking interneurons ($I_1$) in light blue, and slow-spiking interneurons ($I_2$) in dark blue. (**B**) Mean firing rate and CV2 ISI for excitatory and inhibitory populations. Excitatory neurons are divided into two groups, those with (green) and without (purple) external input. The former have a higher firing rate and a lower CV2 ISI. (**C**) The frequency spectrum of the excitatory firing activity across the entire simulation (5 min). The large peak at 20 Hz and 40 Hz reflect the frequency of the phonemic inputs (blue marker). Conversely, the smoother peak at low frequency (orange marker) is due to the network's internal dynamics. The spectrogram is computed for a recall protocol with no silence in between each word, hence the peak is due to word re-activation, Overall the frequency spectrum indicates the network is in the asynchronous firing regime. (**D**) Assembly firing rate autocorrelation averaged over phonemes (blue) and word (orange) assemblies in the associative (dashed) and recall phase (solid). The grey line illustrates the autocorrelation of a population of randomly sampled neurons. The word assemblies (orange) have the longest autocorrelation time in both phases. It indicates that the engram acquired during the association phase can collectively reactivate and sustain firing during retrieval. (**E**) The total inhibitory synaptic weight between $I_1$ neurons and Tripod somatic compartments depends on the average firing rate of the cell. (**F**)Conversely, the total strength of inhibitory synapses targeting the dendrites ($I_2$ neurons) does not depend on the neuron firing rate.

strength correlates with the firing rate of the Tripod neurons. Inhibitory synapses are small for cells that seldomly fire, and strong for those that fire often App. Fig.4.1 *E*. Conversely, the dendritic inhibition does not depend on the rate, all excitatory cells form strong inhibitory synapses on the dendritic compartments App. Fig.4.1 *F*.

## Appendix B: Formation and maintenance of word memories

We can visualize the evolution of the recurrent and forward connections by computing the effective connectivity matrix throughout the simulation ($\mathscr{C}$, App. Fig.4.2 *A*). The panels in App. Fig.4.2 *B* show the average synaptic strength for each element of the effective connectivity matrix during the associative phase. Each row indicates one of the five models tested on the *Overlap* lexicon. Similarly to the main text, the maximum and minimum values refer to the maximum of the effective connectivity matrix and the initial synaptic weight.

All models have strong recurrent word memories, only the dendritic models have distinct phonemes-to-words and word-to-phonemes connections. In addition, the strengthening of synaptic connections on the dendritic compartments is slower than in the soma models. Synapses grow smoothly throughout the asso-

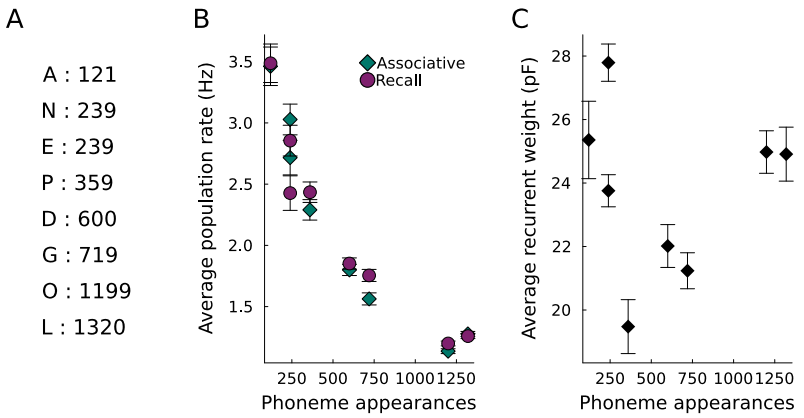



Appendix Figure 4.2: **Formation and maintenance of network memories** (**A**) Effective connectivity matrix for the Tripod Asymmetric model, the scheme illustrates the types of connections analyzed in the remainder of the figure. (**B**) Weights formation during the associative phase, for all the connections in *A* and the five analyzed models. (**C**) Average recurrent and feedforward connection weights during the memory formation phase. The five models are drawn together for comparison. The maximum and minimum refer to the extrema of the EC for each model. (**D**) Weights decay during the recall phase for the three dendritic models. The synaptic strengths remain overall stable with the largest drop happening in the word recurrent connections.

ciative interval. Conversely, in the point neuron models, the recurrent memories are more volatile and prone to be overridden, as indicated by the large fluctuations in the average synaptic weights. A more compact view of the memory formation process is presented in App. Fig.4.2 *C*, with a direct comparison of the recurrent and feedforward connections of the five models. For all the dendritic models, the learning trajectory is exponentially fast, with a timescale in the order of 10 s for the recurrent memories and 100 s for the feedforward connections. In contrast, the two point-neuron models reach high synaptic strength for the recurrent connections within just a few words presentation, but the non-recurrent connections remain close to the initial synaptic weight.

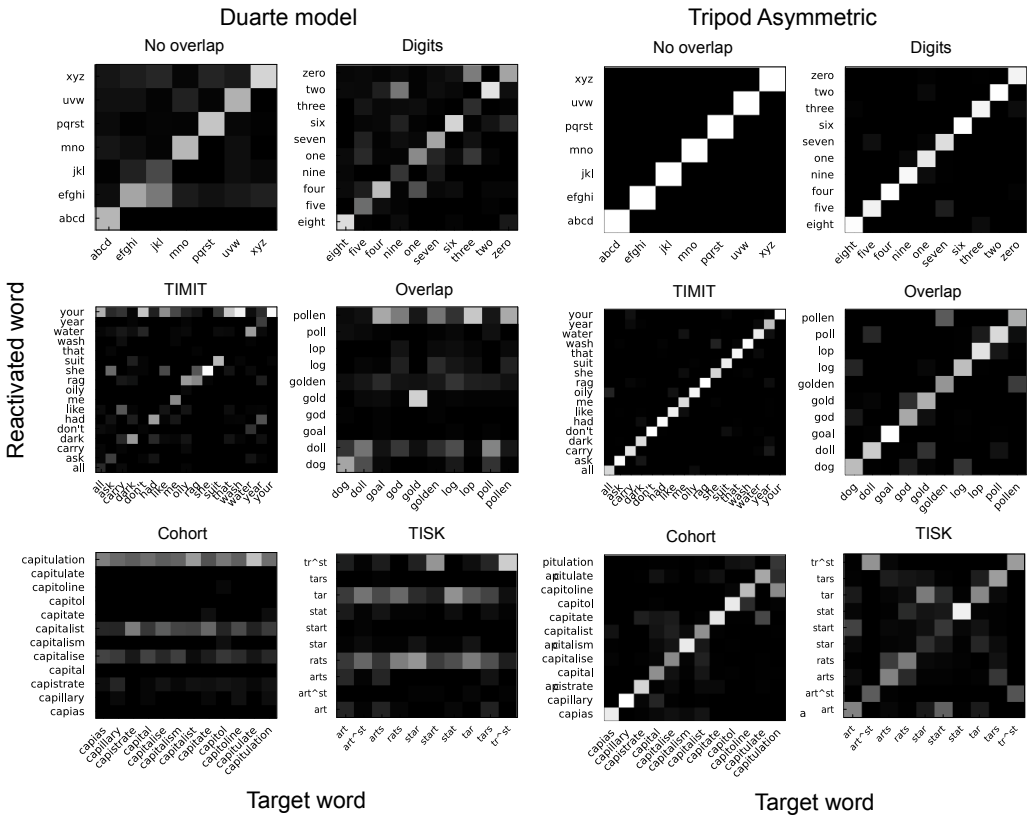
To verify the stability of the memory formed in the associative phase, we run the network with excitatory plasticity active during the recall phase. The panels in App. Fig.4.2 *D* show the evolution of the synaptic strengths over 150 ms during the presentation of the phoneme sequences. The weights do change less than 5 % from those reached at the end of the associative phase; the recognition score, not shown, also remained high, despite a systematic drop of a few percent points. In the main text, the weights were frozen because the integration of the STDP equations is computationally demanding.

## 4.6 Supplementary Material

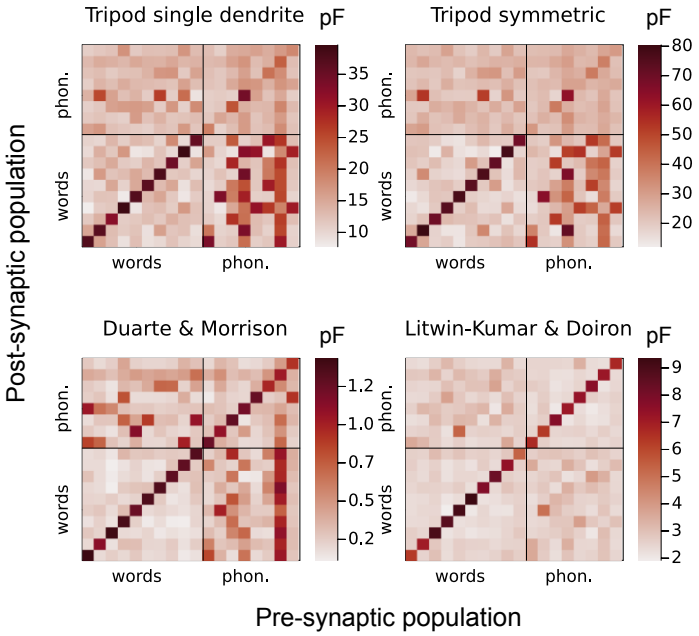



SI Figure 4.1:

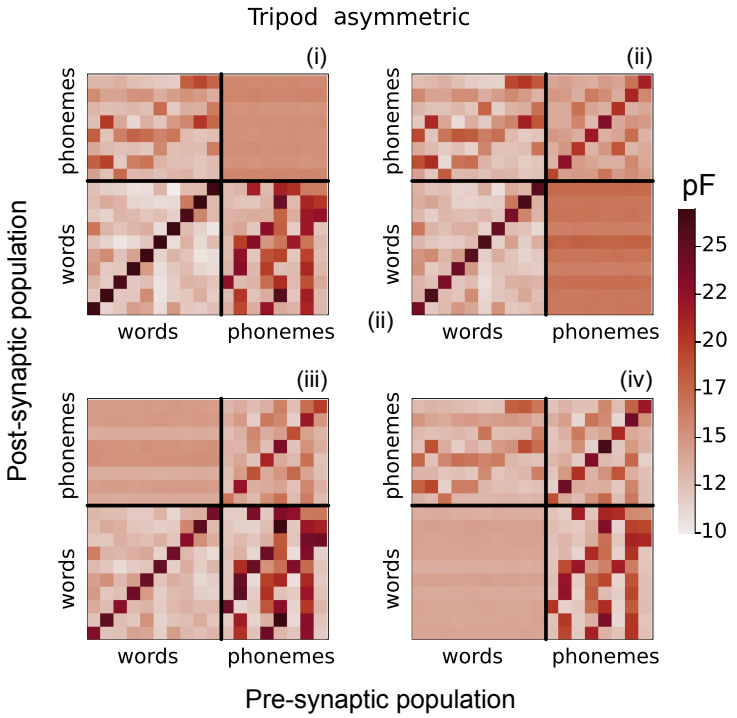
SI Figure 4.2: **Activity and recurrent weights of phoneme assemblies in the Overlap lexicon**
(**A**) Frequency of the phonemes in the input sequence. (**B**) More frequent phonemes have low firing rates, possibly due to neuronal adaptation of the firing rate. (**C**) Additionally, the synaptic scaling applied to glutamatergic synapses tends to reduce the synaptic strength of phonemes that are contained in several words. Because the incoming synaptic weight is fixed, phoneme populations that occur in multiple words must share their synaptic strength among more pre-synaptic pathways, reducing the available synaptic resources for the recurrent connections.



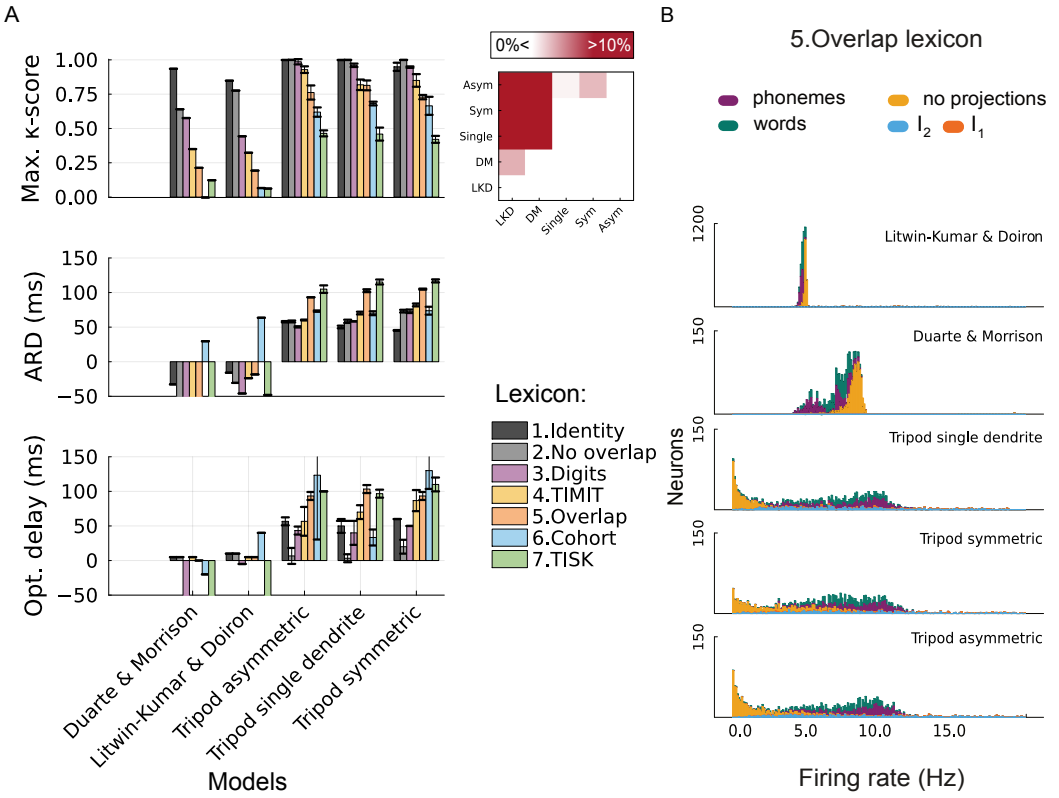

SI Figure 4.3: **Confusion matrices for point neuron and dendritic models** Confusion matrices for six of the lexica that were measured. The left columns show the confusion matrix for the Duarte model, and the right columns are those for the asymmetric tripod network. The elements on the diagonal indicate correct word recognition; conversely, the off-diagonal ones indicate wrong recollection. E.g., in the network with dendrites, when tested in the Cohort lexicon, the word *capitulation* is often confused for the word *capitoline*, but the opposite does not happen. Crucially, the confusion matrices of the point neuron model show correct word recognition for the lexicon with less phonological overlap but are completely random for the three with large phonological overlap. This is not the case for the network with dendrites where most of the words are recognized also in the lexica with large phonological overlap.



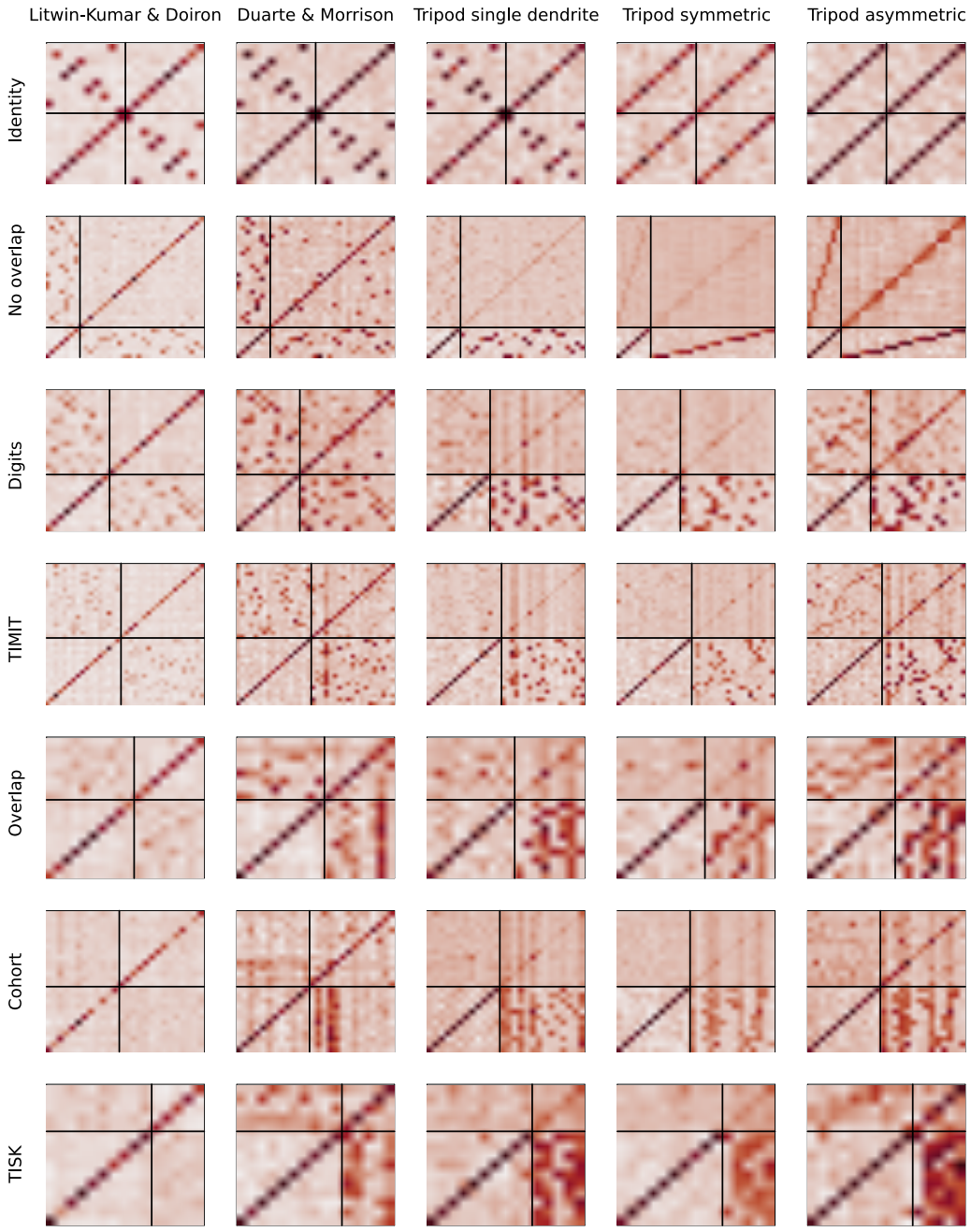

SI Figure 4.4: **Effective connectivity matrices for each model in the *Overlap* lexicon**
Effective connectivity matrices for the symmetric and single dendritic models ( the asymmetric model is portrayed in Fig.4.5)*A* and the two-point neurons models (bottom). The dendritic models develop strong connections between phonemes and word assemblies (bottom-right quadrants). In contrast, the point neuron models do not - as in the LKD network- develop connections that are ineffective in re-activating the word populations - as in the DM.



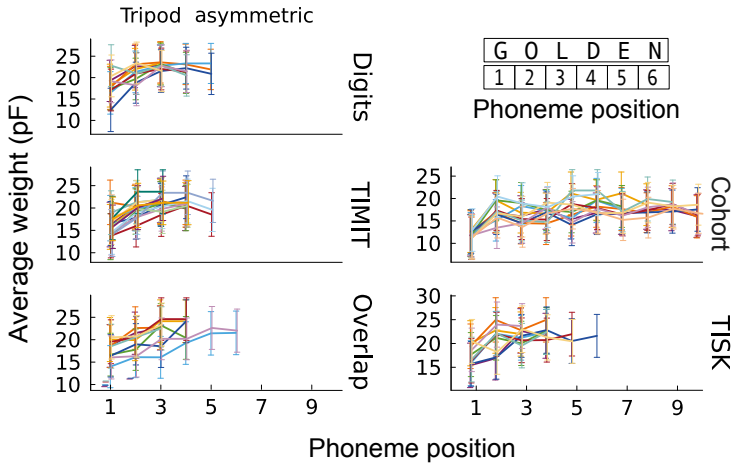

SI Figure 4.5: **Effective connectivity matrices with flattened connections** Effective connectivity matrices for Asymmetric Tripod with flattened connections. Respectively (i) phoneme recurrent connections, (ii) phonemes to word connections, (iii) words to phoneme connections, and (iv) word recurrent connections.



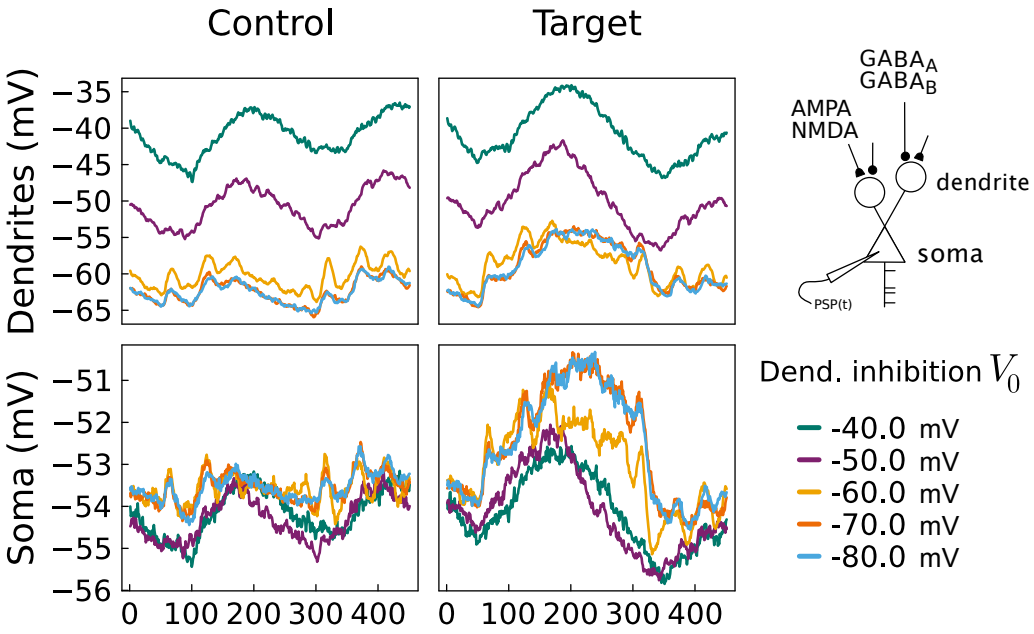

SI Figure 4.6: **Model comparison for the full interval measure firing profiles** (**A**) Comparison of recognition scores and average delay for the full interval measure. The panel up-right shows the average score differences between the five models; asymmetrical and symmetrical models perform slightly better than the single dendrite model. (**B**) Firing rate profile for the five network models during the recall phase of the Lexicon dictionary. The point neuron models have distributions peaked around the firing rate promoted by the inhibitory plasticity; there is a minor difference between the neurons within assemblies and those not. In contrast, the firing rate of dendritic models is largely heterogeneous, and neurons with no direct projections have a low firing rate. Remarkably, the firing rate profile is similar for the three dendritic models. We hypothesize that heterogeneity in the average firing rates of the neurons is an indicator of the network performing the task correctly.




SI Figure 4.7: **Effective connectivity matrices for all models and lexica**
The overall comparison of the effective connectivity matrices indicates that point-neuron models can form hetero-associative connections between phonemes and words (bottom-left quadrants) when the recollected words have no phonological overlap (first two rows, *Identity* and *No overlap* lexica)

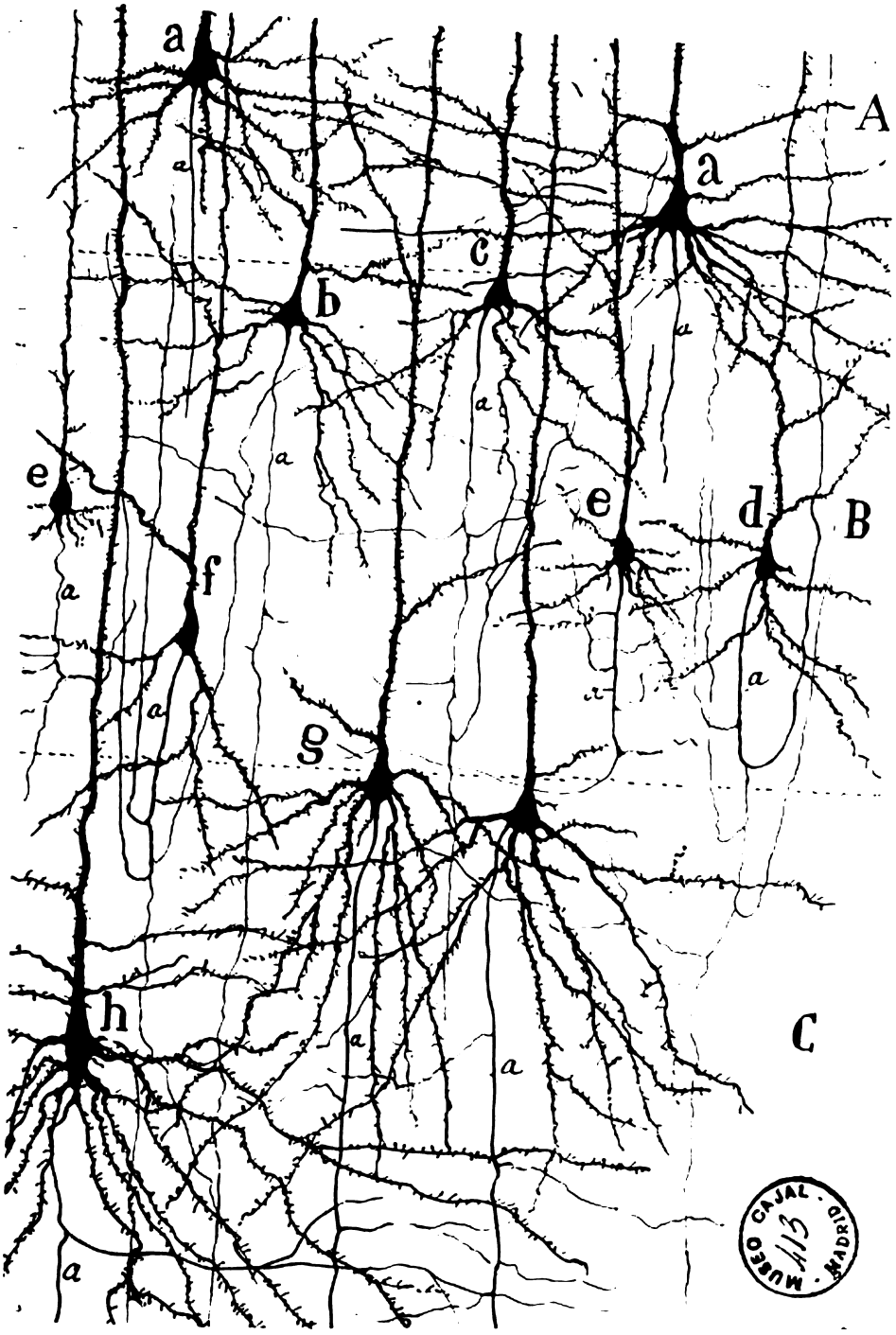



SI Figure 4.8: **Synaptic efficacy of phoneme to word connections**
The synaptic strength between phonemes and word populations is compared with the serial position of the phoneme in the word. The five panels show the phonemes-to-word connectivity for all the lexica that require memory, for the Tripod asymmetric model. The measures are averaged together in Fig. 5 of the main text. Conversely, here, the scale changes for every model.




SI Figure 4.9: **Dendritic inhibition governs signal-to-noise ratio in dendritic membrane potential**
(**A**) The average membrane potential of the dendritic and somatic compartments of a word assembly during the presentation of a phoneme sequence. The left panel shows the average potential of 100 cells randomly selected cells (control), instead, the right panel illustrates the membrane potential of the word assembly corresponding to the phoneme sequence presented (target). The colors indicate five different target membrane potentials ($V_0$) used for the v-iSTDP rule. The plot shows that for more hyperpolarized target potential (−70 mV to −80 mV) the difference between the target and the control population increases.

# 5 | A biological model of word form recognition



## Abstract

Humans efficiently retrieve spoken words from a lifelong learned lexicon in just a few hundred milliseconds. The selection of the lexical candidate follows parallel and incremental access to all the word memories matching the phonological input. This capacity is captured in models of spoken word recognition, such as TRACE, Shortlist B, TISK, and others. However, recognizing words in the face of the phonological overlap of the human lexicon imposes ad-hoc algorithmic solutions (*Temporal Order problem*, TOP). Most computational models assume position-dependent phonemic representations, which conflicts with recent experimental findings. This study proposes a biologically constrained network model with position-invariant encodings. The model recognizes words with phonological overlap and reproduces the dynamics of lexical access.

The model comprises spiking neurons with dendritic structure and represents phonemes and words as sparsely connected cell assemblies. During an association phase, the model learns to connect phoneme sequences to co-activated word assemblies through Hebbian plasticity. In recognition, words are activated from phoneme input alone. Analysis of somatic firing rates and dendritic potentials in word assemblies reveals incremental activation and lexical competition during the presentation of phoneme sequences. Temporal integration occurs in the dendrites, which implement the relevant short-term memory required for solving the TOP. This memory component is due to dendritic plateau potentials induced by NMDA spikes. We also explore the effects of phonological mismatch on a phonemic continuum. By measuring the activity of the phoneme population as a proxy for phonemic categorization, we find that the recurrent network connectivity provides lexical feedback only after lexical selection occurs but not during the early stages of lexical access.

Finally, we develop a simplified abstraction of the spiking neural network, including a hidden variable for the dendritic membrane potential. The model is a small recurrent rate network that solves the TOP thanks to the extended timescale of the hidden nodes. This research attempts to bridge the gap between biophysical models and cognitive computations in word recognition through a biologically constrained implementation of sequence detection networks.

# 5.1 Introduction

Speech is one of the most outstanding human capacities. Within less than a hundred milliseconds, the acoustic information picked in the cochlea arrives in the auditory cortex; there, the sensory inputs are integrated into the listener's cognitive experience. The first stage of speech comprehension involves the identification of words from the acoustic stimuli; in everyday speech, as in controlled settings, this process lasts 150 ms to 300 ms (Costa, Strijkers, Martin, & Thierry, 2009; Marslen-Wilson, 1973, 1987). The speed and robustness with which words are accessed catalyzed researchers' attention for a long time. However, *how* word recognition happens in the human brain remains largely unanswered (McQueen, 2007). Or, more precisely, which is the nature of the brain processes supporting it? (Poeppel & Idsardi, 2022). We attempt to answer this question through the lens of the dendritic network model. Here, we test the model against well-characterized phenomena of word recognition and discuss similarities and differences. Crucially, we show that our biologically constrained network solves a long-standing issue in connectionist models, the Temporal Order Problem, and offers a new perspective to the debate on the role of lexical feedback.

## The cognitive processes in spoken word recognition

Human word recognition (HWR) has been extensively described as a computational process (McQueen, 2007; Scharenborg et al., 2005; Vitevitch et al., 2018). Over decades, this agenda yielded a body of experimental evidence on the computations carried out by the speech recognition system residing in the human brain. First, word recognition is continuous and incremental; all the words matching the phonological evidence must be co-activated in the lexicon until disambiguating sounds are perceived (Allopenna et al., 1998; Marslen-Wilson & Welsh, 1978; Zwitserlood, 1989). Second, the processing of information between the pre-lexical and lexical levels is cascading, such that the goodness of the available phonological information influences lexical processing during the selection of the lexical candidate (Gow & Gordon, 1995; Toscano, McMurray, Dennhardt, & Luck, 2010), where the selected candidate is the word that is recognized. And third, the lexical alternatives compete, affecting the speed and probability of correct lexical selection (McQueen et al., 1994; Vitevitch & Luce, 1998). In addition, human speakers have a lexicon in the order of 10 to 100 thousand words built with only a few dozen phonemes (Maddieson, 1984).

Thus, word recognition requires a fine-grained sensitivity to the disposition of the sounds in the utterance, which is crucial for mapping the speaker's intention to the correct lexical item in the listener's mind (McQueen, Dahan, & Cutler, 2003).

These computational principles have hence guided the derivation of cognitive models of word recognition (Weber & Scharenborg, 2012) in the hypothesis that a detailed specification of the mapping between input-output transformations (computational and algorithmic levels of abstraction, Marr (2010)) might help to infer the cognitive operations supporting word recognition and, eventually, their implementation in the neural substrate (Magnuson & Crinnion, 2022). Unfortunately, neither theoretical nor experimental evidence in HWR has been sufficient to determine the complete picture of the ongoing neural processes. Computational and cognitive models can explain several of the principles described, but they remain largely unconstrained at the implementation level, and cannot pinpoint which neural mechanisms carry out the operations they describe. Two main obstacles hinder the validation of such models. On the one hand, most of the computational models have no clear linking hypothesis. It is virtually impossible to correspond the models' dimension-less entities (nodes, activation, etc.) either with the psychophysical measures, such as the response time measures - pp. 88 in Spivey, Joanisse, and McRae (2012) - or with the constituents of the neurobiological substrate (Poeppel, 2012). On the other hand, there is no agreement on which levels of analysis the models should be compared and with which priority; certain are highly accurate in the behavioral predictions but lack a neurally plausible description (*Shortlist B*, Norris and McQueen (2008)).

## Explanatory power and limits of connectionist models

Computational cognitive models can, however, formalize the wealth of data in simple descriptions and clarify the gaps in our current understanding. For example, the TRACE model, with its successes and limitations, outlines the fundamental issues that connectionist models of spoken word recognition face (McClelland & Elman, 1986). TRACE was developed to formalize the interactive hypothesis of the lexical access process, and its functioning is based on feedback and feedforward connections between lexical and pre-lexical layers. The model is a recurrent network that shares some similarities with the architecture of the human brain, that is, the recurrent network of cortical neurons and conciliates the dynamics of word recognition with general facts about the brain's biology, such as recurrence, facilitation (excitation), and competition (inhibition). Beyond the

lack of an explicit link with the implementation level, the TRACE model — and the whole class of connectionist models — has been criticized for two algorithmic assumptions that do not conciliate with experimental evidence.

The first and most debated issue with TRACE relates to the computational role of feedback connections. In interactive models, feedback makes the model robust to noise and implicitly encodes sub-lexical representations such as diphones and triphones (Elman & McClelland, 1988; Magnuson, Mirman, Luthra, Strauss, & Harris, 2018; McClelland & Elman, 1986). Lexical feedback was initially introduced in models to account for the experimental evidence of online interactions between lexical and sub-lexical representations (Ganong, 1980). In phoneme categorization experiments, listeners tended to associate the same ambiguous sounds with different phonemes, depending on whether they were presented in a word or non-word context. This evidence has long favored the hypothesis of a continuous stream of feedback information from the lexical stage, aiming to sharpen the representations in the pre-lexical one (Magnuson et al., 2018). However, the Ganong effect did not stand the proof of time. Attempts to reproduce the original study showed that pre-lexical statistics and phonotactics played a significant role in phonetic perception (Mann & Repp, 1981; Pitt & McQueen, 1998) and that the lexical bias may instead be a spurious effect (McQueen, Jesse, & Norris, 2009). Moreover, upon closer inspection, the functional principles of online lexical feedback can also be contended. Computationally speaking, feedback cannot improve word selection because it cannot integrate the lack of sensory information; conversely rather, it may cause perceptual hallucinations at the phonemic level (Norris, McQueen, & Cutler, 2000).

The second well-known problem concerns TRACE's mechanism to recognize words over continuous time. In the model, word activity depends on the combination of phonemes presented rather than their temporal disposition; order does not matter. Thus, the sequence of phonemes that belong to multiple words, e.g., phonemes within anagrams (*dog* and *god*), cause the co-activation of multiple lexical items and limit successful word recognition. Considering the sizeable phonetic overlap in human lexica, the problem is structural. In TRACE, as in many other models, the solution is to code for an additional feature that expresses the temporal occurrence of the phoneme, a temporal marker. This framework prescribed that the word *dog* is composed of *D1, O2, G3*, while the word *god* contains *G1, O2, D3*, thus only the phoneme *O2* is shared, while the others are specific to the two sequences. Critically, the approach leads to the reduplication of phonemes and words, such that the model becomes computationally

intractable even for a lexicon of small size. McQueen et al. (1994) shows that the number of necessary connections exceeds 100 billions for a lexicon of 200 words and 15 phonemes, which is *[...] arguably inelegant and implausible* solution (pp. 83, Spivey et al., 2012). A more recent connectionist model, the time-invariant string kernel (TISK) model (Hannagan et al., 2013), presents some improvements concerning the computational resources necessary for the model. The authors *fix* TRACE by introducing a position-invariant diphone matrix between the pre-lexical and lexical layers. The feedforward phonemic connections activate the diphone, which, in turn, activates the word. Although this mechanism solves the explosion of connections, it does not relieve from encoding the order of phonemes via additional features. This underlying assumption is at odds with experimental evidence, which showed time-invariant encoding of phonetic and phonemic features in the superior temporal gyrus (Gwilliams et al., 2022; Mesgarani et al., 2014), rejecting the hypothesis of time-dependent encoding of speech cues.

In synthesis, although the architecture and algorithms of connectionist models bear a resemblance with the causal dynamics of the brain, their predictions and assumptions are at odds with experimental and theoretical evidence. To solve this impasse, we propose a recurrent network model based on dendritic computation and derived bottom-up from known biological principles (Pulvermüller, Tomasello, Henningsen-Schomers, & Wennekers, 2021; Quaresima et al., 2022). The model is summarized in sufficient detail in the Methods section, and a full description is presented in Chapter 4. In the Results sections, we show that the model respects the principles of word recognition discussed and presents a parsimonious solution to the Temporal Order Problem.

## 5.2 Methods

The present chapter studies the dynamics of single-word recognition on the Tripod network model and corroborates the analysis with a minimal dynamical system reduction of the network computations. The following presents a recap of the general properties of the model implemented while a detailed description is available in Section 4.4. The details concerning the implementation of the minimal dynamical system model are presented in the respective section.

### The dendritic spiking neural network model

We investigate word recognition on a biologically constrained neural network composed of neurons with dendrites. The model is a spiking neural network composed of 2000 excitatory cells and 500 inhibitory neurons, of which 175 fast spiking and 325 slow spiking cells. The excitatory cells are Asymmetric Tripod neurons (Quaresima et al., 2022). They are composed of three compartments: two passive dendritic compartments of uniformly distributed length (150 μm to 400 μm) and a somatic AdEx compartment (Brette & Gerstner, 2005). The dendrites integrate external inputs via four synaptic receptors, two with fast timescales, AMPA and $GABA_A$ , and two with slow timescales, NMDA and $GABA_B$ . The dendrites endow the cell with a short-term memory, expressed in the depolarized membrane potential. The network is connected recurrently on the dendritic compartments; each soma-dendrite connection has a binomial distribution with $p$ =0.1. Conversely, inhibitory cells target both the somatic or the dendritic compartments $p$ =0.2. The synaptic connections among the cells are plastic; excitatory plasticity forms bonding in co-activating cells (Bono & Clopath, 2017), inhibitory plasticity aims to control the firing rate and the dendritic membrane potential to homeostatic values (Vogels et al., 2011).

### Stimuli and cell assemblies formation

The model is stimulated with sequences of phonemes and words; it receives external projections over sets of randomly sampled cells. Each linguistic entity (words and phonemes) targets 5% of the network population (approximately 100 cells). The stimulation protocol is divided into two phases; during the associative phase, the network receives pre-synaptic activity from both words and phonemes projections, and, via unsupervised Hebbian synaptic plasticity, it forms homo-associative and hetero-associative synaptic engrams.

The network model was presented with seven lexica. Five of them were already presented in the previous study (*Digits, TIMIT, Overlap, Cohort, TISK* ); in addition, we included the lexica *Labials & Alveolars* and *Velars* from the Ganong (1980) study shown in Table5.1.

Notice that, in all the lexica used, the pre-lexical representations used are not actual phonemes but the letters of the written word form. This choice was made to simplify the analysis and has no consequences on the generality of the results. Words and phonemes were presented in random order, with a 50 ms interval among each. All the items were presented approximately 50 times; the exact

Table 5.1: Lexica from the Ganong (1980) study. The table presents the words that compose each lexicon, the average phonological overlap, and the lexicon size

| Vocabulary | Words | Phon. Overlap | Lexicon size |
|---|---|---|---|
| Labials & Alveolars | dirt, deep, dark, teach, boat, babe, beef, depth, tarp, turf, page, bash, text, peace, pope, past | 4 | 16 |
| Velars | geese, keep, corpse, couch, gorge, gulp, kiss, cult, gift, garb, cars, gout | 3.3 | 12 |

number of word occurrences depended on the lexicon size, the average word length, the phoneme duration, and the simulation time. In the experiments, the network is always simulated for 5 min in the associative phase and 5 min in the recall phase.

### Firing rate and statistics

***Alpha function*** To analyze the firing rate, we convolved the neuronal spike trains with an alpha-function with timescale 10 ms.

$$r(t) = \sum_{n=0}^{N} \Theta(t - t_n) \frac{t - t_n}{\tau} \exp\left(1 - \frac{t - t_n}{\tau}\right) \tag{5.1}$$

The alpha function is skewed forward and ensures that no information on future spikes influences the previous firing states (Roth & van Rossum, 2009).

***Statistical tests*** To measure the statistics of word recognition based on the firing rate measure (Fig.5.1), we used classical p-value statistics leveraging two pre-existing Julia packages, *HyphothesisTest* and *MultipleTesting*. After slicing the word interval in chunks of 10 ms, we used the *UnquelVarianceTTest* to estimate the significance of the target word activity. The null hypothesis is that the word and the competitors have equal mean rates. To combine the word-pairs statistics, we used the Simes procedure (Simes, 1986). It implements an advanced version of the Bonferroni correction but is suitable for highly correlated T-distributions, which is the case for the populations' activity. The same procedure was used to

estimate the p-value across words and lexica. A p-value analysis was also performed in Fig.5.4. The correlation measures performed in Fig.5.3, Fig.5.3, and Fig.5.4, Fig.5.5 were computed with the package *Statistics* and the confidence intervals estimated with *LsqFit*. All the correlations annotated in the pictures are within the 95% confidence interval.

## 5.3 Results

### 5.3.1 Computational principles of spoken word recognition

#### Lexical access occurs at the uniqueness point

In the present study, aiming for comparisons with experimental results in word recognition, we analyze the word activity based on the average across trials of the population firing rate. The average is performed across approximately 70 words and over 10 network instantiations (or network configurations). The latter should be considered as the grand average across different participants. The firing rate was computed by convolving the network spikes with an alpha function (Eq.5.1). The average rate of the assemblies of four words, together with their standard deviation, are portrayed in Fig.5.1A; the remaining words and lexica are illustrated in Appendix C. The dotted and dashed vertical lines in the panels indicate the uniqueness (UP) and the offset point (OFF). In agreement with the time course of single-trial recognition presented in Section 4.2, the traces indicate there is a peak in the firing rate between 50 ms to 100 ms after the uniqueness point.

By comparing the activity of the target word population with those of the remaining words in the lexicon (lexical competitors), we also observe that the difference in the population firing rate is significant at the uniqueness point. We uncover the dynamics of word recognition by computing the firing rate's significance for each word, with a binning window of 10 ms. Hence we aligned each word's p-value measure to the onset, uniqueness point, and offset point of the word and computed the combined p-value for the entire lexicon (Methods). The three panels in Fig.5.1*B* show the time course of the p-values of the lexica for each of the three reference time frames. In the top panel, the zero of the x-axis is the word-onset time, the UP in the central panel, and the OFF in the bottom. The time course of word recognition is similar across the seven lexica, in the face of variations in lexicon properties (word length and average phonological overlap, Appendix A).

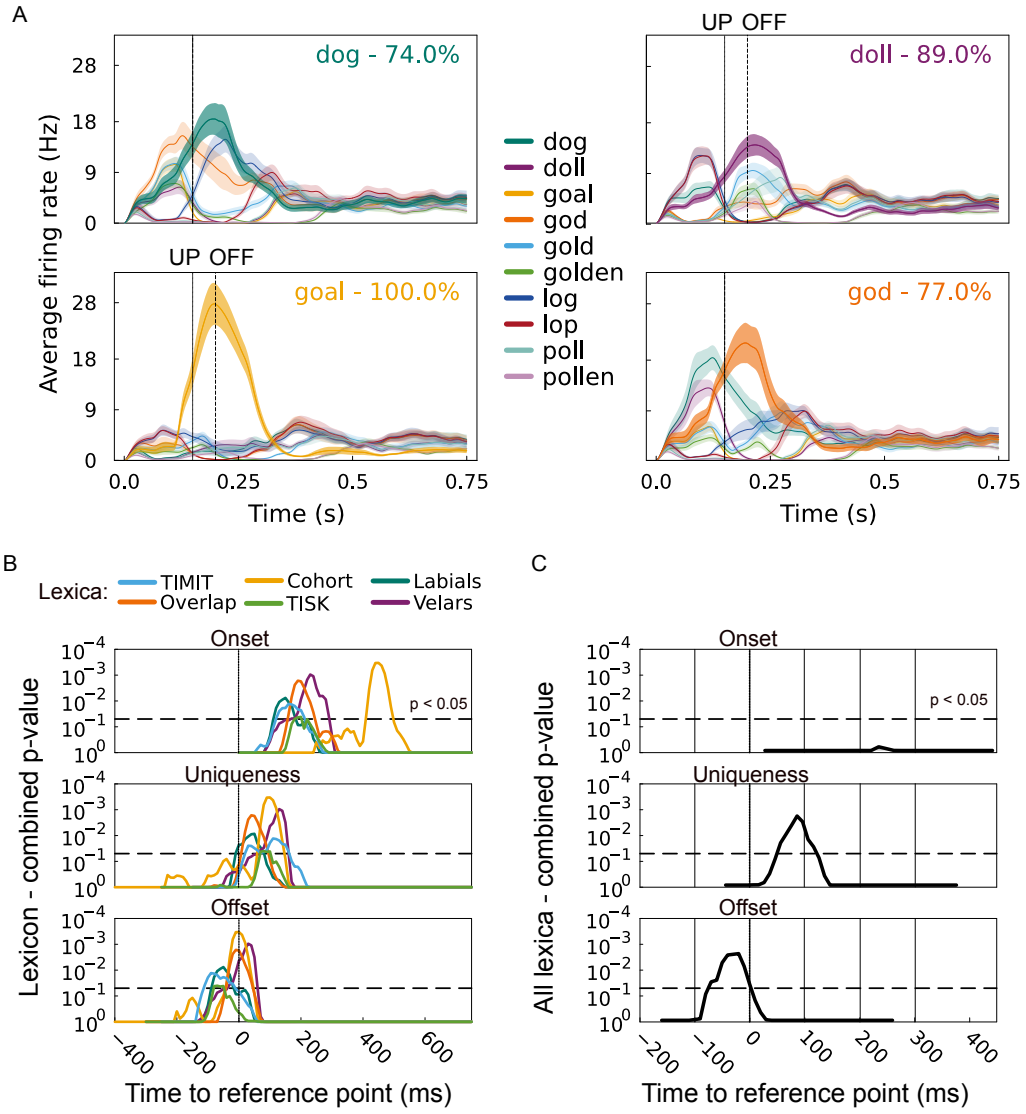



figure 5.1: **Grand-average measures of the firing rate show significant lexical access**

*Figure caption continues on the next page.*

Eventually, we combined the statistical significance of the seven lexica in Fig.5.1*C*. The offset and uniqueness point curves remain significant, but the onset curve does not (top panel). Word's activity peaks before the offset point (bottom panel) and within 0 ms to 100 ms from the UP (middle panel). Considering the network delay between the activation of the phonemes and the word population, we conclude that the phonemes presented at the offset time are not causal to word recognition and that the UP is the crucial time point for lexical access in the grand average measures. The non-significance of the onset curve is because words have different lengths, and the distance between the uniqueness and on-

figure 5.1: **Grand-average measures of the firing rate show significant lexical access**
(**A**) Grand average of word activation based on firing rate; the panel shows four examples for the lexicon *5. Overlap*. The firing rate is obtained by averaging over all the word presentations in each experiment. The ribbon indicates the standard deviation across ten datasets. The activity of the target word population and its competitors is color-coded and the percentage indicates the word recognition $\kappa$-score (Appendix A). The firing rate of the target populations peaks approximately 50 ms after the uniqueness point (UP, dotted line). In the four examples, the offset point (OFF, dashed line) matches the uniqueness point (UP solid line) only in the first and last panels. (**B**) We measure the p-values of the firing rate binning with 10 ms intervals. The panel indicates the combined statistics of words' p-values across each lexicon. When aligned to the onset, uniqueness, and offset points, the curves reveal that lexical access is significant at a fixed distance from the uniqueness and offset points. (**C**) The p-values are also combined across lexica; the three curves represent the significance statistics combined across all the lexica. In the three panels, the curves are aligned to the onset, uniqueness, and offset points. The UP appears to be the critical time point for lexical access because words are significantly recognized *after* the phoneme presented at the UP.

set points is variable. Thus, the intervals in which words are significant do not match. Conversely, the pool of words used presents less variability in the intervals between the uniqueness and the offset points (the OFF follows UP of one or two phonemes on average).

The present results indicate that the Tripod network has a time course of word recognition dynamics compatible with psychophysical (Marslen-Wilson & Welsh, 1978) and electrophysiological (Gwilliams et al., 2018; Winsler, Midgley, Grainger, & Holcomb, 2018) evidence. Further analysis of the dynamics of word recognition is presented in Appendix B, where we look closer at the dendritic membrane potential. The analysis reveals that the peak depolarization of the average membrane potential of the word population correlates with the single word score (Appendix A) while the peak firing rate response does not. We infer that although the firing rate of the word populations can be used as a read-out of word recognition, as shown in this section, the membrane potential is the causal variable for lexical access and selection.

### Effects of lexical neighbors on target word activation

The grand-average analysis of population activity suggests that lexical access occurs after the uniqueness point. However, in our experiments, as in natural languages, some words are more challenging to recognize than others. This

is testified by the variations in single word scores (Fig.5.1*A* and Appendix A). To explain the differences within our lexica, we focus on the impact of lexical neighbors on the word recognition score. Words that share sequences of sounds (phonemes) with the target words are co-activated until sensory or contextual evidence is sufficient for lexical selection. In addition, words compete for activation and, consequently, the number of lexical neighbors impacts the time course and success probability of recognizing the target word (Connine, Blasko, & Titone, 1993; Luce & Pisoni, 1998; Marslen-Wilson & Zwitserlood, 1989; Morton, 1969; Toscano, Anderson, & McMurray, 2013; Vitevitch & Luce, 1998).

We test the co-activation of lexical competitors and target words by comparing the average rate of four groups of words: the target words, the cohort, the rhyme, and the non-related words. The cohort class accounts for the words that share the same onset sequence (two or more phonemes). Conversely, rhyme refers to words with the same ending sequence (two or more phonemes). To compare the activity of these four categories, we estimated the activity of each as the average across all the competitor words in the pool. To align the activity profile of words with different lengths (3 to 10 phonemes), we used the uniqueness point as the center of the word, similarly to how it was done in the previous section. The results are presented in Fig.5.2*A*. The top panel shows the average rate over time of the word categories; as expected, the target word, or referent (green), has a pronounced peak 100 ms after the uniqueness point. The two classes of lexical neighbors, i.e., cohort (purple) and rhyme (yellow), are also activated above the average of the remaining words (orange). The curves show clear signs of the phonemic inputs in small bumps at the frequency of 20 Hz; these bumps are not visible in the target word average for two reasons. First, there are more samples for the target words than the competitors; second, the target word activation via the recurrent connections rather than the phonemic projections.

Because of the sequential order with which phonemes are presented, the cohort group activates at the word onset. In turn, the rhyme peaks after the uniqueness point. To increase the statistical power, we averaged segments of time together in intervals of 100 ms. For each, we computed the average activity of the four groups. The result is the bottom panel in Fig.5.2*A*, starting from the interval (−300 ms to −200 ms) before the UP. The interval averages indicate that the cohort group is significantly activated, above the rhyme and target words, in the two intervals before −100 ms. At the uniqueness point, the cohort and

rhyme words have similar activity, but rhymes remain high in the two following intervals while the cohort group slowly fades.

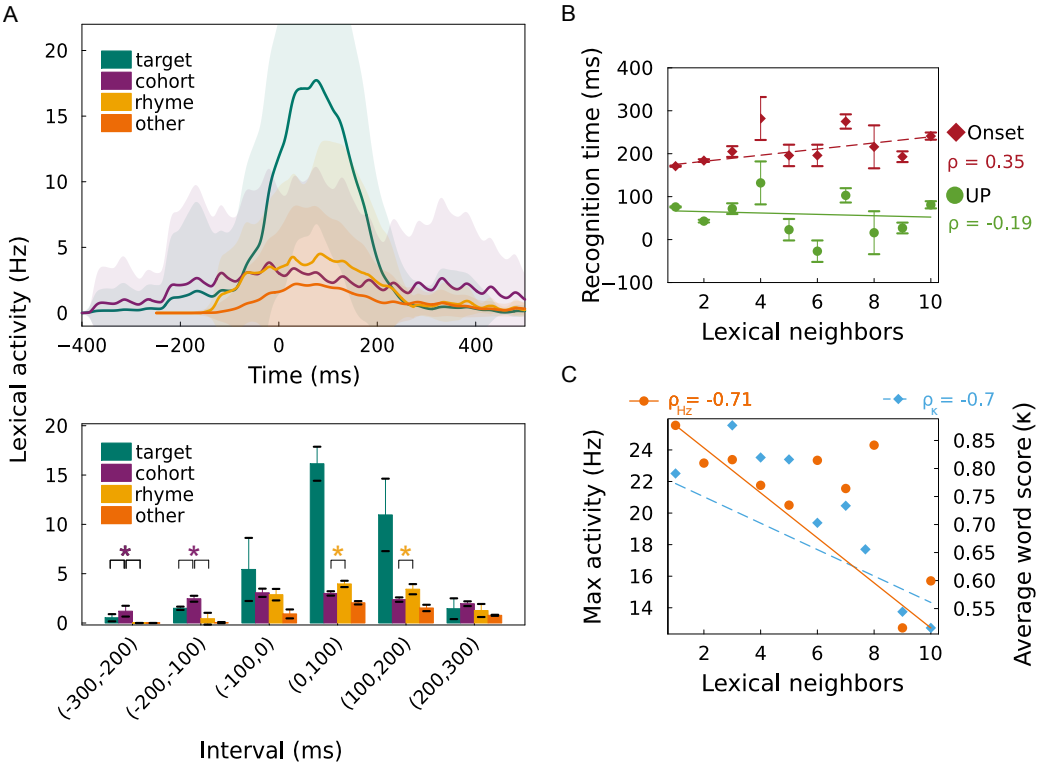



figure 5.2: **Lexical neighbors activate and affect word recognition time and accuracy** (**A**) The top panel shows the average rate of target, cohort, and rhyme words. Words are aligned to their uniqueness point in the origin of the x-axis (0 ms). Each curve was chunked in intervals of 100 ms; the bottom panel shows their average and the standard deviation. The cohort words are significantly more active than the target and the rhyme words in the early intervals (purple stars); similarly, the rhyme words are more active after the target word uniqueness points (yellow stars). (**B**) Measures of the average recognition time as a function of the number of lexical neighbors: words that share two or more phonemes in the lexicon. The recognition time is measured as the peak of the firing rate response of the target word population. The recognition time from the word onset increases significantly with the number of lexical neighbors but decreases when the time is measured from the uniqueness point. (**C**) The strength of the target population response (orange, solid line and circles) and the recognition score (blue, dashed line and diamonds) decreases when the number of lexical neighbors increases.

We probed the ongoing competition between word populations by measuring the impact of lexical neighbors on the time course and accuracy of lexical access. This time, we grouped the target words by the number of lexical neigh-

bors. Fig.5.2*B* shows the time of the firing rate peak measured from the onset and uniqueness points. An increase in the number of neighbors delays lexical selection (positive correlations with onset time, red diamonds) but accelerates recognition after sufficient evidence is presented (negative correlation, UP, green circles). These trends are consistent within a 5% confidence interval - and in line with the average recognition delay measured in Appendix A. Lexical neighbors also affect the recognition score. When the number of competitors augments the score decreases (Fig.5.2*C*). We measured both the population peak rate (Hz, blu circles) and the word recognition score (orange diamonds). The presence of words with shared phonemes hampers recognition significantly; the population response drops by 50% ($\rho_{Hz} = -0.71$) and so does the recognition scores, from 0.85 to 0.5 ($\rho_{\kappa} = -0.7$).

**Early phonemic mismatch reduces, not prevents, lexical access**

The third computational principle of word recognition that must be addressed concerns the impact of phonetic mismatch on word recognition. Because information cascades from the phonemic level to the lexical one in a continuous fashion, the occurrence of phonological misalignment reduces the probability of correct lexical access in a graded manner (McQueen et al., 2003; N. O. Schiller & Meyer, 2003). As a consequence, the accuracy and time course of lexical access depends on the phonological support provided for the target word (Andruski, Blumstein, & Burton, 1994; Dahan, Magnuson, Tanenhaus, & Hogan, 2001).

We investigate the network response to ambiguous phonological information by measuring word accuracy on words with gradually distorted inputs. To this aim, we used the two lexica from (Ganong, 1980), (*Labials & Alveolars* and *Velars*). The words composing them are four or five phonemes long and begin with a voiced or voiceless phoneme, they are paired together such that the acoustic continuum moves from a non-word to a word in one direction and from a word to a non-word in the other direction (e.g., *Dext-Text* and *Dark-Tark*). In the original experiment, the authors mixed the phonemic stimuli by modulating the voice-onset-time (VOT) of a synthetic speech input on a continuum, similarly, we modulated our phonemic input by mixing the external projections on the two target phonemes (*b-p*, and *d-t* in the first lexicon, and *g-c* in the second one). We encode phonetic ambiguity through displacements of the external projections; synaptic afferents are moved between the voiced-voiceless pairs according to the position in the continuum. The population coding of the VOT feature is compatible with electrophysiological evidence on spatial coding of temporal speech cues

(Fox et al., 2020). The mixing protocol is illustrated in Fig.5.3*A*. The continuum was divided into seven steps (the voiced phoneme received the 15 to 85 % of projections activity and the voiceless one its complementary amount). The raster plots in Fig.5.3*B* show that reduced external inputs imply weakened activation of the phoneme populations.

We run ten network simulations for each word pair where the entire phonemic continuum is presented - only one pair receives the ambiguous input - for a total of 140 simulations. The results are aggregated based on the continuum and presented in Fig.5.3*C*. The two curves indicate the average word score over the words starting with the voiced consonants (moving from word to non-word, purple) and those starting with voiceless phonemes (word to non-word, green). In both directions, the words converge to the optimal recognition score (0.90 to 0.95) as an exponential function of the phonological evidence presented; at each step of the continuum, additional evidence on the phoneme identity has less informational value. We should notice that the two groups do not reach the same average recognition score, most likely due to the unbalanced distribution of the phonemes used, which was not controlled.

Further insights on the dynamics of the word populations were gathered by recording the word assemblies' rate activity and membrane potential across the phonemic continuum. The panel in Fig.5.3*D* illustrates the two measures. In this case, we considered only the trials where the target words start with voiced consonants and used the other group as a contrast; we limited the analysis to words with four phonemes. The panel in Fig.5.3*E* shows the average rate of the target populations. Direct inspection of the curves indicates that moving on the phonemic continuum (red to light-blue) decreases the word assembly rate since the early stages 0 ms to 50 ms. The target populations (solid lines) and the contrast (dashed lines) increase their firing rate over the word interval, with graded intensity, depending on the degree of phonetic mismatch. Crucially, 50 ms before the word offset, which corresponds roughly to the average uniqueness point, the two populations set apart; the target one reaches peak activity, and the contrast fades away. In both cases, the traces of the phonemic mismatch converge together, such that the rate does not contain information on the phonemic continuum right after the words offset 200 ms to 250 ms. The target population reaches approximately the same maximum over the entire continuum. Eventually, the average rate decays upon the arrival of the following word. The membrane potential Fig.5.3*F* (blue to yellow color-coded) follows a similar trajectory. However, its variability within the phonemic continuum is smaller than

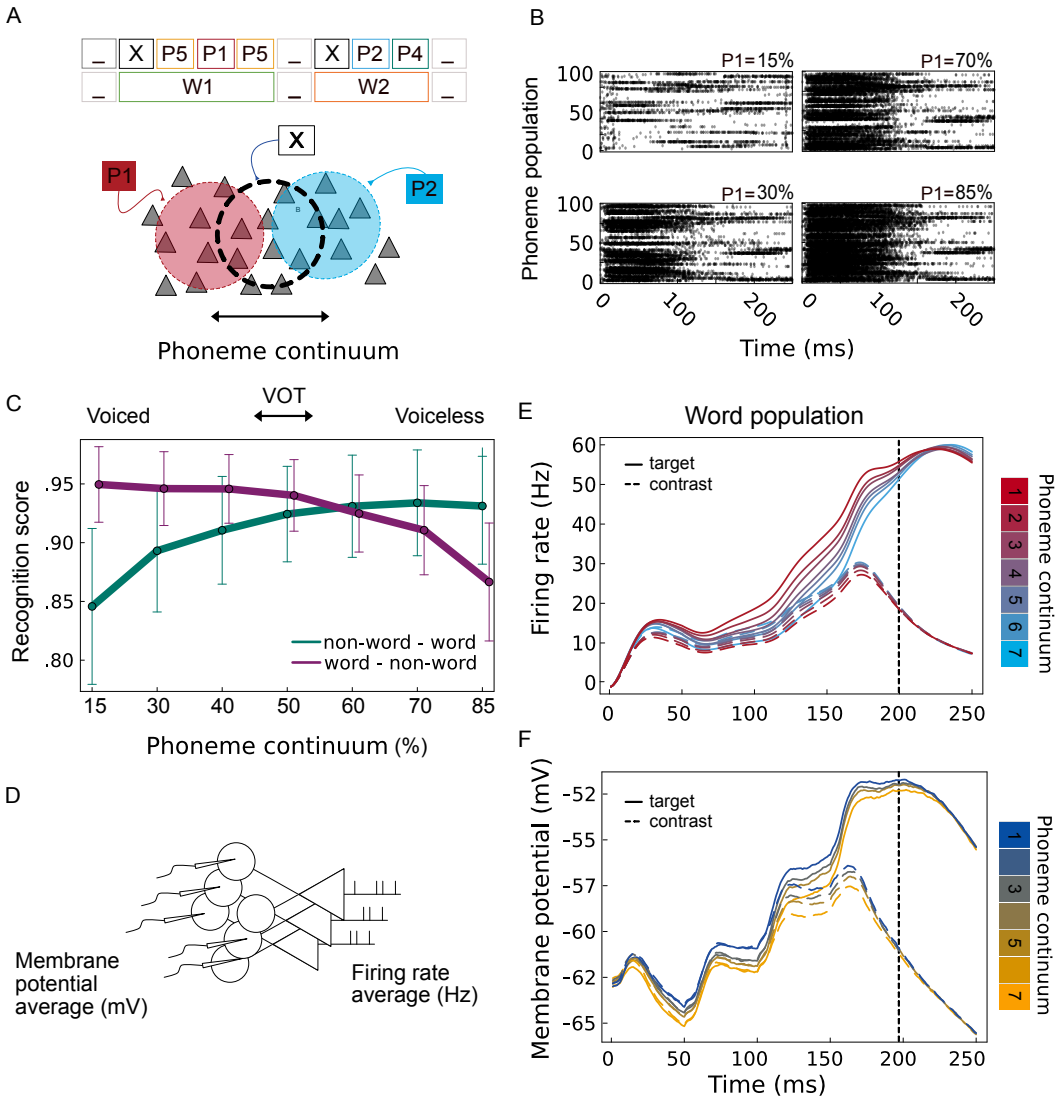



figure 5.3: **Phonemic mismatch has graded effect on word's score and population dynamics**

*Figure caption continues on the next page.*

in the rate case. The differences between target and contrast words are also small in the early word stages. As a consequence, the two sets of conditions cannot be fully distinguished in the early interval (0 ms to 100 ms). Following the uniqueness point, the membrane dynamics present a more neat distinction between the target words with full and partial phonemic match. This difference is noticeable in the peak membrane potential at 200 ms, which is larger when words have more phonological support. We observe gradation in the membrane potential across the phonemic continuum. A weaker membrane depolarization corresponds to larger ambiguity and lower $\kappa$-score . On the other hand, the peak

figure 5.3: **Phonemic mismatch has graded effect on word's score and population dynamics**
(**A**) Scheme of phonemic continuum manipulations. The populations of *P1* and *P2* are interpolated by a weighted sampling among their projections. The resulting phoneme population is indicated by the placeholder phoneme *X*. (**B**) Variations in the continuum result in weaker or stronger phonemic population activity. (**C**) Average $\kappa$-score for the words starting with the voiced and voiceless phonemes, the VOT reduces left to right and the percentage of active projections (x-axis). The score (y-axis) depends on the phonemic continuum and increases exponentially when phonological ambiguity is reduced. (**D**) Schematic of the membrane potential and firing rate measures; the former averages the membrane potential of the dendritic compartments and the latter concerns somatic activity. (**E**) Average firing rate of voiced-onset words (target, solid lines), compared to voiceless-onset ones (contrast, dashed lines). Only words with four phonemes were considered. The average rate is color-coded for its position in the phonemic continuum. Red represents the voiced extreme (15%) and light blue the voiceless one (85%). (**F**) Likewise *E*, here, the measured variable is the dendritic membrane potential. We plot only four continuum points (1,3,5,7) to ease the visualization. The impact of phonemic mismatch on both measures increases throughout the word interval (graded rate curves, 50 ms to 200 ms) but disappears after lexical access.

rate activity is not modulated by the phonemic continuum. The dissociation of rate and membrane potential is a pivotal feature of our model.

### 5.3.2 Late lexical bias in the phonemes' assembly membrane potential

The correspondence between the phonemic continuum and word $\kappa$-score indicates an interaction between the sub-lexical and lexical levels during word recognition. The modes and nature of this interaction have been the center of a long-standing debate since the seminal work by Ganong (1980). The original experiment indicated the presence of a lexical bias in the perception of the phonemic units; ambiguous phonemes are attracted towards the word-like end of the continuum. In the interactive hypothesis, upon presentation of an ambiguous input, the lexical bias, or Ganong effect, is expressed as activation of the phoneme consistent with the lexical context. For example, *Xast* should reactivate *p* over *b*, because the first corresponds to a word in the lexicon and the latter does not (*past* vs.*bast*). We reproduced with our model an analysis similar to the original study and tested the strength and the time course of feedback activity between the lexical and pre-lexical populations. In order to avoid ef-

fects due to phonotactic regularities, e.g., the repetition of one of the measured phonemes (as in *babe*), we removed all the words that contained a repetition of either the voiced or voiceless phoneme, reducing the lexica to 9 pairs of words.

Hence, we measured the time course of the average population rate (Fig.5.4*A*, red to light-blue) and membrane potential (Fig.5.4*B* blue to yellow) of the voiced phonemes during the word (solid) or non-word (dashed) conditions. The intensity of phonemic activity in the early stage (0 ms to 100 ms) is proportional to the percentage of voiced-population's neurons targeted by the mixed projections in the continuum, larger at the left end of the continuum. This is the case for both the rate and the membrane curves. In contrast, at the late stage of the word interval, 150 ms to 200 ms, the rate continues to decay to the homeostatic range 3 Hz to 5 Hz but the membrane potential undergoes a second peak. The depolarization of the phoneme population's dendrites is due to the feedback from the recurrent connections in the lexical stage. The difference between the word and non-word conditions is significant only within this late peak.

To favor the comparison with the experimental literature, we devised a simple measure that resembles the phonetic decision task in the original experiment. We selected two relevant intervals, early ([*ON, ON* + 50 ms]) and late ([*OFF* - 50 ms*, OFF* + 50 ms]), where *ON* and *OFF* indicate the onset and offset points. Thus, we compute the average population rate and membrane and compare it across non-word and word conditions. Following the methods of the original study, in Fig.5.4*C*, we measure the average difference in the activity of the phoneme population, e.g., *p* and *b*, when in the word to non-word continuum (green) and in the non-word to word direction (purple). If present, the difference between the two continuum directions (purple minus green) must be interpreted as a bias effect due to the lexical context; such measure is portrayed in Fig.5.4*D*.

The analysis in Fig.5.4*C* indicates that, in the early intervals (left column), the strength of the phonemic populations depends on the inputs from the projections; inputs on the voiced side of the continuum (x-axis, left) correspond to more robust responses in the voiced population (y-axis, top), and similarly, inputs on the voiceless side cause stronger activation in the voiceless phonemes. Crucially, this appears the same for both directions of the continuum, indicating the absence of the Ganong effect, both in the rate (top-left) and in the membrane potential (bottom-left) in panels *C* and *D*. The lexical bias effect appears later in the word interval. Our measure indicates that although the phoneme rate is close to the baseline activity, the membrane potential is not. The feedback

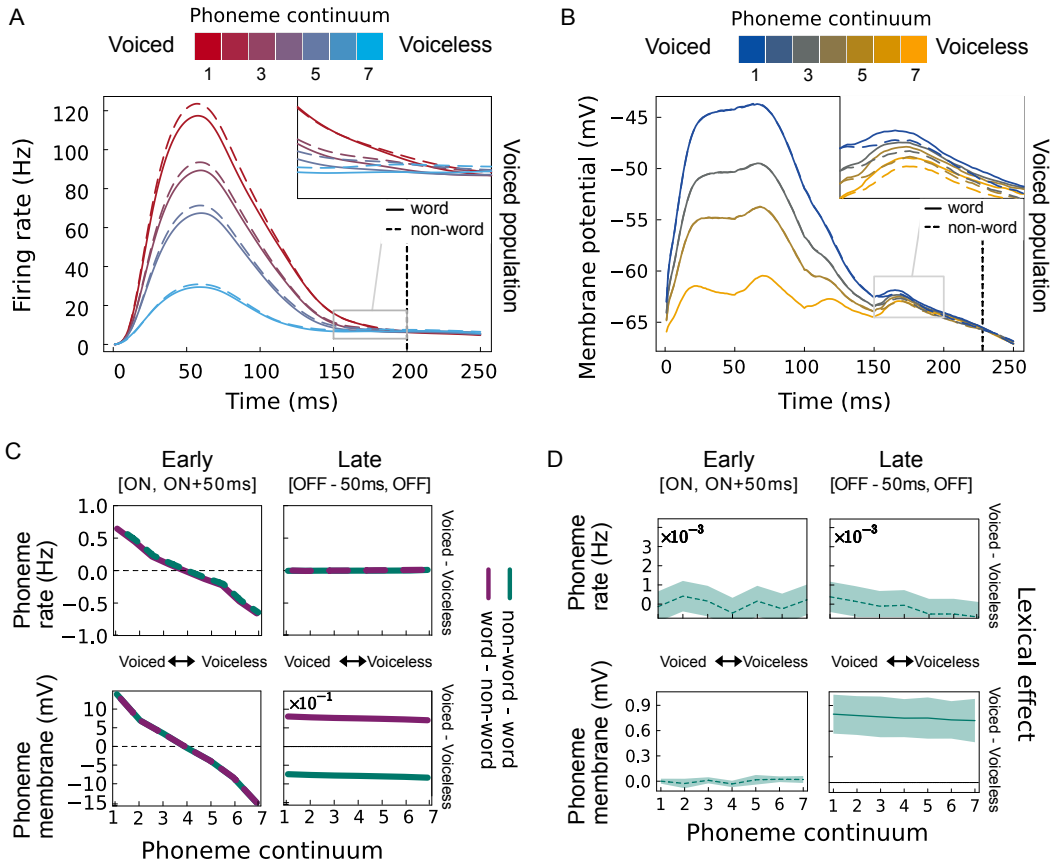



figure 5.4: **Word recognition depends on lexicon properties**
(**A**) Average firing rate of the voiced phoneme populations (*B, D, G*) when presented throughout the phonemic continuum (red to light-blue), in the two lexical conditions (word, solid line; non-word, dashed). The rate response depends on the continuum; voiced phonemes are less active at its right end, where the inputs correspond to 85% of the voiceless phoneme population (cfr. with Fig.5.3*A*). The rate converges to baseline at the late interval stage, 150 ms to 200 ms. (**B**) Likewise, in *A*, the curves represent the average membrane potential. The dynamics are analogous only in the early phase of the dendritic membrane. After lexical access, the potential has a bump, significantly larger in word (solid line) than in non-word (dashed line) conditions. (**C**) Average difference of the activity in the voiced and voiceless populations across the continuum. The four panels illustrate the measure of the rate (top) and membrane potential (bottom) at an early interval (0 ms to 50 ms, left) and at a late interval (offset minus 50 ms to offset). Only the membrane potential measured at the late interval (bottom right) significantly differs between the two traces. (**D**) Size of the lexical bias effect in the four conditions described in *C*. The lexical bias effect in phonemic population activity is present only in the membrane potential at the lexical access time; it does not depend on its position on the continuum, and it remains significant in all the measures.

connections provide a significant increase in the dendritic depolarization of the phoneme population during the late interval (0.8 mV, $p < 0.001$). The effect is approximately the same across the continuum, which means it does not bear memory of the amount of phonemic mismatch at the onset period.

The results provided in this section indicate that lexical access can be achieved upon mismatching input without the help of feedback activity. The activation of word populations bias the phoneme assemblies only after lexical selection and it is insufficient to trigger the burst of the phonemic populations. However, it can be detected in the subthreshold membrane potential. This difference may speak for a different functional scope for the feedback connections; dendritic depolarization can silently act on the dendritic synapses and favor the re-organization of the network connectivity, for example, to adapt to a novel sound context (Norris, 2003). Crucially, the model we implemented has a recurrent component, usually associated with interactive models. We argue that what sets the model apart is the presence of two classes of variables: the observable variables, expressed by the neuron firing activity, and the slow-decaying hidden variables expressed in the dendritic membrane timescale. To prove that including these elements makes a good word recognition model, we show in the following section that a reduced version of the network with dendrites can address one of the fundamental problems of word recognition: the issue with time. Further evaluations of the model response in the lexical-bias experiment, along with the necessary criticisms of the current implementation, are presented in the Discussion.

### 5.3.3 Word recognition in a minimal dynamical system: long timescales and inhibition are enough for recognition

The present network model shows that the word recognition dynamics can be retrieved without explicit encoding of phoneme onset time (position-dependent representations). We hypothesized that such capacity can be isolated from the network and belongs to two simple computational primitives: the implementation of long timescales in the network's hidden states and the presence of inhibitory feedback. To this scope, we explored an extremely simplified computational model obtained from reducing the network with dendrites to its bare-bone components. For each word and phoneme population of the network, we substituted it with a single node in a directed graph. An example of the graph corresponding to the *Overlap* lexicon is portrayed in Fig.5.5*A*. The connections among the nodes correspond to the average synaptic weights in the original dendritic

model. The resulting graph is a recurrent network, and its connections contain all the information necessary to achieve correct lexical retrieval. It is important to notice that such connections were learned through Hebbian plasticity in the spiking network, and we offer no solutions to derive them in the reduced model. Similarly, the dashed arrows express the external projections on the phoneme populations. In order to favor the stability of the network, we normalize the connectivity matrix column-wise ($T_{ij}$). The matrix $T_{ij}$ expresses the activation of the post-synaptic node $i$ after the activity on the pre-synaptic one $j$.

The first question is whether $T_{i,j}$ can compute the correct word starting from the input phonemes when the order of phonemes is not expressed. We assume that the activity on the phonemic projections is homogenous for all the input phonemes and compute a word input vector ($I_i$) Hence, we multiply the input vector times the transition matrix and obtain an output vector ($O_j$), which indicates the graph's output state based on the phonemic input. The operation is a simple matrix multiplication; $O = T \cdot I$. For example, the input vector for the word *golden* contains the phonemes *O, N, L, G, E, D* (right side of Fig.5.5*B*). After the matrix product, the node with the strongest activation corresponds to the word *golden* (left side). We verified that the correspondence between inputs and words holds for all those input vectors with partial phonological overlap. In contrast, when the lexicon contains anagrams, words composed of the same phonemes as in *dog* and *god*, the correct word cannot be activated because the input vectors are intrinsically ambiguous for the two words (*D, G, O*, SI Fig.5.4 *A*).

To overcome the limits posed by the lack of temporal information and retrieve the sequential structure of the input vector, we encode the inputs with an alpha-function (Eq.5.1) with timescale $\Delta = 50$ ms and centered at the phoneme onset time. The function aims to represent the activity over the projected phoneme populations. The phonemic inputs obtained are illustrated in Fig.5.5*C*; they overlap each other, and the only difference between the input vectors of the two anagrams (*dog, god*) is the onset point of the alpha function: the identity of the phonemes is position independent. The input information is thus integrated over time with a simple dynamical system that expresses the three main features of the network with the dendrites: the synaptic configuration, the long dendritic timescales and feedback inhibition. The state update of the dynamical

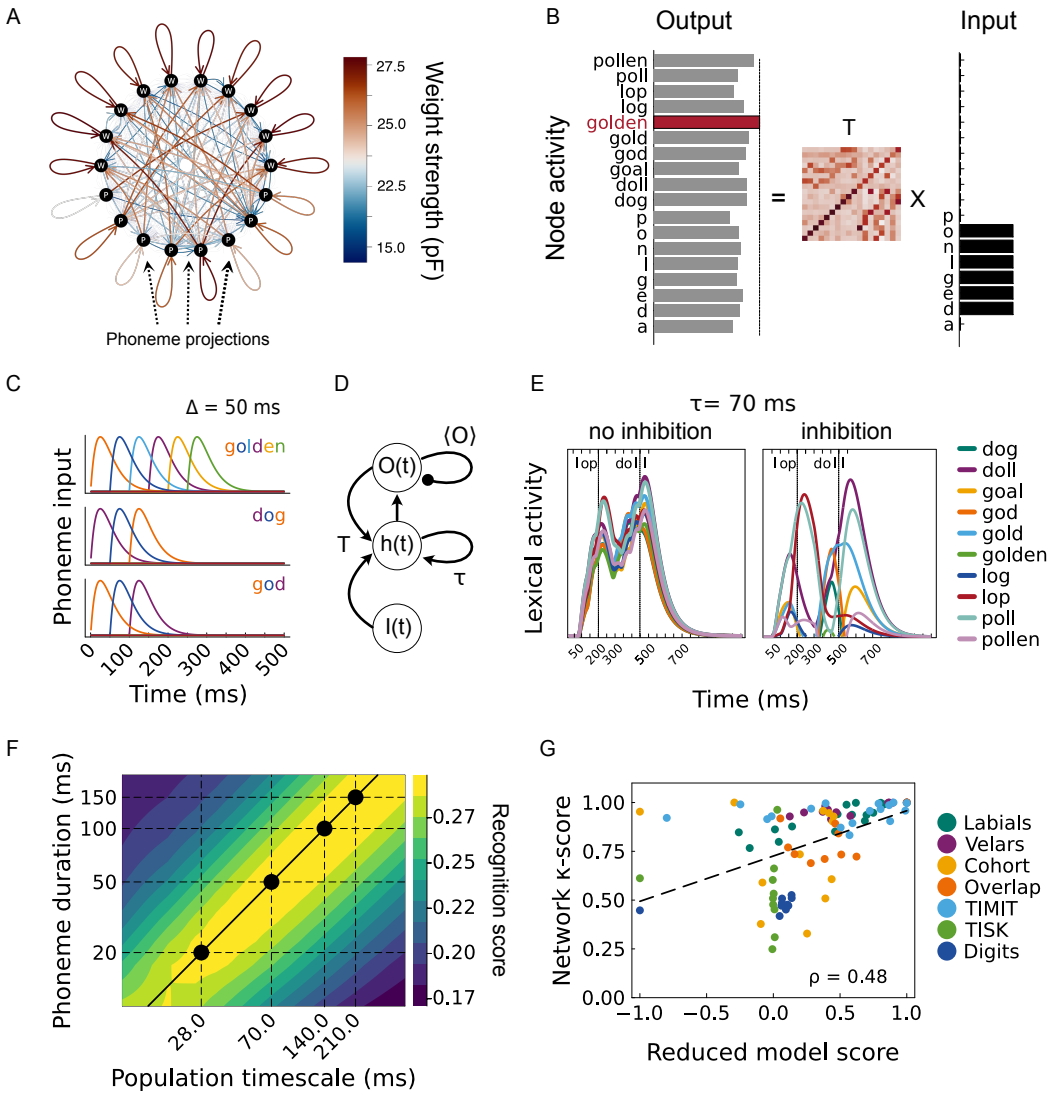



figure 5.5: **Minimal interactive model of word recognition**
(**A**) Network representation of the average synaptic connections among the phonemes and word populations. The network has feedforward and feedback connections. (**B**) The normalized connectivity matrix, or transition matrix $T_{ij}$, projects the phonemic inputs into word nodes' activity. The panel shows the activation of the word *golden* upon presenting a set of phonemes stripped of the temporal information. (**C**) Temporal information is reintroduced in the reduced model by associating an alpha function with the phonological input. The onset time of the function is aligned with the phoneme onset time but the position is not explicitly coded.

figure 5.5: (**D**) The inputs are integrated into a minimal dynamical system. The inputs ($I(t)$) stimulate the hidden state of the system ($h(t)$) via the transition matrix ($T$). The hidden state has a slow decay with timescale $\tau$ and determines the state of the observable ($O(t)$) which expresses the node activity. Finally, $O$ undergoes feedback inhibition, which maintains the activity within a boundary; the inhibition is proportional to the average value of the observable itself. (**E**) The model has scarce word recognition accuracy when inhibition is omitted (left), but performances improve when the control mechanism is in place. Crucially, the model can proficiently process words in a sequence. (**F**) The recognition score of the reduced model depends on the ratio between the phoneme timescale ($\Delta$) and the population timescale ($\tau$). The recognition score peaks when the latter is approximately 1.4 × the former (straight black line). (**G**) Within the same parameter range, the word $\kappa$-score score of the extended network model correlates significantly with the recognition score of the simplified model.

system is also governed by the transition matrix $T_{ij}$. Two equations fully define the dynamical system:

$$\tau \frac{dh_i}{dt} = -h_i(t) + \sum_j T_{ij}(I_j(t) + O_j(t)) \tag{5.2}$$

$$O_i(t) = h_i(t) - \langle h(t) \rangle \tag{5.3}$$

The first equation describes the dynamics of a hidden variable $h$, which loosely represents the membrane potential of the network with dendrites. The hidden state has a decay timescale $\tau$ and receives feedforward ($I_j$) and recurrent ($O_j$) inputs through the transition matrix. The second equation controls the amplitude of the graph's observables; it acts as the homeostatic mechanism generally associated with inhibitory processes. The scaling operation in Eq.5.3 is strictly necessary to implement the recurrent connections; otherwise, the system enters into a run-away activity state, and the dynamical system diverges. The three variables in the system, $h_i; I_i; O_i$, represent the hidden population states, the input vectors, and the observable variable, respectively. They evolve over continuous time and are approximated by discrete equations with the Euler (first-order) finite difference method. The dynamical system is illustrated in Fig.5.5*D*.

Next from the recurrent matrix, the reduced model has only two variables: the timescales of the phonemic projections ($\Delta$) and the hidden variable ($\tau$). For a quick insight into the dynamics of the model, we tested the activation of two successive words, separated by 100 ms of silence. Fig.5.5*E* illustrates the word nodes activity in models with and without inhibition. The comparison indicates that the target words are correctly reactivated in the model with inhibition (right

panel, red and purple peaks) but not when the model is purely feedforward (left panel). In the inhibition-less model, the inputs saturate the state vector, and all word populations get activated. The node activity timescale used in the example is $\tau$ =70 ms, similar to the timescale of dendritic processes. For longer or shorter timescales (e.g., 20 and 150 ms, SI Fig.5.1 *B*) the recognition process is less accurate. We examined this relationship systematically in Fig.5.5*F* comparing the average recognition accuracy among the lexica. To measure the model accuracy, we computed the average difference between the target node's activity and the remaining competitors within 150 ms from the uniqueness point; in this estimate, values above zero indicate that words are recognized above chance. The method is illustrated in greater detail in SI Fig.5.2 . Fig.5.5*F* shows a region of parameters that reaches a significantly higher recognition score. When the population timescale ($\tau$) is approximately 1.4 times the phoneme duration ($\Delta$) word recognition is maximal. The same trend applies to all the lexica measured (SI Fig.5.1 *C*) except the *TISK* and *DIGITS*. Crucially, when the timescales are chosen in agreement with the network's ones (phoneme duration, $\Delta$ =50 ms; dendritic integration, $\tau$ =70 ms), the reduced model's scores have a significant correlation with the word $\kappa$-score of the extended dendritic model ($\rho$ =0.48, Fig.5.5*F*).

The reduced model achieves correct lexical access with minimal ingredients. Thus, the minimal computational principles it relies on are sufficient for the word recognition task. It retrieves correct words based on phonemes whose feature space lacks a temporal marker. Unlike the network with dendrites, constructed with fidelity to known biological processes and dimensions, the reduced model has two loose ends. First, the integration of recurrent and feedforward connectivity is not commensurate. The two streams should be scaled such that the output of the transition matrix remains unitary; in the present case, the streams are summed together, and the model output is not interpretable in biophysical terms. Second, the reduced model does not include non-linearity in the signal integration; so far, our attempts to use the non-linear transformation of the hidden state (e.g., sigmoid function) have not been successful; the model diverges. Because non-linearity is one of the canonical aspects of neuronal dynamics, we think future work should elaborate further on this inconsistency.

In conclusion, the reduced model presented in this section offers an algorithmic-level interpretation of the implementation-level model hitherto analyzed, the network with dendrites. Albeit its recognition is far from perfect, it shows that

simple, interactive models of word recognition can get rid of explicit temporal information for the integration of phoneme sequences.

## 5.4 Discussion

The present work investigated the neural dynamics of spoken word recognition in a network model with dendrites. The study is divided into three sections. First, we show that the model is consistent with the computational principles derived from psychophysical and electrophysiological experiments (Gow & Gordon, 1995; Luce & Pisoni, 1998; Marslen-Wilson, 1987; McMurray et al., 2022; McQueen, 2007; Toscano et al., 2010; Vitevitch & Luce, 1998; Winsler et al., 2018; Zwitserlood, 1989). These are, in short, the following. Lexical representations are incrementally activated during the presentation of phonemic evidence; the presence of lexical neighbors delays and obstacles lexical selection, with both cohort and rhyme lexical classes being re-activated; and partial phonemic mismatch degrades the word recognition score but keeps its time course intact. Later, we derived the model's predictions on the lexical bias effect and observed how it aligns with the perceptual learning hypothesis for feedback between lexical and pre-lexical representations. Finally, we obtained a highly reduced version of the model, which retains similar performances and exposes the core computational mechanisms driving phonemic integration: the long integration timescales of dendritic non-linearity and recurrent inhibition.

The results demonstrate that detailed neurobiological implementations, such as models of the dendritic structure in cortical neurons, lead to a parsimonious mechanistic description of human word recognition. The network responded to phonemic stimuli on a realistic time course (50 ms) with assembly activity on timescales comparable to the actual system, the brain. This was achieved without fitting the network's parameters; they were obtained from bottom-up modeling of human neuronal physiology (Chapter 2) and mammals' cortical networks (Chapter 4). Additionally, we stress that the recognition score achieved for single words may be a lower bound of the network capacity; indeed, the delayed activity of the word assemblies exceeded the timeframe in which the phonemes were presented, causing an interference with the words following. We do not account for this interaction among successive lexical items nor attempt to minimize it with longer pauses in between words. Instead, we deem that such interference proves that activity at the lexical level is robust against

the incoming sensory input, and this property may allow for sentence-level unification on longer timescales.

It is also essential to notice that we tested spoken word recognition based on the average activity of the word populations, with the firing of single cells averaged into a single variable for each word presentation. In contrast, the two electrophysiological studies that inspired the present work (Chan et al., 2014; Vaz et al., 2020) leveraged single-unit recordings. Nonetheless, our model would maintain a high recognition score if we increase the spatial resolution of the rate measure, for example, sub-sampling among the most active cells in the assembly; because the averages correspond to a linear summation, the lower resolution can only diminish the information expressed in each cells' firing rate. Thus, our assembly-level measure also sets a lower bound to the network's read-out performance by giving up fine-grain spatial resolution. We chose to focus on the population scale because it is relatable to the signals measured in the larger set of word recognition studies that use magnetoencephalography (Gwilliams et al., 2022, 2018; Piai, Roelofs, Jensen, Schoffelen, & Bonnefond, 2014), extracranial (McMurray et al., 2022; Winsler et al., 2018) and intracranial electrophysiology (Cibelli et al., 2015; Leonard, Bouchard, Tang, & Chang, 2015; Pasley et al., 2012); as they result from averaging the current flow in large neuronal populations. In the following, we will first discuss the strengths and misses of the model, comparing it with the known facts on biological networks. Hence, we will focus on the implications of our results on the two open debates addressed: the computational nature of lexical feedback and the solution to the temporal extent problem.

### A biological causal model of word recognition

The dendritic model presented aims to conciliate computational descriptions of the nervous tissue with psychological evidence of cognitive functions. Our effort aims to build causal explanations of the lexical access process amenable to being neuroscientifically tested. Such an attempt is in line with the growing necessity of biologically constrained models that can cap the space of possible computational solutions and establish explicit links between cognitive and biological descriptions (Poeppel, 2012; Poeppel & Idsardi, 2022; Pulvermüller et al., 2021). However, our model complies with some but not all the requirements for biological realism proposed in Pulvermüller et al. (2021) and Fitz et al. (2022, under revision). In agreement with the constraints on biological networks, it implements a spiking multi-compartment model with rich receptor dynamics

that undergo unsupervised Hebbian learning and homeostatic inhibitory mechanisms; the model also integrates single-cell dynamics within the cell assembly (100 cells) at a full network scale (2500 cells). Conversely, it lacks inter- and intra-areal connectivity and the layered architecture of the cortical sheet and its parameters are not tuned to the specifics of language regions of the human brain (Palomero-Gallagher & Zilles, 2017). The absence of network structure has consequences for the computations that the model can express. In the following, we will address the severity of these missing components for the cognitive function at stake and, where possible, propose improvements to implement in future work.

Experimental evidence indicates that speech signatures are decodable from the temporal areas (Herschl's Gyrus and Superior Temporal Gyrus) as early as 50 ms from word onset (Chang et al., 2010; Gwilliams et al., 2018) and word forms are retrieved within 250 ms (Costa et al., 2009; Marslen-Wilson & Welsh, 1978), such rapid access is due to the few synaptic relays that separate the cochlea and the language areas. Nonetheless, in natural speech, lexical access relies also on context and word meaning (Brodbeck et al., 2022; Hagoort et al., 2004) so, in such a short time we must account for two additional functional stages besides word recognition, namely, speech normalization and segmentation (Eisner & McQueen, 2018; McQueen, 2005; Nusbaum & Magnuson, 1997). These operations originate from the interactions between both hemispheres and between frontal and temporal areas and imply the integration of lexical cues with semantic, syntactic, and pragmatic information (Formisano et al., 2008; Hagoort, 2017, 2019).

Thus, not modeling intra-areal connectivity, we commit to model lexical access in a strictly bottom-up fashion, where no information is available for recognition beyond the acoustic evidence and the immediate word-form context. The word recognition process described concerns only the access to the phonological word form in the Superior Temporal Gyrus (STG), devoid of its semantic or syntactic features, which would indeed require modeling further brain regions, such as those associated with the Wernicke's and Broca's areas (Garagnani et al., 2009; Tomasello et al., 2018). The underlying hypothesis is that word recognition, via dendritic integration, occurs on a timescale faster than the inter-areal interactions, such that the contextual information appears as a parameter rather than a variable in the model. We propose this based on the spatiotemporally localized dynamics of word-forms neural correlates (Gwilliams et al., 2018; Yi et al., 2019). Also, it is essential to highlight that the lack of pathways for contex-

tual information should not be solved solely by introducing additional synaptic inputs that mimic full-brain afferents. To adequately accommodate inter-areal connectivity, we should implement a spatial scaffold that arranges incoming connections from higher-order areas to the STG model. Indeed, the biological networks of the cortical sheet are organized in layers, with the feedback and feedforward axonal afferents targeting the superior and inferior layers, respectively (Braitenberg & Schüz, 1998b; Senzai et al., 2019). The layered architecture is not present in our model and it seems an inescapable feature of the network if we want to study the interaction between feedforward and feedback streams.

The hypothesis that our model accounts for lexical access in the STG based on sole feedforward information is apparently at odds with the abstract representations we use as input for the model. Indeed, the abstract phonetic categories become available to the system following phonemic normalization. The process of extracting abstract phonemic categories from the contingent speech sounds takes place, arguably, before word recognition (Clarke & Garrett, 2005; Norris, 2003; Poellmann, Bosker, McQueen, & Mitterer, 2014), and it must originate from the interaction between the auditory stimuli and the contextual information; for example, prior expectation on the speakers' pronunciation (Kleinschmidt & Jaeger, 2015). In the following section, we address this discrepancy, motivate our assumptions on the nature of the phonemic representations used, and then postulate a normalization stage that is autonomous from the lexical access one.

**Abstract phonemic representations and lexical feedback**

In the traditional view of linguistic theory, phonemes are the atomic units of language (D. Jones, 1962; Kazanina, Bowers, & Idsardi, 2018; Liberman et al., 1957); accordingly, words' phonological memories in the lexicon are stored and accessed based on sequences of phonemes. The phoneme hypothesis is supported by recent electrophysiological studies that could locate and time-track the neural correlates of phonemes' identity in the temporal lobe (Chan et al., 2014; Chang et al., 2010; Fox et al., 2020; Gwilliams et al., 2022). Consistently, the present work adopts phonemic segmented units as the fundamental code for lexical access. Abstract phonemes are represented in the network by means of external projections on cell assemblies; such projections are intended to reproduce the feedforward streams from the earlier stages in the auditory hierarchy. On the other hand, it is well-known that phonemes assume different phonetic vests when uttered. When placed in a different lexical or pre-lexical context,

the acoustic realization of a phoneme changes (Uppstad & Tønnessen, 2010). Similarly, there are phonetic variations among speakers that require the listeners to adjust their phonological memory to new and, sometimes, unpredictable sounds (Best et al., 2015; Kleinschmidt & Jaeger, 2015). Such a systematic lack of invariance in phoneme realizations requires a dedicated stage that normalizes the input based on the context and maps the acoustic information into abstract phonemic categories. We address the two issues of lexical and speaker phonemic variability separately and show that a feedforward normalization stage can account for them if we assume the two stages are autonomous.

First, we address the case when phonemic variability is structural to the language. There are two main instances. One relates to phonemes that change their acoustic realization across different words; for example, the phoneme /k/ in /kæt/ and /dʌk/ is realized by the aspirated phone [kh] in cat and a plain or unreleased [k'] in duck (Kazanina et al., 2018). The other concerns phonemes that vary dependently on their lexical context; for example, certain phonemes assimilate with the preceding or following sounds (Gaskell & Marslen-Wilson, 1996; Weber, 2001). Such systematic variability may indicate that the suprasegmental allophone information is the fundamental speech unit rather than the phoneme (Mitterer, Reinisch, & McQueen, 2018). In this view, word memories are accessed based on the allophones heard; thus, the correct pre-lexical to lexical mapping should leverage this type of pre-lexical representation rather than the abstract phonemic ones. This richer segmental information can be included in the model's representations and it will not affect its dynamics as long as the bottom-up signals are expressed in spatialized cell assemblies, as it seems to be the case for phonological representations (Voice Onset Time Fox et al., 2020). Indeed, the sort of phonological representations we use in the pre-lexical layer is not linguistically accurate, it just aims to illustrate the computations in the model. An expanded representation space may even alleviate the computational burden on the sequence recognition mechanism and facilitate lexical access in words with overlapping phonemes.

The second issue we must address is the lack of invariance at the speaker level (Liberman et al., 1967; McQueen, 2005). In contrast to the systematical variations discussed, the speaker variability cannot be addressed by expanding the set of pre-lexical representations because the listeners will always face novel speakers whose phonemic categories cannot be learned ahead. Instead, the problem must be solved by considering a normalization pre-processing stage of the speech sounds. Normalization requires contextual information to bias pre-lexical deci-

sions (Sjerps, Fox, Johnson, & Chang, 2019). In our model, we do not account for this stage and assume that phonemes, possibly allophones, are already resolved into abstract categories. Thus, the lexical and recognition stages are functionally encapsulated, such that context must modulate the normalization stage after lexical access (McQueen et al., 2009; Norris et al., 2000) and not during it. The alternative hypothesis to autonomous lexical and pre-lexical layers is that the two work together to determine both the pre-lexical and lexical decisions, that is, they are interactive (McClelland & Elman, 1986; McClelland, Mirman, & Holt, 2006)

Because cortical networks are intrinsically recurrent within and across regions (Braitenberg & Schüz, 1998a), it would be largely implausible that the lexical and pre-lexical populations are anatomically isolated. However, it is crucial to specify the nature of the information passed across the two functional stages during word recognition. Our model offers a double-fold insight into this problem; even though the two stages are interactive by design, we show that the biological constraints and the predictions are more in line with the autonomous rather than interactive hypothesis. First, although there is no privileged direction for the flow of information during the associative phase, the network develops strong feed-forward connections that bind phonemes to words, but the feedback connections remain weak. This network connectivity feature is evidenced in Fig. 5 of Chapter 4 and is also visible in the recurrent network connectivity of Fig.5.5. The directional asymmetry presumably emerges from the plasticity rule combined with synaptic scaling, which limits the fan-in weights in phoneme populations. Second, the model does not reproduce the online lexical bias effect on the pre-lexical representations when we tested with a setup analogous to Ganong (1980)'s experiment. Crucially, the lexical bias effect is one of the strongest pieces of evidence that supports the interactive hypothesis. In our simulations, phoneme populations did not activate more or less depending on the word or non-word condition; conversely, we found a significant effect of lexical bias in the sub-threshold membrane potentials, after the end time of the stimulus interval. The increased membrane depolarization can provide a mechanism for learning a new mapping between the phonetic and phonemic levels based on voltage spike-timing-dependent plasticity. These computational results indicate an interaction between the two functional stages exists, but only after lexical access occurred. The model's prediction offers a causal explanation of lexical to pre-lexical activity observed in neuroimaging studies (Gow, Segawa, Ahlfors, & Lin, 2008; Myers & Blumstein, 2008) and provides a highly plausible

mechanism for perceptual learning (Norris, 2003). Interestingly, the membrane dynamics of pre-lexical populations are in punctual agreement with novel data obtained from magnetoencephalography recordings by Gwilliams et al. (2018) during a phonemic continuum manipulation. The study shows that phonemic ambiguity is strongly represented in the neural signal, and lexical bias does not affect the early processing stages. In their analysis and our model (Fig.5.5), pre-lexical representations are re-activated at the word uniqueness point and encode the amount of phonological ambiguity.

The network with dendrites seems a good model candidate for the sought-after sequence integration mechanism in lexical access (Yi et al., 2019), which is, so far, not fully explained by any of the biologically plausible models available in the literature. The core mechanism of the model is dendritic integration, which, we showed, helps solve one of the core problems in connectionist models. The following section compares the network with the dendrites model to these previous studies and discusses its advantages and limitations.

**Solving the temporal extent problem with long timescales**

The last decades have seen the flourishing of computational models in explaining speech perception. In particular, from the seminal introduction of TRACE (McClelland & Elman, 1986), an increasing number of connectionist (NNs) models achieved word recognition, or more in general sequence labeling (Amodei et al., 2015; Graves, 2012; Hannagan et al., 2013; Scharenborg, van der Gouw, Larson, & Marchiori, 2019). Nowadays, the spoken word recognition problem can be successfully solved by leveraging network recurrence and backpropagation algorithms (Adolfi, Bowers, & Poeppel, 2023). However, the algorithms used to train neural networks are often not plausible, and when the model is constrained to dynamics that match psychological and neural evidence, things change. One of the main difficulties to account for is the temporal structure of words and their staggering level of phonological overlap. This problem was already present in the TRACE model, often referred to as the temporal extent problem or temporal order problem (TOP). The problem is due to the necessity of introducing temporal markers in phonemic inputs such that anagrams and words-within-words can be correctly distinguished. It results in an implausibly large number of nodes and connections the network must include to process a moderate-sized lexicon. Mitigations for the problem have been introduced by Hannagan et al. (2013); You and Magnuson (2018); the model presents a pre-lexical layer that can detect diphone structures in the input. However, the model has two limits that must

yet be addressed; first, it does not provide a strategy to learn the diphone matrix that is biologically plausible; second, it still requires explicit temporal markers in the phoneme inputs, markers that cannot be found in electrophysiological data (Gwilliams et al., 2022).

Our model provides a systematic and biologically plausible solution to the problems mentioned above. We demonstrated that introducing slow-decaying neural states, be it in the dendritic membrane or the hidden states of the reduced model, leads to the successful recognition of anagrams. The network structure is achieved following biologically plausible plasticity rules. Crucially, our work reached a similar weight structure than the one predicted in Sequence Detection networks (Knoblauch & Pulvermüller, 2005); successive phonemes have stronger synaptic weights than early ones. Our truly novel contribution is the introduction of the slow decay of dendritic compartments which allows the phonemes to bind each other upon arrival in the system and makes the word-form populations sensitive to sequences rather than to single phonemic features. In spite of the computational advantages of the long timescale, phonological overlap still degrades the accuracy and the significance of lexical access (Appendix A and Appendix B). The lexicon where the model encountered the most difficulties is borrowed from the original TISK study (Hannagan et al., 2013) and it contains a dozen words composed from only six phonemes.

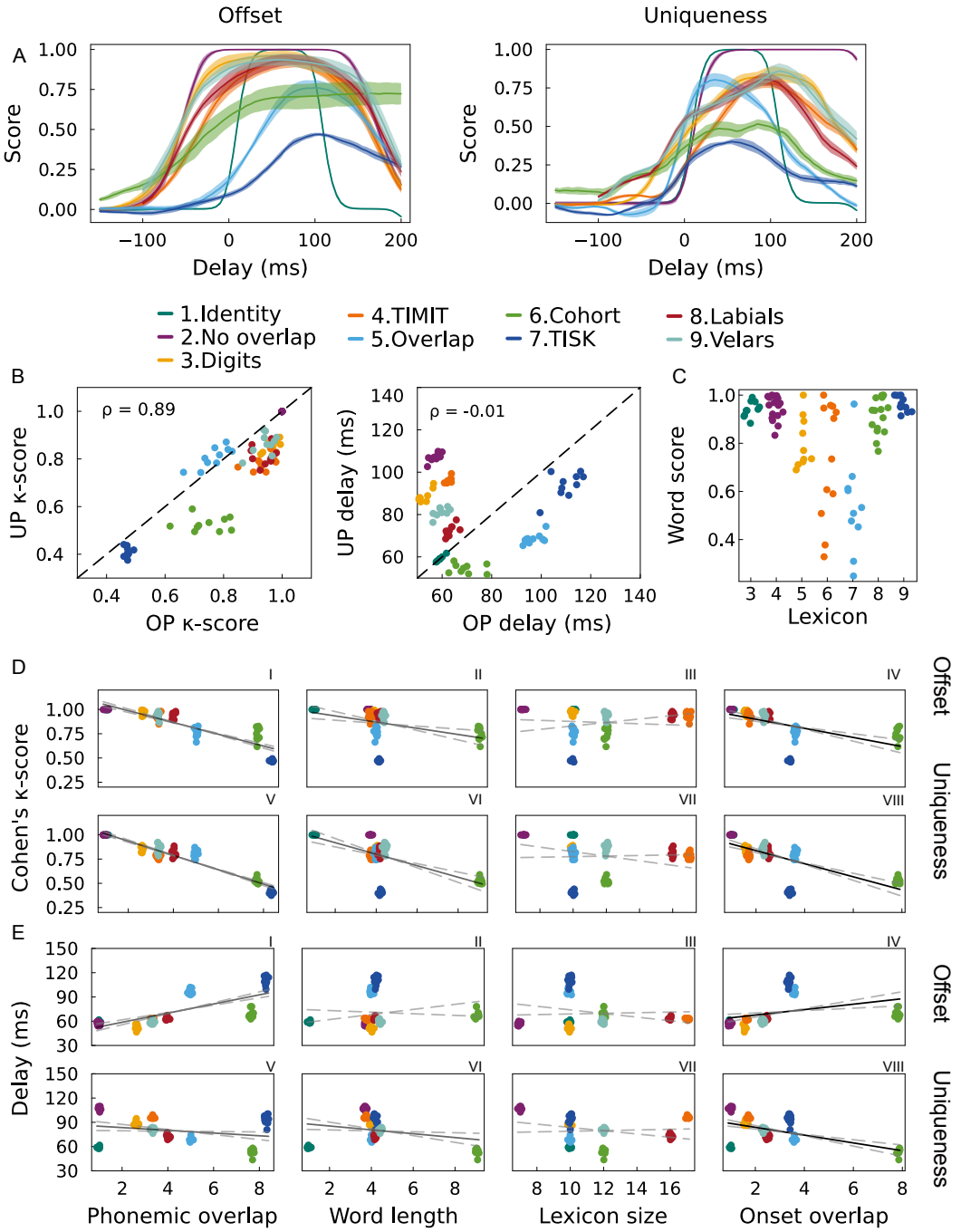
The models' shortcomings in accounting for human performance must still be put in perspective, and the network with dendrites should be evaluated in light of its small scale compared to the real system. We used a network of only 2000 neurons, which has six orders fewer neurons than the STG (a hundred million, if we assume the conservative estimate that only one per thousand cortical neurons belong to the region). Similarly, the Tripod has only two dendritic compartments, while the actual computationally segregated compartments in a real neuron could be one or two orders larger (Hawkins & Ahmad, 2016; Larkum, 2022). In contrast, the dozen words and phonemes tested in the model are three or four orders of magnitude smaller than those of human languages. Thus, if the same neural mechanism has to be implemented in the brain, the signal-to-noise ratio would be much higher than the one obtained in the model. The scaling problem remains crucial to account for human cognitive functions and, so far, neither the connectionist models, TISK and TRACE, have satisfying results when tested in large lexica (Nenadić & Tucker, 2020).

## 5.5 Appendix

### Appendix A: Shared phonemes decrease task accuracy and delay word recognition

In the previous chapter, we introduced the lexicon $\kappa$-score score as a measure of lexical access. Here, we recap the measure and derive the word $\kappa$-score . The lexica tested in the present chapter comprises real English words; each set varies in the number of words and the number of phonemes involved, and they have different recognition scores. An overview of their recognition score ($\kappa$-score) and average recognition delay (ARD) is portrayed in App. Fig.5.1 *A* for the uniqueness and offset points measures. The $\kappa$-score indicates whether words can be recognized based solely on the average population firing rate; the ARD gives an idea of the time it takes from the word onset for optimal recognition. Both values are computed from the score-delay functions in App. Fig.5.1 *A*. Instead, the uniqueness point (UP) and offset point (OP) measures refer to the time interval used to compute the population firing rate. The UP measures one single phoneme interval 50 ms at the uniqueness point of the word; in contrast, the OP measures the entire word interval between the onset and the offset of the word. The two measures achieve similar $\kappa$-score (Pearson correlation $\rho$ =0.89). However, their ARD are independent ($\rho$ =−0.01), as depicted in App. Fig.5.1 *B*, which express different aspects of the word recognition dynamics. The $\kappa$-score is an estimate of recognition over the entire lexicon; however, word accuracy varies within it. Starting from the confusion matrix (Chapter 4, Figure 2), we computed the recognition index associated with each single word. The panel in App. Fig.5.1 shows that the difference in the recognition score presents significant differences among the words of certain lexica.

To understand why certain lexica achieve higher scores or take longer to be recognized, we individuated four lexicon properties and compared the $\kappa$-score and ADR across each. We computed (1) the word *overlap* as the average number of phonemes in a word divided by the number of phonemes in the lexicon; (2) the average word length; (3) the number of words in the lexicon, its size; and (4) the average distance between the onset and the uniqueness point, or onset overlap measure. The panels in App. Fig.5.1 *D* show the recognition score against these dimensions for all the lexica, for both the offset point (top panels, I-IV) and the uniqueness point (bottom panels, V-VIII). The relationship between the lexicon properties and the score was interpolated with a linear fit. We report the interaction as significant if the fit's slope does not change sign within the 5 %



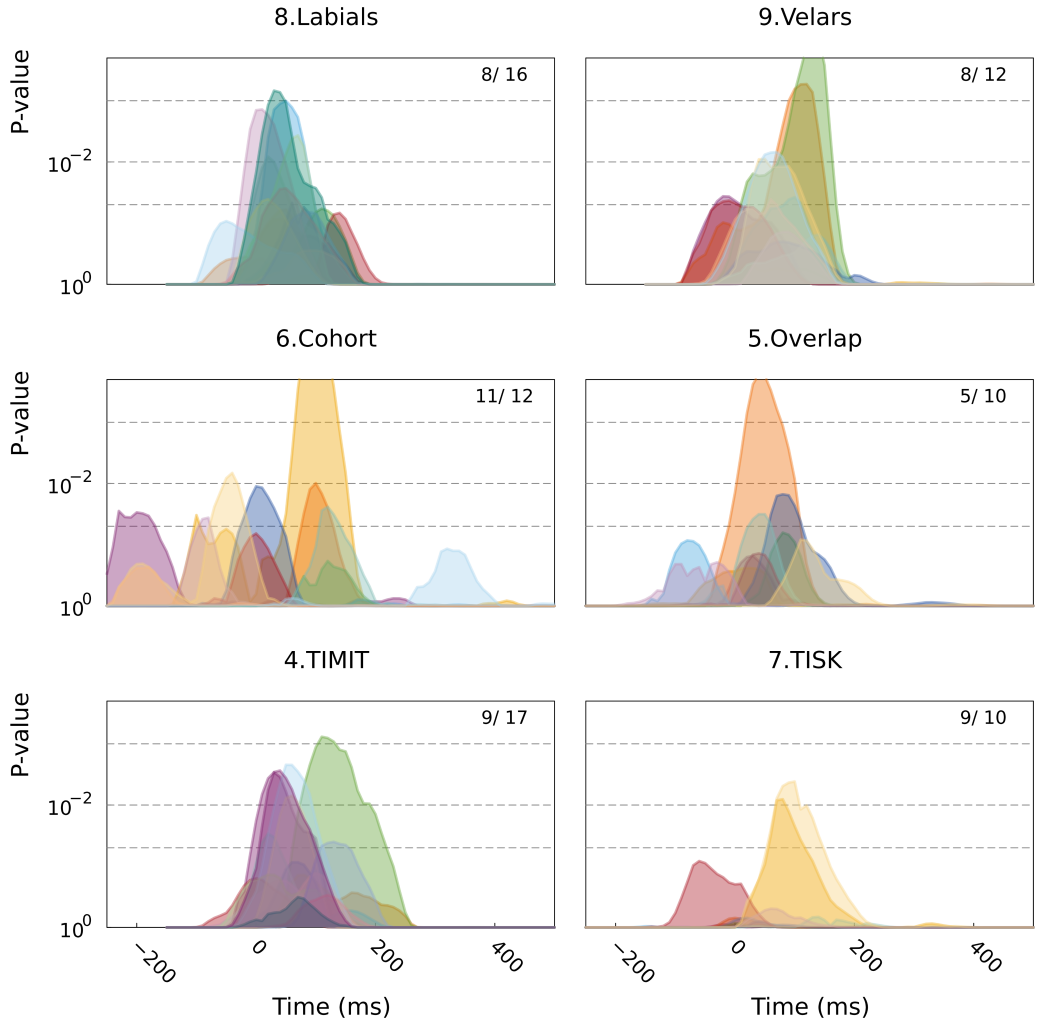

Appendix Figure 5.1: **Word recognition depends on lexicon properties**
(**A**) Cohen's $\kappa$-score measured in function of the delay from the offset and uniqueness points for the nine lexica tested. The offset measure compares the population spike rate on the full word interval, and the uniqueness point measures only the activity at the uniqueness point phoneme (50 ms interval). (**B**) Comparison of Cohen's score ($\kappa$-score, left panel) and average recognition delay (ARD, right panel) in the offset (x-axis) and uniqueness point (y-axis) conditions. The average recognition delay is computed as the $\kappa$-score weighted sum of the delays in the interval −150 ms to 200 ms. The $\kappa$-score of the two measures has a Pearson correlation of 0.89, but their ARDs are uncorrelated ($\rho$=0.01).

Appendix Figure 5.1: (**C**) Recognition score of each word for all the lexica tested. The recognition score is homogeneous within the lexicon. (**D**) $\kappa$-score as a function of word overlap, average word length, lexicon size, and onset overlap. Each of the ten points with the same color represents a different instance of the same lexicon. The straight lines correspond to the linear fit of the data. The recognition score is negatively correlated with the word overlap and the average word length (I, II, IV, V, VI, VIII) but not with the lexicon size. The impact of the lexicon's properties is the same in the offset and uniqueness point conditions. The grey lines indicate the 5 % confidence interval of the linear fit. Black lines are absent when the correlation is not significant. Notice that the four dimensions of each lexicon are not independent; longer words tend to have larger phonemic overlaps. (**E**) Comparison of the lexica properties against the ARD, similarly to *D*. The time between the sequence presentation and the word recognition increases for the offset time measure when the lexicon increases in phoneme and sequence overlap (I, IV). Conversely, the uniqueness point measure becomes more responsive (VIII)

confidence interval (dashed lines); when not significant, the panel only shows the range (absent black line). The recognition score correlates negatively with the word overlap (I, V), the average word length (II, VI), and the onset overlap (IV, VIII) for both the offset and the uniqueness point measures. These correlations indicate that the more words share phonemes, the more difficult it is to recognize them. Conversely, the lexicon size has a null correlation with the recognition score, which suggests that the network's memory capacity is larger than the lexicon sizes tested. The analysis indicates that a network with 2000 cells has enough memory capacity for a lexicon of the order of 10 to 20 words. However, the recognition score is affected by shared sequences and phonemes among the words. Given the limited number of phonemic signs in human language (order of fifty), increasing the number of words in the lexicon would necessarily increase the phonemic overlap.

A clearer picture of the impact of shared phonemes on the recognition dynamics emerges from the analysis of the model's average recognition delay (ARD), App. Fig.5.1 *D*. The contrast between ARDs and $\kappa$-scores is due to the opposite interactions between the variables and the lexicon's characteristics. A picture of the interactions with the lexicon's dimensions (1-4) in the offset and uniqueness point measure is portrayed in App. Fig.5.1 *D*. In this case, only the phonemic (I, V) and onset overlap (IV, VIII) significantly impact the ARD measure. For the full interval measure (OP), increasing the number of shared phonemes, the ARD goes from 60 ms to 120 ms. Conversely, the ARD in the UP measure decreases with the length of the overlapping sequences in the lexicon (IV); the later the

UP is encountered, the faster the word assembly is reactivated after it, which is confirmed by direct inspection of the score delays curves (App. Fig.5.1 *A*). By contrasting the two results, we can infer the dynamics of the word populations. The reactivation of the correct assembly occurs late when the identity of the phonemes carries ambiguity; because of the phoneme overlap, other word assemblies are partially reactivated. Hence, population activity during the early phonemes gives no decisive or misleading information. However, targets are preactivated during the lexical competition, and their firing activity rapidly peaks when the input disambiguates at the uniqueness point, reducing the ARD in the UP measure.




Appendix Figure 5.2: **Combined p-values over word grand-average** The p-values were computed with the Simes procedure starting from the target-control pairs in each lexicon. The annotated fraction indicates the number of words that were recognized with p<0.05

## Appendix B: Dendritic membrane dynamics is causal to correct lexical access

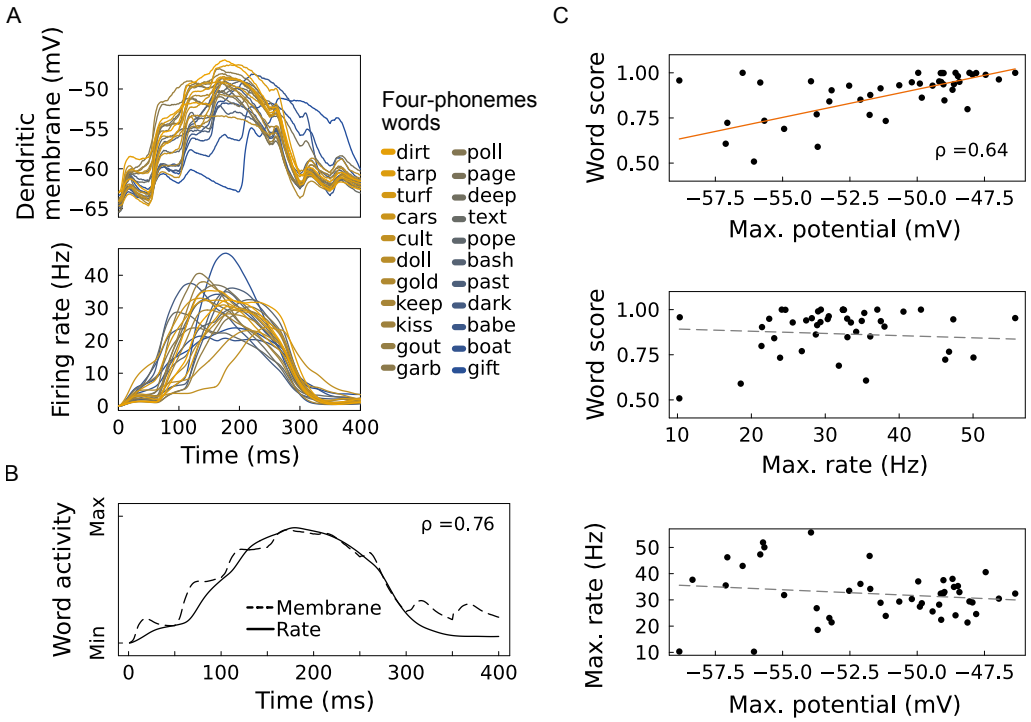
In the neuronal circuitry, the spikes are the main form of intracellular communication; however, the somatic spike accounts for the integration of pre-synaptic spikes only indirectly. Before the soma fires, several processes occur in the cell body that mostly goes unseen to the post-synaptic cells; the spike is *only* a probabilistic expression of the membrane potential (Fiorillo, Kim, & Hong, 2014; Mikulasch et al., 2020). Hitherto, we have only measured word reactivation through a spike-rate measure. In the Tripod model, the hidden processes comprehend the membrane potential in the two dendrites and the synaptic conductances. Hereby, we explore the dynamics of membrane potential in the word population and show that it has larger predictive power on the correct word recognition than not the spike rate itself.

To measure the neurons' membrane potential, we have to record the dynamic variable during the simulation; this procedure is costly for both computation time and memory usage. Hence, we limited the measure to a subset of words and lexica. We chose to record only words composed of four phonemes in the *Overlap, Labials & Alveolars* and *Velars* lexica. The restriction to words of four phonemes facilitates the comparison and limits the possibility of confounds in the analysis. We considered 24 word populations. The average membrane dynamics are shown in the top panel App. Fig.5.3 *A*. The membrane potentials are mediated across the two dendrites and on the entire population; the color scheme sorts the word by the average of the membrane potential in the measured interval, 400 ms. The bottom panel shows the firing rates, convolved via an alpha-function with 10 ms timescales. In panel App. Fig.5.3 *C*, we average across the entire set of words and compare the firing and membrane potential. Because the two variables have different units of measure, we present them together by scaling them to their numeric range, thus their maximum and minimum.

From inspecting the words' membranes, we notice that the potential presents clear signs of the phonetic inputs' time course. Upon the first phoneme (0 ms), the membrane potentials have a peak depolarization, which fades within the first 50 ms. We hypothesize that these initial up and down strokes are due to the inhibitory activity on the word populations. After the 50 ms of silence, the network global rate is diminished, and the sudden activity is not balanced by inhibition. Network activity increases rapidly following the activity of the external phonetic projections. Thus, it leads to more robust and undirected inhibitory

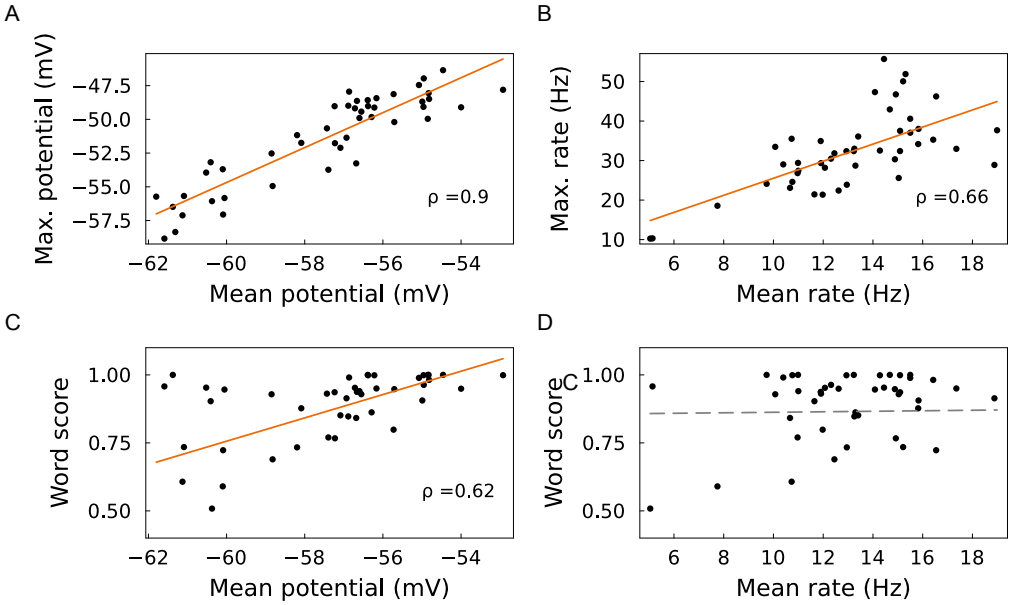
feedback, which re-hyperpolarizes the word populations' dendrites. The arrival of the second phoneme hits the dendrites at 50 ms when the word's average potential is sitting at a slightly larger value than its equivalent at word onset. However, now the membrane potential increases steadily and jumps of approximately 10 mV, the contrast between the input of the first and second phonemes is also visible in the global membranes average (App. Fig.5.3 *B*); this time the membrane goes up and there is no down-stroke. The difference is likely due to the triggering of the NMDARs' regenerative processes, which maintain the membrane depolarized in the face of the inhibition. In addition, the population also starts firing, and the recurrent connections contribute to keeping the membranes depolarized. In this early 100 ms, the membrane anticipates the firing, suggesting that its dynamic is causal to the activation of the word populations. From the third phoneme onwards, the impact of the phoneme inputs diminishes in favor of the depolarization due to the recurrent spikes. In between 150 ms to 200 ms, both the firing rate and the membrane reach their peaks and start fading away. They are hampered by the rising feedback inhibition, the fast time course of somatic inhibitory plasticity, and the adaptive firing threshold of the excitatory cells. Interestingly, the firing rate does not present the step-like dynamics of the membrane potential; this is partly because the alpha-function convolution smoothes it and because the firing is mainly due to recurrent contributions, which have timescales independent from the phonemic inputs' interval.

To clarify the dendritic membrane's causal role, we also measured the correlation of membrane depolarization with the word $\kappa$-score . The first panel in App. Fig.5.3 *C* shows a significant positive correlation ($\rho$ =0.64, within a 95% confidence interval) between the peak membrane potential (x-axis) and the word $\kappa$-score . In contrast, although relying on the same spike measure, the firing rate does not correlate significantly with the recognition score (central panel). Neither does it with the membrane potential (bottom panel). We also tested these correlations by measuring the mean of rate and membrane rather than their peak values, with identical results. This analysis suggests that the membrane potential of the word populations contains more information regarding the network state than the population firing rate. Indeed, the $\kappa$-score does not account for the strength of word activity but rather for the relative intensity of the target population compared to the competitors' firing rate. Thus, we conclude that measuring the population firing rate in isolation is not, strictly speaking, a measure of word recognition because it is not re-scaled against the whole network activity. Conversely, the membrane potential average is a measure of



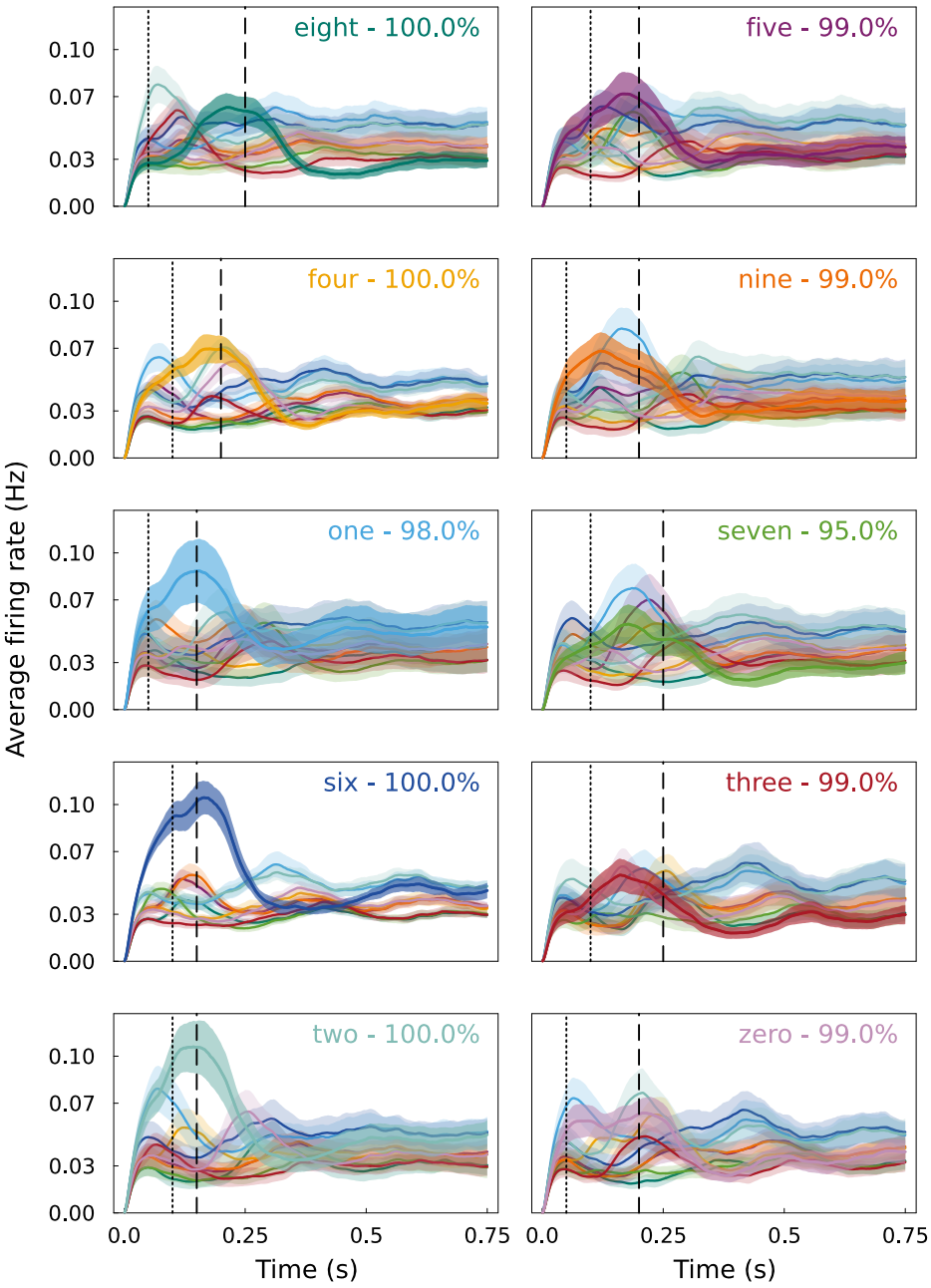

Appendix Figure 5.3: **Dendritic membrane potential anticipates firing activity and explains recognition score**
(**A**) Membrane potential (top) and firing rate (bottom) of 24 words selected across four lexica; the color scheme orders the words by their average membrane potential. We selected only words with four phonemes so the neural dynamics can be successfully compared within an interval of 400 ms after word onset. The membrane potential traces show the signature of the phonemic inputs (upward spikes every 50 ms), but the firing rate does not. (**B**) Comparison of average membrane potential and firing rate for the set of words in *A*, both are scaled by the maximum and minimum values reached in the interval. The membrane potential anticipates the firing rate in the early phonemes 0 ms to 100 ms; the *rho* states the average correlation between each word's membrane-rate pair. (**C**) Correlation between membrane potential, firing rate, and word-recognition score. The maximum membrane potential of the word population correlates with the recognition score (top panel), but the firing rate does not (central panel). The bottom panel also indicates no significant correlation between the membrane potential and the firing rate when compared across words.

recognition that expresses the target activity in relationship with the lexical competitors. The present results may not hold when the lexical pool is expanded to words with heterogeneous phonemic lengths. However, they potentially impact the experimental analysis of spiking activity in animals and humans.

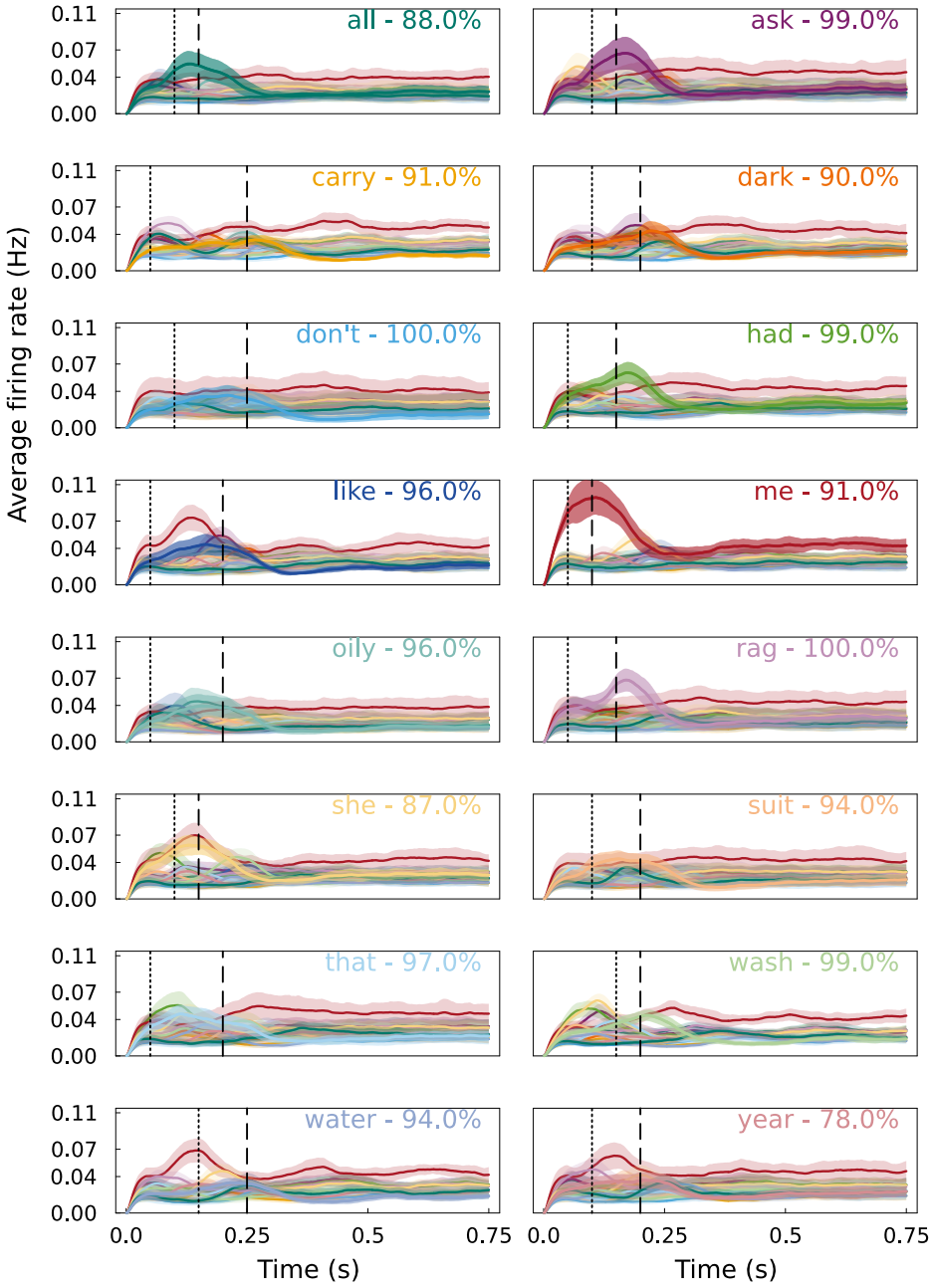



Appendix Figure 5.4: **Membrane potential and firing rate of target populations**
(**A-B**) The mean and maximum value of the membrane and rate correlate among them. (**C-D**) The word $\kappa$-score correlates with the average and peak of the membrane potential (left panel) but does not correlate with the statistics of the firing rate (right panel). Rate and membrane do not correlate in any combinations tested (max-max, mean-max, etc.).

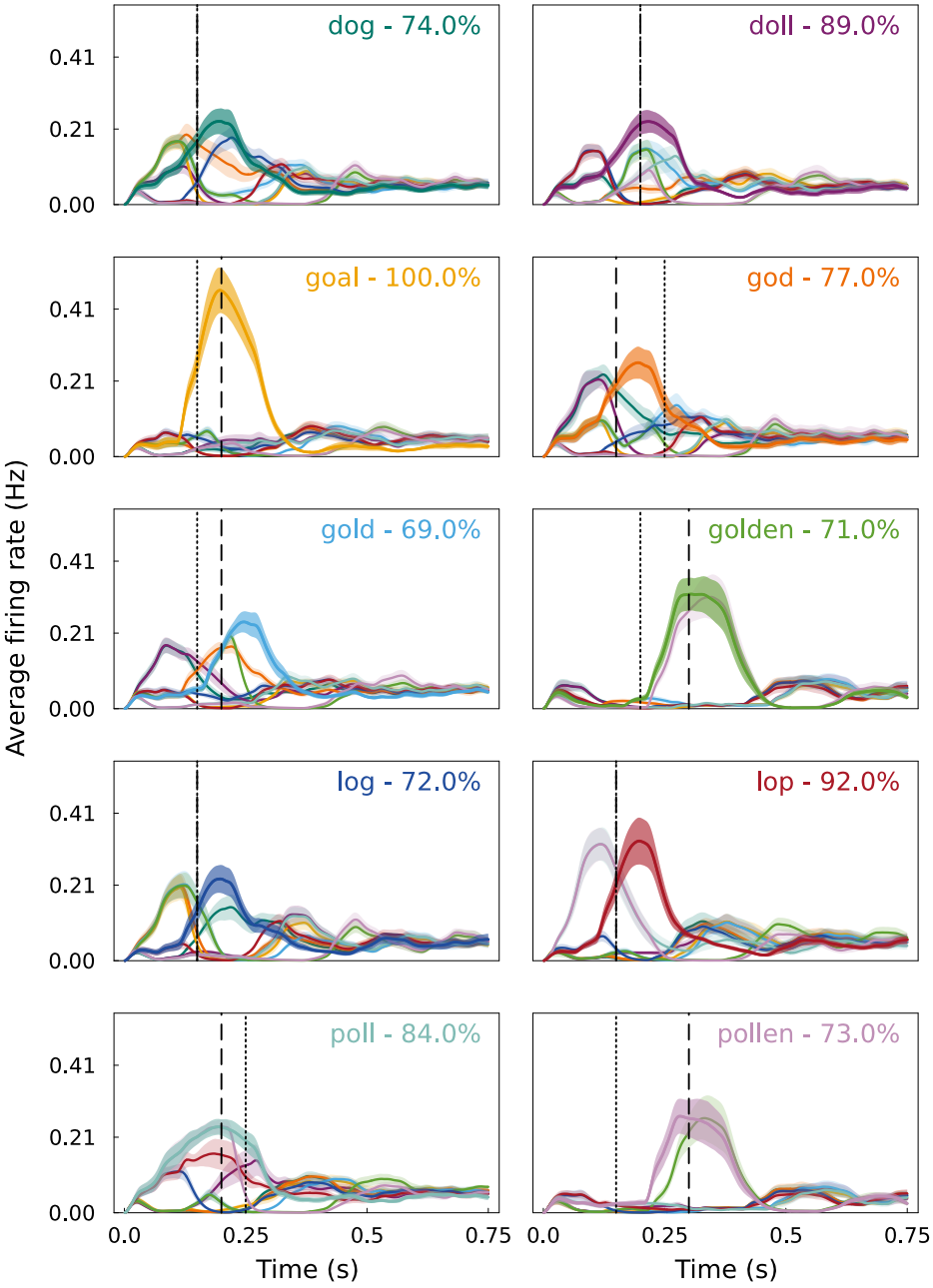
## Appendix C: Firing rate and p-values



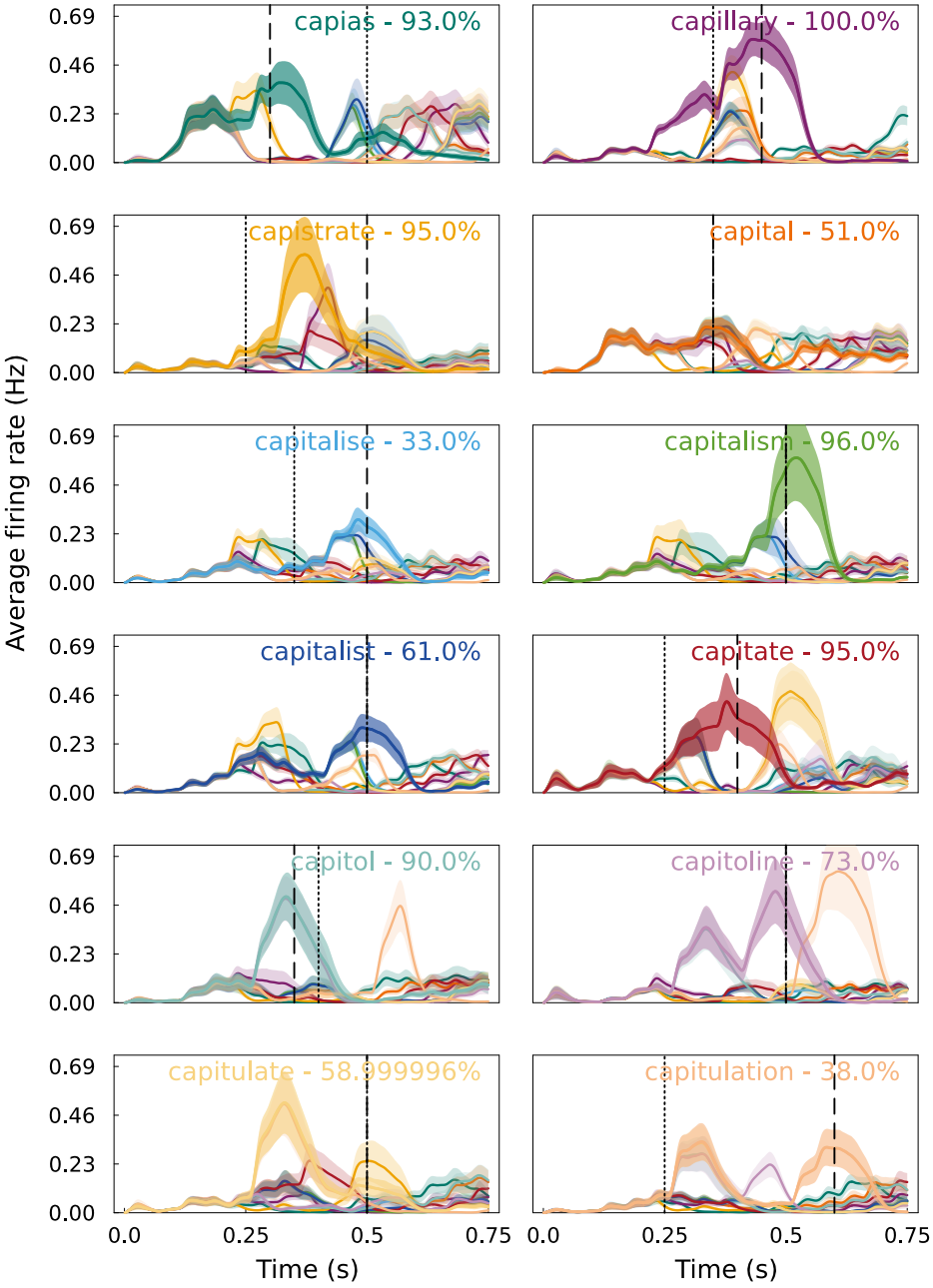

Appendix Figure 5.5: **Firing rates *Digits* lexicon**



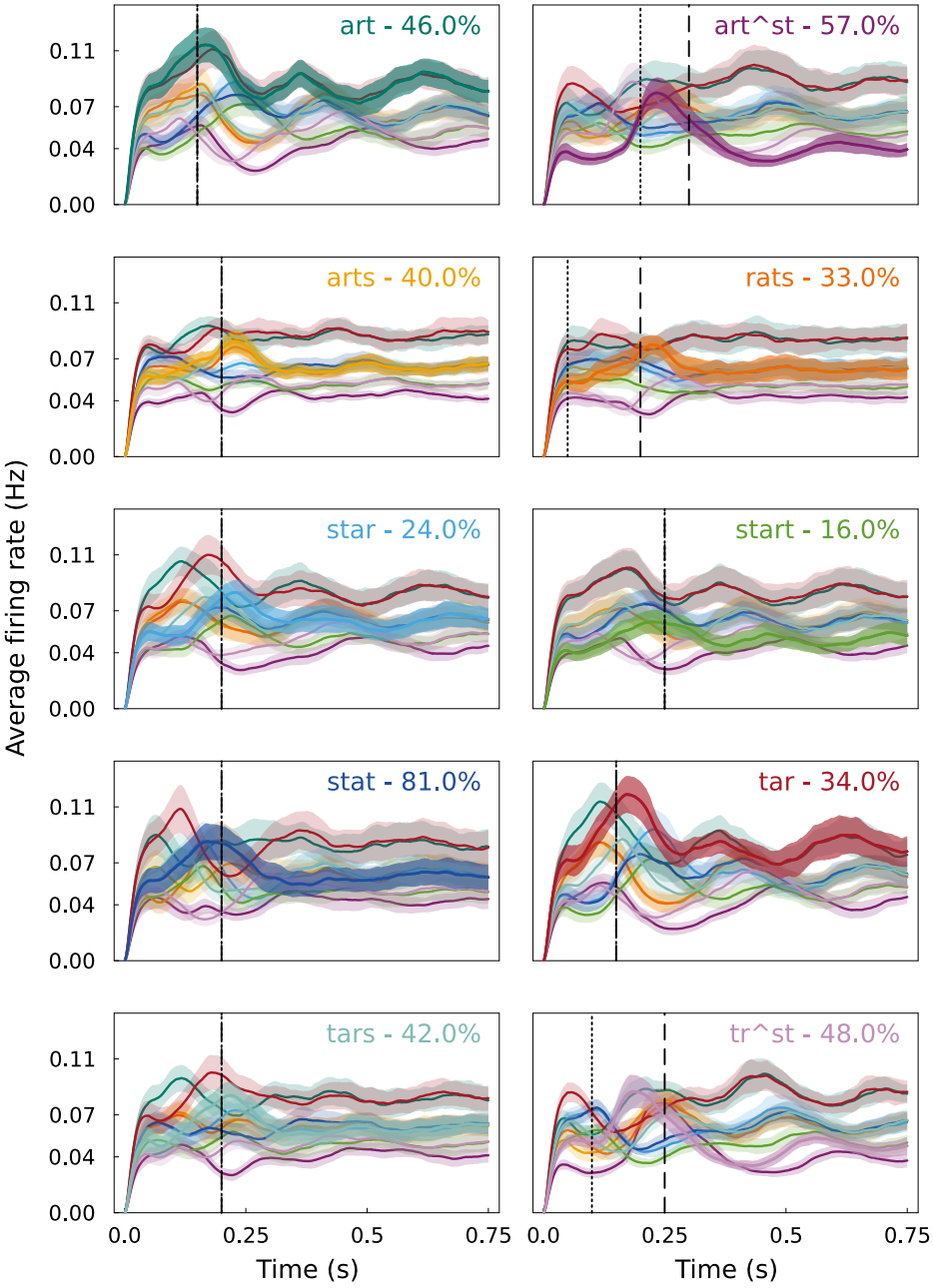

Appendix Figure 5.6: **Firing rates *TIMIT* lexicon**



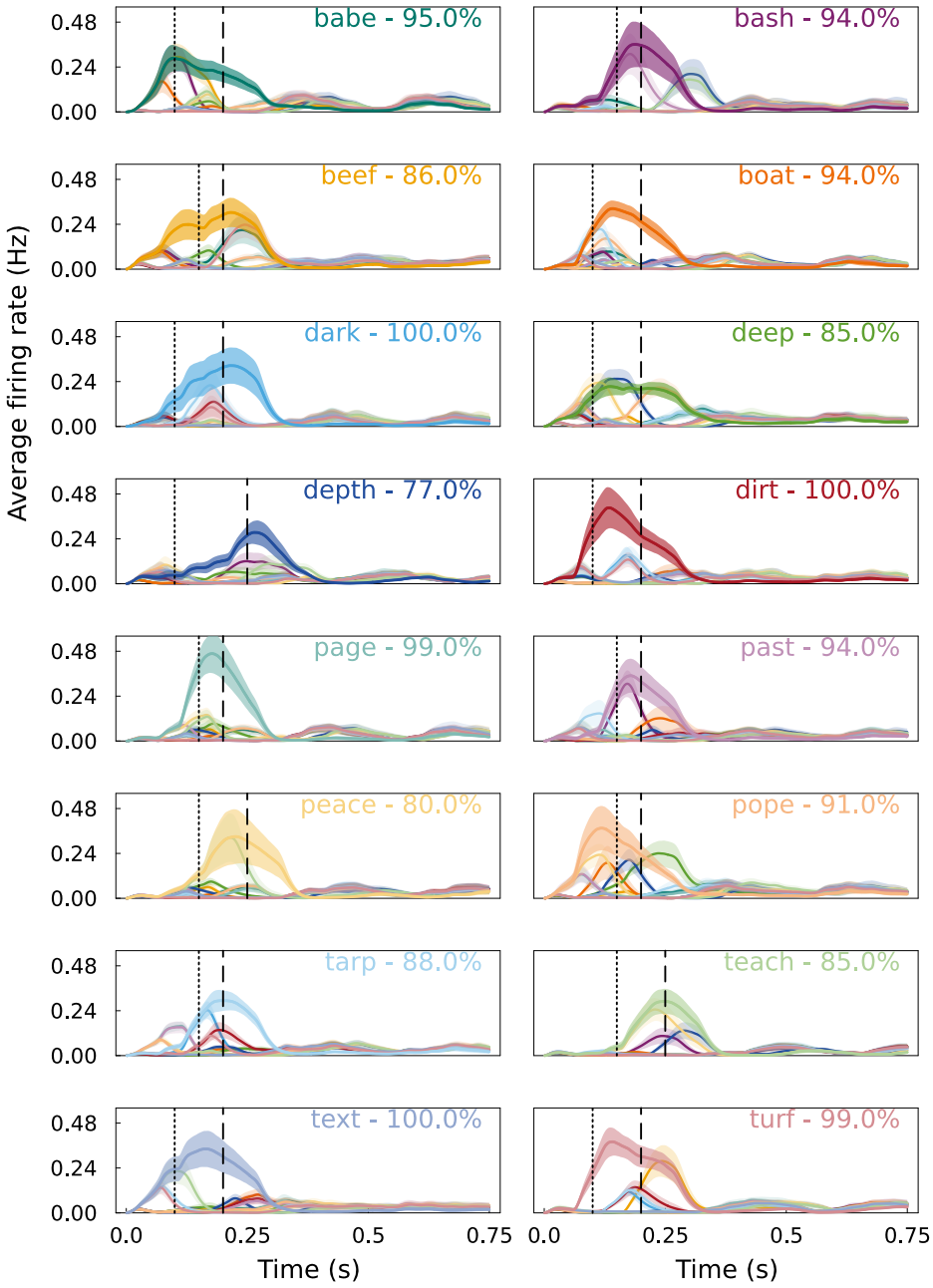

Appendix Figure 5.7: **Firing rates *Overlap* lexicon**



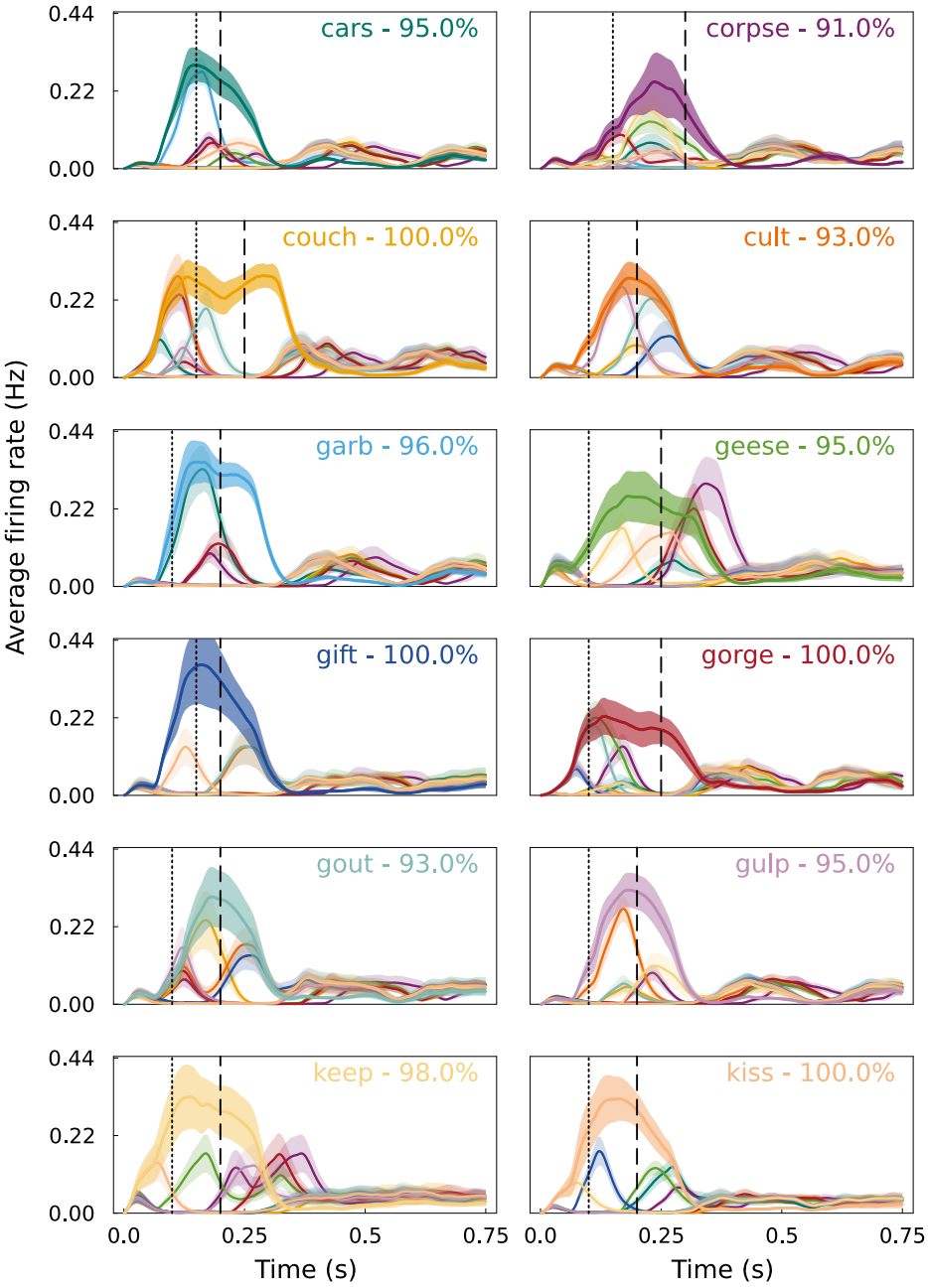

Appendix Figure 5.8: **Firing rates *Cohort* lexicon**




Appendix Figure 5.9: **Firing rates *TISK* lexicon**

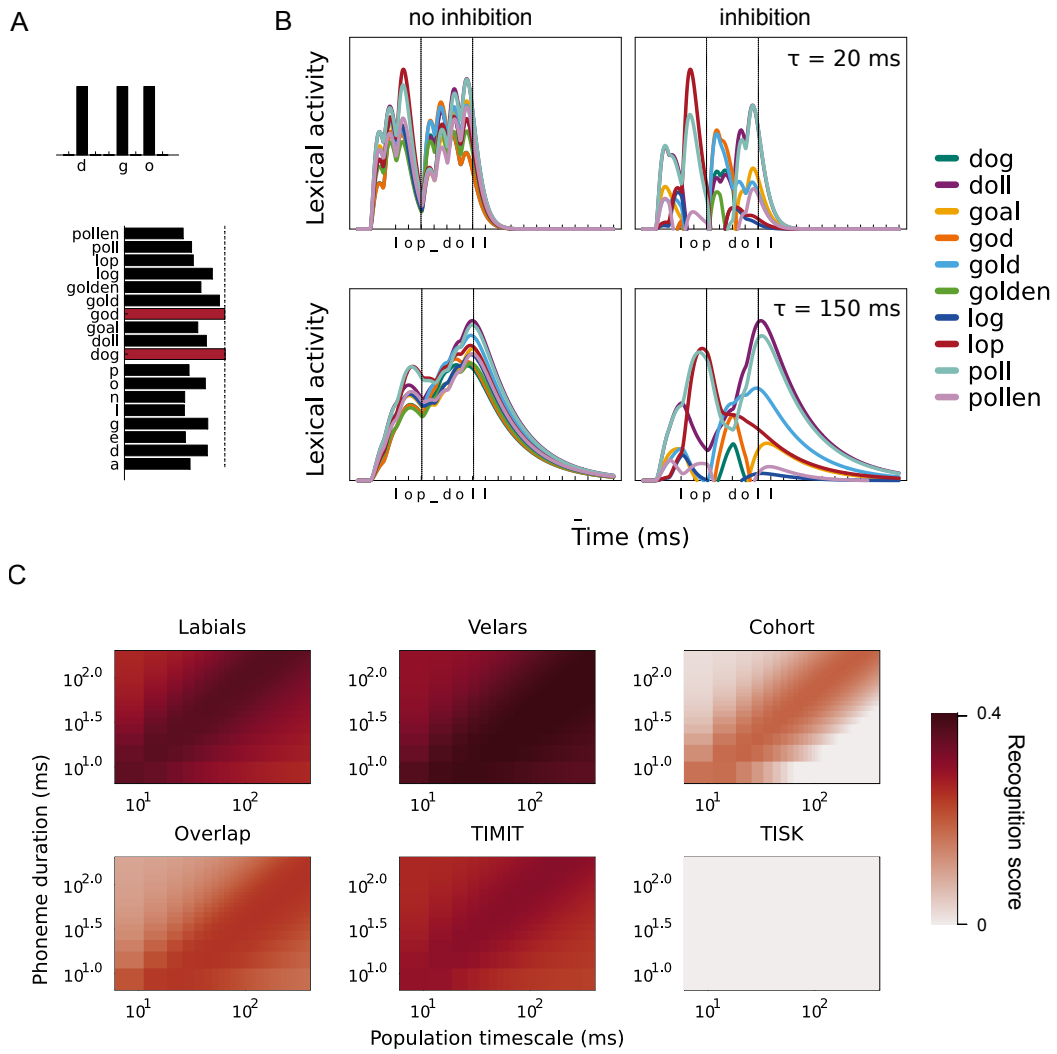



Appendix Figure 5.10: **Firing rates *Labials & Alveolars* lexicon**

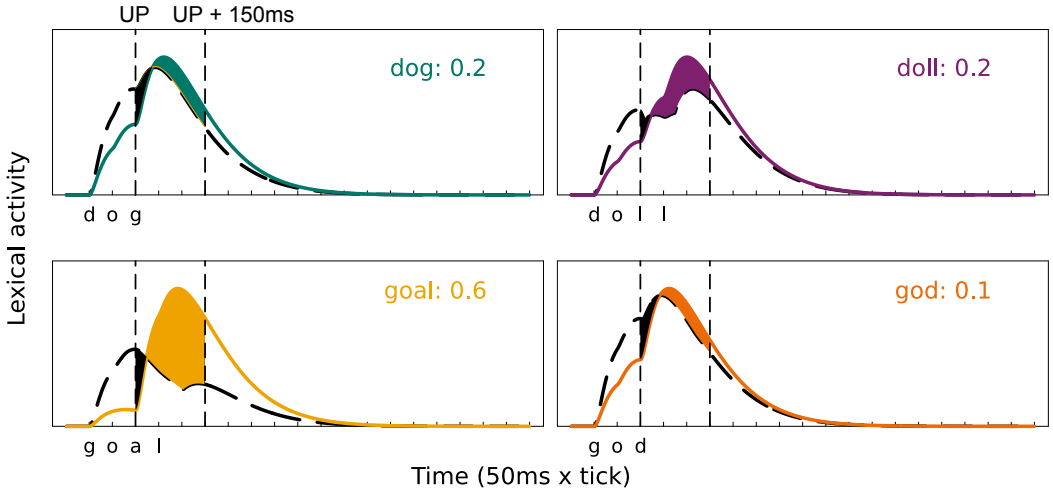



Appendix Figure 5.11: **Firing rates *Velars* lexicon**

## 5.6 Supplementary Material

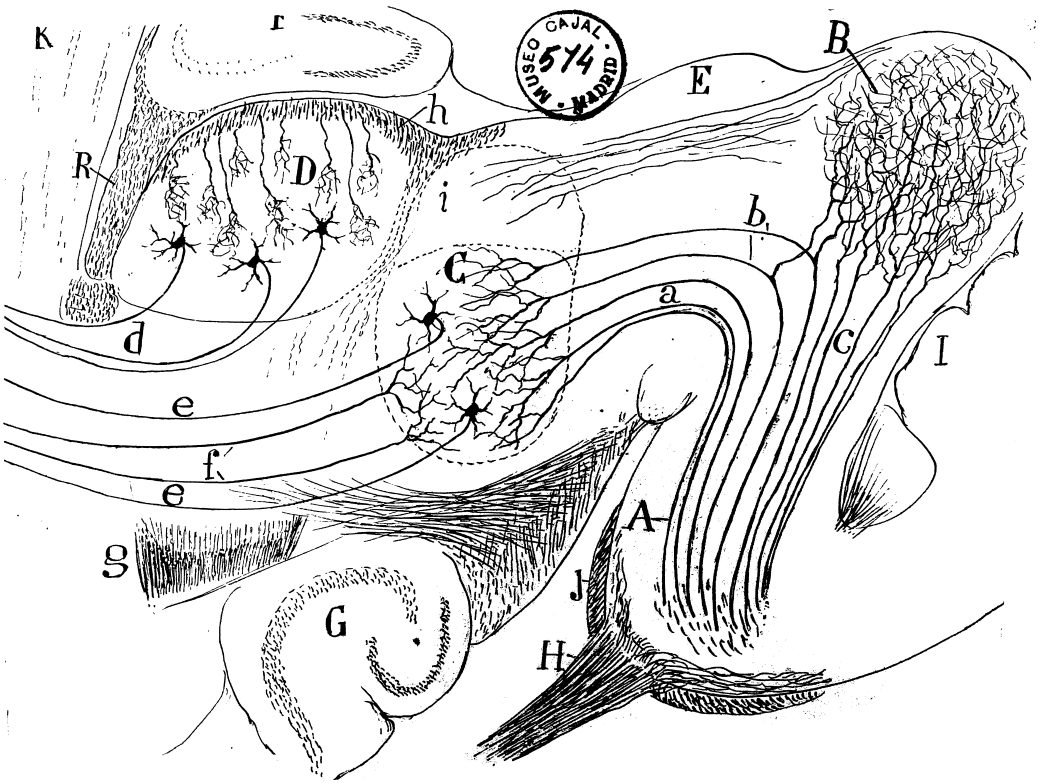



Supplementary figure 5.1: **Minimal interactive model**
(**A**) The transition matrix $T_{ij}$ necessarily fails in distinguishing anagrams because the inputs encode no information about the temporal order. In the example, *god* and *dog* reactivate with the same strength. (**B**) The timescale of the hidden state $\tau$ determines the accuracy in integrating the inputs. Timescales too short (upper panels) lead to memory-less processes that can recognize short words (*lop*) but fail on longer ones (*doll*). If the timescale is too long (bottom panel, $\tau = 150\,\text{ms}$), the activity from the previous input presentation merges in the following word, impairing lexical access. (**C**) Average recognition across the lexicon six of the lexica tested (*DIGITS* is omitted). Except for the *TISK* lexicon, the others present a darker shade on the diagonal; recognition is optimal when the population timescale is 1 to 1.5 times longer than the input timescale. The average of the first five panels is presented in Fig.5.5*E*.




Supplementary figure 5.2: **Word score measure of the dynamical system model**
To measure word recognition, we computed the average difference between the target node's activity and the envelope of the competitor's activity within 150 ms from the uniqueness point; the average is then normalized against the total area, it yields a number in the interval $(-1, 1)$. The colored curves are the target populations, while the black curves account for the competitors. When the targets' activity dominates over the other words in the lexicon (colored areas vs. black areas), recognition is successful, and the score is positive. Comparison of the activity traces of four words from the *Overlap* lexicon. The targets have, on average, larger activity than the competitors. The score is computed as the difference between the colored and black areas divided by the area of the former. Thus, the score is positive when recognition is successful. The score is then divided by the activity envelope average (target plus competitors), resulting in values in the interval (-1,1) TODO: recompute the values with normalization.

# 6 | General discussion



## 6.1 Summary of the experimental chapters

In this dissertation, I proposed a link between the neurobiology of cortical networks and the psychological phenomena observed during lexical access and selection. The work relies on a novel neuron model, the Tripod neuron (Quaresima et al., 2022). It was derived from electrophysiological evidence on dendritic processing and accounts for important features of pyramidal cells such as segregated integration, NMDA spikes, and shunting inhibition. These neuronal properties are not captured by point-neuron models and are expected to play a role in the overall computational capacities of network models (Larkum, 2022; London & Häusser, 2005; Papoutsi et al., 2014; Poirazi & Papoutsi, 2020; Spruston, 2008).

A careful analysis of the Tripod neuron model was presented in **Chapter 2**, where two sets of physiological parameters from mouse and human cortical cells were compared. In the latter case, the neuron showed a strong non-linear dendritic response with prolonged membrane depolarization. I refer to this as *dendritic memory* because it maintains information about previous synaptic inputs on short timescales. Later, I showed that the Tripod neuron achieves various computational tasks that require temporal integration. I conclude that segregated dendritic integration is a computational primitive that introduces a processing memory in the order of a hundred milliseconds in the neuron model and expands its capacity for sequence processing.

The neuron model was further investigated in **Chapter 3**. Here, I analyzed the cell's response to synaptic bombardment, as observed in cortical networks. Following previous results on the balance of excitation and inhibition (E/I) in a single-compartment model (Kuhn et al., 2004), I investigated the conditions for dendritic balance under synaptic activity with a high input rate. In the E/I balanced condition, the Tripod neuron reproduced distinctive cortical dynamics, such as the high-conductance state and the up-down transition. The study revealed that the non-linear integration provided by NMDA receptors in dendrites also increases the model's sensitivity to fluctuations in the input and provides a better fit to the response function of biological neurons.

In **Chapter 4**, I moved from single-cell simulation to investigating a network of Tripod neurons in a simplified word recognition task. I presented the network with sequences of phonemes and the associated word forms. The network received phoneme and word inputs as spatially segregated projections; via Hebbian synaptic plasticity, the projections induced the formation of cell assemblies. The word forms were selected from lexica with eight to sixteen lexical items. Crucially, some of the lexica contained words with large phonological overlaps.

To correctly recognize words in these lexica the word memories must be sensitive to the order of the phonemes presented. Recognition was measured based on the activation of the correct word assembly in response to the sequence of phonemic inputs. Along with the dendritic neuron model, the network also included other important anatomical and physiological constraints (Fitz et al., 2024; Pulvermüller et al., 2021), such as sparse recurrent connectivity, plastic excitatory synapses, background noise, and fast and slow spiking interneurons (Duarte & Morrison, 2019; Litwin-Kumar & Doiron, 2014; Tomasello et al., 2018; Zenke et al., 2015). A novel element was the introduction of voltage-dependent inhibitory synaptic plasticity (v-iSTDP) on the interneurons that target dendritic synapses.

The combination of excitatory and inhibitory plasticity enabled the formation of stable word memories in the dendritic model. Word memories consist of potentiated synapses between phonemes and word assemblies. These heteroassociative connections were necessary for the activation of word assemblies, and their specific arrangements mediated the discrimination of words with high phonological overlap. However, these network properties were not sufficient. To achieve high recognition accuracy, the Tripod neurons had to exhibit strong non-linear integration jointly with tight dendritic inhibition. These two neuronal mechanisms governed the encoding of dendritic memory and thus the temporal integration capacity of the network. Through dendrites, the short-term memory of the input sequence interacted with the long-term memory of the stored word forms, mediated by inhibitory control. In contrast, networks with simpler point neurons did not form phoneme-to-word connections when the lexicon had words with phonological overlaps and thus failed to recognize most of the lexical items. This work demonstrates that dendrites enhance the computational capacities of the network and provide novel insights into the nature of word recognition and the mental lexicon.

The dynamics of word populations were then evaluated against core computational principles of lexical access and selection, such as incremental access, competition among lexical neighbors, and cascading information between the pre-lexical and lexical stages (McQueen, 2007; Vitevitch et al., 2018). This analysis was carried out in **Chapter 5**. The model showed broad agreement with the postulated dynamics of word form access, even though the model parameters were not explicitly tuned to reproduce human behavior: first, the firing rate of target word assemblies becomes larger than that of lexical competitor only when sufficient phonological information is provided, i.e., at the unique-

ness point; Second, stronger lexical competition delays word recognition; and third, corrupted phonemic input causes a decrease in the recognition score.

In the Tripod network model, lexical selection is expressed in the word's firing rate, but lexical access occurs in the assembly's dendritic compartments: the membrane potential integrates the stimuli throughout the word, reaching the peak at the word uniqueness point (Chapter 5, Appendix B). Importantly, the presence of dendrites in the model allowed the network to solve the Temporal Order Problem without the implementation of position-dependent phonemic representations (Hannagan et al., 2013; Magnuson et al., 2013; McClelland & Elman, 1986). In the final section of the chapter, I also reduced the dendritic network model to a smaller dynamical system with only phoneme and word nodes. The reduced model shows that introducing a slow-decaying hidden variable for the membrane potential allows for word recognition with an accuracy that is comparable to the spiking network.

Together, these results indicate that the fundamental computations in word recognition can be achieved in a biologically constrained model of the cortical tissue. The studies *bridge* the neurobiological description of brain networks and the cognitive ontology of word recognition. Starting at the implementation level and with a clear computational goal in mind, I derived the algorithmic solution that the network provides. The bridge is provided by the dynamical systems perspective on brain activity and language processing (Fitz et al., 2024).

The outcome of these studies is a two-fold mapping hypothesis. First, I propose that word-form memories are due to strong, hetero-associative synapses between phonemes and word assemblies. Second, I hypothesize that the incremental activation of word memories observed in human word recognition is due to the rise of the dendritic plateau potential following NMDA spikes in word assemblies. These links are established by systematically reducing the biological substrate to a biophysical dynamical system, via mathematical abstraction and computer simulations. The reductions aim to isolate the bio-physical dynamics that contribute to the sequence detection capacity. First, the Tripod neuron is a reduced model of a biological neuron which incorporates key aspects of dendritic integration and supports temporal integration (Chapter 2). Second, the dynamical system network is a reduced model of the biological network that retains dendritic memory through the addition of hidden variables (Chapter 5).

The simulation work in this thesis spans three subfields of neuroscience; neurophysiology (Chapters 2 and 3), network computation and plasticity (Chapter 4), and cognitive-computational modeling (Chapter 5). Each chapter includes a

discussion of results and their relevance in the field. I will now focus on what I consider the three central contributions of the thesis: the role of dendritic integration in network computation, the synaptic architecture of word memories, and the time course of lexical feedback during word form selection.

## 6.2 Outcomes of the present study

### 6.2.1 Dendritic integration and network computations

One of the contributions of this work concerns the role played by dendrites in the computations carried out in neural networks. The question is now center stage in computational neuroscience due to the extensive experimental and computational evidence indicating that dendrites perform essential functions beyond the linear addition of inputs to the soma (London & Häusser, 2005; Payeur et al., 2019; Poirazi et al., 2003; J. Schiller et al., 2000). Remarkably, human dendrites seem to have unique physiological properties (Eyal et al., 2018, 2016; Gidon et al., 2020), that endow them with additional processing capabilities. The question of whether modeling dendrites is conceptually useful is addressed in a recent perspective article by Larkum (2022). The author argues that dendrites are likely to be an important level of resolution for modeling and understanding cortical networks. The somatocentric view may be too simplistic and not adequately capture the computational function of biological neurons. My results are in line with such a view. I show that the reduced dendritic model presented in Chapter 2—the Tripod neuron—expresses a form of temporal integration that is not available in point-neuron models.

Among the functional primitives provided by segregated dendritic compartments, dendritic memory is the most relevant for the present work. Dendritic memory is a form of short-term, or processing memory (Fitz et al., 2024; Petersson & Hagoort, 2012). Information is maintained in the depolarized state of a dendritic compartment and lasts approximately a hundred milliseconds. The physiological underpinnings of dendritic memory are the electrical segregation of the dendritic compartments from the soma and the regenerative currents due to voltage-gated receptors, i.e., NMDARs. Notably, the possibility that plateau potentials could encode transient memories was also proposed by Major et al. (2008) and Augusto and Gambino (2019). However, it had yet to be tested in a computational model of cognitive function. Only recently it was shown that segregated dendritic integration can be used to process sequential patterns of stim-

uli (Leugering et al., 2023), achieving results that are similar to those presented in Chapter 4. In addition, dendritic memory is not the only short-term memory component available to the cell. Fitz et al. (2020) proposed another neuronal adaptation mechanism for memory maintenance, which is helpful for processing inputs with temporal structure (Pedrelli & Hinaut, 2022; Salaj et al., 2020). Neuronal adaptation can be modeled as a hyper-polarizing current (Brette & Gerstner, 2005), i.e., a slow dynamic variable coupled to the temporal evolution of the membrane potential. However, such a current reduces neuronal excitability and may therefore limit the capacity of neurons to form memories through Hebbian plasticity. Thus, dendritic memory could have a twofold role. On the one hand, it maintains information about previous synaptic activity, and on the other hand, it interacts with long-term plasticity and facilitates synaptic potentiation.

In chapters 2 and 3, I also investigated differences in neuronal integration due to varying dendritic geometry. I showed that long dendritic compartments (distal configurations) endow the neuron with a longer memory span than short ones (proximal configurations). Conversely, the proximal compartments have a stronger impact on the somatic activity than distal ones and can elicit burst firing. The asymmetrical Tripod neuron, with one distal and one proximal dendrite, exhibits several distinct firing patterns while maintaining long-lasting dendritic memories. However, direct comparisons of different dendritic and wiring configurations in the network (Chapter 4) revealed that neither asymmetry nor additional dendritic branches enhanced network capacity for word recognition. The present study does not yet offer a satisfactory explanation of why and how models with one or two dendrites achieve the same recognition score. This result is even more puzzling if one considers that Tripod neurons with two segregated dendritic branches form only half of the synaptic contacts within the assemblies than single–dendrite neurons (Fig.4.5). The reduced number of synaptic connections in neurons with two dendrites suggests that these models are more efficient in encoding memories. This result agrees with previous computational evidence which associates an increase in the number of dendritic branches with larger storage capacity (Baronig & Legenstein, 2024; Bono & Clopath, 2017; Cazé, Jarvis, Foust, & Schultz, 2017; Hawkins & Ahmad, 2016; Kastellakis & Poirazi, 2019; Legenstein & Maass, 2011; Mel, 1992; G. R. Yang et al., 2016) and with the recently formulated hypothesis that *compartmentalized non-linear dendritic branches serve as the fundamental memory unit of the nervous system* (The dendritic engram, Kastellakis, Tasciotti, Pandi, & Poirazi, 2023). However, whether our model implementation can fully leverage the *dendritic engram* to in-

crease its word recognition capacity and increase the recognition scores remains an open empirical question for future studies.

The absence of functional differences between symmetric, asymmetric, and single-dendrite configurations may be due to a more general limitation of the neuron model implemented in this thesis. The Tripod neuron is a model of the basal region of the pyramidal cell (Eyal et al., 2018; Spruston, 2008), and it expresses only one (i.e., the NMDA-spike) of the several dendritic non-linearities observed in cortical cells (Gidon et al., 2020; Larkum et al., 2022; London & Häusser, 2005). Crucially, the Tripod lacks a mechanism for the generation of $Ca^{2+}$ (Calcium) spikes, a type of dendritic action potential that is widely observed in pyramidal cells of the mammalian cortex (Larkum et al., 2007, 2022). Typically, Calcium spikes signal the simultaneous activation of multiple dendritic branches in the apical tuft, and they interact non-linearly with backpropagating action potentials from the soma (Hay, Hill, Schürmann, Markram, & Segev, 2011; Larkum, Nevian, Sandler, Polsky, & Schiller, 2009). Functionally, $Ca^{2+}$ spikes play a role in modulating perception (Takahashi et al., 2020; Takahashi, Oertner, Hegemann, & Larkum, 2016), and have been proposed as a neural correlate of the integration of internal states (expressed on apical dendrites) and sensory experiences (targeting the basal region) (Larkum, 2013). Because the Tripod neuron does not implement this dendritic mechanism, it is possible that it is not equipped with sufficient physiological machinery to distinguish the co-activation of one or multiple branches, limiting its capacity to instantiate the dendritic engram. Modeling $Ca^{2+}$ spikes in the Tripod neuron would allow us to study the cortical mechanism for integrating sensory (e.g., acoustic) and contextual linguistic information (e.g., semantic). I will return to this aspect in the Outlook section below (Section 6.3).

### 6.2.2 Word memories as cell assemblies

The second main contribution of this thesis concerns theoretical insights into the synaptic organization of long-term memories with sequential structure, specifically word-form memories. Recently, Poeppel and Idsardi (2022) argued that *"we don't understand how the brain stores anything, let alone words"*, highlighting the fact that the mechanisms supporting long-term storage of linguistic units are not fully understood (see also Gallistel, 2021). This thesis offers a computational implementation of word memories based on neurobiological constraints. I show that word memories can be stored as cell assemblies where synaptic structure encodes the order relationship among phonemes. Importantly, the network

model had to include segregated dendritic compartments with pronounced non-linearities and strong dendritic inhibition to form memories in the face of phonological overlap. When these conditions were applied, word memories could be acquired rapidly from sparse input and were maintained robustly over time (Chapter 4, Appendix B).

Similar ideas about word memories and cortical engrams were already discussed in Pulvermüller (1999). Previous cortical models of the mental lexicon, however, focused on the acquisition and activation of word memories as distributed neural representations (Garagnani et al., 2009; Tomasello et al., 2018). In addition, these studies did not address lexical access and selection based on the recognition of phoneme sequences. Thus, the present work is novel in that it can achieve word form recognition. One important outcome of this study is that, through vSTDP, the network with dendritic structure converges towards a specific synaptic organization, which is necessary to recollect word memories with phonological overlap. Such a synaptic structure was already proposed in previous theoretical work (Sequence Detector networks, Knoblauch & Pulvermüller, 2005; Pulvermüller, 2003). Although Sequence Detector networks were not tested in simulations, the authors provided a mathematical demonstration that these networks would selectively activate distinct assemblies depending on the sequential order of input stimuli. This capacity relies on a relatively simple synaptic configuration where later inputs had stronger synaptic connections onto engrams than early ones. A similar structure emerged in the Tripod network through vSTDP (Fig 5.6). However, the present work does not yet address exactly *how* this structure between phonemes and words is achieved through STDP. Based on experience with the model, I hypothesize that the slow decay of the dendritic dynamics and the homeostatic synaptic mechanism cooperate to form stronger synapses in later phonemes. Further investigations should manipulate the plasticity rule to test this hypothesis explicitly.

The problem of categorizing inputs with sequential structure is not specific to language processing. In computational neuroscience and biologically inspired machine learning, several spiking network-based algorithmic solutions to sequence recognition can be found. However, these models typically use biologically implausible learning rules. A classical example is the Tempotron (Gütig & Sompolinsky, 2006, 2009), a supervised plasticity rule for leaky-integrate-and-fire point-neuron models, which optimizes the pre-synaptic weights such that a trained neuron fires depending on different classes of spike input patterns. Other approaches leveraged reservoir computing (Buonomano & Maass, 2009;

Duarte et al., 2018; Fitz et al., 2020; Nicola & Clopath, 2017) and more recently gradient-based algorithms for SNN, such as eligibility propagation (Bellec et al., 2019) and surrogate gradient learning (Zenke & Vogels, 2020). In these cases, temporal integration is intertwined in the network latent space and requires supervised read-out algorithms (Duarte & Morrison, 2014; Fitz et al., 2020). Whether the same read-out computation can be performed by a network assembly remains unclear because, besides the supervised algorithm, the read-out layer often violates Dale's principle [1]. In addition, achieving sequence recognition through cell assemblies rather than read-out neurons offers advantages concerning the robustness of the process. Cell assemblies can maintain systematic input-output transformations in the face of synaptic or neuronal turnover (Fauth & van Rossum, 2019; Kossio, Goedeke, Klos, & Memmesheimer, 2021).

Another aspect of the synaptic configuration that requires further investigation is the functional role of recurrent connections among the pre-lexical assemblies. These connections may contribute to the activation of the target word assembly by sharpening the network response to the external stimuli and encoding the lexicon's phonotactic and phonemic transition probabilities. For example, in the *Overlap* lexicon, the phoneme transition ($P \rightarrow O$) is more likely to occur than the transition ($L \rightarrow D$). We have shown that removing these connections caused small, although significant, drops in the recognition capacity. However, the networks were driven by strong external inputs, which mostly override reverberating activity from previously activated phonemes and this might have obscured the functional role of these connections. Future studies should try to balance the magnitude of external stimulation with the internal activity from recurrent connections and quantify the contribution of the phoneme engrams to word recognition. Reducing the external drive would also clarify the impact of word-to-phonemes feedback connections during and after lexical access and selection. At the present stage, these connections have a limited impact on the activity of phoneme populations.

### 6.2.3 Feedback activity from lexical to pre-lexical assemblies

Although feedback connections were weaker than feedforward ones, it was possible to characterize the time course of lexical feedback on the phoneme populations. In the human word recognition system, feedback streams support rec-

[1] Dale's principle states that a given neuron contains and releases only one neurotransmitter and exerts the same functional effects at all of its termination sites. Instead, read-out neurons of different classes associate positive and negative weights to the same pre-synaptic cell.

ollection when phonological features are ambiguous. The degree of *acceptable* mismatch between the phonological evidence and the retrieved word form depends on the linguistic and semantic context of the utterance (Connine, 1994, 2004; Whalen, 1991). For example, if a word is phonologically similar to another word that could be used in the same context, these two words may be confused. Conversely, if the context strongly constrains the interpretation of the phonological information, one of the words is recognized more easily (Connine, Titone, Deelman, & Blasko, 1997; Dahan et al., 2001). Contextual inference of phonological information may rely on feedback streams from the associative cortex to the sensory-perceptual areas (de Lange, Heilbron, & Kok, 2018; Heilbron & Chait, 2018; Heilbron, Richter, Ekman, Hagoort, & de Lange, 2020; Rao & Ballard, 1999) and there is broad agreement that when features are predictable based on previous context, these predictions influence the neural activity of pre-lexical features in the STG (de Lange et al., 2018; Heilbron, Ehinger, Hagoort, & de Lange, 2019; Leonard, Baud, Sjerps, & Chang, 2016). However, neither the exact time course nor the neural underpinnings of lexical feedback are well understood.

In speech comprehension, there are two competing hypotheses concerning the functional scope of feedback from lexical to the pre-lexical stage. The interactive hypothesis posits an online, bidirectional stream of information between the lexical and pre-lexical stages (Magnuson et al., 2018). This hypothesis is expressed in terms of the recurrent structure of connectionist models, such as (McClelland & Elman, 1986, TRACE) and (Hannagan et al., 2013, TISK), in which pre-lexical and lexical nodes interact in both directions. The autonomous hypothesis, in contrast, postulates a degree of functional encapsulation for the pre-lexical stage (Norris et al., 2000). In this view, the feedback information does not sharpen pre-lexical representations online; its contribution is to adapt the perceptual receptive fields to the novel input and to update the priors of pre-lexical categories for future processing (Kleinschmidt & Jaeger, 2015; Norris, 2003; Norris, McQueen, & Cutler, 2016). While the arguments of this debate are discussed in Chapter 5 in greater detail, here I highlight how a neurobiologically grounded model makes explicit the constraints on the time course of online interactions.

In Chapters 4 and 5, I showed that word assemblies were fully activated after the uniqueness point which, on average, occurred 100 ms from the stimulus onset. Accordingly, when the activity of phoneme populations was measured in word vs. non-word conditions, lexical feedback occurred after 150 ms from the

onset of the ambiguous pre-lexical stimulus. Thus, lexical feedback was not immediate. Arguably, online lexical feedback requires the activation of the word form representations. This is possible in interactive models (e.g., TRACE) because the node activity increases throughout the stimulus presentation and can immediately reverberate onto the pre-lexical units. However, in the Tripod networks, the integration process is not observable in the population firing rate because the membrane potential of assemblies depolarizes slowly. The firing rate increases only after sufficient disambiguating information is presented, that is, at the uniqueness point. Therefore, the activity of the word form representations is not accessible to the pre-lexical stages during lexical access. In other words, feedback only occurs after lexical selection. Thus, the outcome of my simulations suggests that sensory information is not *interactively* processed in the lexical and pre-lexical stages. Such temporal constraints are in line with some experimental observations. For example, Gwilliams et al. (2018) has shown that neural activity is sensitive to phonemic ambiguity both in the early stage of the acoustic input and after the word uniqueness point. The phonemic ambiguity *re-emerges* a hundred milliseconds after the acoustic signal has vanished. Similarly, Cibelli et al. (2015) could detect significant differences between words and pseudo-words, such as cohort size, only after hundreds of milliseconds from the stimulus onset.

Although lexical activation in the model does not influence phonological features during the presentation of the stimulus, this does not imply that lexical information will not affect future words. In line with this, the results in Chapter 5 indicate that there was a lexical bias for the phoneme populations connected to a lexical assembly and it occurred after lexical selection. This effect is observed in the depolarization of the membrane potential of the pre-lexical populations connected to the lexical item. Thus, in the model, feedback has a facilitating effect on the activation of pre-lexical assemblies. In addition, dendritic depolarization interacts with the Hebbian plasticity mechanism in the network and promotes the re-organization of synaptic structure. Hence, it may also support perceptual adaptation to acoustic contexts or speaker identities (Clarke & Garrett, 2005; Norris, 2003). However, to leverage feedback information of this type, the network model would require additional plasticity mechanisms that can keep track of the synaptic activity during the stimulus. This form of plasticity has recently been observed in behaving animals during delayed reward tasks (Bittner et al., 2017) and especially involves the $Ca^{2+}$ messenger released during an NMDA spike (Gonzalez, Negrean, Liao, Polleux, & Losonczy, 2023).

This form of synaptic plasticity acts on the timescale of seconds, and thus the name Behavioral Timescale Synaptic Plasticity (BTSP). Future work should test an explicit perceptual adaptation protocol and implement BTSP, or similar plasticity rules, to investigate whether the lexical feedback observed in the model supports the re-organization of the perceptual space.

## 6.3 Thesis outlook

Before moving forward to the research questions stemming from the present results, I would like to highlight that this work is an interdisciplinary effort that ties together computational neurobiology and psycholinguistics. The results demonstrate that this approach can open novel perspectives in all the research areas involved. For instance, the study of temporal integration on the timescale of words contributes to the understanding of the temporal computations occurring in networks with dendrites. The biological models provided rely on neurobiological evidence and clarify the constraints of the adaptive dynamical system that supports language, i.e., the brain. The constraints are determined by neurobiological research and they determine the building blocks from which causal models are synthesized (Fitz et al., 2024; Pulvermüller et al., 2021). Because there is a finite set of processes and mechanisms that are known to occur among brain cells, the model can constructively verify which of them contribute to the linguistic task at them.

The word recognition model outlined in this thesis is a work in progress since causal models need to constantly be refined. Because most of my efforts went into building the neuron model and the functional network, less time was available for crafting the vocabularies. One linguistic imprecision concerns the actual statistics of phonological overlap in human lexica, which is not reflected in the vocabularies used in Chapter 4 and Chapter 5. I also refer to the pre-lexical features as *phonemes*, but they are just letters corresponding to the written word forms. Therefore, the presented model of spoken word form accounts for a language that has English-like words, rather than English. The use of letters as inputs, however, creates *more* ambiguity for the model than phonemes would (e.g., the grapheme /o/ in *doll* and *poll* would be pronounced differently in British English). As I write, I am analyzing further lexica with phonological categories, and the results are fully consistent with the observations in Chapter 5.

Despite these linguistic simplifications, the model is a step towards neurobiological realism. As pointed out by Magnuson and Crinnion (2022) in a recent

review, *"future progress will likely require a deeper understanding of the neurobiological foundations of speech processing guided by innovative, neurally-realistic models" (p. 490)*. Thus, I will focus on three main future projects that leverage biological plausibility to address some relevant questions in spoken word recognition.

### 6.3.1 Phonological variability, suprasegmental information, and incremental learning

The lexica that were used in Chapters 4 and 5 were borrowed from previous studies, or designed ad-hoc, to test the implications of phonological overlap. However, they also lack other important properties of spoken words in English, e.g., variability in the duration of phonemes, stress, and syllabic structure. Precise acoustic information is critical in spoken word recognition (Fox et al., 2020; Gwilliams et al., 2018; McQueen, 2007) and including it in the model is likely to change the outcome of the selection process. In some cases, variability in phoneme duration, or the presence (absence) of stress may contribute to coping with phonological overlap. Thus, testing the model with more realistic phonological units is desirable and may provide further insights. For example, it is currently unclear whether word selection in the network would still work when phonological segments do not have constant length. Secondly, using more realistic stimuli would also allow a direct comparison of modeling results with behavioral and electrophysiological studies. Such comparisons would help to test the model against experimental evidence, beyond the computational principles of word recognition established in Chapter 5. Since the dendritic model evolves in real physical time, it is expected to yield fine-grained temporal predictions that might inspire novel experiments (Magnuson et al., 2013; Vitevitch et al., 2018).

Another change that one could introduce concerns the learning protocol in the associative phase. Phonemes and words were presented together during learning, with no intermediate sub-lexical representations between phonemic segments and words. However, access to representations such as syllables or morphemes has been observed experimentally (Frye et al., 2008; Gwilliams, 2020; Tabossi, Collina, Mazzetti, & Zoppello, 2000). The simultaneous learning of pre-lexical features and words is also at odds with evidence from language acquisition, which indicates that babies learn short sub-lexical units first (Mehler, Segui, & Frauenfelder, 1981). Future simulations should consider suprasegmen-

tal structures and test whether incremental learning of phonemic, syllabic, morphemic, and word-form assemblies contributes to word recognition. Presumably, this could enhance the recognition of long words and also the discrimination of short ones in lexica with large phonological overlap. Moreover, consolidating sub-lexical assemblies first could boost the learning speed for novel words and enable fast mapping of phoneme sequences to novel words (Constant, Pulvermüller, & Tomasello, 2023).

### 6.3.2 Explicit modeling of acoustic and phonetic features

The present model assumes that the transformation from low-level acoustic features to abstract phonemic categories (e.g., phonemes and allophones, Mitterer et al., 2018) occurs in an unspecified upstream circuit. Thus, I omit modeling normalization, which is one of the components of the pre-lexical stage. This operation parses acoustic features into categories (Magnuson et al., 2013; McQueen, 2005) and is believed to occur in the STG (Bidelman, Moreno, & Alain, 2013; Chang et al., 2010). Normalization is necessary to retrieve prototypical word forms from the variable acoustic production of human speakers. In addition, listeners can also adapt their categorical priors to the acoustic context, which is thought to occur via feedback connections from lexical areas (de Lange et al., 2018; Kleinschmidt & Jaeger, 2015).

A model of normalization should account for the mapping of variable speech onto word forms. It would provide insights into how acoustic features are unified into sub-segmental (e.g., voice onset time) and segmental (e.g., phonemes) categories and how brain networks integrate and adapt to novel acoustic inputs. However, to include these computations in the model, one must first identify a meaningful encoding of sensory information reaching the cortex. This can be done through spike encodings that reproduce the physical transformations of acoustic signals into spike trains occurring in the cochlea. I re-implemented two recently published spike encodings (Cramer, Stradmann, Schemmel, & Zenke, 2020; Pan et al., 2020) in a point-neuron network model without dendrites. In these experiments, STDP did not lead to the emergence of categorical representations of the variable acoustic inputs. Future studies that harness the temporal integration properties of dendrites might lead to different outcomes.

Increased linguistic realism would come at the cost of more input features and larger superposition between cell assemblies in the network. Although I have not observed adverse effects of lexicon size on recognition (Chapter 5, Appendix A), the overlap of phonological features may negatively affect network behavior. A

solution suggested by biology is the spatial organization of receptive fields in the STG, tuned to hierarchies of increasingly longer chunks of inputs (Berezutskaya et al., 2017; Sjerps & Chang, 2019). To reproduce this topology, one may consider cortical circuits with spatial structure and non-random synaptic connections. Such a circuit could receive input projections that are organized based on similarities between phonological units.

### 6.3.3 Lexical integration in cortical circuits

Eventually, one should aim to include the last stage of word recognition in the model, lexical integration. This would imply modeling the interface between phonology and semantics (Jackendoff, 2007) and thus the unification of lexical items into sentences (Fitz et al., 2020; Hagoort, 2005; Uhlmann, 2020). Also in this case, knowledge of the human brain biology can guide the model constituents. In neocortex, neurons are organized into six layers, each expressing neuronal types with diverse morphology, electrophysiology, and connectivity (Braitenberg & Schüz, 1998b; Palomero-Gallagher & Zilles, 2017; Zilles & Amunts, 2009). Computational studies indicate that this structure influences both the ways in which information is processed within the local circuit and how it is integrated across brain areas (Haeusler & Maass, 2007; Haeusler et al., 2009; Heinzle, Hepp, & Martin, 2007; Markram et al., 2015).

These studies are based on the principle that the layer structure and its internal connectivity, known as canonical microcircuit, is a fundamental computational unit of the neocortex (Bastos et al., 2012; Douglas & Martin, 2007). My network model, which features Tripod neurons with only the basal region of pyramidal cells, would loosely correspond to the superior layers of a microcircuit. It lacks the large pyramidal cells of the inferior layers from which cortical-cortical connections depart (Harris & Shepherd, 2015), the rich apical dendritic tufts that collect feedback from the top layers (Larkum et al., 2007), and the highly interconnected granular cells of the central layer, which gather most of the incoming connections from the thalamus (Atencio, Sharpee, & Schreiner, 2009).

In the present model, the lack of specificity in the populations that receive, process, and relay the incoming signals may reduce the present model's capacity to integrate acoustic and contextual information and prevent it from accounting for language processing beyond single-word recognition. A neurobiologically grounded model of word recognition in context would require the implementation of the rich layer structure of the neocortex. Adding two more layers, i.e., the granular and infragranular ones, together with the relevant neuron types and

connectivity, would already offer new insights into the neuronal dynamics supporting lexical integration. To this aim, our group has started developing a Julia software package that embeds the Tripod neuron in a reduced layer structure. Modeling this structure would greatly increase biological realism and therefore take the model one step closer to exhibiting human-like language behavior.

## 6.4 Final remarks

This dissertation has established a computational bridge between the cellular biology of the brain and the psychological operations carried out during word recognition. The model accounts for some aspects of word recognition, but much has yet to be done to understand the biological foundations of our capacity for language. Nevertheless, the dynamical system presented here instantiates a constructive model that connects physical variables, dendrites, and synapses to linguistic notions such as phonemes, and words. It constitutes a mapping hypothesis between the units of language and basic neurobiological processes. To validate this hypothesis, it is necessary to complement the simulation work with experiments that can test the model's predictions. This can only be achieved with interdisciplinary collaborations between theorists and experimenters. I hope the work presented here will foster future interaction between researchers in language modeling, experimental psycholinguistics, and computational neurobiology.

# References


Adesnik, H., & Naka, A. (2018). Cracking the Function of Layers in the Sensory Cortex. *Neuron*, *100*(5), 1028–1043. doi: 10.1016/j.neuron.2018.10.032

Adolfi, F., Bowers, J. S., & Poeppel, D. (2023). Successes and critical failures of neural networks in capturing human-like speech recognition. *Neural Networks*, *162*, 199–211. doi: 10.1016/j.neunet.2023.02.032

Ahmad, S., & Hawkins, J. (2016). How do neurons operate on sparse distributed representations? A mathematical theory of sparsity, neurons and active dendrites. *ArXiv160100720 Cs Q-Bio*.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models. *J. Mem. Lang.*, *38*(4), 419–439. doi: 10.1006/jmla.1997.2558

Almeida-Filho, D. G., Lopes-dos-Santos, V., Vasconcelos, N. A. P., Miranda, J. G. V., Tort, A. B. L., & Ribeiro, S. (2014). An investigation of Hebbian phase sequences as assembly graphs. *Front Neural Circuits*, *8*, 34. doi: 10.3389/fncir.2014.00034

Amit, D. J. (1995). The Hebbian paradigm reintegrated: Local reverberations as internal representations. *Behav. Brain Sci.*, *18*(4), 617–657. doi: 10.1017/S0140525X00040164

Amodei, D., Anubhai, R., Battenberg, E., Case, C., Casper, J., Catanzaro, B., … Zhu, Z. (2015). Deep Speech 2: End-to-End Speech Recognition in English and Mandarin. *ArXiv151202595 Cs*.

Anderson, J. S., Lampl, I., Gillespie, D. C., & Ferster, D. (2000). The contribution of noise to contrast invariance of orientation tuning in cat visual cortex. *Science*, *290*(5498), 1968–1972. doi: 10.1126/science.290.5498.1968

Anderson, J. S., Lampl, I., Reichova, I., Carandini, M., & Ferster, D. (2000). Stimulus dependence of two-state fluctuations of membrane potential in cat visual cortex. *Nat Neurosci*, *3*(6), 617–621. doi: 10.1038/75797

Andruski, J. E., Blumstein, S. E., & Burton, M. (1994). The effect of subphonetic differences on lexical access. *Cognition*, *52*(3), 163–187. doi: 10.1016/0010-0277(94)90042-6

Antic, S. D., Zhou, W.-L., Moore, A. R., Short, S. M., & Ikonomu, K. D. (2010). The decade of the dendritic NMDA spike. *J Neurosci Res*, *88*(14), 2991–3001. doi: 10.1002/jnr.22444

Armeni, K., Willems, R. M., van den Bosch, A., & Schoffelen, J.-M. (2019). Frequency-specific brain dynamics related to prediction during language comprehension. *NeuroImage*, *198*, 283–295. doi: 10.1016/j.neuroimage.2019.04.083

Ascher, U. M., & Petzold, L. R. (1998). *Computer methods for ordinary differential equations and differential-algebraic equations* (1st ed.). USA: Society for Industrial and Applied Mathematics.

Atencio, C. A., Sharpee, T. O., & Schreiner, C. E. (2009). Hierarchical computation in the canonical auditory cortical circuit. *Proc Natl Acad Sci U S A*, *106*(51), 21894–21899. doi: 10.1073/pnas.0908383106

Augustinaite, S., Kuhn, B., Helm, P. J., & Heggelund, P. (2014). NMDA spike/plateau potentials in dendrites of thalamocortical neurons. *Journal of Neuroscience*, *34*(33), 10892–10905. doi: 10.1523/JNEUROSCI.1205-13.2014

Augusto, E., & Gambino, F. (2019). Can NMDA spikes dictate computations of local networks and behavior? *Front. Mol. Neurosci.*, *12*. doi: 10.3389/fnmol.2019.00238

Avermann, M., Tomm, C., Mateo, C., Gerstner, W., & Petersen, C. C. H. (2012). Microcircuits of excitatory and inhibitory neurons in layer 2/3 of mouse barrel cortex. *J. Neurophysiol.*, *107*(11), 3116–3134. doi: 10.1152/jn.00917.2011

Baggio, G., & Hagoort, P. (2011). The balance between memory and unification in semantics: A dynamic account of the n400. *Language and Cognitive Processes*, *26*(9), 1338–1367. doi: 10.1080/01690965.2010.542671

Bagur, S., Bourg, J., Kempf, A., Tarpin, T., Bergaoui, K., Guo, Y., … Bathellier, B. (2022). *Emergence of a time-independent population code in auditory cortex enables sound categorization and discrimination learning* (preprint). Neuroscience. doi: 10.1101/2022.12.14.520391

Baker, C., Zhu, V., & Rosenbaum, R. (2020). Nonlinear stimulus representations in neural circuits with approximate excitatory-inhibitory balance. *PLoS Comput Biol*, *16*(9), e1008192. doi: 10.1371/journal.pcbi.1008192

Baronig, M., & Legenstein, R. (2024). Context association in pyramidal neurons through local synaptic plasticity in apical dendrites. *Front Neurosci*, *17*, 1276706. doi: 10.3389/fnins.2023.1276706

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, *76*(4), 695–711. doi: 10.1016/j.neuron.2012.10.038

Beaulieu-Laroche, L., Brown, N. J., Hansen, M., Toloza, E. H. S., Sharma, J., Williams, Z. M., ... Harnett, M. T. (2021). Allometric rules for mammalian cortical layer 5 neuron biophysics. *Nature*, *600*(7888), 274–278. doi: 10.1038/s41586-021-04072-3

Beaulieu-Laroche, L., Toloza, E. H., van der Goes, M.-S., Lafourcade, M., Barnagian, D., Williams, Z. M., ... Harnett, M. T. (2018). Enhanced dendritic compartmentalization in human cortical neurons. *Cell*, *175*(3), 643-651.e14. doi: 10/cvwp

Bellec, G., Scherr, F., Subramoney, A., Hajek, E., Salaj, D., Legenstein, R., & Maass, W. (2019). A solution to the learning dilemma for recurrent networks of spiking neurons. *Nat. Commun.*(1). doi: 10.1101/738385

Beniaguev, D., Segev, I., & London, M. (2021). Single cortical neurons as deep artificial neural networks. *Neuron*, *109*(17), 2727-2739.e3. doi: 10/gmhz55

Benucci, A., Verschure, P. F. M. J., & König, P. (2004). Two-state membrane potential fluctuations driven by weak pairwise correlations. *Neural Comput.*, *16*(11), 2351–2378. doi: 10/fr47c6

Berezutskaya, J., Freudenburg, Z. V., Güçlü, U., van Gerven, M. A., & Ramsey, N. F. (2017). Neural tuning to low-level features of speech throughout the perisylvian cortex. *J. Neurosci.*, *37*(33), 7906–7920. doi: 10.1523/JNEUROSCI.0238-17.2017

Best, C. T., Shaw, J., Docherty, G., Evans, B. G., Foulkes, P., Hay, J., ... Wood, S. (2015). From Newcastle MOUTH to Aussie ears : Australians' perceptual assimilation and adaptation for Newcastle UK vowels. *INTERSPEECH 2015 Speech Speech Better Underst. Most Important Biosignal Sept. 6–10 2015 Int. Congr. Cent. Dresd. Ger.*, 1932–1936.

Bhalla, U. (2017). Synaptic input sequence discrimination on behavioral timescales mediated by reaction-diffusion chemistry in dendrites. *eLife*, *6*. doi: 10.7554/eLife.25827

Bidelman, G. M., Moreno, S., & Alain, C. (2013). Tracing the emergence of categorical speech perception in the human auditory system. *NeuroImage*, *79*, 201–212. doi: 10/f466vp

Bittner, K. C., Grienberger, C., Vaidya, S. P., Milstein, A. D., Macklin, J. J., Suh, J., ... Magee, J. C. (2015). Conjunctive input processing drives feature

selectivity in hippocampal CA1 neurons. *Nat Neurosci*, *18*(8), 1133–1142. doi: 10.1038/nn.4062

Bittner, K. C., Milstein, A. D., Grienberger, C., Romani, S., & Magee, J. C. (2017). Behavioral time scale synaptic plasticity underlies CA1 place fields. *Science*, *357*(6355), 1033–1036. doi: 10/gbwz54

Bono, J., & Clopath, C. (2017). Modeling somatic and dendritic spike mediated plasticity at the single neuron and network level. *Nat. Commun.*, *8*(1). doi: 10.1038/s41467-017-00740-z

Bono, J., Wilmes, K. A., & Clopath, C. (2017). Modelling plasticity in dendrites: from single cells to networks. *Curr. Opin. Neurobiol.*, *46*, 136–141. doi: 10.1016/j.conb.2017.08.013

Booker, S. A., & Wyllie, D. J. (2021). NMDA receptor function in inhibitory neurons. *Neuropharmacology*, *196*, 108609. doi: 10.1016/j.neuropharm.2021.108609

Bouhadjar, Y., Wouters, D. J., Diesmann, M., & Tetzlaff, T. (2022). Sequence learning, prediction, and replay in networks of spiking neurons. *PLOS Computational Biology*, *18*(6), e1010233. doi: 10.1371/journal.pcbi.1010233

Braitenberg, V., & Schüz, A. (1998a). Cortico-Cortical Connections. In V. Braitenberg & A. Schüz (Eds.), *Cortex: Statistics and Geometry of Neuronal Connectivity* (pp. 129–134). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-662-03733-1_26

Braitenberg, V., & Schüz, A. (1998b). Layers. In V. Braitenberg & A. Schüz (Eds.), *Cortex: Statistics and Geometry of Neuronal Connectivity* (pp. 139–150). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-662-03733-128

Braitenberg, V., & Schüz, A. (2013). *Cortex: Statistics and Geometry of Neuronal Connectivity*. Springer Science & Business Media.

Branco, T., Clark, B. A., & Häusser, M. (2010). Dendritic discrimination of temporal input sequences in cortical neurons. *Science*, *329*(5999), 1671–1675. doi: 10.1126/science.1189664

Branco, T., & Häusser, M. (2010). The single dendritic branch as a fundamental functional unit in the nervous system. *Curr. Opin. Neurobiol.*, *20*(4), 494–502. doi: 10.1016/j.conb.2010.07.009

Branco, T., & Häusser, M. (2011). Synaptic integration gradients in single cortical pyramidal cell dendrites. *Neuron*, *69*(5), 885–892. doi: 10.1016/j.neuron.2011.02.006

Brette, R., & Gerstner, W. (2005). Adaptive exponential integrate-and-fire model

as an effective description of neuronal activity. *J. Neurophysiol.*, *94*(5), 3637–3642. doi: 10.1152/jn.00686.2005

Brodbeck, C., Bhattasali, S., Cruz Heredia, A. A., Resnik, P., Simon, J. Z., & Lau, E. (2022). Parallel processing in speech perception with local and global representations of linguistic context. *eLife*, *11*, e72056. doi: 10.7554/eLife.72056

Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J. Comput. Neurosci.*, 26.

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The Word Frequency Effect in Word Processing: An Updated Review. *Curr Dir Psychol Sci*, *27*(1), 45–50. doi: 10.1177/0963721417727521

Buonomano, D. V., & Maass, W. (2009). State-dependent computations: spatiotemporal processing in cortical networks. *Nat. Rev. Neurosci.*, *10*(2), 113–125. doi: 10.1038/nrn2558

Buzsáki, G. (2010). Neural Syntax: Cell Assemblies, Synapsembles, and Readers. *Neuron*, *68*(3), 362–385. doi: 10.1016/j.neuron.2010.09.023

Cajal, S. R. (1954). *Neuron Theory Or Reticular Theory?* Editorial CSIC - CSIC Press.

Callan, A. R., Heß, M., Felmy, F., & Leibold, C. (2021). Arrangement of excitatory synaptic inputs on dendrites of the medial superior olive. *J Neurosci*, *41*(2), 269–283. doi: 10.1523/JNEUROSCI.1055-20.2020

Caucheteux, C., & King, J.-R. (2022). Brains and algorithms partially converge in natural language processing. *Commun Biol*, *5*(1), 1–10. doi: 10.1038/s42003-022-03036-1

Cazé, R. D., Humphries, M., & Gutkin, B. (2013). Passive Dendrites Enable Single Neurons to Compute Linearly Non-separable Functions. *PLOS Computational Biology*, *9*(2), e1002867. doi: 10.1371/journal.pcbi.1002867

Cazé, R. D., Jarvis, S., Foust, A. J., & Schultz, S. R. (2017). Dendrites enable a robust mechanism for neuronal stimulus selectivity. *Neural Comput*, *29*(9), 2511–2527. doi: 10/ghzkjj

Cazé, R. D., & Stimberg, M. (2021). Neurons with dendrites can perform linearly separable computations with low resolution synaptic weights. *F1000Res*, *9*, 1174. doi: 10.12688/f1000research.26486.3

Chan, A. M., Dykstra, A. R., Jayaram, V., Leonard, M. K., Travis, K. E., Gygi, B., … Cash, S. S. (2014). Speech-Specific Tuning of Neurons in Human Superior Temporal Gyrus. *Cereb. Cortex*, *24*(10), 2679–2693. doi: 10/f6mg55

Chang, E. F., Rieger, J. W., Johnson, K., Berger, M. S., Barbaro, N. M., & Knight,

R. T. (2010). Categorical speech representation in human superior temporal gyrus. *Nat. Neurosci.*, *13*(11), 1428–1432. doi: 10/bksczr

Chaudhuri, R., & Fiete, I. (2016). Computational principles of memory. *Nat. Neurosci.*, *19*(3), 394–403. doi: 10.1038/nn.4237

Chen, X., Rochefort, N. L., Sakmann, B., & Konnerth, A. (2013). Reactivation of the same synapses during spontaneous up states and sensory stimuli. *Cell Reports*, *4*(1), 31–39. doi: 10.1016/j.celrep.2013.05.042

Chiu, C. Q., Martenson, J. S., Yamazaki, M., Natsume, R., Sakimura, K., Tomita, S., ... Higley, M. J. (2018). Input-specific NMDAR-dependent potentiation of dendritic gabaergic inhibition. *Neuron*, *97*(2), 368-377.e3. doi: 10/gcv3j8

Cibelli, E. S., Leonard, M. K., Johnson, K., & Chang, E. F. (2015). The influence of lexical statistics on temporal lobe cortical dynamics during spoken word listening. *Brain and Language*, *147*, 66–75. doi: 10.1016/j.bandl.2015.05.005

Clarke, C., & Garrett, M. (2005). Rapid adaptation to foreign-accented English. *The Journal of the Acoustical Society of America*, *116*, 3647–58. doi: 10.1121/1.1815131

Clopath, C., Büsing, L., Vasilaki, E., & Gerstner, W. (2010). Connectivity reflects coding: a model of voltage-based STDP with homeostasis. *Nat. Neurosci.*, *13*(3), 344–352. doi: 10.1038/nn.2479

Cohen, Y., Shen, J., Semu, D., Leman, D. P., Liberti, W. A., Perkins, L. N., ... Gardner, T. J. (2020). Hidden neural states underlie canary song syntax. *Nature*, 1–6. doi: 10/gg2wnc

Compte, A., Sanchez-Vives, M. V., McCormick, D. A., & Wang, X.-J. (2003). Cellular and network mechanisms of slow oscillatory activity (<1 hz) and wave propagations in a cortical network model. *J. Neurophysiol.*, *89*(5), 2707–2725. doi: 10.1152/jn.00845.2002

Cone, I., & Shouval, H. Z. (2021). Learning precise spatiotemporal sequences via biophysically realistic learning rules in a modular, spiking network. *eLife*, *10*, e63751. doi: 10.7554/eLife.63751

Connine, C. M. (1994). Vertical and horizontal similarity in spoken-word recognition. In *Perspectives on sentence processing* (pp. 107–120). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

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*(6), 1084–1089. doi:

10.3758/BF03196741

Connine, C. M., Blasko, D. G., & Titone, D. (1993). Do the Beginnings of Spoken Words Have a Special Status in Auditory Word Recognition? *Journal of Memory and Language*, *32*(2), 193–210. doi: 10.1006/jmla.1993.1011

Connine, C. M., Titone, D., Deelman, T., & Blasko, D. (1997). Similarity Mapping in Spoken Word Recognition. *Journal of Memory and Language*, *37*(4), 463–480. doi: 10.1006/jmla.1997.2535

Constant, M., Pulvermüller, F., & Tomasello, R. (2023). Brain-constrained neural modeling explains fast mapping of words to meaning. *Cereb. Cortex*, *33*(11), 6872–6890.

Cossart, R., Aronov, D., & Yuste, R. (2003). Attractor dynamics of network UP states in the neocortex. *Nature*, *423*(6937), 283–288. doi: 10.1038/nature01614

Costa, A., Strijkers, K., Martin, C., & Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proc. Natl. Acad. Sci.*, *106*(50), 21442–21446. doi: 10.1073/pnas.0908921106

Cowan, R. L., & Wilson, C. J. (1994). Spontaneous firing patterns and axonal projections of single corticostriatal neurons in the rat medial agranular cortex. *J Neurophysiol*, *71*(1), 17–32. doi: 10/gnj5tm

Cramer, B., Stradmann, Y., Schemmel, J., & Zenke, F. (2020). The heidelberg spiking datasets for the systematic evaluation of spiking neural networks. *IEEE Trans. Neural Netw. Learning Syst.*, 1–14. doi: 10/gkc8zx

Cunningham, M. O., Pervouchine, D. D., Racca, C., Kopell, N. J., Davies, C. H., Jones, R. S. G., … Whittington, M. A. (2006). Neuronal metabolism governs cortical network response state. *Proc. Natl. Acad. Sci. U. S. A.*, *103*(14), 5597–5601.

Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes*, *16*(5-6), 507–534. doi: 10.1080/01690960143000074

Dasika, V. K., White, J. A., & Colburn, H. S. (2007). Simple models show the general advantages of dendrites in coincidence detection. *J. Neurophysiol.*, *97*(5), 3449–3459. doi: 10/dc8tr4

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How Do Expectations Shape Perception? *Trends Cogn. Sci.*, *22*(9), 764–779. doi: 10.1016/j.tics.2018.06.002

Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., & Pallier, C. (2015). The Neural Representation of Sequences: From Transition Probabilities to Algebraic Patterns and Linguistic Trees. *Neuron*, *88*(1), 2–19. doi: 10/ggdjtd

Dembrow, N. C., & Spain, W. J. (2022). Input rate encoding and gain control in dendrites of neocortical pyramidal neurons. *Cell Reports*, *38*(7), 110382. doi: 10.1016/j.celrep.2022.110382

Denève, S., & Machens, C. K. (2016). Efficient codes and balanced networks. *Nat. Neurosci.*, *19*(3), 375–382. doi: 10.1038/nn.4243

Destexhe, A., Hughes, S. W., Rudolph, M., & Crunelli, V. (2007). Are corticothalamic 'up' states fragments of wakefulness? *Trends Neurosci*, *30*(7), 334–342. doi: 10/ctjgqb

Destexhe, A., Rudolph, M., Fellous, J. M., & Sejnowski, T. J. (2001). Fluctuating synaptic conductances recreate in vivo-like activity in neocortical neurons. *Neuroscience*, *107*(1), 13–24. doi: 10.1016/s0306-4522(01)00344-x

Destexhe, A., Rudolph, M., & Paré, D. (2003). The high-conductance state of neocortical neurons in vivo. *Nat Rev Neurosci*, *4*(9), 739–751. doi: 10.1038/nrn1198

Digby, R. J., Bravo, D. S., Paulsen, O., & Magloire, V. (2017). Distinct mechanisms of up state maintenance in the medial entorhinal cortex and neocortex. *Neuropharmacology*, *113*, 543–555. doi: 10/f9hsjd

Dong, M., Huang, X., & Xu, B. (2018). Unsupervised speech recognition through spike-timing-dependent plasticity in a convolutional spiking neural network. *PLoS ONE*, *13*(11), e0204596. doi: 10.1371/journal.pone.0204596

Doron, M., Chindemi, G., Muller, E., Markram, H., & Segev, I. (2017). Timed synaptic inhibition shapes NMDA spikes, influencing local dendritic processing and global I/O properties of cortical neurons. *Cell Rep.*, *21*(6). doi: 10.1016/j.celrep.2017.10.035

Douglas, R. J., & Martin, K. A. C. (2007). Mapping the Matrix: The Ways of Neocortex. *Neuron*, *56*(2), 226–238. doi: 10.1016/j.neuron.2007.10.017

Dringenberg, H. C. (2020). The history of long-term potentiation as a memory mechanism: Controversies, confirmation, and some lessons to remember. *Hippocampus*, *30*(9), 987–1012. doi: 10.1002/hipo.23213

Droste, F., & Lindner, B. (2017). Up-down-like background spiking can enhance neural information transmission. *eNeuro*, *4*(6). doi: 10.1523/eneuro.0282-17.2017

Duarte, R. C. F., & Morrison, A. (2014). Dynamic stability of sequential stimulus representations in adapting neuronal networks. *Front. Comput. Neurosci.*,

*8*. doi: 10.3389/fncom.2014.00124

Duarte, R. C. F., & Morrison, A. (2019). Leveraging heterogeneity for neural computation with fading memory in layer 2/3 cortical microcircuits. *PLoS Comput Biol*, *15*(4), e1006781. doi: 10/gksng8

Duarte, R. C. F., Uhlmann, M., den van Broek, D., Fitz, H., Petersson, K. M., & Morrison, A. (2018). Encoding symbolic sequences with spiking neural reservoirs. In *2018 Int. Jt. Conf. Neural Netw. IJCNN* (pp. 1–8). Rio de Janeiro: IEEE. doi: 10.1109/IJCNN.2018.8489114

Ebsch, C., & Rosenbaum, R. (2018). Imbalanced amplification: a mechanism of amplification and suppression from local imbalance of excitation and inhibition in cortical circuits. *PLoS Comput Biol*, *14*(3), e1006048. doi: 10.1371/journal.pcbi.1006048

Eisner, F., & McQueen, J. M. (2018). Speech perception. In J. T. Wixted (Ed.), *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (pp. 1–46). Hoboken, NJ, USA: John Wiley & Sons, Inc. doi: 10.1002/9781119170174.epcn301

Elman, J. L., & McClelland, J. L. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language*, *27*(2), 143–165. doi: 10.1016/0749-596X(88)90071-X

Ermentrout, B. (1992). Complex dynamics in winner-take-all neural nets with slow inhibition. *Neural Networks*, *5*(3), 415–431. doi: 10.1016/0893-6080(92)90004-3

Evans, S., & Davis, M. H. (2015). Hierarchical Organization of Auditory and Motor Representations in Speech Perception: Evidence from Searchlight Similarity Analysis. *Cereb Cortex*, *25*(12), 4772–4788. doi: 10/f73hrz

Eyal, G., Verhoog, M. B., Testa-Silva, G., Deitcher, Y., Benavides-Piccione, R., DeFelipe, J., ... Segev, I. (2018). Human cortical pyramidal neurons: from spines to spikes via models. *Front. Cell. Neurosci.*, *12*. doi: 10/ggztwh

Eyal, G., Verhoog, M. B., Testa-Silva, G., Deitcher, Y., Lodder, J. C., Benavides-Piccione, R., ... Segev, I. (2016). Unique membrane properties and enhanced signal processing in human neocortical neurons. *eLife*, *5*, e16553. doi: 10/f9q5wm

Fauth, M. J., & van Rossum, M. C. (2019). Self-organized reactivation maintains and reinforces memories despite synaptic turnover. *eLife*, *8*, e43717. doi: 10.7554/eLife.43717

Fedorenko, E., & Thompson-Schill, S. L. (2014). Reworking the language net-

work. *Trends Cogn Sci*, *18*(3), 120–126. doi: 10.1016/j.tics.2013.12.006

Fiete, I. R., Senn, W., Wang, C. Z. H., & Hahnloser, R. H. R. (2010). Spike-time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity. *Neuron*, *65*(4), 563–576. doi: 10.1016/j.neuron.2010.02.003

Fiorillo, C. D., Kim, J. K., & Hong, S. Z. (2014). The meaning of spikes from the neuron's point of view: predictive homeostasis generates the appearance of randomness. *Front Comput Neurosci*, *8*, 49. doi: 10.3389/fncom.2014.00049

Fişek, M., & Häusser, M. (2020). Are human dendrites different? *Trends in Cognitive Sciences*, *24*(6), 411–412. doi: 10/ggzhbx

Fitz, H., Hagoort, P., & Petersson, K. M. (2024). Neurobiological causal models of language processing. *Neurobiology of Language*, 1–49. doi: 10.1162/nol_a_00133

Fitz, H., Uhlmann, M., van den Broek, D., Duarte, R. C. F., Hagoort, P., & Petersson, K. M. (2020). Neuronal spike-rate adaptation supports working memory in language processing. *PNAS*, *117*(34), 20881–20889. doi: 10/gmv8cs

Fleidervish, I. A., Binshtok, A. M., & Gutnick, M. J. (1998). Functionally distinct NMDA receptors mediate horizontal connectivity within layer 4 of mouse barrel cortex. *Neuron*, *21*(5), 1055–1065. doi: 10.1016/s0896-6273(00)80623-6

Formisano, E. (2019). The "Speech ready" auditory cortex. In *Human language: From genes and brains to behavior* (pp. 481–499). The MIT Press. doi: 10.7551/mitpress/10841.003.0042

Formisano, E., De Martino, F., Bonte, M., & Goebel, R. (2008). "Who" is saying "what"? Brain-based decoding of human voice and speech. *Science*, *322*(5903), 970–973. doi: 10.1126/science.1164318

Fox, N. P., Leonard, M., Sjerps, M. J., & Chang, E. F. (2020). Transformation of a temporal speech cue to a spatial neural code in human auditory cortex. *eLife*, *9*, e53051. doi: 10/ghd8kd

Frye, R. E., Fisher, J. M., Witzel, T., Ahlfors, S. P., Swank, P., Liederman, J., & Halgren, E. (2008). Objective phonological and subjective perceptual characteristics of syllables modulate spatiotemporal patterns of superior temporal gyrus activity. *NeuroImage*, *40*(4), 1888–1901. doi: 10/bbgrff

Fuster, J. M. (1997). Network memory. *Trends in Neurosciences*, *20*(10), 451–459. doi: 10.1016/S0166-2236(97)01128-4

Gallistel, C. R. (2021). The physical basis of memory. *Cognition*, *213*, 104533. doi: 10.1016/j.cognition.2020.104533

Gallistel, C. R., & King, A. P. (2011). *Memory and the Computational Brain: Why Cognitive Science will Transform Neuroscience*. John Wiley & Sons.

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *J. Exp. Psychol. Hum. Percept. Perform.*, *6*(1), 110–125. doi: 10.1037/0096-1523.6.1.110

Garagnani, M., Wennekers, T., & Pulvermüller, F. (2009). Recruitment and Consolidation of Cell Assemblies for Words by Way of Hebbian Learning and Competition in a Multi-Layer Neural Network. *Cogn Comput*, *1*(2), 160–176. doi: 10.1007/s12559-009-9011-1

Gaskell, M. G., & Marslen-Wilson, W. D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(1), 144–158. doi: 10.1037/0096-1523.22.1.144

Gastaldi, C., Schwalger, T., De Falco, E., Quiroga, R. Q., & Gerstner, W. (2021). When shared concept cells support associations: Theory of overlapping memory engrams. *PLoS Comput Biol*, *17*(12), e1009691. doi: 10.1371/journal.pcbi.1009691

Gerstner, W., Kistler, W. M., Naud, R., & Paninski, L. (2014). *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge University Press.

Gidon, A., Zolnik, T. A., Fidzinski, P., Bolduan, F., Papoutsi, A., Poirazi, P., ... Larkum, M. E. (2020). Dendritic action potentials and computation in human layer 2/3 cortical neurons. *Science*, *367*(6473), 83–87. doi: 10.1126/science.aax6239

Gillett, M., Pereira, U., & Brunel, N. (2020). Characteristics of sequential activity in networks with temporally asymmetric Hebbian learning. *Proc Natl Acad Sci U S A*, *117*(47), 29948–29958. doi: 10.1073/pnas.1918674117

Gjorgjieva, J., Clopath, C., Audet, J., & Pfister, J.-P. (2011). A triplet spike-timing-dependent plasticity model generalizes the Bienenstock-Cooper-Munro rule to higher-order spatiotemporal correlations. *Proceedings of the National Academy of Sciences*, *108*(48), 19383–19388. doi: 10/bnksgs

Gjorgjieva, J., Drion, G., & Marder, E. (2016). Computational implications of biophysical diversity and multiple timescales in neurons and synapses for circuit performance. *Curr Opin Neurobiol*, *37*, 44–52. doi: 10.1016/j.conb.2015.12.008

Gonzalez, K. C., Negrean, A., Liao, Z., Polleux, F., & Losonczy, A. (2023). *Synaptic Basis of Behavioral Timescale Plasticity*. bioRxiv. doi: 10.1101/2023.10.04.560848

Górski, T., Depannemaecker, D., & Destexhe, A. (2021). Conductance-based adaptive exponential integrate-and-fire model. *Neural Computation*, *33*(1), 41–66. doi: 10.1162/neco_a_01342

Gow, D. W., & Gordon, P. C. (1995). Lexical and prelexical influences on word segmentation: Evidence from priming. *J. Exp. Psychol. Hum. Percept. Perform.*, *21*(2), 344–359. doi: 10/bd5xsk

Gow, D. W., Segawa, J. A., Ahlfors, S. P., & Lin, F.-H. (2008). Lexical influences on speech perception: a Granger causality analysis of MEG and EEG source estimates. *Neuroimage*, *43*(3), 614–623. doi: 10.1016/j.neuroimage.2008.07.027

Granger, R., Whitson, J., Larson, J., Lynch, G., & Lynch, G. (1994). Non-Hebbian properties of long-term potentiation enable high-capacity encoding of temporal sequences. *Proc. Natl. Acad. Sci. U. S. A.*, *91*(21), 10104–10108. doi: 10.1073/pnas.91.21.10104

Graves, A. (2012). *Supervised Sequence Labelling with Recurrent Neural Networks* (Vol. 385). Berlin, Heidelberg: Springer. doi: 10.1007/978-3-642-24797-2

Grossberg, S. (2003). Resonant neural dynamics of speech perception. *Journal of Phonetics*, *31*(3-4), 423–445. doi: 10.1016/S0095-4470(03)00051-2

Guerguiev, J., Lillicrap, T. P., & Richards, B. A. (2017). Towards deep learning with segregated dendrites. *eLife*, *6*, e22901. doi: 10.7554/eLife.22901

Gütig, R., & Sompolinsky, H. (2006). The tempotron: a neuron that learns spike timing–based decisions. *Nat Neurosci*, *9*(3), 420–428. doi: 10.1038/nn1643

Gütig, R., & Sompolinsky, H. (2009). Time-warp-invariant neuronal processing. *PLoS Biol.*, *7*(7), e1000141. doi: 10.1371/journal.pbio.1000141

Gwilliams, L. (2020). How the brain composes morphemes into meaning. *Phil. Trans. R. Soc. B*, *375*(1791), 20190311. doi: 10.1098/rstb.2019.0311

Gwilliams, L., King, J.-R., Marantz, A., & Poeppel, D. (2022). Neural dynamics of phoneme sequences reveal position-invariant code for content and order. *Nat Commun*, *13*(1), 6606. doi: 10.1038/s41467-022-34326-1

Gwilliams, L., Linzen, T., Poeppel, D., & Marantz, A. (2018). In Spoken Word Recognition, the Future Predicts the Past. *J. Neurosci.*, *38*(35), 7585–7599. doi: 10.1523/JNEUROSCI.0065-18.2018

Haeusler, S., & Maass, W. (2007). A Statistical Analysis of Information-Processing Properties of Lamina-Specific Cortical Microcircuit Models. *Cereb Cortex*, *17*(1), 149–162. doi: 10.1093/cercor/bhj132

Haeusler, S., Schuch, K., & Maass, W. (2009). Motif distribution, dynamical properties, and computational performance of two data-based cortical microcircuit templates. *J. Physiol.-Paris*, *103*(1-2), 73–87. doi: 10.1016/j.jphysparis.2009.05.006

Haga, T., & Fukai, T. (2018). Dendritic processing of spontaneous neuronal sequences for single-trial learning. *Sci Rep*, *8*(1), 15166. doi: 10.1038/s41598-018-33513-9

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in Cognitive Sciences*, *9*(9), 416–423. doi: 10.1016/j.tics.2005.07.004

Hagoort, P. (2013). MUC (Memory, Unification, Control) and beyond. *Front. Psychol.*, *4*.

Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's region and beyond. *Current Opinion in Neurobiology*, *28*, 136–141. doi: 10.1016/j.conb.2014.07.013

Hagoort, P. (2017). The core and beyond in the language-ready brain. *Neuroscience & Biobehavioral Reviews*, *81*, 194–204. doi: 10.1016/j.neubiorev.2017.01.048

Hagoort, P. (2019). The neurobiology of language beyond single-word processing. *Science*, *366*(6461), 55–58.

Hagoort, P. (2020). The meaning-making mechanism(s) behind the eyes and between the ears. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1791), 20190301. doi: 10/gg6xhm

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of Word Meaning and World Knowledge in Language Comprehension. *Science*, *304*(5669), 438–441. doi: 10.1126/science.1095455

Hannagan, T., Magnuson, J., & Grainger, J. (2013). Spoken word recognition without a TRACE. *Front. Psychol.*, *4*. doi: 10/gbfpw9

Harris, K. D., & Shepherd, G. M. G. (2015). The neocortical circuit: themes and variations. *Nat. Neurosci.*, *18*(2), 170–181. doi: 10.1038/nn.3917

Häusser, M. (2001). Synaptic function: dendritic democracy. *Current Biology*, *11*(1), R10-R12. doi: 10.1016/S0960-9822(00)00034-8

Hawkins, J., & Ahmad, S. (2016). Why neurons have thousands of synapses, a theory of sequence memory in neocortex. *Front Neural Circuits*, *10*. doi: 10.3389/fncir.2016.00023

Hay, E., Hill, S., Schürmann, F., Markram, H., & Segev, I. (2011). Models of Neocortical Layer 5b Pyramidal Cells Capturing a Wide Range of Dendritic and Perisomatic Active Properties. *PLOS Computational Biology*, *7*(7), e1002107. doi: 10.1371/journal.pcbi.1002107

Hebb, D. O. (1949). *The organization of behavior; a neuropsychological theory*. Oxford, England: Wiley.

Heilbron, M., & Chait, M. (2018). Great Expectations: Is there Evidence for Predictive Coding in Auditory Cortex? *Neuroscience*, *389*, 54–73. doi: 10.1016/j.neuroscience.2017.07.061

Heilbron, M., Ehinger, B., Hagoort, P., & de Lange, F. P. (2019). Tracking Naturalistic Linguistic Predictions with Deep Neural Language Models. *2019 Conf. Cogn. Comput. Neurosci.*. doi: 10.32470/CCN.2019.1096-0

Heilbron, M., Richter, D., Ekman, M., Hagoort, P., & de Lange, F. P. (2020). Word contexts enhance the neural representation of individual letters in early visual cortex. *Nat. Commun.*, *11*(1), 1–11. doi: 10.1038/s41467-019-13996-4

Heinzle, J., Hepp, K., & Martin, K. A. C. (2007). A microcircuit model of the frontal eye fields. *J. Neurosci.*, *27*(35), 9341–9353. doi: 10.1523/JNEUROSCI.0974-07.2007

Hemberger, M., Shein-Idelson, M., Pammer, L., & Laurent, G. (2019). Reliable Sequential Activation of Neural Assemblies by Single Pyramidal Cells in a Three-Layered Cortex. *Neuron*. doi: 10.1016/j.neuron.2019.07.017

Herstel, L. J., & Wierenga, C. J. (2021). Network control through coordinated inhibition. *Curr Opin Neurobiol*, *67*, 34–41. doi: 10.1016/j.conb.2020.08.001

Higley, M., & Contreras, D. (2006). Balanced excitation and inhibition determine spike timing during frequency adaptation. *J. Neurosci.*. doi: 10/d73cbf

Hines, M. (1989). A program for simulation of nerve equations with branching geometries. *International Journal of Bio-Medical Computing*, *24*(1), 55–68. doi: 10.1016/0020-7101(89)90007-X

Hines, M., & Carnevale, N. (2001). NEURON: a tool for neuroscientists. *The Neuroscientist*, *7*, 123–135.

Hiratani, N., & Fukai, T. (2017). Detailed dendritic Excitatory/Inhibitory balance through heterosynaptic spike-timing-dependent plasticity. *J. Neurosci.*, *37*(50), 12106–12122. doi: 10.1523/JNEUROSCI.0027-17.2017

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Phys-*

*iol.*, *117*(4), 500–544. doi: 10.1113/jphysiol.1952.sp004764

Holt, G. R., Softky, W. R., Koch, C., & Douglas, R. J. (1996). Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons. *Journal of Neurophysiology*, *75*(5), 1806–1814. doi: 10.1152/jn.1996.75.5.1806

Huang, Z. J., & Paul, A. (2019). The diversity of GABAergic neurons and neural communication elements. *Nat Rev Neurosci*, *20*(9), 563–572. doi: 10.1038/s41583-019-0195-4

Hullett, P. W., Hamilton, L. S., Mesgarani, N., Schreiner, C. E., & Chang, E. F. (2016). Human Superior Temporal Gyrus Organization of Spectrotemporal Modulation Tuning Derived from Speech Stimuli. *J. Neurosci.*, *36*(6), 2014–2026. doi: 10.1523/JNEUROSCI.1779-15.2016

Iascone, D. M., Li, Y., Sümbül, U., Doron, M., Chen, H., Andreu, V., ... Polleux, F. (2020). Whole-neuron synaptic mapping reveals spatially precise excitatory/inhibitory balance limiting dendritic and somatic spiking. *Neuron*, *106*(4), 566-578.e8. doi: 10/ggztwk

Jackendoff, R. (2007). A Parallel Architecture perspective on language processing. *Brain Res.*, *1146*, 2–22. doi: 10.1016/j.brainres.2006.08.111

Jadi, M. P., Behabadi, B. F., Poleg-Polsky, A., Schiller, J., & Mel, B. W. (2014). An augmented two-layer model captures nonlinear analog spatial integration effects in pyramidal neuron dendrites. *Proc. IEEE*, *102*(5), 782–798. doi: 10.1109/JPROC.2014.2312671

Jadi, M. P., Polsky, A., Schiller, J., & Mel, B. W. (2012). Location-dependent effects of inhibition on local spiking in pyramidal neuron dendrites. *PLoS Comput. Biol.*, *8*(6). doi: 10.1371/journal.pcbi.1002550

Jahr, C. E., & Stevens, C. F. (1990). Voltage dependence of NMDA-activated macroscopic conductances predicted by single-channel kinetics. *J. Neurosci.*, *10*(9), 3178–3182. doi: 10.1523/JNEUROSCI.10-09-03178.1990

Jercog, D., Roxin, A., Barthó, P., Luczak, A., Compte, A., & De La Rocha, J. (2017). UP-DOWN cortical dynamics reflect state transitions in a bistable network. *eLife*, *6*. doi: 10.7554/elife.22425

Jones, D. (1962). *The Phoneme: Its Nature and Use*. W. Heffer.

Jones, I. S., & Kording, K. P. (2021). Might a Single Neuron Solve Interesting Machine Learning Problems Through Successive Computations on Its Dendritic Tree? *Neural Computation*, *33*(6), 1554–1571. doi: 10.1162/neco_a_01390

Kaiser, J., Billaudelle, S., Müller, E., Tetzlaff, C., Schemmel, J., & Schmitt, S. (2022). Emulating dendritic computing paradigms on analog neuromor-

phic hardware. *Neuroscience*, *489*, 290–300. doi: 10.1016/j.neuroscience.2021.08.013

Kamondi, A., Acsády, L., & Buzsáki, G. (1998). Dendritic spikes are enhanced by cooperative network activity in the intact hippocampus. *J. Neurosci.*, *18*(10), 3919–3928.

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A., & Hudspeth, A. J. (2012). *Principles of neural science* (5th ed.). McGraw-Hill.

Kaplan, D. M., & Craver, C. F. (2011). The Explanatory Force of Dynamical and Mathematical Models in Neuroscience: A Mechanistic Perspective*. *Philos. Sci.*, *78*(4), 601–627. doi: 10.1086/661755

Kastellakis, G., & Poirazi, P. (2019). Synaptic Clustering and Memory Formation. *Front. Mol. Neurosci.*, *12*. doi: 10/gjqbfc

Kastellakis, G., Silva, A. J., & Poirazi, P. (2016). Linking memories across time via neuronal and dendritic overlaps in model neurons with active dendrites. *Cell Reports*, *17*(6), 1491–1504. doi: 10.1016/j.celrep.2016.10.015

Kastellakis, G., Tasciotti, S., Pandi, I., & Poirazi, P. (2023). The dendritic engram. *Front. Behav. Neurosci.*, *17*.

Kazanina, N., Bowers, J. S., & Idsardi, W. (2018). Phonemes: Lexical access and beyond. *Psychon Bull Rev*, *25*(2), 560–585. doi: 10.3758/s13423-017-1362-0

Kee, T., Sanda, P., Gupta, N., Stopfer, M., & Bazhenov, M. (2015). Feed-Forward versus Feedback Inhibition in a Basic Olfactory Circuit. *PLoS Comput Biol*, *11*(10), e1004531. doi: 10.1371/journal.pcbi.1004531

Kirchner, J. H., & Gjorgjieva, J. (2021). Emergence of local and global synaptic organization on cortical dendrites. *Nat Commun*, *12*(1), 4005. doi: 10.1038/s41467-021-23557-3

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, *122*(2), 148–203. doi: 10.1037/a0038695

Knoblauch, A., & Pulvermüller, F. (2005). Sequence Detector Networks and Associative Learning of Grammatical Categories. In D. Hutchison et al. (Eds.), *Biomimetic Neural Learning for Intelligent Robots* (Vol. 3575, pp. 31–53). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/11521082_3

Koch, C. (1998). *Biophysics of computation: information processing in single neurons, by* (Oxford University Press. ed., Vol. 22). Elsevier.

Koch, C. (1999). *Biophysics of computation: Information processing in single*

*neurons*. Oxford University Press.

Koch, C., Poggio, T., & Torre, V. (1983). Nonlinear interactions in a dendritic tree: localization, timing, and role in information processing. *Proc. Natl. Acad. Sci.*, *80*(9), 2799–2802. doi: 10.1073/pnas.80.9.2799

Kossio, Y. F. K., Goedeke, S., Klos, C., & Memmesheimer, R.-M. (2021). Drifting assemblies for persistent memory: Neuron transitions and unsupervised compensation. *PNAS*, *118*(46). doi: 10/gnpb7b

Kuhn, A., Aertsen, A., & Rotter, S. (2004). Neuronal integration of synaptic input in the fluctuation-driven regime. *J. Neurosci.*, *24*(10), 2345–2356. doi: 10.1523/JNEUROSCI.3349-03.2004

Kumar, A., Schiff, O., Barkai, E., Mel, B. W., Poleg-Polsky, A., & Schiller, J. (2018). NMDA spikes mediate amplification of inputs in the rat piriform cortex. *Elife*, *7*. doi: 10.7554/elife.38446

Lafourcade, M., van der Goes, M.-S. H., Vardalaki, D., Brown, N. J., Voigts, J., Yun, D. H., ... Harnett, M. T. (2022). Differential dendritic integration of long-range inputs in association cortex via subcellular changes in synaptic AMPA-to-NMDA receptor ratio. *Neuron*, *110*(9), 1532-1546.e4. doi: 10.1016/j.neuron.2022.01.025

Langille, J. J., & Brown, R. E. (2018). The Synaptic Theory of Memory: A Historical Survey and Reconciliation of Recent Opposition. *Front Syst Neurosci*, *12*, 52. doi: 10.3389/fnsys.2018.00052

Langille, J. J., & Gallistel, C. R. (2020). Locating the engram: Should we look for plastic synapses or information-storing molecules? *Neurobiol Learn Mem*, *169*, 107164. doi: 10.1016/j.nlm.2020.107164

Larkum, M. E. (2013). A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. *Trends Neurosci.*, *36*(3), 141–151. doi: 10.1016/j.tins.2012.11.006

Larkum, M. E. (2022). Are dendrites conceptually useful? *Neuroscience*, *489*, 4–14. doi: 10.1016/j.neuroscience.2022.03.008

Larkum, M. E., Nevian, T., Sandler, M., Polsky, A., & Schiller, J. (2009). Synaptic Integration in Tuft Dendrites of Layer 5 Pyramidal Neurons: A New Unifying Principle. *Science*, *325*(5941), 756–760. doi: 10.1126/science.1171958

Larkum, M. E., Waters, J., Sakmann, B., & Helmchen, F. (2007). Dendritic Spikes in Apical Dendrites of Neocortical Layer 2/3 Pyramidal Neurons. *J. Neurosci.*, *27*(34), 8999–9008. doi: 10/c3z532

Larkum, M. E., Wu, J., Duverdin, S. A., & Gidon, A. (2022). The guide to

dendritic spikes of the mammalian cortex in vitro and in vivo. *Neuroscience*, *489*, 15–33. doi: 10.1016/j.neuroscience.2022.02.009

Legenstein, R., & Maass, W. (2011). Branch-Specific Plasticity Enables Self-Organization of Nonlinear Computation in Single Neurons. *J. Neurosci.*, *31*(30), 10787–10802. doi: 10.1523/JNEUROSCI.5684-10.2011

Leonard, M. K., Baud, M. O., Sjerps, M. J., & Chang, E. F. (2016). Perceptual restoration of masked speech in human cortex. *Nat Commun*, *7*, 13619. doi: 10.1038/ncomms13619

Leonard, M. K., Bouchard, K. E., Tang, C., & Chang, E. F. (2015). Dynamic Encoding of Speech Sequence Probability in Human Temporal Cortex. *J. Neurosci.*, *35*(18), 7203–7214. doi: 10.1523/JNEUROSCI.4100-14.2015

Leugering, J., Nieters, P., & Pipa, G. (2023). Dendritic plateau potentials can process spike sequences across multiple time-scales. *Front. Cogn.*, *2*.

Leuze, C. W. U., Anwander, A., Bazin, P.-L., Dhital, B., Stüber, C., Reimann, K., … Turner, R. (2014). Layer-Specific Intracortical Connectivity Revealed with Diffusion MRI. *Cereb. Cortex N. Y. NY*, *24*(2), 328. doi: 10.1093/cercor/bhs311

Li, S., Liu, N., Zhang, X., McLaughlin, D. W., Zhou, D., & Cai, D. (2019). Dendritic computations captured by an effective point neuron model. *Proc Natl Acad Sci USA*, *116*(30), 15244–15252. doi: 10/ggcrwf

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychol. Rev.*, *74*(6), 431–461. doi: 10.1037/h0020279

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *J. Exp. Psychol.*, *54*(5), 358–368. doi: 10.1037/h0044417

Litwin-Kumar, A., & Doiron, B. (2014). Formation and maintenance of neuronal assemblies through synaptic plasticity. *Nat. Commun.*, *5*(1). doi: 10.1038/ncomms6319

Liu, G. (2004). Local structural balance and functional interaction of excitatory and inhibitory synapses in hippocampal dendrites. *Nat. Neurosci.*, *7*(4), 373–379. doi: 10.1038/nn1206

Loewenstein, Y., Mahon, S., Chadderton, P., Kitamura, K., Sompolinsky, H., Yarom, Y., & Häusser, M. (2005). Bistability of cerebellar Purkinje cells modulated by sensory stimulation. *Nat Neurosci*, *8*(2), 202–211. doi: 10.1038/nn1393

London, M., & Häusser, M. (2005). Dendritic computation. *Annu. Rev. Neurosci.*,

*28*(1), 503–532. doi: 10.1146/annurev.neuro.28.061604.135703

Luce, P. A., & Pisoni, D. B. (1998). Recognizing Spoken Words: The Neighborhood Activation Model. *Ear Hear*, *19*(1), 1–36.

Luczak, A., Bartho, P., Marguet, S. L., Buzsaki, G., & Harris, K. D. (2007). Sequential structure of neocortical spontaneous activity in vivo. *Proc. Natl. Acad. Sci.*, *104*(1), 347–352. doi: 10.1073/pnas.0605643104

Luo, L. (2015). *Principles of neurobiology*. Garland Science.

Lytton, W. W., Arle, J., Bobashev, G., Ji, S., Klassen, T. L., Marmarelis, V. Z., … Sanger, T. D. (2017). Multiscale modeling in the clinic: diseases of the brain and nervous system. *Brain Inform*, *4*(4), 219–230. doi: 10.1007/s40708-017-0067-5

Maddieson, I. (1984). *Patterns of Sounds*. Cambridge University Press.

Maes, A., Barahona, M., & Clopath, C. (2020). Learning spatiotemporal signals using a recurrent spiking network that discretizes time. *PLOS Computational Biology*, *16*(1), e1007606. doi: 10.1371/journal.pcbi.1007606

Magee, J. C., & Cook, E. P. (2000). Somatic EPSP amplitude is independent of synapse location in hippocampal pyramidal neurons. *Nat Neurosci*, *3*(9), 895–903. doi: 10.1038/78800

Magnuson, J. S., & Crinnion, A. M. (2022). Spoken word recognition. In A. Papafragou, J. C. Trueswell, & L. R. Gleitman (Eds.), *The Oxford Handbook of the Mental Lexicon* (p. 0). Oxford University Press. doi: 10.1093/oxfordhb/9780198845003.013.23

Magnuson, J. S., Mirman, D., Luthra, S., Strauss, T., & Harris, H. D. (2018). Interaction in Spoken Word Recognition Models: Feedback Helps. *Front. Psychol.*, *9*. doi: 10.3389/fpsyg.2018.00369

Magnuson, J. S., Mirman, D., & Myers, E. (2013). Spoken word recognition. In *The Oxford handbook of cognitive psychology* (pp. 412–441). New York, NY, US: Oxford University Press. doi: 10.1093/oxfordhb/9780195376746.013.0027

Major, G., Polsky, A., Denk, W., Schiller, J., & Tank, D. W. (2008). Spatiotemporally graded NMDA Spike/Plateau potentials in basal dendrites of neocortical pyramidal neurons. *Journal of Neurophysiology*, *99*(5), 2584–2601. doi: 10/fqp336

Maksimov, A., Diesmann, M., & van Albada, S. J. (2018). Criteria on balance, stability, and excitability in cortical networks for constraining computational models. *Front. Comput. Neurosci.*, *12*.

Mann, V. A., & Repp, B. H. (1981). Influence of preceding fricative on stop

consonant perception. *J. Acoust. Soc. Am.*, *69*(2), 548–558. doi: 10.1121/1.385483

Markram, H., Muller, E., Ramaswamy, S., Reimann, M. W., Abdellah, M., Sanchez, C. A., … Schürmann, F. (2015). Reconstruction and Simulation of Neocortical Microcircuitry. *Cell*, *163*(2), 456–492. doi: 10.1016/j.cell.2015.09.029

Marr, D. (2010). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. MIT Press.

Marslen-Wilson, W. D. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, *244*, 522–523. doi: 10.1038/244522a0

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1-2), 71–102. doi: 10.1016/0010-0277(87)90005-9

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*(1), 29–63. doi: 10.1016/0010-0285(78)90018-X

Marslen-Wilson, W. D., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 576–585. doi: 10.1037/0096-1523.15.3.576

Maxwell, D. J., Belle, M. D., Cheunsuang, O., Stewart, A., & Morris, R. (2007). Morphology of inhibitory and excitatory interneurons in superficial laminae of the rat dorsal horn. *J Physiol*, *584*(Pt 2), 521–533. doi: 10.1113/jphysiol.2007.140996

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86. doi: 10/cpmfpm

McClelland, J. L., Mirman, D., & Holt, L. L. (2006). Are there interactive processes in speech perception? *Trends Cogn Sci*, *10*(8), 363–369. doi: 10.1016/j.tics.2006.06.007

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, *5*(4), 115–133. doi: 10.1007/BF02478259

McMurray, B., Sarrett, M. E., Chiu, S., Black, A. K., Wang, A., Canale, R., & Aslin, R. N. (2022). Decoding the temporal dynamics of spoken word and nonword processing from EEG. *NeuroImage*, *260*, 119457. doi: 10.1016/j.neuroimage.2022.119457

McQueen, J. M. (2005). Speech Perception. In *Handbook of Cognition* (pp. 256–

276). London: SAGE Publications Ltd. doi: 10.4135/9781848608177

McQueen, J. M. (2007). Eight questions about spoken word recognition. In M. G. Gaskell (Ed.), *The Oxford Handbook of Psycholinguistics* (pp. 37–54). Oxford University Press.

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cogn Sci*, *30*(6), 1113–1126. doi: 10.1207/s15516709cog0000_79

McQueen, J. M., Dahan, D., & Cutler, A. (2003). Continuity and gradedness in speech processing. In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and Phonology in Language Comprehension and Production* (pp. 39–78). Berlin, New York: DE GRUYTER MOUTON. doi: 10.1515/9783110895094.39

McQueen, J. M., Jesse, A., & Norris, D. (2009). No lexical–prelexical feedback during speech perception or: Is it time to stop playing those Christmas tapes? *Journal of Memory and Language*, *61*(1), 1–18. doi: 10.1016/j.jml.2009.03.002

McQueen, J. M., Norris, D., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(3), 621–638. doi: 10.1037/0278-7393.20.3.621

Mehler, J., Segui, J., & Frauenfelder, U. (1981). The Role of the Syllable in Language Acquisition and Perception. In T. Myers, J. Laver, & J. Anderson (Eds.), *Advances in Psychology* (Vol. 7, pp. 295–305). North-Holland. doi: 10.1016/S0166-4115(08)60205-8

Mel, B. W. (1992). NMDA-based pattern discrimination in a modeled cortical neuron. *Neural Computation*, *4*(4), 502–517. doi: 10.1162/neco.1992.4.4.502

Mesgarani, N., David, S. V., Fritz, J. B., & Shamma, S. A. (2014). Mechanisms of noise robust representation of speech in primary auditory cortex. *PNAS*, *111*(18), 6792–6797. doi: 10/f53cp9

Metzger, S. L., Littlejohn, K. T., Silva, A. B., Moses, D. A., Seaton, M. P., Wang, R., … Chang, E. F. (2023). A high-performance neuroprosthesis for speech decoding and avatar control. *Nature*, *620*(7976), 1037–1046. doi: 10.1038/s41586-023-06443-4

Meyer, G., González-Hernández, T. H., & Ferres-Torres, R. (1989). The spiny stellate neurons in layer IV of the human auditory cortex. A golgi study. *Neuroscience*, *33*(3), 489–498. doi: 10.1016/0306-4522(89)90401-6

Miehl, C., & Gjorgjieva, J. (2022). *Stability and learning in excitatory synapses*

*by nonlinear inhibitory plasticity*. doi: 10.1101/2022.03.28.486052

Mikulasch, F. A., Rudelt, L., & Priesemann, V. (2020). Local dendritic balance enables learning of efficient representations in networks of spiking neurons. *ArXiv201012395 Q-Bio*.

Miles, R., Tóth, K., Gulyás, A. I., Hájos, N., & Freund, T. F. (1996). Differences between somatic and dendritic inhibition in the hippocampus. *Neuron*, *16*(4), 815–823. doi: 10.1016/S0896-6273(00)80101-4

Milojkovic, B. A., Radojicic, M. S., & Antic, S. D. (2005). A strict correlation between dendritic and somatic plateau depolarizations in the rat prefrontal cortex pyramidal neurons. *J Neurosci*, *25*(15), 3940–3951. doi: 10.1523/JNEUROSCI.5314-04.2005

Mirman, D., & Britt, A. E. (2014). What we talk about when we talk about access deficits. *Philos Trans R Soc Lond B Biol Sci*, *369*(1634), 20120388. doi: 10.1098/rstb.2012.0388

Mitterer, H., Reinisch, E., & McQueen, J. M. (2018). Allophones, not phonemes in spoken-word recognition. *Journal of Memory and Language*, *98*, 77–92. doi: 10.1016/j.jml.2017.09.005

Mohan, H., Verhoog, M. B., Doreswamy, K. K., Eyal, G., Aardse, R., Lodder, B. N., ... de Kock, C. P. (2015). Dendritic and Axonal Architecture of Individual Pyramidal Neurons across Layers of Adult Human Neocortex. *Cereb Cortex*, *25*(12), 4839–4853. doi: 10/f73g7f

Molnár, Z., & Pollen, A. (2014). How unique is the human neocortex? *Development*, *141*(1), 11–16. doi: 10.1242/dev.101279

Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic Theory of Working Memory. *Science*, *319*(5869), 1543–1546. doi: 10.1126/science.1150769

Morton, J. (1969). Interaction of information in word recognition. *Psychol. Rev.*, *76*(2), 165–178. doi: 10.1037/h0027366

Mountcastle, V. B. (1957). Modality and topographic properties of single neurons of cat's somatic sensory cortex. *J. Neurophysiol.*, *20*(4), 408–434. doi: 10.1152/jn.1957.20.4.408

Mountcastle, V. B. (1997). The columnar organization of the neocortex. *Brain*, *120*(4), 701–722. doi: 10.1093/brain/120.4.701

Myers, E. B., & Blumstein, S. E. (2008). The neural bases of the lexical effect: an fMRI investigation. *Cereb. Cortex*, *18*(2), 278–288. doi: 10.1093/cercor/bhm053

Myme, C. I. O., Sugino, K., Turrigiano, G. G., & Nelson, S. B. (2003). The NMDA-to-AMPA ratio at synapses onto layer 2/3 pyramidal neurons is conserved

across prefrontal and visual cortices. *Journal of Neurophysiology*, *90*(2), 771–779. doi: 10.1152/jn.00070.2003

Nenadić, F., & Tucker, B. V. (2020). Computational modelling of an auditory lexical decision experiment using jTRACE and TISK. *Language, Cognition and Neuroscience*, *35*(10), 1326–1354. doi: 10.1080/23273798.2020.1764600

Nicola, W., & Clopath, C. (2017). Supervised learning in spiking neural networks with FORCE training. *Nat. Commun.*, *8*(1), 2208. doi: 10/gcr4j2

Nordlie, E., Tetzlaff, T., & Einevoll, G. T. (2010). Rate Dynamics of Leaky Integrate-and-Fire Neurons with Strong Synapses. *Frontiers in Computational Neuroscience*, *4*. doi: 10.3389/fncom.2010.00149

Norris, D. (2003). Perceptual learning in speech. *Cognit. Psychol.*, *47*(2), 204–238. doi: 10.1016/S0010-0285(03)00006-9

Norris, D. (2017). Short-Term Memory and Long-Term Memory are Still Different. *Psychol Bull*, *143*(9), 992–1009. doi: 10.1037/bul0000108

Norris, D., & McQueen, J. M. (2008). Shortlist B: a Bayesian model of continuous speech recognition. *Psychol Rev*, *115*(2), 357–395. doi: 10/dbshsk

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: feedback is never necessary. *Behav Brain Sci*, *23*(3), 299-325; discussion 325-370. doi: 10.1017/s0140525x00003241

Norris, D., McQueen, J. M., & Cutler, A. (2016). Prediction, Bayesian inference and feedback in speech recognition. *Lang. Cogn. Neurosci.*, *31*(1), 4–18. doi: 10.1080/23273798.2015.1081703

Nusbaum, H., & Magnuson, J. (1997). Talker normalization: Phonetic constancy as a cognitive process. *Talker variability in speech processing*.

Oh, W. C., Parajuli, L. K., & Zito, K. (2015). Heterosynaptic structural plasticity on local dendritic segments of hippocampal CA1 neurons. *Cell Reports*, *10*(2), 162–169. doi: 10.1016/j.celrep.2014.12.016

Oikonomou, K. D., Singh, M. B., Sterjanaj, E. V., & Antic, S. D. (2014). Spiny neurons of amygdala, striatum, and cortex use dendritic plateau potentials to detect network UP states. *Front. Cell. Neurosci.*, *8*, 292. doi: 10/gndxqb

Ojemann, G. A. (2013). Human Temporal Cortical Single Neuron Activity during Language: A Review. *Brain Sci*, *3*(2), 627–641. doi: 10.3390/brainsci3020627

Ojemann, G. A., & Schoenfield-McNeill, J. (1998). Neurons in human temporal cortex active with verbal associative learning. *Brain Lang.*, *64*, 317–327. doi: 10.1006/brln.1998.1982

Ojemann, G. A., Schoenfield-McNeill, J., & Corina, D. (2009). The roles of human lateral temporal cortical neuronal activity in recent verbal memory encoding. *Cereb. Cortex*, *19*(1), 197–205. doi: 10.1093/cercor/bhn071

Okun, M., & Lampl, I. (2008). Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nat Neurosci*, *11*(5), 535–537. doi: 10.1038/nn.2105

Ostojic, S. (2011). Interspike interval distributions of spiking neurons driven by fluctuating inputs. *J Neurophysiol*, *106*(1), 361–373. doi: 10.1152/jn.00830.2010

Otor, Y., Achvat, S., Cermak, N., Benisty, H., Abboud, M., Barak, O., ... Schiller, J. (2022). Dynamic compartmental computations in tuft dendrites of layer 5 neurons during motor behavior. *Science*, *376*(6590), 267–275. doi: 10.1126/science.abn1421

Pagkalos, M., Chavlis, S., & Poirazi, P. (2022). *Dendrify: A new framework for seamless incorporation of dendrites in Spiking Neural Networks.* bioRxiv. doi: 10.1101/2022.05.03.490412

Palmer, L. M., Shai, A. S., Reeve, J. E., Anderson, H. L., Paulsen, O., & Larkum, M. E. (2014). NMDA spikes enhance action potential generation during sensory input. *Nat Neurosci*, *17*(3), 383–390. doi: 10.1038/nn.3646

Palomero-Gallagher, N., & Zilles, K. (2017). Cortical layers: Cyto-, myelo-, receptor- and synaptic architecture in human cortical areas. *NeuroImage*. doi: 10.1016/j.neuroimage.2017.08.035

Pan, Z., Chua, Y., Wu, J., Zhang, M., Li, H., & Ambikairajah, E. (2020). An efficient and perceptually motivated auditory neural encoding and decoding algorithm for spiking neural networks. *Front. Neurosci.*, *13*. doi: 10.3389/fnins.2019.01420

Papoutsi, A., Sidiropoulou, K., & Poirazi, P. (2014). Dendritic nonlinearities reduce network size requirements and mediate ON and OFF states of persistent activity in a PFC microcircuit model. *PLoS Comput Biol*, *10*(7), e1003764. doi: 10/gn8rb5

Paré, D., Shink, E., Gaudreau, H., Destexhe, A., & Lang, E. J. (1998). Impact of spontaneous synaptic activity on the resting properties of cat neocortical pyramidal neurons in vivo. *J. Neurophysiol.*, *79*(3), 1450–1460. doi: 10.1152/jn.1998.79.3.1450

Park, Y., & Geffen, M. N. (2020). A circuit model of auditory cortex. *PLOS Computational Biology*, *16*(7), e1008016. doi: 10.1371/journal.pcbi.1008016

Pasley, B. N., David, S. V., Mesgarani, N., Flinker, A., Shamma, S. A., Crone, N. E.,

... Chang, E. F. (2012). Reconstructing Speech from Human Auditory Cortex. *PLOS Biology*, *10*(1), e1001251. doi: 10.1371/journal.pbio.1001251

Payeur, A., Béïque, J.-C., & Naud, R. (2019). Classes of dendritic information processing. *Current Opinion in Neurobiology*, *58*, 78–85. doi: 10.1016/j.conb.2019.07.006

Payeur, A., Guerguiev, J., Zenke, F., Richards, B. A., & Naud, R. (2021). Burst-dependent synaptic plasticity can coordinate learning in hierarchical circuits. *Nat Neurosci*, *24*(7), 1010–1019. doi: 10/gj2n5h

Pedrelli, L., & Hinaut, X. (2022). Hierarchical-Task Reservoir for Online Semantic Analysis From Continuous Speech. *IEEE Trans Neural Netw Learn Syst*, *33*(6), 2654–2663. doi: 10.1109/TNNLS.2021.3095140

Pedrosa, V., & Clopath, C. (2020). Voltage-based inhibitory synaptic plasticity: network regulation, diversity, and flexibility. *bioRxiv*, 2020.12.08.416263. doi: 10/gh62md

Petersson, K. M., & Hagoort, P. (2012). The neurobiology of syntax: beyond string sets. *Philos. Trans. R. Soc. B Biol. Sci.*, *367*(1598), 1971–1983. doi: 10.1098/rstb.2012.0101

Petralia, R. S., Yokotani, N., & Wenthold, R. J. (1994). Light and electron microscope distribution of the NMDA receptor subunit NMDAR1 in the rat nervous system using a selective anti-peptide antibody. *J. Neurosci.*, *14*(2), 667–696.

Piai, V., Roelofs, A., Jensen, O., Schoffelen, J.-M., & Bonnefond, M. (2014). Distinct Patterns of Brain Activity Characterise Lexical Activation and Competition in Spoken Word Production. *PLoS ONE*, *9*(2), e88674. doi: 10.1371/journal.pone.0088674

Pitt, M. A., & McQueen, J. M. (1998). Is Compensation for Coarticulation Mediated by the Lexicon? *Journal of Memory and Language*, *39*(3), 347–370. doi: 10.1006/jmla.1998.2571

Plotnikov, D., Rumpe, B., Blundell, I., Ippen, T., Martin, J., & Morrison, A. (2016). NESTML: a modeling language for spiking neurons. doi: arXiv:1606.02882

Poellmann, K., Bosker, H. R., McQueen, J. M., & Mitterer, H. (2014). Perceptual adaptation to segmental and syllabic reductions in continuous spoken Dutch. *Journal of Phonetics*, *46*, 101–127. doi: 10.1016/j.wocn.2014.06.004

Poeppel, D. (2012). The maps problem and the mapping problem: Two challenges for a cognitive neuroscience of speech and language. *Cogn. Neu-*

*ropsychol.*, *29*(1-2), 34–55. doi: 10.1080/02643294.2012.710600

Poeppel, D., & Embick, D. (2005). Defining the Relation Between Linguistics and Neuroscience. In *Twenty-first century psycholinguistics: Four cornerstones* (pp. 103–118). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Poeppel, D., & Idsardi, W. (2022). We don't know how the brain stores anything, let alone words. *Trends in Cognitive Sciences*, S1364661322002066. doi: 10.1016/j.tics.2022.08.010

Poirazi, P., Brannon, T., & Mel, B. W. (2003). Pyramidal neuron as two-layer neural network. *Neuron*, *37*(6), 989–999. doi: 10.1016/S0896-6273(03)00149-1

Poirazi, P., & Papoutsi, A. (2020). Illuminating dendritic function with computational models. *Nat Rev Neurosci*, *21*(6), 303–321. doi: 10/ggvqtn

Polsky, A., Mel, B. W., & Schiller, J. (2004). Computational subunits in thin dendrites of pyramidal cells. *Nat Neurosci*, *7*(6), 621–627. doi: 10.1038/nn1253

Pongracz, F., Poolos, M., Kocsis, J. D., & Shepherd, G. M. (1992). A model of NMDA receptor-mediated activity in dendrites of hippocampal CA1 pyramidal neurons. *J Neurophysiol*, *68*(6), 2248–2259.

Poo, M.-m., Pignatelli, M., Ryan, T. J., Tonegawa, S., Bonhoeffer, T., Martin, K. C., … Stevens, C. (2016). What is memory? The present state of the engram. *BMC Biology*, *14*(1), 40. doi: 10.1186/s12915-016-0261-6

Potjans, T. C., & Diesmann, M. (2014). The cell-type specific cortical microcircuit: relating structure and activity in a full-scale spiking network model. *Cereb. Cortex*, *24*(3), 785–806. doi: 10.1093/cercor/bhs358

Povel, D.-J., & Essens, P. (1985). Perception of Temporal Patterns. *Music Percept. Interdiscip. J.*, *2*(4), 411–440. doi: 10.2307/40285311

Pulvermüller, F. (1999). Words in the brain's language. *Behav. Brain Sci.*, *22*(2), 253-+. doi: 10.1017/S0140525X9900182X

Pulvermüller, F. (2003). Sequence detectors as a basis of grammar in the brain. *Theory Biosci.*, *122*(1), 87–103. doi: 10.1007/s12064-003-0039-6

Pulvermüller, F., Tomasello, R., Henningsen-Schomers, M. R., & Wennekers, T. (2021). Biological constraints on neural network models of cognitive function. *Nat Rev Neurosci*, *22*(8), 488–502. doi: 10/gkzr83

Quaresima, A., Fitz, H., Duarte, R., van den Broek, D., Hagoort, P., & Petersson, K. M. (2022). The tripod neuron: a minimal structural reduction of the dendritic tree. *J. Physiol.*, *n/a*(n/a). doi: 10.1113/JP283399

Rackham, O., Tsaneva-Atanasova, K., Ganesh, A., & Mellor, J. (2010). A Ca2+-based computational model for NMDA receptor-dependent synaptic plasticity at individual post-synaptic spines in the hippocampus. *Front. Synaptic Neurosci.*, *2*. doi: 10.3389/fnsyn.2010.00031

Rajan, K., Harvey, C. D., & Tank, D. W. (2016). Recurrent Network Models of Sequence Generation and Memory. *Neuron*, *90*(1), 128–142. doi: 10.1016/j.neuron.2016.02.009

Rall, W. (2011). Core conductor theory and cable properties of neurons. In R. Terjung (Ed.), *Comprehensive Physiology* (p. cp010103). Hoboken, NJ, USA: John Wiley & Sons, Inc. doi: 10.1002/cphy.cp010103

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.*, *2*(1), 79–87. doi: 10.1038/4580

Rapp, M., Yarom, Y., & Segev, I. (1996). Modeling back propagating action potential in weakly excitable dendrites of neocortical pyramidal cells. *PNAS*, *93*(21), 11985–11990. doi: 10.1073/pnas.93.21.11985

Reifenstein, E. T., Bin Khalid, I., & Kempter, R. (2021). Synaptic learning rules for sequence learning. *eLife*, *10*, e67171. doi: 10.7554/eLife.67171

Renart, A., de la Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., & Harris, K. D. (2010). The Asynchronous State in Cortical Circuits. *Science*, *327*(5965), 587–590. doi: 10.1126/science.1179850

Renart, A., Moreno-Bote, R., Wang, X.-J., & Parga, N. (2007). Mean-driven and fluctuation-driven persistent activity in recurrent networks. *Neural Comput.*, *19*(1), 1–46. doi: 10.1162/neco.2007.19.1.1

Rescorla, M. (2020). The computational theory of mind. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2020 ed.). Metaphysics Research Lab, Stanford University.

Riquelme, J. L., Hemberger, M., Laurent, G., & Gjorgjieva, J. (2023). Single spikes drive sequential propagation and routing of activity in a cortical network. *eLife*, *12*, e79928. doi: 10.7554/eLife.79928

Roth, A., & van Rossum, M. C. W. (2009). Modeling synapses. In E. De Schutter (Ed.), *Computational Modeling Methods for Neuroscientists* (pp. 139–160). The MIT Press. doi: 10.7551/mitpress/9780262013277.003.0007

Rowley, D. A., Rogish, M., Alexander, T., & Riggs, K. J. (2017). Cognitive correlates of pragmatic language comprehension in adult traumatic brain injury: A systematic review and meta-analyses. *Brain Inj*, *31*(12), 1564–1574. doi: 10.1080/02699052.2017.1341645

Russell, L. E., Dalgleish, H. W. P., Nutbrown, R., Gauld, O. M., Herrmann, D., Fişek, M., … Häusser, M. (2022). All-optical interrogation of neural circuits in behaving mice. *Nat Protoc*, *17*(7), 1579–1620. doi: 10.1038/s41596-022-00691-w

Sakurai, Y., Osako, Y., Tanisumi, Y., Ishihara, E., Hirokawa, J., & Manabe, H. (2018). Multiple Approaches to the Investigation of Cell Assembly in Memory Research—Present and Future. *Front. Syst. Neurosci.*, *12*.

Salaj, D., Subramoney, A., Kraisnikovic, C., Bellec, G., Legenstein, R., & Maass, W. (2020). Spike frequency adaptation supports network computations on temporally dispersed information. *eLife*, *10*, e65459. doi: 10.7554/eLife.65459

Sanders, H., Berends, M., Major, G., Goldman, M. S., & Lisman, J. E. (2013). NMDA and GABAB (KIR) conductances: the "perfect couple" for bistability. *J. Neurosci.*, *33*(2), 424–429. doi: 10.1523/JNEUROSCI.1854-12.2013

Scharenborg, O., Norris, D., Bosch, L., & McQueen, J. M. (2005). How should a speech recognizer work? *Cogn Sci*, *29*(6), 867–918. doi: 10/fpk2zg

Scharenborg, O., van der Gouw, N., Larson, M., & Marchiori, E. (2019). The Representation of Speech in Deep Neural Networks. In I. Kompatsiaris, B. Huet, V. Mezaris, C. Gurrin, W.-H. Cheng, & S. Vrochidis (Eds.), *Multimed. Model.* (pp. 194–205). Cham: Springer International Publishing. doi: 10/gpfgws

Schiller, J., Major, G., Koester, H. J., & Schiller, Y. (2000). NMDA spikes in basal dendrites of cortical pyramidal neurons. *Nature*, *404*(6775), 285–289. doi: 10.1038/35005094

Schiller, N. O., & Meyer, A. S. (Eds.). (2003). *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*. DE GRUYTER MOUTON. doi: 10.1515/9783110895094

Schulz, J. M., Knoflach, F., Hernandez, M.-C., & Bischofberger, J. (2018). Dendrite-targeting interneurons control synaptic NMDA-receptor activation via nonlinear $\alpha$5-GABA A receptors. *Nat. Commun.*, *9*(1), 3576. doi: 10/ghzkjr

Scott, D. N., & Frank, M. J. (2023). Adaptive control of synaptic plasticity integrates micro- and macroscopic network function. *Neuropsychopharmacol.*, *48*(1), 121–144. doi: 10.1038/s41386-022-01374-6

Senzai, Y., Fernandez-Ruiz, A., & Buzsáki, G. (2019). Layer-specific physiological features and interlaminar interactions in the primary visual cortex of the mouse. *Neuron*, *101*(3), 500-513.e5. doi: 10.1016/j.neuron.2018.12.009

Sezener, E., Grabska-Barwińska, A., Kostadinov, D., Beau, M., Krishnagopal, S., Budden, D., … Latham, P. E. (2021). *A rapid and efficient learning rule for biological neural circuits*. bioRxiv. doi: 10.1101/2021.03.10.434756

Shinomoto, S., Sakai, Y., & Funahashi, S. (1999). The ornstein-uhlenbeck process does not reproduce spiking statistics of neurons in prefrontal cortex. *Neural Computation*, *11*(4), 935–951. doi: 10.1162/089976699300016511

Shouval, H. Z. (2011). What is the appropriate description level for synaptic plasticity? *Proceedings of the National Academy of Sciences*, *108*(48), 19103–19104. doi: 10/bndqcg

Shu, Y. S., Hasenstaub, A., & McCormick, D. A. (2003). Turning on and off recurrent balanced cortical activity. *Nature*, *423*(6937), 288–293. doi: 10.1038/nature01616

Silverman, B. W. (1981). Using kernel density estimates to investigate multimodality. *J. R. Stat. Soc. Ser. B Methodol.*, *43*(1), 97–99. doi: 10/ghz7pg

Simes, R. J. (1986). An improved Bonferroni procedure for multiple tests of significance. *Biometrika*, *73*(3), 751–754. doi: 10.1093/biomet/73.3.751

Sjerps, M. J., & Chang, E. F. (2019). The cortical processing of speech sounds in the temporal lobe. In *Human language: From genes and brains to behavior* (pp. 361–378). The MIT Press.

Sjerps, M. J., Fox, N. P., Johnson, K., & Chang, E. F. (2019). Speaker-normalized sound representations in the human auditory cortex. *Nat. Commun.*, *10*(1), 2465. doi: 10/gf4n43

Smith, S. L., Smith, I. T., Branco, T., & Häusser, M. (2013). Dendritic spikes enhance stimulus selectivity in cortical neurons in vivo. *Nature*, *503*(7474). doi: 10.1038/nature12600

Spivey, M., Joanisse, M., & McRae, K. (2012). *The Cambridge Handbook of Psycholinguistics*. Cambridge University Press.

Spruston, N. (2008). Pyramidal neurons: dendritic structure and synaptic integration. *Nat. Rev. Neurosci.*, *9*(3), 206–221. doi: 10.1038/nrn2286

Sterling, P., & Laughlin, S. (2015). *Principles of neural design*. Cambridge, MA: MIT Press.

Stokes, M. G. (2015). Activity-silent' working memory in prefrontal cortex: A dynamic coding framework. *Trends Cogn. Sci.*, *19*(7), 394–405.

Strube, C., Gackière, F., Saliba, L., Tell, F., & Kessler, J.-P. (2017). Variability of quantal NMDA to AMPA current ratio in nucleus tractus solitarii neurons. *bioRxiv*, 110569. doi: 10.1101/110569

Stuart, G. J., & Spruston, N. (2015). Dendritic integration: 60 years of progress. *Nat. Neurosci.*, *18*(12), 1713–1721. doi: 10.1038/nn.4157

Tabone, C. J., & Ramaswami, M. (2012). Is NMDA receptor-coincidence detection required for learning and memory? *Neuron*, *74*(5), 767–769. doi: 10/gfv75v

Tabossi, P., Collina, S., Mazzetti, M., & Zoppello, M. (2000). Syllables in the processing of spoken Italian. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 758–775. doi: 10.1037/0096-1523.26.2.758

Takahashi, N., Ebner, C., Sigl-Glöckner, J., Moberg, S., Nierwetberg, S., & Larkum, M. E. (2020). Active dendritic currents gate descending cortical outputs in perception. *Nat Neurosci*, *23*(10), 1277–1285. doi: 10/gg6v3c

Takahashi, N., Oertner, T. G., Hegemann, P., & Larkum, M. E. (2016). Active cortical dendrites modulate perception. *Science*, *354*(6319), 1587–1590. doi: 10.1126/science.aah6066

Tepper, J. M., Wilson, C. J., & Koós, T. (2008). Feedforward and feedback inhibition in neostriatal GABAergic spiny neurons. *Brain Res Rev*, *58*(2), 272–281. doi: 10/fmvg6f

Tetzlaff, C., Kolodziejski, C., Timme, M., & Wörgötter, F. (2011). Synaptic Scaling in Combination with Many Generic Plasticity Mechanisms Stabilizes Circuit Connectivity. *Front. Comput. Neurosci.*, *5*, 47. doi: 10/cjbp36

Tomasello, R., Garagnani, M., Wennekers, T., & Pulvermüller, F. (2018). A Neurobiologically Constrained Cortex Model of Semantic Grounding With Spiking Neurons and Brain-Like Connectivity. *Front. Comput. Neurosci.*, *12*. doi: 10.3389/fncom.2018.00088

Toscano, J. C., Anderson, N. D., & McMurray, B. (2013). Reconsidering the role of temporal order in spoken word recognition. *Psychon Bull Rev*, *20*(5), 10.3758/s13423-013-0417-0. doi: 10.3758/s13423-013-0417-0

Toscano, J. C., McMurray, B., Dennhardt, J., & Luck, S. J. (2010). Continuous Perception and Graded Categorization: Electrophysiological Evidence for a Linear Relationship Between the Acoustic Signal and Perceptual Encoding of Speech. *Psychol. Sci.*, *21*(10), 1532–1540. doi: 10.1177/0956797610384142

Tremblay, R., Lee, S., & Rudy, B. (2016). GABAergic interneurons in the neocortex: from cellular properties to circuits. *Neuron*, *91*(2), 260–292. doi: 10.1016/j.neuron.2016.06.033

Triesch, J., Vo, A. D., & Hafner, A.-S. (2018). Competition for synaptic building

blocks shapes synaptic plasticity. *eLife*, *7*, e37836. doi: 10.7554/eLife.37836

Tucker, B. V., Brenner, D., Danielson, D. K., Kelley, M. C., Nenadić, F., & Sims, M. (2019). The Massive Auditory Lexical Decision (MALD) database. *Behav Res*, *51*(3), 1187–1204. doi: 10.3758/s13428-018-1056-1

Tukker, J. J., Beed, P., Schmitz, D., Larkum, M. E., & Sachdev, R. N. S. (2020). Up and down states and memory consolidation across somatosensory, entorhinal, and hippocampal cortices. *Front. Syst. Neurosci.*, *14*, 22. doi: 10/gnkb8p

Turrigiano, G. (2011). Too many cooks? Intrinsic and synaptic homeostatic mechanisms in cortical circuit refinement. *Annu. Rev. Neurosci.*, *34*(1), 89–103. doi: 10/cxt3fb

Uhlmann, M. (2020). *Neurobiological models of sentence processing* (Unpublished doctoral dissertation). Radboud University Nijmegen.

Ujfalussy, B. B., & Makara, J. K. (2019). Impact of functional synapse clusters on neuronal response selectivity. *bioRxiv*(1), 634220. doi: 10.1101/634220

Ujfalussy, B. B., Makara, J. K., Lengyel, M., & Branco, T. (2018). Global and multiplexed dendritic computations under in vivo-like conditions. *Neuron*, *100*(3), 579-592.e5. doi: 10.1016/j.neuron.2018.08.032

Uppstad, P. H., & Tønnessen, F. E. (2010). The Status of the Concept of 'Phoneme' in Psycholinguistics. *J Psycholinguist Res*, *39*(5), 429–442. doi: 10/cdng84

van den Broek, D., Uhlmann, M., Fitz, H., Duarte, R., Hagoort, P., & Petersson, K. M. (2017). *The best spike filter kernel is a neuron.* Extended abstract *Cognitive Computational Neuroscience* conference, NYC, September 6–8.

Vaz, A. P., Wittig, J. H., Inati, S. K., & Zaghloul, K. A. (2020). Replay of cortical spiking sequences during human memory retrieval. *Science*, *367*(6482), 1131–1134. doi: 10.1126/science.aba0672

Vitevitch, M. S., & Luce, P. A. (1998). When Words Compete: Levels of Processing in Perception of Spoken Words. *Psychol Sci*, *9*(4), 325–329. doi: 10.1111/1467-9280.00064

Vitevitch, M. S., Siew, C. S. Q., & Castro, N. (2018). Spoken Word Recognition. In *The Oxford Handbook of Psycholinguistics* (pp. 30–47). Oxford University Press.

Vogels, T. P., Froemke, R. C., Doyon, N., Gilson, M., Haas, J. S., Liu, R., ... Sprekeler, H. (2013). Inhibitory synaptic plasticity: spike timing-dependence and putative network function. *Front Neural Circuits*, *7*, 119.

doi: 10/gfzd7c

Vogels, T. P., Sprekeler, H., Zenke, F., Clopath, C., & Gerstner, W. (2011). Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. *Science*, *334*(6062), 1569–1573. doi: 10.1126/science.1211095

Von Neumann, J. (1958). *The computer and the brain*. Oxford, England: Yale Univer. Press.

Wang, X. J. (1999). Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. *J. Neurosci.*, *19*(21), 9587–9603.

Wang, X.-J. (2021). 50 years of mnemonic persistent activity: quo vadis? *Trends in Neurosciences*, *44*(11), 888–902. doi: 10.1016/j.tins.2021.09.001

Warren, R. D. (1970). Perceptual Restoration of Missing Speech Sounds. *Science*, *167*.

Weber, A. (2001). Help or Hindrance: How Violation of Different Assimilation Rules Affects Spoken-Language Processing. *Lang Speech*, *44*(1), 95–118. doi: 10.1177/00238309010440010401

Weber, A., & Scharenborg, O. (2012). Models of spoken-word recognition. *Wiley Interdiscip. Rev. Cogn. Sci.*, *3*(3), 387–401. doi: 10.1002/wcs.1178

Whalen, D. H. (1991). Subcategorical phonetic mismatches and lexical access. *Perception & Psychophysics*, *50*(4), 351–360. doi: 10.3758/BF03212227

Wilmes, K. A., & Clopath, C. (2023). Dendrites help mitigate the plasticity-stability dilemma. *Sci Rep*, *13*(1), 6543. doi: 10.1038/s41598-023-32410-0

Wilson, C. (2008). Up and down states. *Scholarpedia J*, *3*(6), 1410.

Wilson, C., & Kawaguchi, Y. (1996). The origins of two-state spontaneous membrane potential fluctuations of neostriatal spiny neurons. *J. Neurosci.*, *16*(7), 2397–2410. doi: 10/ghzkjz

Winnubst, J., Cheyne, J. E., Niculescu, D., & Lohmann, C. (2015). Spontaneous activity drives local synaptic plasticity in vivo. *Neuron*, *87*(2), 399–410. doi: 10.1016/j.neuron.2015.06.029

Winnubst, J., & Lohmann, C. (2012). Synaptic clustering during development and learning: the why, when, and how. *Front. Mol. Neurosci.*, *5*. doi: 10.3389/fnmol.2012.00070

Winsler, K., Midgley, K. J., Grainger, J., & Holcomb, P. J. (2018). An electrophysiological megastudy of spoken word recognition. *Lang. Cogn. Neurosci.*, *33*(8), 1063–1082. doi: 10.1080/23273798.2018.1455985

Wu, Y. K., Hengen, K. B., Turrigiano, G. G., & Gjorgjieva, J. (2020). Homeostatic

mechanisms regulate distinct aspects of cortical circuit dynamics. *Proc Natl Acad Sci USA*, *117*(39), 24514–24525. doi: 10/ghqd84

Wybo, W. A., Jordan, J., Ellenberger, B., Marti Mengual, U., Nevian, T., & Senn, W. (2021-01-26, 2021). Data-driven reduction of dendritic morphologies with preserved dendro-somatic responses. *eLife*, *10*, e60936. doi: 10.7554/eLife.60936

Yang, G. R., Murray, J. D., & Wang, X. J. (2016). A dendritic disinhibitory circuit mechanism for pathway-specific gating. *Nat Commun*, *7*, 12815–12815. doi: 10.1038/ncomms12815

Yang, S., Gao, T., Wang, J., Deng, B., Lansdell, B., & Linares-Barranco, B. (2021). Efficient spike-driven learning with dendritic event-based processing. *Front. Neurosci.*, *15*. doi: 10.3389/fnins.2021.601109

Yi, H. G., Leonard, M. K., & Chang, E. F. (2019). The encoding of speech sounds in the superior temporal gyrus. *Neuron*, *102*(6), 1096–1110. doi: 10/gf39xq

You, H., & Magnuson, J. S. (2018). TISK 1.0: An easy-to-use Python implementation of the time-invariant string kernel model of spoken word recognition. *Behav Res*, *50*(3), 871–889. doi: 10/gdqsfx

Zajzon, B., Duarte, R., Mahmoudian, S., Morrison, A., & Duarte, R. (2019). Passing the Message: Representation Transfer in Modular Balanced Networks. *Front. Comput. Neurosci.*, *13*(December), 79. doi: 10.3389/fncom.2019.00079

Zenke, F., Agnes, E. J., & Gerstner, W. (2015). Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. *Nat Commun*, *6*(1), 6922. doi: 10.1038/ncomms7922

Zenke, F., & Gerstner, W. (2017). Hebbian plasticity requires compensatory processes on multiple timescales. *Phil. Trans. R. Soc. B*, *372*(1715), 20160259. doi: 10/ggsb47

Zenke, F., & Vogels, T. P. (2020). The remarkable robustness of surrogate gradient learning for instilling complex function in spiking neural networks. *bioRxiv*, 2020.06.29.176925. doi: 10/gh448m

Zhou, S., & Yu, Y. (2018). Synaptic e-i balance underlies efficient neural coding. *Front. Neurosci.*, *12*. doi: 10/ghzkjw

Zilles, K., & Amunts, K. (2009). Receptor mapping: architecture of the human cerebral cortex:. *Curr. Opin. Neurol.*, *22*(4), 331–339. doi: 10.1097/WCO.0b013e32832d95db

Zucca, S., D'Urso, G., Pasquale, V., Vecchia, D., Pica, G., Bovetti, S., ... Fellin, T.

(2017). An inhibitory gate for state transition in cortex. *eLife*, *6*, e26177. doi: 10.7554/eLife.26177.001

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, *32*(1), 25–64. doi: 10.1016/0010-0277(89)90013-9

# Nederlandse samenvatting

Wat gebeurt er in onze hersenen als we een woord herkennen? Het beantwoorden van deze vragen vereist het doorgronden van het woordherkenningsproces en hoe onze hersenfysiologie dit mogelijk maakt. De psycholinguïstiek gaat vooruit in het begrijpen van hoe woorden in menselijke taal worden gebruikt, maar we moeten nog veel leren over welke delen van de hersenen dit ondersteunen. Dit proefschrift verbindt de psychologie van woordherkenning met neurobiologie; het stelt computermodellen voor die de fysiologie van zenuwcellen associëren met het leren en onthouden van gesproken woorden. De modellen implementeren vergelijkingen die rekening houden met de biologische processen die plaatsvinden in de neuronen, zoals de elektrische stromen in of door cellen. De vier experimentele hoofdstukken hoofdstukken gaan van een model van neuronfysiologie naar het bestuderen van de rekencapaciteit van een corticaal netwerk. In de hoofdstukken 2 en 3 wordt een nieuw model van de zenuwcel voorgesteld en geanalyseerd, het Tripod-neuron, dat nauwkeuriger is in het reproduceren van functies en dynamiek van neuronen die in in-vivostudies zijn waargenomen. Het model introduceert vergelijkingen voor de dendrieten, enkele substantiële anatomische delen van de zenuwcellen die vaak zijn verwaarloosd in computationele studies. Computersimulaties geven aan dat menselijke dendrieten het geheugen van het ontvangen ingangssignaal gedurende een korte tijd behouden, ongeveer 0.1 s. Hoofdstuk 5 bewijst dat dit vermogen in netwerken van neuronen kan worden gebruikt om sequenties op de tijdschaal van menselijke taal te herkennen, zoals woorden met overlappende klanken (*god* en *dog*). In hoofdstuk 6 wordt het model met succes vergeleken met klassieke resultaten in de psycholinguïstiek. Het netwerk houdt zich aan de rekenprincipes die zijn afgeleid van menselijk gedrag en geeft voorspellingen over verschijnselen die moeilijk experimenteel te onderzoeken zijn. Omdat het model sterk is gebaseerd op de werkelijke hersenbiologie, biedt het huidige onderzoek een fysiologische verklaring voor hoe woorden worden geleerd, onthouden en opgeroepen tijdens spraak. De resultaten geven aan dat dendrieten, langzame excitatoire receptoren en remmende controle coördineren voor het vormen en heractiveren van netwerkgeheugens met sequentiële structuren, zoals gesproken woorden.

# English Summary

What happens in our brains when we recognize a word? How do humans differ from animals in understanding speech sounds? What goes wrong when we don't grasp it? Answering these questions requires fathoming the word recognition process and how our brain physiology enables it. Psycholinguistics strides in understanding how words are used in human language, but we still have much to learn about which parts of the brain support this. This thesis bridges the psychology of word recognition with neurobiology; it proposes computational models that associate the physiology of nervous cells with the learning and recall of spoken words. The models implement equations that account for the biological processes occurring in the neurons, such as the passage of electric currents within or across cells. The four experimental chapters build up from a model of neuron physiology to studying the computational capacity of a cortical network. Chapters 2 and 3 propose and analyze a novel model of the nervous cell, the Tripod neuron, that is more accurate in reproducing functions and dynamics of neurons observed in *in-vivo* studies. The model introduces equations for the dendrites, some substantial anatomical parts of the nervous cells that have often been neglected in computational studies. Computer simulations indicate that human dendrites maintain the memory of the input signal received for a short time, approximately 0.1 s. Chapter 5 proves this capacity can be leveraged in networks of neurons to recognize sequences on the timescale of human language, such as words with overlapping sounds (*god* and *dog*). In Chapter 6, the model is successfully compared to classical results in psycholinguistics. The network adheres to the computational principles derived from human behavior and provides predictions on phenomena that are hard to investigate experimentally. Because the model is highly constrained to actual brain biology, the present study offers a physiological explanation of how words are learned, memorized, and recalled during speech. The results indicate that dendrites, slow excitatory receptors, and inhibitory control coordinate for forming and reactivating network memories with sequential structures, such as spoken words.

# Sommario italiano

Cosa succede nel nostro cervello quando riconosciamo una parola? Perché gli esseri umani differiscono dagli animali nella comprensione dei suoni del parlato? Cosa va storto quando una parola ci sfugge? Per rispondere a queste domande dobbiamo sapere come le parole vengono riconosciute nella nostra fisiologia cerebrale. La psicolinguistica ha largamento progredito nello spiegare come vengono utilizzate le parole nel linguaggio umano, ma abbiamo ancora molto da imparare su quali parti del cervello supportano questa capacità. Questa tesi mira a conciliare le teorie psicologiche sul riconoscimento delle parole con le più recenti scoperte in neurobiologia; propone modelli computazionali che associano la fisiologia dei neuroni all'apprendimento e al richiamo delle parole pronunciate. I modelli implementano equazioni che tengono conto dei processi biologici che si verificano nelle cellule, come il passaggio di correnti elettriche all'interno o tra neuroni. I quattro capitoli sperimentali procedono da un modello di fisiologia del neurone allo studio della capacità computazionale di una rete della corteccia. I capitoli 2 e 3 introducono un nuovo modello della cellula nervosa, il Tripod neuron, più accurato nella riproduzione delle funzioni e della dinamica dei neuroni osservati negli studi *in vivo*. Il modello introduce equazioni per i dendriti, alcune parti anatomiche delle cellule nervose che sono spesso trascurate negli studi computazionali. Le simulazioni indicano che i dendriti umani mantengono la memoria del segnale elettrico ricevuto per un breve periodo, circa 0.1 s. Il capitolo 5 dimostra che questa capacità può essere sfruttata nelle reti di neuroni per distinguere sequenze sulla scala temporale del linguaggio umano, ad esempio riconoscere parole con suoni sovrapposti (*Roma* e *amor*). Nel Capitolo 6, il modello viene confrontato con successo con i risultati classici della psicolinguistica. Aderisce ai principi computazionali derivati dal comportamento umano e fornisce previsioni su fenomeni difficili da indagare sperimentalmente. Poiché il modello è vincolato alla biologia cerebrale, il presente studio propone una spiegazione fisiologica di come le parole vengono apprese, memorizzate e richiamate nel parlato. I risultati indicano che i dendriti, insieme ad alcuni ricettori e al controllo inibitorio, coordinano la formazione e la riattivazione di memorie di rete con struttura sequenziale, come le parole.

# Acknowledgements

Writing a Ph.D. thesis is a once-in-a-lifetime experience. You don't really know how it goes until you are here writing the acknowledgments. Naturally, you write them in a rush, like all the rest. After five years of abstracts, presentations, and article submissions, you can make it under time pressure. To the best of my abilities, I will walk through the last five years of my life and thank all the people who have seen this Ph.D. project grow, and become a manuscript.

The first remarkable passage of this Ph.D. was an October night in Florence, when Limerick read me the tarots and suggested to accept this position. The last card was the World. *The World represents an ending to a cycle of life, a pause in life before the next big cycle beginning with the fool. It is an indicator of a major and inexorable change, of tectonic breadth.* Thank you Limerick, for your clairvoyance; thank you the World for this new cycle. I don't remember if the card was upright or reversed. It was a question as everything in my life at that point. A Ph.D. in the Netherlands was a challenging, unexpected turn, but I had solid ground to build on. I want to thank the friends who were in my life at that moment and encouraged me to take a step forward. We walked together on endless mountain paths and this was just one of them. Thank you Simone, Alessandro, Flaminia, Jacopo, Claudia, Boyska, Elisa, Giacomo, Adriano, and Scacco. Thanks to The Science Zone, AvANa, and Officina for what we learned together. I am also immensely glad for my family who stayed by my side in these five long years - although they often had no idea of what was going on. Thank you Morena, Francesco, Martina, Luca, Elisa, Sabrina, Stefano, Giuseppina, Raffaele, Antonina, Maria, Ernesto, and Cesira. Thank you mamma e papà for your unconditional trust and support. Thanks to my roots that made me strong to face the Dutch wind.

Upon my arrival in Nijmegen, the World came to me in distinct forms. The peace of the forest surrounding the Max Planck Institute for Psycholinguistics, the white snow of February 2019, and the biting wind while biking with my hands in my pockets. On this unknown stage, several actors and actresses were preparing to be new friends, colleagues, and mentors. The person who most contributed to bringing this manuscript to touchdown is a wise, knowledgeable, and

joyful man, my thesis promotor, Peter Hagoort. Peter, I am deeply grateful for the support, sympathy, and understanding you have granted me over the years. I wish you many years of prolific scientific production, and I will strive to carry a sparkle of your geniality with me. This thesis would not have been possible without the science gang that supervised and accompanied me during the Ph.D. Thanks to my supervisors, Hartmut and Karl Magnus, and my collaborators Renato, Petros, and Dick, for the insightful conversations, the difficulties we faced together, and the discoveries we made. You have shaped my view on how to do honest scientific research. A particular thank you goes to Hartmut for being a picky reader; your attention to style and clarity is an inestimable quality that will always accompany me. A supportive and vibrant community of researchers at the Neurobiology of Language department has surrounded my Ph.D. experience. Among the many friends and colleagues, I would like to thank Laura for her brilliant smile, often shining in the corridor's darkness on a rainy day. Ambra, for the laughs and the encouragement; Guillermo, for your deep thinking and elegance; Daniel, for the insightful conversations on the brain and the brilliant perseverance. Thank Laura, Cas, Teun, Rowan, Nienke, Sophie, Margot, Micha, Ksenija, and Sara for growing up in science together. You helped me navigate some of the toughest moments of this academic commencement, I shared with you the joys and the lows of the Ph.D. Thanks to my friends and colleagues, Linda, Ellie, Noor, Francesca, Ambra, Rinus, Eva, Anny, Ashley, Lukas, Jakub, Daniel, Filiz, Marlou, Kuhn and all the others for being the wonderful curious crowd that animated the life in the Institute and my Friday morning 10:15. Thanks to Michaela, Carolin, Karin, Meggie and Angela who helped to make organization and bureaucracy bearable. Thank you, Ellie, for the Het Talige Brein blog and the fun we had making it a successful story!

Five years in the Netherlands were not all about the MPI. Outside the hedge of the campus, Nijmegen smiled at me with its Saturday market, the bridges, the lake, and the beautiful colors of Javastraat in spring - let's forgive it for the rest of the year. I shared my time in the city with several friends, who I sincerely hope will remain close until the end of time. The deepest appreciation goes to Ana. During this last year of writing, you offered me love, support and encouragement. We became a team, and I look forward to supporting each other and celebrating our successes in the years to come. Thanks to my acquired family, Sascha and Ania, thanks to Vincent, Dieke, Harry, Emma, Jaco and Merel. I am grateful for the love and friendship you offered me, the raves, the chill evenings in Timorstraat, the breakfasts with arab bread and eggs, the craft beers and all

the rest. Thank you Lili, for the evenings on the couch with vegan burgers and yummy beers. Thank you, Charlotte, for your energy, laugh, and lovely conversations about art and science. Thank you, Marianne. That morning, we watched the sunrise on a bench with birds chirping around. Thank you Adrian, for your inner peace and kindness. Thank you Darius for your investigative stories and the sweet care, thank you Eli and Freya for the last months of laughs. Thank you Esther, for that walk in the forest in a cloud of mosquitos. A special, profound thanks go to the group of friends who gave sense to Nijmegen, thanks Marco, Nora, Noor, Hugo, Elie, Elena, Joanna, Michele, Alejandro, Laura, Ambra, Sealvia, Joanna, Oija, Jill, Filiz, Lorenzo, Achille, Mora, Antigoni. Thank you for the swimmings at the lake, the fires, the concerts, and most of all, for your sincere friendship, you made the last three years shine. I've crossed paths with so many people over these years, and now the cycle is closing and a new World, le Monde, is coming. I am, sadly, aware that several of you won't follow in the upcoming one. To those who will stay, looking forward, to those who will go, wish you a happy, meaningful life. I'm grateful I met you.

Eventually, my acknowledgments go to the Klinker, the Onderbroek, the Plak, and all those who keep alive the flame of social and climate justice, of antiracism, feminism, and antifascism. The title of this thesis calls back to a war movie from 1977. The movie portrays a failed operation of the U.S. army in capturing the bridge over the Northern Rheine in Arhnem, although the shootings were filmed on the Waalbrug. This movie was everything I knew about Nijmegen before starting the Ph.D.. I picked this name because it blinked an eye to my father, who was so hyped about it. However, it was also an excuse to pronounce myself about the mounting militarism facing our society. Academics are too often trapped in their ivory tower, busy on their books, to realize that the world surrounding them is on fire. During my years in Nijmegen, having the opportunity to manifest my voice and presence in the demonstrations organized at the Klinker was vital. It gave me a sense of reality, of being alive and present in these harsh, violent times. In my opinion, advocating for social justice is among the duties of a scientist, as much as communicating their science to the general public. There is no possibility of a free science in an unfree society.

In conclusion, I want to thank the four friends who will help me navigate this last bit until my doctoral defense, Ellie, Filiz, Hugo, and Laura; thank you, Margherita, for kindly accepting my proposal of making a cover art for the thesis. Thank you all! I look forward to the day we will celebrate this journey together.

Sincerely, Alessio

# Curriculum Vitae

Credits: Julia vd Fuhr

Alessio Quaresima was born in Rome, on the 26th of January 1992. He grew up in a medieval village in Rome's countryside, Genazzano. After high school, he moved to Rome to follow the curriculum at the Physics faculty in Sapienza University, Rome. In these years he learned about quantum mechanics while repairing bikes at local Ciclo Officina di Fisica. He attended courses by G. Parisi, P. del Giudice, and V. Loreto, who fostered his scientific interests in complexity, statistics, and neuroscience. Meanwhile, he participated in The Science Zone, a science education association fostering experimental science in school, and in AvANa / Le Dita Nella Presa, a hacklab striving for digital freedom, technological awareness and justice. His first academic achievement was a BSc thesis in Complexity Theory with V. Servedio, analyzing the information entropy of novels. After completing the master's courses, he pursued an Erasmus thesis abroad at Diderot University with S. Bottani, in Paris. In these months he studied models of the spatial growth of neurites, which ignited his fascination with biology. The common interest in human language and neuroscience steered him towards a Ph.D. in the Neurobiology of Language Department at the Max Planck Institute for Psycholinguistics, in Nijmegen. The present doctoral thesis speaks for the last five years. Presently, he is working at the Institut de l'Audition in Paris, he is investigating auditory working memory in an animal model of hearing. His interest lies in the physiological understanding of how the human brain supports language. Alessio is also invested in science communication. In Nijmegen, he kept himself busy organizing public events on the verge of arts and science. Among the others, he co-organized the opening of the BioMa lab and talked about simulation in neuroscience at the In Science Nijmegen Film Festival. He also wrote several pieces for the MPI Talkling and Het Talige Brein outreach blogs. In his free time, he enjoys walking uphill in the Alps, swimming in lakes, attending expositions, and more recently he became keen on video projections and live performances.

# Publications

Quaresima, A., Fitz, H., Duarte, R., Broek, D. van den, Hagoort, P., & Petersson, K. M. (2022). The tripod neuron: A minimal structural reduction of the dendritic tree. *The Journal of Physiology*, 601, 15, 3265-3295.

Quaresima, A., Fitz, H., Duarte, R., Hagoort, P., & Petersson, K. M. (2023). Dendrites support formation and reactivation of sequential memories through Hebbian plasticity. *bioRxiv*, (p. 2023.09.26.559322).

Quaresima, A., Duarte, R., Fitz, H., Hagoort, P., & Petersson, K. M. (*in preparation*). Dendritic non-linearities enable up-down states in single neurons.

# Additional Information

***Competing interests*** I, Alessio Quaresima, declare that the research presented in this thesis was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

***Chapter front-page figures:*** The images are original drawings of Santiago Ramón y Cajal, obtained from the online catalog of the Instituto Cajal (CSIC).

- **Chapter 1**: Extending growth cones. Growth cones are dynamic extensions of a developing neuron. The variety of the growth cones shown is due to the complexity of the paths they were to navigate.
- **Chapter 2**: Portray of a pyramidal neuron of the cerebral cortex. Pyramidal cells exhibit a characteristic cone-shaped cell body, a single apical dendrite extending upwards to the cortical surface, basal dendrites, and basal axons.
- **Chapter 3**: Lateral spines of dendrites. The preparation was obtained by a modified version of the Ehrlich method.
- **Chapter 4**: Inhibitory interneurons of the neurogliaform class. Found in many diverse regions of the brain, including the hippocampus, cerebral cortex, and visual cortex, the interneurons shown here are from the auditory cortex.
- **Chapter 5**: Pyramidal cells from the superior temporal gyrus (STG). The STG contains the auditory cortex, which is responsible for processing sound, and Wernicke's area, which is necessary for the processing of speech to be understood as language rather than simply sounds. Shown here are several layers of pyramidal cells in the superior temporal gyrus, which is layered similarly to other areas of the temporal cortex.
- **Chapter 6**: The mammalian auditory midbrain, shown in the figure, is part of the ascending auditory pathway, responsible for relaying sensory signals from the ear into the primary auditory cortex deep in the brain.

MAX PLANCK INSTITUTE
FOR PSYCHOLINGUISTICS

**VISITING ADDRESS**
Wundtlaan 1
6525 XD Nijmegen
The Netherlands

**POSTAL ADDRESS**
P.O. Box 310
6500 AH Nijmegen
The Netherlands

**CONTACT**
T +31(0)24 3521 911
F +31(0)24 3521 213
E info@mpi.nl
Twitter @MPI_NL
www.mpi.nl