*Annual Review of Linguistics*

# Spoken Word Recognition: A Focus on Plasticity

Efthymia C. Kapnoula,[1,2] Mina Jevtović,[1]
and James S. Magnuson[1,2,3]

[1] Basque Center on Cognition, Brain and Language (BCBL), Donostia–San Sebastián, Spain; email: kapnoula@gmail.com

[2] IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

[3] Department of Psychological Sciences, University of Connecticut, Storrs, Connecticut, USA

Annu. Rev. Linguist. 2024. 10:233–56

First published as a Review in Advance on October 31, 2023

The *Annual Review of Linguistics* is online at linguistics.annualreviews.org

https://doi.org/10.1146/annurev-linguistics-031422-113507

**Keywords**

spoken word recognition, plasticity, development, multilingualism, training, adaptation

## Abstract

Psycholinguists define spoken word recognition (SWR) as, roughly, the processes intervening between speech perception and sentence processing, whereby a sequence of speech elements is mapped to a phonological wordform. After reviewing points of consensus and contention in SWR, we turn to the focus of this review: considering the limitations of theoretical views that implicitly assume an idealized (neurotypical, monolingual adult) and static perceiver. In contrast to this assumption, we review evidence that SWR is plastic throughout the life span and changes as a function of cognitive and sensory changes, modulated by the language(s) someone knows. In highlighting instances of plasticity at multiple timescales, we are confronted with the question of whether these effects reflect changes in content or in processes, and we consider the possibility that the two are inseparable. We close with a brief discussion of the challenges that plasticity poses for developing comprehensive theories of spoken language processing.

## 1. INTRODUCTION

Spoken word recognition (SWR) refers to the processes that map sequences of sublexical elements (e.g., speech sounds) to phonological wordforms that index lexical representations. Wordforms allow access to lexical knowledge from speech (e.g., grammar, semantics; Pustejovsky 1995). Lexical knowledge in long-term memory is commonly assumed to compose the mental lexicon (for an alternative view, see Elman 2004, 2009).

Theories and models of SWR tend to have a restricted scope: Most implicitly or explicitly assume that SWR starts with the products of speech perception (acoustic-phonetic features or even phonemes) and ends with the activation and/or selection of the wordform (or series of wordforms) that best integrates bottom-up input with prior knowledge (e.g., known words) and context. This "simplified mapping perspective" (Magnuson & Crinnion 2022) constrains SWR to a more tractable problem than the full path from acoustics to semantics and beyond. Although unresolved challenges remain within this limited scope, it has allowed researchers to discover many details of the fine-grained time course of (this aspect of) spoken language processing.

As Magnuson (2008) argues, while simplifying assumptions may be necessary, they tend to function like (pre)theoretical assumptions when they persist for decades, as they have in SWR. While the limitations of this scope and the need to escape it are increasingly acknowledged in SWR research, other implicit assumptions are arguably less prominent in current research but also constrain the way theorists construe the challenges of SWR and thus how we study it empirically. Specifically, most SWR theories and models implicitly (or explicitly) assume a static, crystallized system used by a single, idealized, young adult monolingual speaker of a language spoken in WEIRD (Western, educated, industrialized, rich, and democratic) cultures (Henrich et al. 2010; most prominently English, less frequently languages such as French, Dutch, and German, even less frequently languages such as Japanese, Chinese, Arabic, Hebrew, and Russian, and only rarely ones such as African languages or minoritized or otherwise low-status languages and dialects within WEIRD cultures; Kirk 2023). We hold that these assumptions are highly problematic and actively impede deeper understanding of SWR. To illustrate why, we review the growing body of research that reveals SWR to be a plastic, dynamic system, with representations and mechanisms changing constantly over the life span, as well as research that suggests that representations and processes supporting SWR are strongly affected by crosslinguistic differences and by multilingualism.

A plausible expectation for the science of SWR might be that there are core, universal algorithms that are separable from the "contents" of language—the number of words a person knows, the number of languages they know and their proficiency in each, or aspects of the specific language(s) they know. The goal of language scientists concerned with SWR would be to identify those core algorithms. One might also expect that the algorithms could be discoverable by studying just one or a few languages. We take an alternative view: Language processes are not separable from linguistic content. Changing the contents (learning more words, the structure and content of a language, learning more languages, or changes in sensory acuity or cognitive capacities) alters the entire system in ways that cannot be predicted from current models. In this review, we marshal evidence by examining how SWR changes over the life span, how it is affected by changes in capacities (as a function of learning to read, or declines in sensory acuity with aging), and how SWR is modulated by multilingualism and by details of the languages a person knows. Even if we do not convince readers of this theoretical position, we shall see that the field is not yet able to identify core algorithms even if they exist. The common denominator that emerges is the principle that our experiences with language shape the structure of the system that supports SWR—on scales ranging from milliseconds to years. Experience-driven changes may affect different aspects of SWR both directly and indirectly. Our main goal is to describe a wide range of examples of plasticity in SWR and discuss the potential mechanisms underlying these changes. For example, plasticity may take the form of change in the way lexical representations interact with each other in real time, but it can also reflect changes in related cognitive systems (e.g., working memory). Importantly, even for changes within the SWR system itself, the exact locus of change can be difficult to pinpoint given that the boundaries between SWR structure, mechanisms, and content are often blurred (Elman 2004). Thus, in this review we largely refrain from making strong theoretical claims about the exact locus and nature of plasticity. Before turning to these issues, we begin with a brief outline of the main points of consensus and debate in the current literature.

## 2. POINTS OF CONSENSUS AND CONTENTION

Decades of research have been devoted to understanding the specifics of the mechanisms and the representations underlying SWR. While there are still many open debates and unresolved challenges, progress has resulted in consensus on core principles.

### 2.1. Consensus

The current consensus boils down to three propositions and has changed only slightly since Marslen-Wilson's (1993) description. First, SWR is continuous: As a word is heard, the mechanisms supporting SWR attempt to map the unfolding signal to known words incrementally. Second, there is graded, parallel consideration of multiple candidates: As the signal unfolds, multiple words are activated in proportion to their similarity to the ongoing signal, weighted by prior probability (typically only with respect to frequency of occurrence). Third, there must be a selection mechanism that allows the system to rank or choose among candidate words, explicitly or emergently (as we discuss below).

### 2.2. Contention

Despite the consensus on fundamental principles, several aspects remain unsettled. Here we focus on thorny points that, despite extensive debate and empirical exploration, remain unresolved.

**2.2.1. Inputs.** Models differ in the SWR inputs they assume. Some use pseudospectral, over-time acoustic-phonetic features (TRACE; McClelland & Elman 1986), while others use human diphone confusion probabilities to approximate noisy, phonologically constrained uptake of acoustics (Norris & McQueen 2008), and others simply use phonemes (Hannagan et al. 2013, Luce & Pisoni 1998). Any simplifying assumption that divorces SWR from hearing and speech perception has the potential to complicate rather than simplify scientific understanding by hiding potential constraints (cf. Magnuson 2008). For example, Salverda et al. (2003) found that the "embedding problem" (McQueen et al. 1995)—the fact that most words have other words embedded within them (e.g., *catalog* potentially "contains" *cat*, *at*, *a*, *cattle*, *law*, and *log*)—is mitigated significantly by prosodic cues (e.g., listeners are sensitive to durational differences that predict word length, such that *hammer* competes more strongly with *hamster* than with *ham*). Thus, assuming any kind of speech perception product that does not preserve surface detail may be problematic. An exception is EARSHOT, a recent recurrent neural network model (Magnuson et al. 2020) that maps spectral slices to semantics via hidden recurrent units, preserving surface details. While phonetically organized responses emerge in EARSHOT's hidden units that resemble responses in the human cortex (e.g., Mesgarani et al. 2014), the model has not yet been tested for sensitivity to surface details such as prosody. Nonetheless, an opportunity for advancing SWR theories is developing models that operate directly on speech signals.

**2.2.2. Nature of lexical representations.** Many theories propose abstract lexical representations that do not contain nonlinguistic information, such as accent, speech rate, and loudness, or talker-specific details (Bowers 2000, McQueen et al. 2006). Some posit underspecified phonological representations that only include nondefault or unpredictable features (Lahiri & Marslen-Wilson 1991). However, there is robust evidence that listeners encode and use surface details when recognizing words (Bradlow et al. 1999, Goldinger 1996, Kapnoula & Samuel 2019, Salverda et al. 2003, Schacter & Church 1992), consistent with a detail-preserving exemplar-based or episodic view (e.g., Goldinger 1996, 1998; Luce et al. 2012). Pierrehumbert (2016) has argued compellingly for a hybrid approach that includes both abstract and episodic representations.

**2.2.3. Similarity.** While there is consensus that greater similarity predicts more difficult processing, theories differ in how they define similarity. The Cohort Model (e.g., Marslen-Wilson 1987) posits that listeners exploit the sequential nature of speech. If one hears a word beginning "s. . . ," this is both evidence that the first segment is /s/ and evidence against any other segment (in graded fashion, as /s/ is more similar to /f/ than to /m/). The Cohort algorithm attempts to add each incoming phoneme onto the preceding series until silence is encountered or adding a phoneme would not result in a possible word. Those conditions identify word boundaries (e.g., /d/ of *drank* in *cat drank* implies a boundary) or a need for reanalysis (e.g., *cat alarm* could be parsed as *cattle arm*, but it might not be compatible with context). On this view, the recognition "cohort" will include words that overlap in the first few phonemes, as only these words will become substantially activated.

Many studies support the onset cohort hypothesis. In gating, where listeners hear progressively larger snippets of words starting from word onset and must provide a completion, listeners only provide onset cohort completions (e.g., Grosjean 1980). Clever paradigms have yielded evidence that cohort items are sufficiently activated to drive detectable semantic priming. Hearing a word like *beaker* can prime *insect*, presumably because *beetle* is strongly activated. However, there is almost no evidence for activation of words that mismatch at onset in these paradigms, even when global similarity is high (e.g., hearing *beaker* will not activate *speaker* sufficiently to prime *stereo*; Marslen-Wilson & Zwitserlood 1989). Shortlist (Norris 1994) and Shortlist B (Norris & McQueen 2008) also privilege onset similarity.

In stark contrast, the Neighborhood Activation Model (NAM; Luce & Pisoni 1998) proposes that global similarity is paramount, and two words will be neighbors (sufficiently similar that hearing one will activate the other) if they differ by no more than a single phonemic <u>d</u>eletion, <u>a</u>ddition, or <u>s</u>ubstitution (commonly called the DAS rule). For example, *cat* has neighbors due to deletion (*at*), addition (*scat*, *cast*), and substitution (e.g., *bat*, *cot*, *cab*). Crucially, cohort items that differ by more than one phoneme are not neighbors (e.g., *cab–cabin*, *cabin–cabinet*). Neighborhood size accounts for modest but significant variance in lexical decision and naming beyond that explained by word frequency alone, with larger neighborhoods predicting more difficult processing (slower, or less accuracy for speech in noise), at least for monosyllabic words in English (Luce & Pisoni 1998).

This leads to an apparent discrepancy (cf. Magnuson 2016): "Pairwise" approaches (those that examine interactions between specific pairs of words) tend only to find evidence for cohort activation, while the NAM DAS metric accounts for significant variance in performance on a "lexical dimensions" (regression) approach despite excluding many cohorts and including many noncohorts. Allopenna et al. (1998) used the Visual World Paradigm (VWP; Tanenhaus et al. 1995) to examine fixation proportions over time to targets, cohorts, rhymes, and unrelated items (e.g., *beaker–beetle–speaker–carriage*) as participants followed spoken instructions to interact with a visual display containing images of the items (e.g., "click on the beaker"). They found early, relatively high fixation proportions for cohorts and later, smaller proportions for rhymes (though significantly elevated relative to the unrelated baseline). This suggests that both similarity metrics (onset cohort and global neighborhood) are relevant to SWR.[1]

**2.2.4. Selection.** A proposition shared by many theories and models[2] is that candidate words compete for recognition. Accounts of SWR ranging from the Cohort Model (e.g., Marslen-Wilson 1987; Marslen-Wilson & Welsh 1978) to interactive activation theory (Hannagan et al. 2013, McClelland & Elman 1986) to the original Shortlist model (Norris 1994) to NAM (Luce & Pisoni 1998) assume that competition is a central aspect of SWR because it affords emergent selection.

In TRACE (McClelland & Elman 1986), node activations change due to excitatory bottom-up and/or top-down connections between levels of representation and inhibitory connections within levels. TRACE inputs are pseudospectral acoustic-phonetic feature patterns over time. Features activate phonemes, phonemes activate words that contain them, and words send feedback to constituent phonemes. Crucially, there is lateral inhibition within layers. Word nodes have direct inhibitory links to other words, and phonemes have inhibitory connections to other phonemes, yielding competitive dynamics that emergently govern activation. The assumption of an explicit or implicit mechanism for competition is consistent with results showing that having more neighbors predicts slower recognition (e.g., Luce & Pisoni 1998).

Competition is eschewed by the Bayesian model Shortlist B (Norris & McQueen 2008), along with the concept of activation. Instead, the model assumes explicit selection of the word (or word sequence) with highest probability. If one assumes that confidence (and inversely, reaction time) varies with the value of the highest probability, words with few neighbors and/or high prior probability (frequency of occurrence) will be predicted to be recognized more readily than words with many neighbors and/or low frequency.

While there have been debates about whether some models assume too much inhibition (Marslen-Wilson & Warren 1994 versus Dahan et al. 2001), in general, empirical results demand an account that can predict that words in denser similarity neighborhoods will be more difficult to recognize (Luce & Pisoni 1998). We return to lexical competition below, considering evidence that interlexical inhibition not only exists but also appears to be a plastic aspect of SWR in development.

**2.2.5. Feedback.** While linguistic interpretation obviously depends on integrating bottom-up signals with top-down knowledge and context, a long-standing debate concerns whether integration occurs within perceptual pathways (e.g., via feedback from lexical to sublexical representations) or postperceptually. Norris et al. (2000) claimed that all valid top-down effects could be explained by postperceptual integration, that feedback is not necessary, and that it could not help word recognition. Space does not permit a full treatment of this topic, so we refer readers

---

[1] Alternative/additional similarity metrics have been proposed. We cannot review them all due to space constraints, but we note that it seems likely that we have yet to discover a truly comprehensive similarity metric.

[2] "Model" is sometimes used as a synonym for "theory," and in the philosophy and practice of science, both words have many senses. On Marr's (1982) taxonomy of information processing theories, when cognitive scientists present a "model," they often mean a mechanistic theory at Marr's algorithmic level [e.g., Marslen-Wilson's (1987) Cohort Model]. We maintain a distinction between theories and models, where models are a means of making theories concrete and testable, whether as a verbal model (a set of propositions or a flowchart), a mathematical model (equations that compress or abstract observations), or a simulation model (typically implemented as a computer program that simulates mechanisms proposed by a theory). For more discussion, we refer readers to Magnuson et al. (2012).

to recent publications in this debate that include substantial reviews (Magnuson et al. 2018 versus Norris et al. 2018) as well as a recent replicable experiment that provides strong evidence in favor of feedback (Luthra et al. 2021).

**2.2.6. Summary.** Despite ongoing debates, there is relatively broad consensus that in SWR, multiple lexical candidates are incrementally and gradiently activated in parallel depending on degree of similarity to the input and prior probability, and that a selection mechanism is needed to (implicitly or explicitly) choose among them. With these fundamental principles as a foundation, let us now consider a somewhat neglected aspect of SWR: plasticity.

## 3. PLASTICITY OVER THE LIFE SPAN

In this section we present examples of developmental plasticity in SWR, from infancy throughout the life span.

### 3.1. Early Spoken Word Recognition

Children start speaking around the age of 1 year. However, they understand and recognize some words substantially earlier (around 6 months). How (prelingual) infants perceive spoken language has been studied extensively in psycholinguistics. Substantial research suggests that speech perception in prelingual infants is surprisingly effective, even though it differs from adult SWR in interesting ways.

For instance, infants as young as 1 month are sensitive to subtle differences in phonetic contrasts (e.g., /p/–/b/). In fact, infants are better than adults at discriminating non-native contrasts (e.g., Best & McRoberts 2003, Kuhl et al. 1992). Even though sensitivity to within-category information is maintained into adulthood (Andruski et al. 1994; McMurray et al. 2002; Miller 1997; Samuel 1977, 1982; Sarrett et al. 2020; Toscano et al. 2010), around 6 months, infants begin to lose sensitivity to non-native contrasts, presumably as they attune to native-language phonology (Kuhl et al. 1992, Werker & Tees 1984). This shift reflects one of the earliest forms of plasticity in the system, as perception is reorganized based on experience with the distributions of phonetic features relevant for native-language phoneme contrasts.

Another key concern is the nature of infants' representations. For instance, despite good sound discrimination, 11-month-old infants appear not to apply this discrimination fully in SWR. In one study (Hallé & De Boysson-Bardies 1996), French-learning infants showed similar preferences for unaltered and altered familiar words (e.g., when the French word *canard* 'duck' was presented in its unaltered form /kanaʁ/ and its altered form /ganaʁ/). Similarly, Stager & Werker (1997) found that 14-month-old infants failed to discriminate similar sounds (e.g., *bih* and *dih*) when they were paired with images of objects (a potential labeling context that implied that the syllables could be object names) but showed significant discrimination in a change-detection task (when the sounds were paired with visual checkerboards). Interestingly, 8-month-old infants showed robust discrimination in the labeling context. Stager & Werker (1997) suggested that this apparent decline indexed a developmental change where older infants are learning what details are relevant for lexical discriminations in their native language.

These studies suggest that before around 12 months, lexical representations are not fully specified. Indeed, according to the density hypothesis, infants only need underspecified lexical representations, mainly because they know very few words and, thus, do not need high phonetic and temporal precision. As their vocabulary grows, they are forced to use finer phonetic information to uniquely identify similar words (Charles-Luce & Luce 1990, Storkel 2002). Nevertheless, other work suggests that this may not be the case. For example, Swingley & Aslin (2000) presented 18- to 24-month-old infants with sentences containing either real words (e.g.,

/dɔg/ <*dog*>, /kɑr/ <*car*>) or mispronounced versions (e.g., /tɔg/, /kʌr/). While infants listened to sentences containing target words (either canonical or mispronounced versions), they saw two pictures, one matching the target (e.g., *dog* or *car*). The authors compared infants' eye-movement patterns across the two conditions (canonical versus mispronounced words). Both accuracy (target fixations) and latency to fixate the target indicated that infants recognized canonical pronunciations more easily (with shorter latency and higher accuracy). This suggests that infants' lexical representations are approaching adult-like specification during the second year of life.

Other aspects of SWR also change during the first years of life. For example, Singh and colleagues (2004; see also Singh 2008) reported differences in eye-movement patterns of 7.5- and 10.5-month-old infants during repeated encounters with spoken words presented with neutral or valenced (e.g., happy) affect. When 7.5-month-old infants learned new words, if those words were always presented with one affect (happy or sad), they did not generalize to productions with different affect. However, by 10.5 months, infants readily generalized beyond affect. Infants in this age range also have difficulty generalizing from one talker to another (Houston & Jusczyk 2000). This suggests that it takes several months for infants to learn which aspects of acoustic patterns are relevant for SWR. Finally, Hollich et al. (2000) showed that 15- to 17-month-old infants are better at learning words with fewer neighbors (i.e., sparse neighborhoods), unlike adults, who are better at learning novel words from denser neighborhoods (Storkel et al. 2006).

In sum, infants' receptive lexical representations develop significant specificity even before speech production is fully mastered. These representations undergo substantial changes before becoming as well specified and stable as adult representations. How are changes in representations linked to changes in processes? If the two are separable, it is possible that there are core algorithmic processes that are already operational in infancy. In that case, differences from adult performance would be due to differences in contents (e.g., words known) rather than mechanisms. To our knowledge, this has not been addressed directly. Simulations with computational models could be useful; one could ask, for example, whether infant-like behavior would emerge in successful adult models endowed with infant-like lexicons. If so, this would suggest that the model's algorithms are strong candidates for both infant and adult SWR (and challenge our proposal that processes and contents are inseparable). If not, this could suggest that algorithms are qualitatively different in infancy or that the algorithms assumed by the model are incorrect.

## 3.2. Changes in Spoken Word Recognition Related to Reading Acquisition

The rate of learning novel words is highest during preschool and early elementary school years (Anglin et al. 1993), when children learn on average nine words per day (Markman 1987). This expansion in vocabulary size at this stage is thought to be largely due to children acquiring reading skills; in addition to auditory input, they are now exposed to the written modality as well. Although the potential impact of orthography in spoken language processing is beyond the scope of prominent models of SWR (e.g., Luce & Pisoni 1998, Marslen-Wilson & Welsh 1978, McClelland & Elman 1986, Norris & McQueen 2008), there are several important changes in speech perception and SWR related to reading acquisition.

First, learning to read has been linked to improved ability to distinguish individual speech sounds (phonemes) within spoken words (i.e., phonological awareness; Liberman et al. 1989, Morais et al. 1979). Morais et al. (1979) compared literate and illiterate Portuguese-speaking adults on tasks involving manipulation of individual speech sounds within spoken words (e.g., phoneme deletion, addition, and substitution). While literate adults completed all tasks easily (accuracy was close to ceiling), illiterate adults' performance was low. Better readers consistently show better performance in tasks involving manipulation of sounds within words (Pratt & Brady 1988, Swank & Catts 1994). These findings are often taken as evidence that learning to read shifts the perception of speech from larger units (e.g., syllables) to smaller units [e.g., phonemes (see Port 2007) or other sublexical elements, depending on the writing system (Frost 2012, Ziegler & Goswami 2005)].

Seidenberg & Tanenhaus (1979) examined orthographic effects in a language with fairly opaque spelling-to-sound mapping (English). Rhyme judgment decision times in adult skilled readers were affected by spelling: Word pairs sharing both spelling and pronunciation (e.g., *pie–tie*) yielded faster decision times compared to those sharing pronunciation only (e.g., *pie–eye*). Further work with adults has shown orthographic effects in online auditory tasks. For example, in a lexical decision task, words containing rhymes with unique orthographic representations (e.g., /oʊb/ in *globe* can only be written as <obe>) are recognized faster than words with multiple possible spellings [e.g., /eɪn/ in *sane* can be written as <ane>, <ain> (e.g., *rain*), <ein> (e.g., *rein*), <eign> (e.g., *reign*), or <ayne> (e.g., *Wayne*) (Ziegler & Ferrand 1998; for a review, see Ziegler et al. 2008)].

Ventura et al. (2007) investigated the developmental trajectory of orthographic effects in speech perception. In their cross-sectional study, three groups of Portuguese-speaking children (from second to fourth grade) were tested on both auditory lexical decision and shadowing (where participants repeat an auditorily presented word). Surprisingly, in all three groups, orthographic effects (i.e., differences between words with consistent versus inconsistent spellings, e.g., *globe* versus *sane*) emerged not only in words but also in pseudowords—a result that was not observed for adults. The same facilitatory effect for consistent items was observed in both auditory lexical decision and in speeded and unspeeded shadowing. These results suggest that in developing readers, who still rely on phonological decoding when reading (Share 1995, 2004), the influence of orthography on speech processing is stronger than in skilled adult readers.

To further investigate the developmental trajectory of these effects and identify the turning point where consistency effects become adult-like, Ventura et al. (2008) tested Portuguese-speaking children in grades 4–9. The child-like pattern (orthographic consistency effects in both words and pseudowords) persisted to grade 6. Subsequently, the effects were only seen in words, suggesting that once reading has been sufficiently mastered, the locus of orthographic consistency effects shifts to the lexical level. This turning point is not universal and seems to be strongly modulated by the opacity of the writing system. While the Portuguese children needed at least 5 years of reading experience to exhibit adult-like orthographic consistency effects, French children—who are used to processing larger units from very early in reading acquisition (because French is relatively more opaque, and pronunciations can depend on contexts spanning several letters; see Ziegler & Goswami 2005)—required only about half that time; the adult-like pattern was observed after approximately 2.5 years of reading experience (Pattamadilok et al. 2009). Thus, in developing readers, effects of orthographic consistency depend on the grain size that is relevant in a particular orthography.

Two accounts have been proposed to explain the origins of orthographic impact in spoken word perception. According to the phonological restructuring hypothesis, reading acquisition modifies the nature of existing phonological representations (Muneaux & Ziegler 2004, Perre et al. 2009). Representations of words with a unique possible spelling become more salient during reading acquisition as compared to words with phonological forms that could be spelled in multiple ways and are hence recognized faster. By contrast, the automatic coactivation account proposes that spoken words activate orthographic counterparts regardless of task (Chéreau et al. 2007, Perre & Ziegler 2008). Spoken words with multiple possible spellings activate several orthographic representations, which compete, leading to slower recognition times. However, to date, there is no conclusive evidence adjudicating between the two accounts.

In sum, learning to read alters speech perception and SWR. As children become aware of the components that link phonology and orthography in their language, speech perception reorganizes around the pertinent grain size for the writing system. Moreover, the time course of SWR is affected by the orthographic patterns associated with words during reading acquisition.

## 3.3. Ongoing Changes in Children and Adolescents

In this subsection, we consider development of SWR between infancy and adulthood. First, there is plasticity in the way listeners map speech sounds to lexical representations. Even though experience with native-language phonology shapes listeners' ability to discriminate speech sounds early on (Eimas et al. 1971, Kuhl et al. 1992, Werker & Tees 1984), sensitivity to within-category information (speech perception gradiency) is maintained into adulthood (Andruski et al. 1994; McMurray et al. 2002; Miller 1997; Samuel 1977, 1982; Sarrett et al. 2020; Toscano et al. 2010; for a review, see McMurray 2022), although this gradiency undergoes change over development.

Evidence for plasticity comes from a study by McMurray et al. (2018). They tested children from three age groups (7–8, 12–13, and 17–18 years) using /b/–/p/ and /s/–/ʃ/ word–word continua (e.g., *beach–peach*, *ship–sip*). Pictures of items were presented on the screen, and continuum items were presented auditorily. Eye movements to the competitor were used as an index of gradiency. Sensitivity to within-category differences slowly increased over development (gradiency continues to increase into adulthood; Kapnoula & Samuel 2023). Furthermore, Kutlu et al. (2022) presented preliminary results from a study examining the role of linguistic diversity (indexed by social network information) on the development of gradiency in school-aged children (6–11 years). In line with distributional accounts of speech perception development (e.g., Flege & Bohn 2021), the results suggested that linguistic diversity had a positive impact on speech perception gradiency (i.e., children exposed to higher variability of speech sounds showed higher gradiency).

Moving beyond the segmental level, there is also evidence that lexical competition changes in later childhood. For instance, Rigler and colleagues (2015) recruited two groups of children (9 and 16 years), whom they tested on SWR using the VWP. All participants were aurally presented with target words (e.g., *bees*) along with corresponding images, as well as pictures of cohort competitors (e.g., *bean*), rhyme competitors (e.g., *peas*), and unrelated words (e.g., *cap*). Fixation patterns suggested that 9-year-olds were slower to activate the target and experienced more early competition from similar words than 16-year-olds. This suggests that younger children take more time to suppress competitors.

It is possible, though, that when lexical activation is inferred from fixation patterns, developmental differences could reflect changes in eye-movement control. However, a recent study provides more support for developmental change in lexical competition. Blomquist & McMurray (2023; see also Apfelbaum et al. 2023) replicated Rigler et al.'s (2015) study and further showed that previous experience with items that should induce lexical competition (due to misleading coarticulation inserted by cross-splicing sound files) leads to more effective SWR (via stronger inhibition of competitor lexical items), particularly in younger children. They employed a subcategorical mismatch paradigm tailored to measure competition between specific word pairs [e.g., *cat* and *cap* (Dahan et al. 2001)—a stimulus manipulation originally introduced by Marslen-Wilson & Warren (1994) based on work by Streeter & Nigro (1979)] and examined whether competition changes throughout an experimental session. Importantly, their analyses of the time course of lexical competition (buildup, peak, and resolution) are arguably largely independent of overall differences in looking behavior (Oleson et al. 2017, Seedorff et al. 2018). While older children (12–13 years) showed robust lexical competition throughout the entire experiment, younger children (7–8 years) only did so in the second half of the session. This suggests that language experience (here, repeated trials with the same items) drives the observed differences. Finally, Hendrickson and colleagues (2021) demonstrated that 8- to 11-year-old children modulate lexical competition resolution to adapt to situations of increased uncertainty (e.g., speech at lower intensity, known as soft speech). Since exposure to soft speech is something children usually experience when they start school, this finding suggests that word recognition mechanisms are fine-tuned by language experience, including context such as auditory conditions.

To sum up, although scarce (compared to work done with infants), research on SWR in older children and adolescents points to the same conclusion: The development of mechanisms that underlie different aspects of SWR is tightly linked to listeners' language experience. We suggest that the results reviewed here support the view that processes and contents are not separable aspects of development. For example, if the same processes operated in younger and older children, younger children should show faster resolution of lexical competition because they know fewer words. Instead, development in mechanisms (e.g., inhibitory processes hypothesized to provide lexical selection) occurs in parallel with development of acoustic-phonetic mapping and content (e.g., vocabulary).

## 3.4. Cognitive Aging

Normal aging leads to various changes in cognitive functioning; some aspects of language processing improve, while others decline. For instance, while vocabulary continues to grow throughout the life span (Harada et al. 2013, Kavé 2022), perception of familiar words usually degrades (McLaughlin et al. 2022, Revill & Spieler 2012, Sommers 1996, Wingfield et al. 1991).

Even in the case of preserved hearing, older adults show changing impact of lexical competitors (Revill & Spieler 2012, Wingfield et al. 1991) and are affected more by neighborhood density (i.e., more difficulties recognizing words from denser neighborhoods; McLaughlin et al. 2022, Sommers 1996). Sommers & Danielson (1999) found that older adults have more difficulty recognizing words from high-density neighborhoods in low-predictability sentences. Poorer performance on SWR tasks was linked to older adults' poorer inhibitory processing skills, and the authors inferred that increasing difficulties stem from age-related declines in inhibitory mechanisms. Older adults also tend to rely more on lexical frequency (Revill & Spieler 2012), which apparently slows recognition of lower-frequency words relative to younger adults. Finally, resolving lexical competition in older adults has also been linked to increased cognitive load (McLaughlin et al. 2022). McLaughlin and colleagues (2022) observed larger pupillary responses (associated with greater cognitive load) in older adults compared to younger adults. In addition, this difference was marginally larger for words from denser neighborhoods, pointing to older adults' higher difficulty in resolving lexical competition (see also Cheimariou & Kapnoula 2022).

Differences between younger and older adults tend to be magnified in complex listening situations such as speech in noise (Gosselin & Gagné 2011) or attending to input produced by multiple speakers (Getzmann et al. 2015). Explanations usually emphasize reduced attentional resources, less efficient inhibition of irrelevant and concurrent information, and/or reduced efficiency in attentional control (Tun et al. 2002). To investigate effects of noise on SWR, Ben-David et al. (2011) tested older (mean age 70) and younger (mean age 20) participants on a variant of the VWP. In this task, participants heard spoken sentences ending with target words (e.g., *candle*) while competitor pictures (e.g., *candy*, *sandal*) were displayed. With increased noise, older participants fixated similar-sounding competitors more than younger participants did, suggesting greater difficulty differentiating targets and competitors. The authors attributed the differences to sensory acuity declines. However, we see at least three alternative interpretations. First, as in the findings of some developmental work described above (e.g., Rigler et al. 2015), differences in older adults could stem (in part) from declines in eye-movement control. Second, adverse listening conditions may interfere with the resolution of lexical competition via lateral inhibition, leading to increased residual activation of competitor items. Third, consistent with inhibitory hypotheses (e.g., Sommers & Danielson 1999), adverse listening conditions particularly affect late stages of lexical competition, making older individuals less efficient in handling residual competitors (see Cheimariou & Kapnoula 2022).

Converging evidence for age-related differences under difficult hearing conditions comes from Getzmann and colleagues (2015). They examined brain responses in younger and older adults to spoken words in multispeaker ("cocktail party") environments using event-related potentials (ERPs, where electroencephalography is used to measure neural responses time-locked to stimuli). Apart from higher error rates in recognizing word pairs, older adults also showed weaker responses in N2 and N400 ERP components, suggesting less effective inhibition of concurrent information.

## 3.5. Summarizing Plasticity over the Life Span

Studies reviewed in this section show that SWR is influenced by external factors (e.g., reading acquisition) as well as by age-related internal changes to perceptual and cognitive systems. From fast changes in developing lexical representations in infants and younger children to slower changes occurring with aging in older adults, the reviewed studies thus point to the necessity for SWR processes to adapt to the perceptual and cognitive profiles of the listener over their lifetime. It is possible that core algorithms for SWR remain unchanged throughout adulthood and that age-related changes stem from changes in nonlinguistic characteristics of listeners (declines in sensory acuity, or generalized slowing). If this is the case, one should be able to predict SWR performance after assessing individual differences in such characteristics. While associations have been found (as reviewed above), most variance remains unexplained, challenging this simple possibility.

# 4. CROSSLINGUISTIC CONSIDERATIONS

There is evidence that the language(s) one knows shapes spoken language processing (among other capacities). We focus on two cases: how word recognition is shaped by someone's first language (L1) (i.e., plasticity in the form of crosslinguistic differences) and how knowing multiple languages affects SWR.

## 4.1. Plasticity Revealed by Crosslinguistic Differences

While there is considerable evidence that larger neighborhoods predict slower SWR in English (Luce & Pisoni 1998), Dutch (Vroomen & de Gelder 1995), French (Ziegler et al. 2003), and Japanese (Amano & Kondo 2000), there are reports that larger neighborhoods are associated with faster processing in Spanish (Vitevitch & Rodríguez 2005; but see Gür et al. 2023) and Russian (Arutiunian & Lopukhina 2020). One hypothesis is that more morphologically complex languages will tend to have many forms with high phonological and semantic overlap (e.g., inflected forms), which may promote a more cooperative dynamic.

Languages also vary in constraints that may facilitate SWR. For example, listeners can leverage gender-marked determiners to predict upcoming words in languages such as French (Dahan et al. 2000) and Spanish (Gussow et al. 2019). Similarly, when discourse supports strong expectations for an upcoming noun and listeners encounter a determiner with unexpected gender, robust neural surprise responses are observed (Wicha et al. 2004).

These are just two examples of how language-specific details modulate SWR. Further work is needed to uncover the exact mechanisms by which SWR is shaped by a listener's native language. For example, computational modeling can help assess whether language-specific differences in the direction of neighborhood effects emerge naturally from any extant models when they are applied to languages such as Spanish and Russian.

### 4.2. Effects of Bi/Multilingualism

In addition to differences between speakers of different languages, there is evidence that SWR in one language may be affected by other languages a listener knows. This is an important consideration, given that multilingualism is the modal human experience: Grosjean (2022) estimates that most of the world (50–70%) is at least bilingual, and approximately 15% of speakers know three or more languages. Knowing more than one language raises questions for SWR: Are both/all languages a person knows always "operational"? If so, this potentially increases the phonological and semantic competition environment dramatically. If not, what governs language control and switching?

Many behavioral studies are consistent with parallel activation in both/all languages (Hayakawa & Marian 2020). For example, Russian–English bilinguals participating in a VWP session conducted completely in English fixated *stamp* (/marku/ in Russian) when a stamp was in the display when the instruction was "click on the marker" (Spivey & Marian 1999). Shook & Marian (2019) contrasted this "overt coactivation" (looking at a depicted item with a name similar in the "background" language to a "foreground" language phonological form) with a "covert coactivation" paradigm. When Spanish–English bilinguals were asked in English to click on a picture of a *duck*, they were significantly more likely to fixate *shovel* than an unrelated item, revealing phonological activation in Spanish (*duck* = *pato* and *shovel* = *pala*). Phonological effects are also influenced by phonetic similarity. For example, Spanish–English bilinguals processing Spanish were more likely to fixate items corresponding to potential English competitors when voice onset times (VOTs) were adjusted to be more English-like (Ju & Luce 2004), pointing again to the need for SWR models to accommodate surface details.

Language proficiency and balance also affect SWR in bilinguals. For example, coactivation is more likely when listening to a nondominant language [i.e., there is more impact of the dominant language (L1) on the nondominant (L2) language than vice versa; Shook & Marian 2016]. Sometimes there is complete asymmetry with no evidence for activation of (noncognate) words from L2 when listeners process a dominant L1 (Blumenfeld & Marian 2007). To complicate things even more, Dijkstra et al. (2015) found facilitation from L1 cognates for word recognition in L2, but interference in L1 word recognition from L2 cognates (in a lexical decision task with L1 or L2 sentence contexts preceding words from L1 or L2). Lastly, Sarrett et al. (2022) tested early L2 learners of Spanish (with English as their L1) and found that, in addition to crosslinguistic competition, higher L2 proficiency predicted faster L2 word recognition and better competition resolution (both within L2 and across languages).

### 4.3. Summarizing Crosslinguistic Considerations

In this section, we have presented several representative examples that demonstrate how SWR can be shaped by language-specific factors. These studies suggest that the way SWR unfolds in real time depends on the following: properties of the languages a listener knows, whether a listener is bi/multilingual, and their relative proficiency in each of the languages they know. Together, these results challenge the view that SWR depends on static, universal processes that apply across languages and populations.

## 5. SHORT-TERM PLASTICITY

Change is expected over development, but SWR also exhibits considerable plasticity on short timescales of minutes or seconds. Indeed, short-term plasticity has been a vigorous area of investigation over the past two decades.

## 5.1. Learning New Words

Perhaps the most intuitive form of plasticity that impacts SWR throughout the lifespan is learning new words. Studies using self-contained artificial lexica have been useful for isolating aspects of postlearning processing. For example, Magnuson et al. (2003) used a feedback training paradigm to teach listeners an artificial lexicon. They reported an intriguing result (which has not been further explored) that links back to questions in early lexical development and SWR: As adults learned new neighborhoods of novel words, rhyme competition (e.g., between *pibu* and *dibu*) was stronger than cohort competition (e.g., *pibu* versus *piba*), although with more training the pattern shifted to that observed with well-established words (where cohort competition is earlier and stronger than rhyme competition). They pointed out that if this were seen in development, it might be viewed as a developmental shift to gradually more precise sequential encoding. On the basis of simulations with a computational model, Magnuson et al. (2003) speculated that such a shift reflects noisy (weakly learned) representations rather than a developmental change. Magnuson et al. (2003) also reported that newly learned words appeared to be encapsulated from listeners' preexisting lexicon, in that there was not a significant impact of English-based neighborhood size. However, this could be because the artificial lexicon was presented in an encapsulated way, as though it were relevant only for the laboratory study.

Gaskell & Dumay (2003) examined word learning from incidental exposure rather than from explicit training. Participants performed a phoneme-monitoring task on pseudowords that were designed to overlap with real words until near word offset (e.g., *cathedruke*, based on *cathedral*). They asked whether recurring patterns without semantics would become integrated into listeners' lexicons by examining whether responses to the original words (e.g., *cathedral*) would be slowed (presumably due to competition with the pseudoword, which would shift the uniqueness point of the original word later). Initially, processing of the original words was facilitated, as though the pseudowords had activated the original words. After sufficient exposure and an opportunity for sleep-based consolidation, the recognition of the original words was slowed as predicted by changes in uniqueness points.

Word-learning paradigms continue to be used to uncover more details about learning and processing. Another example of such work is that of Kapnoula et al. (2015), who examined competition between existing and newly learned words using a subcategorical mismatch paradigm with eye tracking similar to that used by Dahan et al. (2001). In this paradigm, cross-splicing is used to create items with misleading coarticulation on vowels (e.g., by splicing the offset of *neck* onto the onset and nucleus of *net*). When misleading coarticulation is consistent with a word, processing is slowed substantially, but when the misleading coarticulation is not consistent with a word (e.g., when the onset and nucleus come from **nep*), there is much less slowing. Kapnoula et al. (2015) found that word-like slowing emerged quickly for a subset of nonwords they trained listeners to recognize (see also Kapnoula & McMurray 2016a).

In general, word-learning studies with adults provide clear evidence for plasticity without speaking directly to the issue of separability between content and processes; adding more words to the lexicon changes processing in intuitive ways without suggesting that processing changes qualitatively. This is not inconsistent with the view that content and structure are inseparable, but it helps define potential boundary conditions for that view.

## 5.2. Everyday Instances of Processing Malleability for Familiar Words

Other findings are more challenging for the separability view. In particular, the mapping from acoustics to linguistic encoding/representations is surprisingly plastic and changes with various forms of context and learning from everyday language experience, as we review in this subsection.

**5.2.1. Adaptive plasticity.** Adaptive plasticity refers to listeners' ability to adjust speech processing to cope with systematic patterns of distortion in the speech signal—for instance, due to the speaker's accent/dialect or background noise. For example, Guediche et al. (2016) examined the way listeners use internal compared with external sources of information to adapt to distorted (noise-vocoded, spectrally shifted) speech signals. At pretest, all participants were asked to identify auditorily presented words that were severely distorted (recognition accuracy at 11%). During training, participants were asked to identify these words; critically, half of them would see the printed form of the word as they heard the distorted stimulus. The other half did not see any visual stimuli. After training, participants did the first task again with different stimuli. Participants who received feedback during training were able to adapt to the signal distortion, which allowed them to perform significantly better at posttest. This shows that external information (in this case, seeing the word) helps listeners tune their system to adapt to distortions.

**5.2.2. Lexically mediated perceptual learning.** Listeners are often exposed to systematically shifted boundaries between phoneme categories (e.g., a talker who has an unusual VOT boundary, producing their voiceless stops with a VOT that overlaps with the canonical range for voiced stops). In the lab, we can simulate this experience by exposing listeners to a talker whose /p/ and /b/ productions have been shifted to noncanonical distributions. To make this shift apparent to listeners, shortened VOTs are encountered only in lexical contexts that have to resolve to /p/ (e.g., *?lum* or *?uzzle*, where *?* is the ambiguous sound). After brief exposure, listeners learn to adjust their VOT boundary, a phenomenon called phonemic recalibration (also called perceptual learning and lexically mediated perceptual learning; Kraljic & Samuel 2005, 2006, 2011; Norris et al. 2003). Note that experiments typically shift the boundary in both directions for different groups of participants (in our example of /p/–/b/, we would present the ambiguous items near the typical boundary only in contexts that resolve to /b/—e.g., *?lanket*, *?arrel*).

In their seminal paper, Norris et al. (2003) exposed listeners to words in which a particular phoneme was replaced with an ambiguous sound. For example, whenever a word should have had an /s/, it actually had an ambiguous mixture of /s/ and /f/. Other listeners heard the same ambiguous sounds, but the sounds were embedded in words that should have had an /f/ in that position. The results showed clear evidence for perceptual recalibration. For example, if listeners heard the ambiguous sound in forms like *sheri?*, they learned to perceive *?* sounds as variants of /f/. These results point to an adaptive mechanism that listeners can use to adjust SWR to the idiosyncrasies of different speakers, and could provide a basis to mitigate the lack-of-invariance problem (the apparent absence of consistent mappings between acoustic patterns and perceptual categories). Finally, another point that speaks to the plasticity of the system is that even newly learned words are able to drive this kind of recalibration (Kapnoula & Samuel 2022, Leach & Samuel 2007).

**5.2.3. Malleability of context effects.** Listeners also leverage various forms of context to facilitate SWR and spoken language comprehension more broadly, including semantics (Altmann & Kamide 1999, Ferretti et al. 2001), local thematic information (Kukona et al. 2011), visual context (Chambers et al. 2004, Eberhard et al. 1995, Sedivy et al. 1999, Tanenhaus et al. 1995), and world knowledge (Hagoort et al. 2004, Metusalem et al. 2012). Notably, the way context is used is not fixed. For example, one way listeners appear to use context is to generate predictions about upcoming content (e.g., Wicha et al. 2003; for a review regarding predictive processes in language comprehension, see Kutas et al. 2011). Heyselaar et al. (2021) examined the plasticity of predictive behavior by measuring listeners' anticipatory eye movements to words' referents while manipulating the proportion of predictable nouns. They found that decreasing predictability led to a decrease in anticipatory eye movements. These data are in line with the idea that listeners' predictive behavior depends on the distributional properties of recent linguistic input (Pickering & Gambi 2018), which highlights the plasticity of the system. It also suggests another challenge for finding core algorithms, since SWR appears to be strongly influenced by (higher-level) linguistic and nonlinguistic context.

**5.2.4. Malleability of priming effects.** Multiple studies have shown that listeners are better at recognizing a word masked in noise when that word has been preceded by another presentation of itself several minutes earlier (long-term identity priming; Ellis 1982, Jackson & Morton 1984, Kempley & Morton 1982). These effects are compatible with a logogen-type model (Morton 1964), according to which different events can cause an increase at the level of activation of a given lexical unit (or a decrease of its recognition threshold). However, priming appears to be outside the scope of current models such as TRACE (McClelland & Elman 1986) and Shortlist B (Norris & McQueen 2008).[3] Independently of the exact mechanism that drives such effects, the critical point is that the process that allows us to recognize familiar words is far from fixed.

**5.2.5. Processing-induced changes.** Kapnoula & McMurray (2016b) tested whether interlexical inhibition changes with experience. They used the subcategorical mismatch VWP of Dahan et al. (2001) mentioned above (see Section 3.3), to which they added a critical manipulation: One group (High Competition) was exposed to target and competitor items used by Dahan and colleagues (2001) (e.g., *net–neck*, *cart–carp*) under conditions that promoted coactivation and competition, and another group (Low Competition) was presented with the same items but under conditions that minimized competition. Subsequently, in a VWP task, participants in the High Competition group were better able to resolve competition among the pairs. Kapnoula & McMurray (2016b) proposed an explanation on which experience resolving competition between a pair of words strengthens inhibitory links between them, making selection via competition resolution more efficient. Such a result is challenging for accounts that maintain separation of contents and processes since it suggests that specific experiences with specific items under specific contextual conditions change system parameters that govern processing.

## 5.3. Summarizing Short-Term Plasticity

In this section we have reviewed studies showing different ways in which SWR processes can be modulated by recent learning, context, and expectations. Even though this kind of plasticity may appear qualitatively different from changes emerging over longer periods of time (e.g., those reviewed in Sections 3 and 4), long-term effects may emerge as the result of accumulated short-term changes. Note that experience-dependent plasticity in the earliest aspects of the processing pipeline (e.g., perceptual learning) is particularly challenging because it suggests that fundamental aspects of processing cannot be predicted without knowing an individual's recent and long-term linguistic history. Evidence that processing dynamics are affected by conditions that promote or diminish competition further suggests that parameters governing processing are directly influenced by linguistic experience. Such challenges are familiar to psycholinguists, as we routinely use corpus-based word frequencies as our best estimate of central tendencies in a population, and we recognize that individuals' own internal (implicit or explicit) estimates of frequency depend on their unique experience.

---

[3] One might be able to tweak these models to accommodate priming (e.g., by temporarily incrementing resting activation levels in TRACE, or prior probabilities in Shortlist B, following any encounter with a specific word), but a comprehensive account that could accommodate both priming and long-term learning is lacking.

## 6. IMPLICATIONS FOR SPOKEN WORD RECOGNITION MODELS

In this review we have presented representative examples showing different forms of plasticity in SWR: changes linked to early development and throughout the life span, differences linked to the structure and content of language (and to the languages a listener knows), and rapid and/or short-term changes linked to recent and/or concurrent context. We expect that virtually all of these findings would be difficult for current models of SWR to accommodate. However, to our knowledge, few have been explored with current models.

Considering again the search for static/universal processes supporting SWR, crosslinguistic differences in fundamental aspects of SWR (e.g., the impact of neighborhood size) point to the need for computational modeling to assess whether current models can simulate crosslinguistic differences simply by changing the contents of the lexicon appropriately. Multilingualism raises greater challenges: While there are excellent models of bilingual word recognition that extend monolingual modeling approaches (for a recent example, see Dijkstra et al. 2019), to our knowledge all posit additional architecture to allow both separation and interaction between languages a listener knows. Thus, based on current understanding, multilingualism implies the need for qualitatively different and more complex cognitive architectures than those put forward by current models and—importantly—for architectures that can also account for the observed plasticity of SWR (e.g., how a listener's L1 and their experience with multiple languages can fundamentally shape the system underlying SWR).

More broadly, what appears to be the most crucial gap is that most current models of SWR are not learning models [for an exception to both gaps, see Magnuson et al. 2020; for an example of how the TRACE model (McClelland & Elman 1986) can be modified to allow context-dependent learning to adjust default mappings, see Mirman et al. 2006]. This issue also highlights a difference between models like TRACE and Shortlist B (Norris & McQueen 2008)—which have discrete levels of representation (such as phonemes and words) and explicit parameters governing activation flow (with distinct parameters governing interlevel excitation and intralevel inhibition in TRACE) or likelihood updates—and learning models such as Simple Recurrent Networks (Elman 1990; for examples extended to SWR, see Gaskell & Marslen-Wilson 1997, Magnuson et al. 2003) and EARSHOT (Magnuson et al. 2020). In learning models, parameters are not conveniently organized or discrete. Instead, complex patterns of positive and negative weights within and between levels are continuously modified by learning. Unpacking those complex connectivity patterns is an inherent challenge in learning models, but doing so may provide one of the best opportunities to advance our theories and models.

In principle, any form of plasticity could be accounted for in the context of a model that incorporates learning appropriately. We believe that part of the reason why incorporating learning in SWR models has not been a priority so far is that the role of plasticity has been largely underestimated.

## 7. CONCLUSIONS

The goal of this contribution was to bring together evidence from research with different populations and different methods that speak to the plasticity of SWR. This review is far from exhaustive; we were only able to review representative patterns of findings that compellingly illustrate that SWR is constantly shaped by a wide variety of factors related to the input itself, characteristics of the listener, learning, and the broader context of communication (e.g., languages a person knows, or the particular pressures of a specific orthography).

Even though the results we have presented emerge across different timescales, it is reasonable to assume that longitudinal changes must stem from repeated short-term effects accumulated over extended periods of time (Protopapas & Kapnoula 2016). In that sense, the mechanisms driving both short-term and long-term adaptations are intricately related or even likely the same—a conclusion consistent with our view that boundaries between processing and structure are not as clearly defined as one might assume. Indeed, as suggested in Section 1, we hold that evidence for constant plasticity implies a system where contents and processes are not separable. Changes in contents and experience with specific items in varying contexts (e.g., contexts that promote or diminish competition; Kapnoula & McMurray 2016b) can modify the parameters that govern processing. In sum, our review includes multiple examples of changes that point to the need for the field to move beyond the perspective of an idealized, static system and to embrace plasticity as a core concern in theories and models of SWR.

## DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

## ACKNOWLEDGMENTS

This research was supported in part by the Basque Government through the BERC 2022–2025 program and by the Spanish State Research Agency through BCBL Severo Ochoa excellence accreditation CEX2020-001010/AEI/10.13039/501100011033 and through project PID2020-119131GB-I0 (BLIS) awarded to J.S.M. This work was also supported by the Spanish Ministry of Economy and Competitiveness (via project PID2020-113348GB-I00 awarded to E.C.K. and Arthur G. Samuel). M.J. was supported by a predoctoral fellowship (associated with project PSI2017 82941-P; grant PRE-2018-083946) from the Spanish Ministry of Science, Innovation and Universities and the Fondo Social Europeo. This research was also supported in part by US National Science Foundation grants 1754284 and 2043903 to J.S.M.

## LITERATURE CITED

Allopenna PD, Magnuson J, Tanenhaus MK. 1998. Tracking the time course of spoken word recognition using eye movements: evidence for continuous mapping models. *J. Mem. Lang.* 38(4):419–39. **https://doi.org/10.1006/jmla.1997.2558**

Altmann GT, Kamide Y. 1999. Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition* 73(3):247–64

Amano S, Kondo T. 2000. *Neighborhood and cohort in lexical processing of Japanese spoken words*. Paper presented at the ISCA Tutorial and Research Workshop (ITRW) on Spoken Word Access Processes, Nijmegen, Neth., May 29–31

Andruski JE, Blumstein SE, Burton M. 1994. The effect of subphonetic differences on lexical access. *Cognition* 52(3):163–87. **https://doi.org/10.1016/0010-0277(94)90042-6**

Anglin JM, Miller GA, Wakefield PC. 1993. Vocabulary development: a morphological analysis. *Monogr. Soc. Res. Child Dev.* 58(10):1–186. **https://doi.org/10.2307/1166112**

Apfelbaum KS, Goodwin C, Blomquist C, McMurray B. 2023. The development of lexical competition in written- and spoken-word recognition. *Q. J. Exp. Psychol.* 76(1):196–219. **https://doi.org/10.1177/17470218221090483**

Arutiunian V, Lopukhina A. 2020. The effects of phonological neighborhood density in childhood word production and recognition in Russian are opposite to English. *J. Child Lang.* 47(6):1244–62. **https://doi.org/10.1017/S0305000920000112**

Ben-David BM, Chambers CG, Daneman M, Pichora-Fuller MK, Reingold EM, Schneider BA. 2011. Effects of aging and noise on real-time spoken word recognition: evidence from eye movements. *J. Speech Lang. Hearing Res.* 54(1):243–62. **https://doi.org/10.1044/1092-4388(2010/09-0233)**

Best CC, McRoberts GW. 2003. Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Lang. Speech* 46(2–3):183–216. **https://doi.org/10.1177/00238309030460020701**

Blomquist C, McMurray B. 2023. The development of lexical inhibition in spoken word recognition. *Dev. Psychol.* 59(1):186–206. **https://doi.org/10.1037/DEV0001457**

Blumenfeld H, Marian V. 2007. Constraints on parallel activation in bilingual spoken language processing: examining proficiency and lexical status using eye-tracking. *Lang. Cogn. Process.* 22(5):633–60. **https://doi.org/10.1080/01690960601000746**

Bowers JS. 2000. In defense of abstractionist theories of repetition priming and word identification. *Psychon. Bull. Rev.* 7(1):83–99

Bradlow AR, Nygaard L, Pisoni D. 1999. Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Percept. Psychophys.* 61(2):206–19

Chambers CG, Tanenhaus MK, Magnuson J. 2004. Actions and affordances in syntactic ambiguity resolution. *J. Exp. Psychol. Learn. Mem. Cogn.* 30(3):687–96. **https://doi.org/10.1037/0278-7393.30.3.687**

Charles-Luce J, Luce PA. 1990. Similarity neighbourhoods of words in young children's lexicons. *J. Child Lang.* 17(1):205–15. **https://doi.org/10.1017/S0305000900013180**

Cheimariou S, Kapnoula EC. 2022. *Age effects on lexical competition are external to the mental lexicon: evidence from eye-movements*. Poster presented at the 63rd Annual Meeting of the Psychonomic Society, Boston, Nov. 17–20

Chéreau C, Gaskell MG, Dumay N. 2007. Reading spoken words: orthographic effects in auditory priming. *Cognition* 102(3):341–60. **https://doi.org/10.1016/J.COGNITION.2006.01.001**

Dahan D, Magnuson J, Tanenhaus MK, Hogan EM. 2001. Subcategorical mismatches and the time course of lexical access: evidence for lexical competition. *Lang. Cogn. Process.* 16(5–6):507–34. **https://doi.org/10.1080/01690960143000074**

Dahan D, Swingley D, Tanenhaus MK, Magnuson J. 2000. Linguistic gender and spoken-word recognition in French. *J. Mem. Lang.* 42(4):465–80. **https://doi.org/10.1006/jmla.1999.2688**

Dijkstra T, Van Hell JG, Brenders P. 2015. Sentence context effects in bilingual word recognition: cognate status, sentence language, and semantic constraint. *Biling. Lang. Cogn.* 18(4):597–613

Dijkstra T, Wahl A, Buytenhuijs F, Van Halem N, Al-Jibouri Z, et al. 2019. Multilink: a computational model for bilingual word recognition and word translation. *Biling. Lang. Cogn.* 22(4):657–79

Eberhard KM, Spivey-Knowlton MJ, Sedivy JC, Tanenhaus MK. 1995. Eye movements as a window into real-time spoken language comprehension in natural contexts. *J. Psycholinguist. Res.* 24(6):409–36

Eimas PD, Siqueland ER, Jusczyk P, Vigorito J. 1971. Speech perception in infants. *Science* 171(3968):303–6. **https://doi.org/10.1126/SCIENCE.171.3968.303**

Ellis AW. 1982. Modality-specific repetition priming of auditory word recognition. *Curr. Psychol. Res.* 2(1–3):123–27. **https://doi.org/10.1007/BF03186752**

Elman JL. 1990. Finding structure in time. *Cogn. Sci.* 14(2):179–211. **https://doi.org/10.1207/s15516709cog1402_1**

Elman JL. 2004. An alternative view of the mental lexicon. *Trends Cogn. Sci.* 8(7):301–6. **https://doi.org/10.1016/j.tics.2004.05.003**

Elman JL. 2009. On the meaning of words and dinosaur bones: lexical knowledge without a lexicon. *Cogn. Sci.* 33(4):547–82. **https://doi.org/10.1111/j.1551-6709.2009.01023.x**

Ferretti TR, McRae K, Hatherell A. 2001. Integrating verbs, situation schemas, and thematic role concepts. *J. Mem. Lang.* 44(4):516–47. **https://doi.org/10.1006/jmla.2000.2728**

Flege JE, Bohn O. 2021. The revised Speech Learning Model (SLM-r). In *Second Language Speech Learning: Theoretical and Empirical Progress*, ed. R Wayland, pp. 3–83. Cambridge, UK: Cambridge Univ. Press

Frost R. 2012. Towards a universal model of reading. *Behav. Brain Sci.* 35(5):263–79. **https://doi.org/10.1017/S0140525X11001841**

Gaskell MG, Dumay N. 2003. Lexical competition and the acquisition of novel words. *Cognition* 89(2):105–32. **https://doi.org/10.1016/S0010-0277(03)00070-2**

Gaskell MG, Marslen-Wilson WD. 1997. Integrating form and meaning: a distributed model of speech perception. *Lang. Cogn. Process.* 12(5–6):613–56

Getzmann S, Hanenberg C, Lewald J, Falkenstein M, Wascher E. 2015. Effects of age on electrophysiological correlates of speech processing in a dynamic "cocktail-party" situation. *Front. Neurosci.* 9:341. **https://doi.org/10.3389/fnins.2015.00341**

Goldinger SD. 1996. Words and voices: episodic traces in spoken word identification and recognition memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 22(5):1166–83

Goldinger SD. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychol. Rev.* 105(2):251–79

Gosselin PA, Gagné JP. 2011. Older adults expend more listening effort than young adults recognizing speech in noise. *J. Speech Lang. Hear. Res.* 54(3):944–58. **https://doi.org/10.1044/1092-4388(2010/10-0069)**

Grosjean F. 1980. Spoken word recognition processes and the gating paradigm. *Percept. Psychophys.* 28:267–83

Grosjean F. 2022. *The Mysteries of Bilingualism: Unresolved Issues*. Hoboken, NJ: Wiley

Guediche S, Fiez JA, Holt LL. 2016. Adaptive plasticity in speech perception: effects of external information and internal predictions. *J. Exp. Psychol. Hum. Percept. Perform.* 42(7):1048–59. **https://doi.org/10.1037/XHP0000196**

Gür E, Samuel AG, Kapnoula EC. 2023. *Phonological neighborhood density effects on Spanish spoken word recognition and word learning*. Poster presented at the 64th Annual Meeting of the Psychonomic Society, San Francisco, Nov. 16–19

Gussow AE, Kapnoula E, Molinaro N. 2019. Any leftovers from a discarded prediction? Evidence from eye-movements during sentence comprehension. *Lang. Cogn. Neurosci.* 34(8):1041–58. **https://doi.org/10.1080/23273798.2019.1617887**

Hagoort P, Hald L, Bastiaansen M, Petersson KM. 2004. Integration of word meaning and world knowledge in language comprehension. *Science* 304(5669):438–41. **https://doi.org/10.1126/SCIENCE.1095455**

Hallé PA, de Boysson-Bardies B. 1996. The format of representation of recognized words in infants' early receptive lexicon. *Infant Behav. Dev.* 19(4):463–81. **https://doi.org/10.1016/S0163-6383(96)90007-7**

Hannagan T, Magnuson J, Grainger J. 2013. Spoken word recognition without a TRACE. *Front. Psychol.* 4:563. **https://doi.org/10.3389/fpsyg.2013.00563**

Harada CN, Natelson Love MC, Triebel KL. 2013. Normal cognitive aging. *Clin. Geriatr. Med.* 29(4):737–52. **https://doi.org/10.1016/j.cger.2013.07.002**

Hayakawa S, Marian V. 2020. Studying bilingualism through eye-tracking and brain imaging. In *Bilingual Lexical Ambiguity Resolution*, ed. RR Heredia, AB Cieślicka, pp. 273–99. Cambridge, UK: Cambridge Univ. Press

Henrich J, Heine SJ, Norenzayan A. 2010. The weirdest people in the world? *Behav. Brain Sci.* 33(2–3):61–83

Hendrickson K, Oleson J, Walker E. 2021. School-age children adapt the dynamics of lexical competition in suboptimal listening conditions. *Child Dev.* 92(2):638–49. **https://doi.org/10.1111/CDEV.13530**

Heyselaar E, Peeters D, Hagoort P. 2021. Do we predict upcoming speech content in naturalistic environments? *Lang. Cogn. Neurosci.* 36(4):440–61. **https://doi.org/10.1080/23273798.2020.1859568/SUPPL_FILE/PLCP_A_1859568_SM1317.DOCX**

Hollich GJ, Jusczyk PW, Luce PA. 2000. Of words, birds, worms, and weeds: infant word learning and lexical neighborhoods. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, Vol. 22. Merced, CA: Cogn. Sci. Soc. **https://escholarship.org/uc/item/13n9q7rj**

Houston DM, Jusczyk PW. 2000. The role of talker-specific information in word segmentation by infants. *J. Exp. Psychol. Hum. Percept. Perform.* 26(5):1570–82. **https://doi.org/10.1037/0096-1523.26.5.1570**

Jackson A, Morton J. 1984. Facilitation of auditory word recognition. *Mem. Cogn.* 12(6):568–74. **https://doi.org/10.3758/BF03213345/METRICS**

Ju M, Luce PA. 2004. Falling on sensitive ears. *Psychol. Sci.* 15(5):314–18. **https://doi.org/10.1111/J.0956-7976.2004.00675.X**

Kapnoula EC, McMurray B. 2016a. Newly learned word forms are abstract and integrated immediately after acquisition. *Psychon. Bull. Rev.* 23:491–99. **https://doi.org/10.3758/s13423-015-0897-1**

Kapnoula EC, McMurray B. 2016b. Training alters the resolution of lexical interference: evidence for plasticity of competition and inhibition. *J. Exp. Psychol. Gen.* 145(1):8–30. **https://doi.org/10.1037/xge0000123**

Kapnoula EC, Packard S, Gupta P, McMurray B. 2015. Immediate lexical integration of novel word forms. *Cognition* 134:85–99. **https://doi.org/10.1016/j.cognition.2014.09.007**

Kapnoula EC, Samuel AG. 2019. Voices in the mental lexicon: Words carry indexical information that can affect access to their meaning. *J. Mem. Lang.* 107:111–27. **https://doi.org/10.1016/J.JML.2019.05.001**

Kapnoula EC, Samuel AG. 2022. Wait long and prosper! Delaying production alleviates its detrimental effect on word learning. *Lang. Cogn. Neurosci.* 38:724–44. **https://doi.org/10.1080/23273798.2022.2144917**

Kapnoula EC, Samuel AG. 2023. *Sensitivity to L1 subphonemic differences predicts vocabulary size in a foreign language*. Work. Pap., BCBL, Donostia–San Sebastián, Spain. **https://osf.io/fytvh?view_only=7cc55f53d19342cc867b953a39b9403e**

Kavé G. 2022. Vocabulary changes in adulthood: main findings and methodological considerations. *Int. J. Lang. Commun. Disord.* **https://doi.org/10.1111/1460-6984.12820**

Kempley ST, Morton J. 1982. The effects of priming with regularly and irregularly related words in auditory word recognition. *Br. J. Psychol.* 73(Part 4):441–54. **https://doi.org/10.1111/J.2044-8295.1982.TB01826.X**

Kirk NW. 2023. MIND your language (s): recognizing Minority, Indigenous, Non-standard(ized), and Dialect variety usage in "monolinguals." *Appl. Psycholinguist.* 44(3):358–64

Kraljic T, Samuel AG. 2005. Perceptual learning for speech: Is there a return to normal? *Cogn. Psychol.* 51(2):141–78

Kraljic T, Samuel AG. 2006. Generalization in perceptual learning for speech. *Psychon. Bull. Rev.* 13(2):262–68

Kraljic T, Samuel AG. 2011. Perceptual learning evidence for contextually-specific representations. *Cognition* 121(3):459–65

Kuhl PK, Williams KA, Lacerda F, Stevens KN, Lindblom B. 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255(5044):606–8. **https://doi.org/10.1126/science.1736364**

Kukona A, Fang S-Y, Aicher KA, Chen H, Magnuson J. 2011. The time course of anticipatory constraint integration. *Cognition* 119(1):23–42. **https://doi.org/10.1016/j.cognition.2010.12.002**

Kutas M, Delong KA, Smith NJ. 2011. A look around at what lies ahead: prediction and predictability in language processing. In *Predictions in the Brain: Using Our Past to Generate a Future*, ed. M Bar, pp. 190–207. Oxford, UK: Oxford Univ. Press

Kutlu E, Fell A, Apfelbaum K, McMurray B. 2022. Effects of multilingual and monolingual social networks on speech perception. *J. Acoust. Soc. Am.* 152(4):A236. **https://doi.org/10.1121/10.0016126**

Lahiri A, Marslen-Wilson W. 1991. The mental representation of lexical form: a phonological approach to the recognition lexicon. *Cognition* 38(3):245–94. **https://doi.org/10.1016/0010-0277(91)90008-R**

Leach L, Samuel AG. 2007. Lexical configuration and lexical engagement: when adults learn new words. *Cogn. Psychol.* 55(4):306–53. **https://doi.org/10.1016/j.cogpsych.2007.01.001**

Liberman IY, Shankweiler D, Liberman AM. 1989. *The Alphabetic Principle and Learning to Read*. Washington, DC: US Dep. Health Hum. Serv.

Luce P, McLennan C, Charles-Luce J. 2012. Abstractness and specificity in spoken word recognition: indexical and allophonic variability in long-term repetition priming. In *Rethinking Implicit Memory*, ed. JS Bowers, CJ Marsolek, pp. 197–214. Oxford, UK: Oxford Univ. Press. **https://doi.org/10.1093/acprof:oso/9780192632326.003.0009**

Luce P, Pisoni D. 1998. Recognizing spoken words: the neighborhood activation model. *Ear Hear.* 19(1):1–36

Luthra S, Peraza-Santiago G, Beeson KN, Saltzman D, Crinnion AM, Magnuson JS. 2021. Robust lexically mediated compensation for coarticulation: Christmash time is here again. *Cogn. Sci.* 45(4):e12962. **https://doi.org/10.1111/cogs.12962**

Magnuson JS. 2008. Nondeterminism, pleiotropy, and single-word reading: theoretical and practical concerns. In *Single-Word Reading: Behavioral and Biological Perspectives*, ed. EL Grigorenko, AJ Naples, pp. 377–404. Mahwah, NJ: Lawrence Erlbaum

Magnuson JS. 2016. Mapping spoken words to meaning. In *Speech Perception and Spoken Word Recognition*, ed. MG Gaskell, J Mirkovic, pp. 86–106. New York: Psychol. Press

Magnuson JS, Crinnion AM. 2022. Spoken word recognition. In *The Oxford Handbook of the Mental Lexicon*, ed. A Papafragou, JC Trueswell, LR Gleitman, pp. 461–90. Oxford, UK: Oxford Univ. Press. **https://doi.org/10.1093/oxfordhb/9780198845003.013.23**

Magnuson JS, Mirman D, Harris HD. 2012. Computational models of spoken word recognition. In *The Cambridge Handbook of Psycholinguistics*, ed. M Spivey, K McRae, M Joanisse, pp. 76–103. Cambridge, UK: Cambridge Univ. Press

Magnuson JS, Mirman D, Luthra S, Strauss T, Harris HD. 2018. Interaction in spoken word recognition models: Feedback helps. *Front. Psychol.* 9:369. **https://doi.org/10.3389/fpsyg.2018.00369**

Magnuson JS, Tanenhaus MK, Aslin RN, Dahan D. 2003. The time course of spoken word learning and recognition: studies with artificial lexicons. *J. Exp. Psychol. Gen.* 132(2):202–27. **https://doi.org/10.1037/0096-3445.132.2.202**

Magnuson JS, You H, Luthra S, Li M, Nam H, et al. 2020. EARSHOT: a minimal neural network model of incremental human speech recognition. *Cogn. Sci.* 44(4):e12823. **https://doi.org/10.1111/cogs.12823**

Markman M. 1987. How children constrain the possible meanings of words. In *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*, ed. U Neisser, pp. 255–87. Cambridge, UK: Cambridge Univ. Press

Marr D. 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: W.H. Freeman

Marslen-Wilson WD. 1987. Functional parallelism in spoken word-recognition. *Cognition* 25:71–102

Marslen-Wilson WD. 1993. Issues of process and representation in lexical access. In *Cognitive Models of Speech Processing: The Second Sperlonga Meeting*, ed. GTM Altmann, R Shillcock, pp. 187–210. Hove, UK: Lawrence Erlbaum

Marslen-Wilson WD, Warren P. 1994. Levels of perceptual representation and process in lexical access: words, phonemes, and features. *Psychol. Rev.* 101(4):653–75

Marslen-Wilson WD, Welsh A. 1978. Processing interactions and lexical access during word recognition in continuous speech. *Cogn. Psychol.* 10(1):29–63. **https://doi.org/10.1016/0010-0285(78)90018-X**

Marslen-Wilson WD, Zwitserlood P. 1989. Accessing spoken words: the importance of word onsets. *J. Exp. Psychol. Hum. Percept. Perform.* 15(3):576–85

McClelland JL, Elman JL. 1986. The TRACE model of speech perception. *Cogn. Psychol.* 18(1):1–86. **https://doi.org/10.1016/0010-0285(86)90015-0**

McLaughlin DJ, Zink ME, Gaunt L, Spehar B, Van Engen KJ, et al. 2022. Pupillometry reveals cognitive demands of lexical competition during spoken word recognition in young and older adults. *Psychon. Bull. Rev.* 29(1):268–80. **https://doi.org/10.3758/S13423-021-01991-0**

McMurray B. 2022. The myth of categorical perception. *J. Acoust. Soc. Am.* 152(6):3819–42. **https://doi.org/10.31234/OSF.IO/DQ7EJ**

McMurray B, Danelz A, Rigler H, Seedorff M. 2018. Speech categorization develops slowly through adolescence. *Dev. Psychol.* 54(8):1472–91

McMurray B, Tanenhaus MK, Aslin RN. 2002. Gradient effects of within-category phonetic variation on lexical access. *Cognition* 86(2):B33–42. **https://doi.org/10.1016/S0010-0277(02)00157-9**

McQueen JM, Cutler A, Briscoe T, Norris D. 1995. Models of continuous speech recognition and the contents of the vocabulary. *Lang. Cogn. Process.* 10:309–31

McQueen JM, Cutler A, Norris D. 2006. Phonological abstraction in the mental lexicon. *Cogn. Sci.* 30(6):1113–26

Mesgarani N, Cheung C, Johnson K, Chang EF. 2014. Phonetic feature encoding in human superior temporal gyrus. *Science* 343(6174):1006–10. **https://doi.org/10.1126/science.1245994**

Metusalem R, Kutas M, Urbach TP, Hare M, McRae K, Elman JL. 2012. Generalized event knowledge activation during online sentence comprehension. *J. Mem. Lang.* 66(4):545–67. **https://doi.org/10.1016/j.jml.2012.01.001**

Miller JL. 1997. Internal structure of phonetic categories. *Lang. Cogn. Process.* 12(5–6):865–70. **https://doi.org/10.1080/016909697386754**

Mirman D, McClelland JL, Holt LL. 2006. An interactive Hebbian account of lexically guided tuning of speech perception. *Psychon. Bull. Rev.* 13(6):958–65

Morais J, Cary L, Alegria J, Bertelson P. 1979. Does awareness of speech as a sequence of phones arise spontaneously? *Cognition* 7(4):323–31. **https://doi.org/10.1016/0010-0277(79)90020-9**

Morton J. 1964. A preliminary functional model for language behaviour. *Int. Audiol.* 3(2):216–25. **https://doi.org/10.3109/05384916409074089**

Muneaux M, Ziegler JC. 2004. Locus of orthographic effects in spoken word recognition: novel insights from the neighbour generation task. *Lang. Cogn. Process.* 19(5):641–60. **https://doi.org/10.1080/01690960444000052**

Norris D. 1994. Shortlist: a connectionist model of continuous speech recognition. *Cognition* 52(3):189–234. **https://doi.org/10.1016/0010-0277(94)90043-4**

Norris D, McQueen JM. 2008. Shortlist B: a Bayesian model of continuous speech recognition. *Psychol. Rev.* 115(2):357–95. **https://doi.org/10.1037/0033-295X.115.2.357**

Norris D, McQueen JM, Cutler A. 2000. Merging information in speech recognition: Feedback is never necessary. *Behav. Brain Sci.* 23(3):299–325. **https://doi.org/10.1017/S0140525X00003241**

Norris D, McQueen JM, Cutler A. 2003. Perceptual learning in speech. *Cogn. Psychol.* 47(2):204–38. **https://doi.org/10.1016/S0010-0285(03)00006-9**

Norris D, McQueen JM, Cutler A. 2018. Commentary on "Interaction in spoken word recognition models." *Front. Psychol.* 9:1568. **https://doi.org/10.3389/fpsyg.2018.01568**

Oleson JJ, Cavanaugh JE, McMurray B, Brown G. 2017. Detecting time-specific differences between temporal nonlinear curves: analyzing data from the visual world paradigm. *Stat. Methods Med. Res.* 26(6):2708–25. **https://doi.org/10.1177/0962280215607411**

Pattamadilok C, Morais J, De Vylder O, Ventura P, Kolinsky R. 2009. The orthographic consistency effect in the recognition of French spoken words: an early developmental shift from sublexical to lexical orthographic activation. *Appl. Psycholinguist.* 30(3):441–62. **https://doi.org/10.1017/S0142716409090225**

Perre L, Pattamadilok C, Montant M, Ziegler JC. 2009. Orthographic effects in spoken language: on-line activation or phonological restructuring? *Brain Res.* 1275:73–80. **https://doi.org/10.1016/J.BRAINRES.2009.04.018**

Perre L, Ziegler JC. 2008. On-line activation of orthography in spoken word recognition. *Brain Res.* 1188(1):132–38. **https://doi.org/10.1016/J.BRAINRES.2007.10.084**

Pickering MJ, Gambi C. 2018. Predicting while comprehending language: a theory and review. *Psychol. Bull.* 144(10):1002–44. **https://doi.org/10.1037/BUL0000158**

Pierrehumbert JB. 2016. Phonological representation: beyond abstract versus episodic. *Annu. Rev. Linguist.* 2:33–52

Port R. 2007. How are words stored in memory? Beyond phones and phonemes. *New Ideas Psychol.* 25(2):143–70. **https://doi.org/10.1016/J.NEWIDEAPSYCH.2007.02.001**

Pratt AC, Brady S. 1988. Relation of phonological awareness to reading disability in children and adults. *J. Educ. Psychol.* 80(3):319–23. **https://doi.org/10.1037/0022-0663.80.3.319**

Protopapas A, Kapnoula E. 2016. Short-term and long-term effects on visual word recognition. *J. Exp. Psychol. Learn. Mem. Cogn.* 42(4):542–65. **https://doi.org/10.1037/xlm0000191**

Pustejovsky J. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press

Revill K, Spieler D. 2012. The effect of lexical frequency on spoken word recognition in young and older listeners. *Psychol. Aging* 27(1):80–87

Rigler H, Farris-Trimble A, Greiner L, Walker J, Tomblin JB, McMurray B. 2015. The slow developmental time course of real-time spoken word recognition. *Dev. Psychol.* 51(12):1690–703. **https://doi.org/10.1037/dev0000044**

Salverda AP, Dahan D, McQueen JM. 2003. The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition* 90:51–89

Samuel AG. 1977. The effect of discrimination training on speech perception: noncategorical perception. *Percept. Psychophys.* 22(4):321–30. **https://doi.org/10.3758/BF03199697**

Samuel AG. 1982. Phonetic prototypes. *Percept. Psychophys.* 31(4):307–14. **https://doi.org/10.3758/BF03202653**

Sarrett ME, McMurray B, Kapnoula E. 2020. Dynamic EEG analysis during language comprehension reveals interactive cascades between perceptual processing and sentential expectations. *Brain Lang.* 211:104875. **https://doi.org/10.1016/j.bandl.2020.104875**

Sarrett ME, Shea C, McMurray B. 2022. Within- and between-language competition in adult second language learners: implications for language proficiency. *Lang. Cogn. Neurosci.* 37(2):165–81

Schacter DL, Church BA. 1992. Auditory priming: implicit and explicit memory for words and voices. *J. Exp. Psychol. Learn. Mem. Cogn.* 18(5):915–30. **https://doi.org/10.1037/0278-7393.18.5.915**

Sedivy JC, Tanenhaus MK, Chambers CG, Carlson GN. 1999. Achieving incremental semantic interpretation through contextual representation. *Cognition* 71(2):109–47

Seedorff M, Oleson J, McMurray B. 2018. Detecting when timeseries differ: using the Bootstrapped Differences of Timeseries (BDOTS) to analyze Visual World Paradigm data (and more). *J. Mem. Lang.* 102:55–67. **https://doi.org/10.1016/j.jml.2018.05.004**

Seidenberg MS, Tanenhaus MK. 1979. Orthographic effects on rhyme monitoring. *J. Exp. Psychol. Hum. Learn. Mem.* 5(6):546–54. **https://doi.org/10.1037/0278-7393.5.6.546**

Share DL. 1995. Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition* 55(2):151–218. **https://doi.org/10.1016/0010-0277(94)00645-2**

Share DL. 2004. Orthographic learning at a glance: on the time course and developmental onset of self-teaching. *J. Exp. Child Psychol.* 87(4):267–98. **https://doi.org/10.1016/J.JECP.2004.01.001**

Shook A, Marian V. 2016. The influence of native-language tones on lexical access in the second language. *J. Acoust. Soc. Am.* 139(6):3102–9. **https://doi.org/10.1121/1.4953692**

Shook A, Marian V. 2019. Covert co-activation of bilinguals' non-target language: phonological competition from translations. *Linguist. Approaches Biling.* 9(2):228–52. **https://doi.org/10.1075/LAB.17022.SHO**

Singh L. 2008. Influences of high and low variability on infant word recognition. *Cognition* 106(2):833–70

Singh L, Morgan JL, White KS. 2004. Preference and processing: the role of speech affect in early spoken word recognition. *J. Mem. Lang.* 51(2):173–89

Sommers MS. 1996. The structural organization of the mental lexicon and its contribution to age-related declines in spoken-word recognition. *Psychol. Aging* 11(2):333–41. **https://doi.org/10.1037/0882-7974.11.2.333**

Sommers MS, Danielson SM. 1999. Inhibitory processes and spoken word recognition in young and older adults: the interaction of lexical competition and semantic context. *Psychol. Aging* 14(3):458–72

Spivey M, Marian V. 1999. Cross talk between native and second languages: partial activation of an irrelevant lexicon. *Psychol. Sci.* 10(3):281–84

Stager C, Werker J. 1997. Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature* 388(6640):381–82

Storkel H. 2002. Restructuring of similarity neighbourhoods in the developing mental lexicon. *J. Child Lang.* 29(2):251–74

Storkel H, Armbrüster J, Hogan TP. 2006. Differentiating phonotactic probability and neighborhood density in adult word learning. *J. Speech Lang. Hear. Res.* 49(6):1175–92. **https://doi.org/10.1044/1092-4388(2006/085**)

Streeter L, Nigro G. 1979. The role of medial consonant transitions in word perception. *J. Acoust. Soc. Am.* 65(6):1533–41

Swank LK, Catts HW. 1994. Phonological awareness and written word decoding. *Lang. Speech Hear. Serv. Sch.* 25(1):9–14. **https://doi.org/10.1044/0161-1461.2501.09**

Swingley D, Aslin R. 2000. Spoken word recognition and lexical representation in very young children. *Cognition* 76:147–66

Tanenhaus MK, Spivey-Knowlton MJ, Eberhard KM, Sedivy JC. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268(5217):1632–34

Toscano JC, McMurray B, Dennhardt J, Luck SJ. 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):1532–40. **https://doi.org/10.1177/0956797610384142**

Tun PA, O'Kane G, Wingfield A. 2002. Distraction by competing speech in young and older adult listeners. *Psychol. Aging* 17(3):453–67. **https://doi.org/10.1037//0882-7974.17.3.453**

Ventura P, Kolinsky R, Pattamadilok C, Morais J. 2008. The developmental turnpoint of orthographic consistency effects in speech recognition. *J. Exp. Child Psychol.* 100(2):135–45. **https://doi.org/10.1016/J.JECP.2008.01.003**

Ventura P, Morais J, Kolinsky R. 2007. The development of the orthographic consistency effect in speech recognition: from sublexical to lexical involvement. *Cognition* 105(3):547–76. **https://doi.org/10.1016/J.COGNITION.2006.12.005**

Vitevitch M, Rodríguez E. 2005. Neighborhood density effects in spoken word recognition in Spanish. *J. Multiling. Commun. Disord.* 3(1):64–73. **https://doi.org/10.1080/14769670400027332**

Vroomen J, de Gelder B. 1995. Metrical segmentation and lexical inhibition in spoken word recognition. *J. Exp. Psychol. Hum. Percept. Perform.* 21(1):98–108. **https://doi.org/10.1037/0096-1523.21.1.98**

Werker JF, Tees RC. 1984. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behav. Dev.* 7:49–63

Wicha NYY, Bates EA, Moreno EM, Kutas M. 2003. Potato not Pope: human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neurosci. Lett.* 346(3):165–68

Wicha NYY, Moreno EM, Kutas M. 2004. Anticipating words and their gender: an event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *J. Cogn. Neurosci.* 16(7):1272–88. **https://doi.org/10.1162/0898929041920487**

Wingfield A, Aberdeen JS, Stine EA. 1991. Word onset gating and linguistic context in spoken word recognition by young and elderly adults. *J. Gerontol.* 46(3):P127–29. **https://doi.org/10.1093/GERONJ/46.3.P127**

Ziegler JC, Ferrand L. 1998. Orthography shapes the perception of speech: the consistency effect in auditory word recognition. *Psychon. Bull. Rev.* 5(4):683–89. **https://doi.org/10.3758/BF03208845/METRICS**

Ziegler JC, Goswami U. 2005. Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychol. Bull.* 131(1):3–29. **https://doi.org/10.1037/0033-2909.131.1.3**

Ziegler JC, Muneaux M, Grainger J. 2003. Neighborhood effects in auditory word recognition: phonological competition and orthographic facilitation. *J. Mem. Lang.* 48(4):779–93. **https://doi.org/10.1016/S0749-596X(03)00006-8**

Ziegler JC, Petrova A, Ferrand L. 2008. Feedback consistency effects in visual and auditory word recognition: Where do we stand after more than a decade? *J. Exp. Psychol. Learn. Mem. Cogn.* 34(3):643–61. **https://doi.org/10.1037/0278-7393.34.3.643**

Annual Review
of Linguistics

Volume 10, 2024

# Contents

Can't Believe It Went By So Fast
*Hans Kamp* ........ 1

Linguistic Encounters: A Scholar's Journey
*Edward L. Keenan* ........ 19

Linguistic Variation and Linguistic Inclusion in the US Educational Context
*Christine Mallinson* ........ 37

The Size of Clausal Complements
*Susanne Wurmbrand* ........ 59

Intensionality and Propositionalism
*Kristina Liefke* ........ 85

Types and Type Theories in Natural Language Analysis
*Peter R. Sutton* ........ 107

Advances in Understanding Stuttering as a Disorder of Language Encoding
*Nan Bernstein Ratner and Shelley B. Brundage* ........ 127

Philippine Psycholinguistics
*Jed Sam Pizarro-Guevara and Rowena Garcia* ........ 145

Semantics of Gesture
*Cornelia Ebert* ........ 169

Phonetics of Early Bilingualism
*Mark Amengual* ........ 191

Analogy in Inflection
*Emily Lindsay-Smith, Matthew Baerman, Sacha Beniamine, Helen Sims-Williams, and Erich R. Round* ........ 211

Spoken Word Recognition: A Focus on Plasticity
*Efthymia C. Kapnoula, Mina Jevtović, and James S. Magnuson* ........ 233

Interjections at the Heart of Language
*Mark Dingemanse* ........ 257

Biased Polar Questions
*Maribel Romero* . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 279

## Errata

An online log of corrections to *Annual Review of Linguistics* articles may be found at http://www.annualreviews.org/errata/linguistics