Article:

Mattys, S.L., O'Leary, R.M., McGarrigle, R.A. orcid.org/0000-0003-1704-1135 et al. (1 more author) (Accepted: 2025) Reconceptualizing Cognitive Listening. Trends in Cognitive Sciences. ISSN: 1364-6613 (In Press)

Reconceptualizing Cognitive Listening

Sven L. Mattys[1], Ryan M. O'Leary[2], Ronan A. McGarrigle[3], & Arthur Wingfield[4]

[1]Department of Psychology, University of York, York YO10 5DD, UK

[2]Department of Communication Sciences and Disorders and Center for Cognitive and Brain Health, Northeastern University, Boston MA 02115, USA

[3]School of Psychology, University of Leeds, Leeds LS2 9JT, UK

[4]Department of Psychology and Volen National Center for Complex Systems, Brandeis University, Waltham, MA 02453, USA

Corresponding author: Sven Mattys (sven.mattys@york.ac.uk)

Abstract

Research on “cognitive listening” has grown exponentially in recent years. Lacking, however, is a conceptual framework to organize the abundance of data coming from the hearing, cognitive, and linguistic sciences. We offer the Data-Resource-Language (DRL) framework that draws from the notions of data-limited processes, resource-limited processes, and effort as a roadmap to understanding the interaction between auditory sensitivity, cognitive resources, and linguistic knowledge during speech perception, especially in adverse conditions. The DRL framework explains how these three sets of abilities predict performance and resource engagement as a function of signal quality. It also provides a platform for characterizing similarities and differences in how normal-hearing, impaired-hearing, and non-native listeners process speech in challenging conditions.

Keywords: speech perception; effort; masking; cognitive resources; audiology

Listing as a cognitive activity

We live in a world of noise, whether it is the sound of traffic or competing speech in a crowded restaurant. Some listeners can transition seamlessly between these environments unencumbered by the noise and distraction. Others may struggle with even moderate acoustic challenges, finding them effortful and hard to overcome.

Attempts have been made to characterize the contribution of listener characteristics to speech perception in challenging conditions. For instance, in the Ease of Language Understanding model (ELU [1-4]), **working memory** capacity is thought to support degraded speech perception through a linking process between the signal and long-term linguistic memory. Similarly, the Framework for Understanding Effortful Listening (FUEL [5]) maps out the dynamic relations between task demands, motivation, and listening effort within a demand-capacity framework inspired by the Capacity Model of Attention [6]. The Model of Listening Engagement (MoLE [7]) adds a focus on subjective experiences such as enjoyment and boredom and emphasizes how these experiences interact with **cognitive resources** and executive control to determine whether an individual successfully engages with listening. These conceptualizations and others have benefitted from, and contributed to, fields known as Auditory Cognitive Science [8], Cognitive Hearing Science [9,10], and Cognitive Audiology [11]. Common to these fields is the assumption that cognition "kicks in" [12] when listening conditions are challenging, an approach encapsulated by the term **cognitive listening**. Here, we take a broad definition of cognition as a set of mental operations that include the sub-components of working memory (short-term phonological storage and executive control) and attention control (selective attention and inhibition).

Despite advances made by the above propositions, both real-world experience and empirical data indicate that, when **signal degradation** is severe, there may not be enough acoustic information that "gets through" to rescue comprehension no matter how much cognitive resource is applied to the task [13-15]. Likewise, cognition may not be substantially engaged when the signal is minimally degraded and performance is high [3,16]. These observations underscore the need for a framework that both captures the operational parameters of cognitive engagement across the continuum of signal quality and offers alternative contributors to performance in conditions where cognitive abilities only have a secondary role.

Some 50 years ago, Norman and Bobrow [17] introduced two terms that are relevant to this objective and are key components of our proposal. They used the term **resource-limited process** to describe conditions in which the application of additional cognitive resources can bring improvement in performance. By contrast, the term **data-limited process** was used to describe conditions in which the input is so degraded that no amount of additional cognitive resource can improve performance. Although speech perception was not the focus of Norman and Bobrow's concern at the time, their terminology can help us formalize the types of challenges listeners encounter on a daily basis, such as acoustic masking, accented speech, and spectrally degraded speech experienced through a cochlear implant.

Alongside Norman and Bobrow's terms, we draw upon the notions of resource availability (cognitive abilities that an individual possesses in a finite amount) and resource engagement (the extent to which those abilities are allocated to a task) to offer a framework aimed at characterizing the relationship between perceptual abilities, cognitive resources, and linguistic knowledge across the full range of signal quality, from severely degraded to intact—a continuum also often referred to as listening demand. The notions of resource availability and engagement are essential to understand how cognitive processes interact with listening demands (Box 1).

The goals of the framework are to (1) identify the listener-specific abilities (perceptual, cognitive, and linguistic) that best predict speech understanding, (2) specify the range of signal quality in which each of these abilities is most likely to predict performance, and (3) differentiate between conditions where resource engagement is likely to be associated with improved performance and conditions where it is not.

Relevance of the data-limited/resource-limited framework for speech perception research

A theory of speech processing in natural listening conditions must provide an account of how perception, cognition, and linguistic knowledge interact to achieve an observed level of performance. To date, no existing theory offers an account that fully integrates all three components. Norman and Bobrow's framework is well suited to describing how two of those components, perception and cognition, constrain performance as a function of signal quality. Each specific degree of signal degradation can be represented by a unique proportion of data-limited and resource-limited processes. When speech is severely degraded, recognition performance is unlikely to exceed our auditory system's ability to decode the impoverished signal. Dedicating more cognitive resources to the task is unlikely to yield further improvement. In its purest form, this data-limited scenario can be illustrated by the relative robustness of pure-tone audiometry tests; below-threshold tones are unlikely to be detected whether or not the listener applies additional cognitive resources to the task—the signal (data) is simply not strong enough to benefit from enhanced attention or memory processes [18].

When signal quality is moderate, allocating additional cognitive resources to the task can improve performance, at least to the extent that those resources are available. These are cases in which there is enough sensory data for cognitive processes such as working memory and attentional focus to play a supporting role in integrating and interpreting the degraded speech fragments [19]. This resource-limited scenario can be illustrated by the observation that individuals with good working memory capacity are generally better at coping with moderate noise than individuals with poorer working memory capacity [20,21].

The effect of signal quality on the trade-off between data-limited and resource-limited processes is supported by empirical evidence. For example, data show that working memory capacity positively correlates with a listener's ability to track two simultaneous talkers when the talkers are spectrally or spatially separated (resource-limited), but not when they are spectrally and spatially overlapped (data-limited) [13]. Likewise, older adults' hearing acuity better predicts lowpass-filtered (data-limited) than unfiltered (resource-limited) speech perception, whereas working memory capacity shows the opposite pattern [14]. The contribution of cognitive processes to degraded speech perception can also be tested by manipulating the amount of processing time made available to the listeners during the task. For instance, when speech is moderately degraded (resource-limited) through noise-vocoding, inserting silent pauses at linguistically salient points within rapid (compressed) speech improves recall performance, presumably because the additional processing time allows listeners to use cognitive processes to "catch up" with the impoverished input. However, when the speech is heavily degraded (data-limited), inserting silent pauses has a smaller impact on recall [15].

The distinction between data-limited and resource-limited regions along the signal-quality continuum is important because it establishes a symbolic boundary between what is reducible to auditory perception and what can be genuinely construed as cognitive listening [22]. The above evidence shows that the data-limited/resource-limited framework can explain a wide range of listening behaviors when signal quality varies from severely to moderately degraded. However, this two-component distinction is silent about the drivers of performance at the upper end of signal quality, where speech is intelligible and performance, while still variable, is high [23-26]. In that region, data converge in showing a decreased contribution of cognitive abilities to performance [3,16]. This pattern is also evident in individuals with mild cognitive impairment, who, relative to typically developing listeners, are less impacted by their impaired cognition when processing intelligible speech than when processing speech in noise [27], suggesting a smaller contribution of cognition to intact than degraded speech processing. However, the factors that *do* predict performance variability in high-intelligibility conditions remain unspecified.

A tripartite Data-limited, Resource-limited, Language-limited (DRL) framework

Our claim is that, at the high end of the signal-quality continuum, listening performance is primarily determined by individual differences in linguistic abilities—it is

language-limited. On the assumption that substantial variability in speech perception performance persists even in favorable signal-quality conditions [15,28,29], we predict that performance will reach an asymptote at a higher level of accuracy for individuals with more complex and nuanced knowledge of the language. While differences in auditory acuity and cognitive abilities may still afford some explanatory power, their role in accounting for performance differences would be smaller than that of individual differences in, e.g., vocabulary knowledge and syntactic fluency. Linguistic factors are predicted to be particularly significant drivers of performance in naturalistic tasks (e.g., narrative comprehension), as performance asymptotes on such tasks are likely to occur well below ceiling level. In these conditions, which represent the majority of real-life communication, more complex processes are required, giving rise to a critical source of individual differences [30].

The link between individual differences in linguistic abilities and speech comprehension has been documented over decades of language-processing research [31,32]. However, this literature is rarely used to inform speech-in-noise research because the evidence is derived primarily from experiments using speech heard in quiet. Conversely, the speech-in-noise literature has deliberately focused on non-linguistic determinants of performance such as hearing sensitivity and cognitive capacities [33], assuming a linguistic level-field for convenience (at least within a native-language population). We are not claiming that linguistic abilities can account for performance differences only when the signal is highly intelligible. Indeed, we know that vocabulary knowledge [34,35] and semantic context [36-37] play a critical role in the resource-limited region, as the use of contextual cues from linguistic information draws upon predictive processes that require cognitive resources [38]. Our claim, rather, is that linguistic abilities are comparatively better predictors of performance under optimal listening conditions, with non-native language users serving as a prominent example (see later section). In very poor signal-quality conditions, listeners might very well attempt to recruit linguistic abilities, but the DRL predicts that such abilities would not have enough signal to work with to contribute to performance in a significant way.

Likewise, we are not claiming that cognitive abilities do not play any role in intelligible speech processing, since a link between working memory and syntactic parsing has been reported [39-41], but rather that cognitive abilities play a comparatively greater role when the signal is moderately degraded.

[Figure 1]

Figure 1A illustrates our tripartite conceptualization of the individual drivers of speech perception. A critical aspect of this framework is the identification of prominent processing regions (perceptual, cognitive, linguistic) at different levels of signal quality (low, moderate, high). In the DRL framework, the signal-quality dimension represents an objective, quantifiable stimulus characteristic (e.g., sound level, signal-to-noise ratio [SNR], number of vocoded bands), whereas the performance function is an approximation of successful perception averaged across participants. The absolute values and boundaries between processing regions are not specified, because they are likely to depend on the nature of the degradation and listener characteristics (see Box 2 and later sections for examples). However, they are arranged in a predictable order as signal quality changes from low to high.

This fixed order, which is the defining feature of the DRL framework, allows us to generate testable predictions about the best-fitting constellation of performance predictors as a function of speech degradation. In low signal-quality conditions, the DRL posits that an individual's results on a battery of basic auditory perception tests (e.g., pure-tone audiometry, gap detection, temporal discrimination) should better predict listening performance than their results on cognitive tests (e.g., working memory, attention, processing speed) or linguistic tests (e.g., vocabulary, syntactic fluency). In moderate signal-quality conditions, cognitive tests should be the dominant predictors and, in high signal-quality conditions, linguistic tests should be the dominant predictors.

An important consideration for assessing the above predictions is the choice of appropriate tests for each set of abilities [20,42-44]. Since psychometric tests often load onto more than just the dimension they are designed to measure, structural equation modelling could be used to extract a latent variable for each set of tests, and those latent variables, rather than the test scores themselves, would be used as predictors of performance for various points on the signal-quality continuum. This procedure, which has been used to compare auditory and cognitive predictors of speech recognition performance in various age groups [45-47], would maximize distinctiveness between predictors and guard against the challenges of reducing broad constructs (perception, cognition, language) to the narrower scope of individual tests.

Operating essentially as a dynamic sliding scale between three dominant processing modes imposed on the listeners by changes in signal quality, the framework can also be used to test whether meaningful discontinuities on the speech-quality continuum (e.g., a positive vs negative SNR) might coincide with tipping points between processing modes and how those tipping points shift as a function of stimulus characteristics (see examples in Box 2). Likewise, the DRL can be used to re-interpret existing data on hearing-impaired and non-native listeners, as described in later sections, and serve as a catalyst for novel questions (see, e.g., Outstanding Questions).

What sets the DRL apart from other models is fourfold: (1) DRL considers the combined influence of individual differences in perceptual, cognitive, and linguistic abilities on performance, (2) DRL regards the degree of signal degradation as a primary factor shaping how such individual differences play out, (3) While other models emphasize recognition performance (e.g., ELU) or resource engagement (e.g., FUEL, MoLE), DRL makes predictions about how *both* recognition performance and resource engagement vary as a function of signal quality, and (4) DRL makes specific predictions about the conditions in which resource engagement is mostly likely to affect speech recognition performance (as described below).

Resource engagement within the DRL framework

An inherent assumption of the DRL framework is that operations performed on the sensory input must compete within the bounds of a limited capacity system. A degraded input requires a greater draw on the available capacity (or resources) than would be needed if the input was clear, resulting in fewer resources being available to conduct higher-level operations on that input or on a concurrent task [5,6,48-50]. The postulated relation between resource engagement and signal clarity is depicted in Figure 1A (grey curve, right Y-axis). As illustrated, resource engagement and intelligibility do not covary in a linear fashion [16,51,52] but, rather, follow an inverted U-shaped curve [16,53,54]. The DRL formalizes this relation by predicting that the link between resource engagement and performance should be strongest in the resource-limited region, whereas performance would be less dependent on cognitive resources when a task is data-limited or language-limited.

The **task-evoked pupil response (TEPR)** has been proposed as a near real-time physiological index of resource engagement that can be measured independently of task performance [55]. Note that TEPR may not map directly onto everyday listening difficulties, such as perceived effort or fatigue [56,57] and care should be taken when interpreting individual differences using this method [58]. However, TEPR is now widely used as a laboratory index of resource engagement during listening owing to its sensitivity to task demand and capacity limits [59,60], especially when combined with other measures [61].

Resource engagement in the resource-limited region

The DRL framework predicts that resource engagement is likely to be highest at the center of the resource-limited region. This claim is supported by extant literature that reveals a peak in TEPR at approximately 50% intelligibility [16,58,60], indicating application of resources as listeners attempt to process a signal within a moderately degraded range. The resource-limited region is also the one in which motivation is most likely to positively translate into increased resource engagement and better performance [5,62]. In other words, this is a region where effort invested ‘pays off’. A listener’s motivation to understand a severely degraded signal may be high at first but is likely to decline if no improvement in performance is achieved. Therefore, interventions seeking to modulate resource engagement via motivation (e.g., reward) should be most effective when applied in conditions of moderate signal quality compared with low or high signal quality.

Resource engagement in the language-limited region

Research on the contribution of cognitive resources to the perception of intelligible speech shows mixed results. One the one hand, TEPR and TEPR variability are shown to decrease in favorable SNRs [53,54]. Similarly, TEPR no longer co-varies with SNR once performance becomes asymptotic [16]. On the other hand, there is some evidence for continued changes in resource engagement in that region. For example, TEPRs continue to decrease as signal clarity improves even when analyses are restricted to trials with 100% accuracy [63], possibly reflecting the reduced cost of revisiting and repairing the input as signal quality improves [64,65]. Evidence for reduced resource engagement with increasing clarity of supra-threshold speech has also been obtained using the dual-task paradigm [66,67].

Despite the evidence of a link between TEPR and signal quality within the region of asymptotic performance, it is important to note that TEPR changes in that region are relatively small compared to changes in the resource-limited region [63]. This suggests that, although changes in signal quality within the language-limited region may affect resource engagement, they do so to a lesser extent than in the resource-limited region, as postulated by the DRL framework. TEPR changes within the language-limited region are also likely to reflect differences in the ease with which linguistic and discourse processes are completed at a supra-threshold level [68,69].

Resource engagement in the data-limited region

When the signal quality is so poor that no amount of effort can restore comprehension, studies have correspondingly shown smaller TEPRs than in the resource-limited region [53,54]. The low predictive power of cognitive abilities in that region can reflect either a listener's unsuccessful attempt to use their cognitive resources or disengagement from a task perceived to be too hard to be worth the effort. The latter option is in accord with the claim made within the field of neuroeconomics that individuals will engage effort to perform a task only if they believe that this effort is likely to yield some degree of success or "return on investment" [33]. Therefore, research examining motivational factors in speech perception [58,70-75] offers a promising avenue for understanding the relationship between resource engagement and task performance across the signal-quality continuum [76]. For example, listeners who report giving up (low motivation) when the signal is highly degraded have smaller pupil dilations than listeners who report not giving up (high motivation) [77]. The DRL provides a basis for contextualizing the debate on whether the link between motivation, resource engagement, and performance applies across all levels of task difficulty or primarily in moderately difficult listening conditions [5,71,78].

Applying the DRL framework to specific populations

Hearing-impaired listeners

It is well documented that individuals with hearing impairment struggle with speech in noise [36]. In this population, poorer access to data due to reduced hearing sensitivity means that the data-limited region should extend to the right along the signal-quality continuum (Figure 1B, upper panel). In turn, DRL predicts greater involvement of cognitive and linguistic abilities at higher (moderate to high) levels of signal quality, as depicted by a rightward shift of those two regions. Cognitive resources have indeed been shown to be engaged even in favorable signal-quality conditions in individuals with hearing loss [53], representing a rightward shift of the tipping point between the data-limited and resource-limited regions. A consequence of this shift is that only in highly clear signal-quality conditions can hearing-impaired listeners successfully operate within the language-limited region.

The above predictions imply that listening conditions that would be resource-limited for normal-hearing individuals may be data-limited for hearing-impaired individuals [79] and listening conditions that would be language-limited for normal-hearing individuals may be resource-limited for hearing-impaired individuals. This could explain why working memory capacity is often found to be a better predictor of speech perception in noise among hearing-impaired than normal-hearing individuals [43]. Given that highly degraded listening conditions are rarely encountered by older adults with hearing loss in everyday life [80], these listeners may be confined to a more permanent state of operating within an effortful, resource-limited region. This situation may be contrasted with listeners with normal hearing, who can afford to be more sparing in their use of cognitive resources in service of higher-level linguistic processing of the input [81-84]. It is thus not surprising to hear reports of exhaustion and mental fatigue by individuals with impaired hearing. For them, the everyday communicative world is one of sustained resource-intensive listening [5,85,86].

For hearing-impaired listeners who use hearing aids, the DRL proposes that the resulting boost in signal quality could propel listeners from operating within a data-limited region, where cognitive support is ineffective, to operating within a resource-limited region, where cognitive resources can contribute meaningfully to speech recognition performance. This prediction is supported by data showing a significant contribution of cognitive abilities to aided speech understanding [47] compared to a dominance of hearing factors, such as hearing thresholds, in unaided speech understanding [46,87]. In other words, by promoting effective mapping between the improved sensory input and lexical representations through **learning**, aided hearing could shift the boundary between those two regions to the left, and hence, decrease instances of unrewarded effort and subsequent fatigue. Indeed, compared to unaided listeners, hearing-impaired listeners who are provided with a hearing aid show improvement in cognitive function, especially working memory capacity [88]. By implication, we predict stronger contributions of cognitive support to speech recognition for hearing-impaired listeners who successfully acclimate to a hearing device.

Non-native listeners

A special case of language-limited processes must be considered when investigating the challenges experienced by individuals with non-native knowledge of the language. Non-native listeners often show greater vulnerability to signal degradation, the so-called non-native speech-in-noise disadvantage [89-92]. In this group, incomplete linguistic knowledge makes it hard to successfully fill in the gaps created by signal degradation using linguistic top-down knowledge, a process often seen as a hallmark of native listening [93]. This is particularly problematic for older non-native adults where reduced working memory and age-related hearing impairment further challenge cognition and perception [94,95]. In the DRL framework, we argue that the incomplete linguistic knowledge that characterizes non-native listeners results in a language-limited region that extends leftward (Figure 1B, lower panel). Thus, for non-native listeners, individual differences in linguistic knowledge (of the non-native language) should be a stronger predictor of performance across a much broader range of signal quality conditions than for native listeners. The resource-limited region, too, will likely cover a broader range, as cognitive abilities such as working memory and attentional control are recruited to compensate for the lack of linguistic support. This claim is supported by evidence for widespread cognitive resource engagement in high-intelligibility conditions in non-native listeners [96-99]. A potential consequence of the greater overlap between the resource-limited region and the language-limited regions could be more cases where resource engagement fails to translate into improved performance [100].

Given the high degree of variability in language proficiency among non-native listeners, there is ample scope to evaluate individual differences within this population. Based on the region shifts proposed above, we predict that non-native speakers with superior linguistic knowledge will show a narrower language-limited region than speakers with lower proficiency. Thus, there will be a reduced overlap between the resource-limited and language-limited regions relative to their less-proficient counterparts. Indeed, analogous to the impact of wearing a hearing aid on cognitive enhancement in hearing-impaired listeners, a prediction of the DRL is that progressing from lower to higher proficiency in the course of learning a second language would amount to narrowing the relative involvement of linguistic processes and engaging cognitive resources where it is most impactful (i.e., in moderate signal-quality conditions). These predictions, as well as those pertaining to hearing-impaired listeners, could be tested using the latent-variable approach described earlier, with contrasted constellations of predictors expected to be found at different levels of signal quality for normal-hearing, impaired hearing, and non-native listeners.

Concluding remarks

The field of hearing science can benefit from reconceptualizing cognitive listening by embracing and developing the notions of data-limited and resource-limited processes postulated by Norman and Bobrow [17] within an account that considers a full range of signal quality. We offer the DRL framework, which expands upon those notions and brings linguistic abilities to the fore, focusing on current theoretical challenges in speech perception research. The framework partitions the listening experience into three zones of preferential processes (perceptual, cognitive, linguistic) as a function of signal quality (low, moderate, high). In doing so, it provides testable predictions about performance and resource engagement that can be used to reinterpret existing data, generate hypotheses, and ask novel questions (Table 1).

The DRL framework also presents opportunities for further exploration in clinical practice and training. From the perspective of the listener, the DRL emphasizes learning as a means of shifting boundaries between processing regions to strategically allocate cognitive resources whenever such a shift is likely to pay off. If a listener is made aware through training that a process is data-limited, that person may learn to restrain from investing resources, thereby limiting the cumulative toll of sustained effortful listening and re-directing resources to other activities. Conversely, if a listener is aware that the task is resource-limited, that person may increase their engagement of cognitive resources, resulting in greater comprehension of the spoken content. The DRL conceptualization also presents a mechanism (narrowing reliance on the language-limited region) by which non-native listeners can improve speech perception through honing their linguistic abilities.

Box 1. Resource availability and engagement within the DRL framework

The dual concepts of *effort* and *resources*, central to Norman and Bobrow's formulation [17], are captured by the definition of effort as *the intentional allocation of cognitive resources to overcome obstacles for successful completion of a listening task* [5]. In this regard, we note that effort is not a unitary concept, but that it encompasses both the subjective sense of effort, tied closely with an individual's judgement of the difficulty of accomplishing a task, and the objective sense of effort as measured by, e.g., the size of the Task-Evoked Pupil Response (TEPR). Although the two senses of effort are closely aligned [7,101], it is the objective sense that is intended by most researchers, and it is this sense that we use in the DRL framework.

We consider *resources* in terms of working memory, processing speed, and executive function [102], with these functions collectively defining *resource availability*. As often argued [6,48,49], resource availability (i.e., capacity) is limited, such that allocation of resources to one demanding task or set of operations leaves fewer resources available for the simultaneous conduct of other demanding tasks or operations. This principle underlies Kahneman's articulation of his general resource model of attention [6] upon which Norman and Bobrow's definitions are implicitly based. It is also a principle adopted by the DRL framework. When the DRL framework references increased resource engagement within the resource-limited region of task performance, we thus postulate that resource engagement is constrained by both the individual's resource availability *per se* and the individual's decision as to the necessary allocation of these limited resources to the task. Factors affecting this decision include the importance of the task to the individual, the individual's motivation to perform the task as well as they can (perhaps based on a reward), and the belief that the task is manageable and, therefore, engaging additional resources is likely to bring a successful return on the investment [71,103].

While the DRL considers perceptual, cognitive, and linguistic abilities as separable sources of individual differences, linguistic challenges may draw on domain-general cognitive resources when speech is particularly complex [68]. That is, while resource capacity and linguistic ability are fixed, cognitive resources may be dynamically recruited to support linguistic abilities.

Box 2. Signal degradation and adaptation within the DRL framework

For the sake of simplicity, the DRL framework treats signal quality as a unitary concept. However, signal degradation can take different forms [104], and these may affect the relative contributions of the DRL processing regions. For instance, compared to the mostly energetic nature of broad-band steady-state maskers, fluctuating noise often affords glimpses of the target signal. Successfully exploiting such glimpses has been shown to involve attention control and working memory [10,105]. Likewise, maskers with an informational content are likely to engage both attention control and linguistic abilities [106]. Thus, boundaries between DRL regions are likely to be modulated by the type of signal degradation in ways that can be tested empirically.

The nature of the degradation also has implications for how perceptual processes and cognitive resources interact during learning. Speech degraded intrinsically (e.g., accented, noise-vocoded, disordered speech) often provides systematic distortions that are learnable through knowledge-driven perceptual adaptation and acoustic-to-phonetic remapping [19,107-110]. In contrast, speech degraded extrinsically, by a competing sound source (e.g., noise, competing talkers), involves more random distortions and is therefore less readily learnable [108,111]. Distinct learning curves have been found for the two types of degradation [112,113]. Differences in learnability should affect the DRL in predictable ways. For example, intrinsic degradation should have a greater impact than extrinsic degradation on regions shifts during learning, with the data-limited region shrinking more markedly in the course of learning speech with intrinsic than extrinsic degradation.

Less clear is whether exposure to intrinsic degradation should lead to a greater role for the resource-limited region. If perceptual adaptation requires only limited involvement of attentional processes [114], learning should be relatively impervious to individual differences in attention control. Therefore, adaptation to intrinsic degradation should be relatively independent of individual differences in attention control within the resource-limited region. However, if adaptation is strongly underpinned by attentional processes [115], individual differences in attention control should influence how well listeners adapt to intrinsic degradation, hence increasing the relative contribution of the resource-limited region. Thus, whether the involvement of the resource-limited region changes during exposure to intrinsic degradation depends on theoretical assumptions about the mechanisms underlying perceptual adaptation. This question has clinical implications as well. Since the signal produced by a cochlear implant constitutes a paradigmatic case of intrinsic degradation, understanding the role of attention (and cognition in general) is crucial to establish the possible contribution of cognitive resources to auditory plasticity and reorganization in cochlear-implanted users.

Glossary

Working Memory: a limited-capacity system that temporarily holds and manipulates auditory and linguistic information and plays a crucial role in active listening, learning, and reasoning.

Cognitive resources: cognitive fundamentals such as working memory, processing speed, and executive functions that an individual possesses in a finite amount (resource availability) and that can be allocated to a listening task. Controlling resource allocation is volitional and usually effortful.

Cognitive listening: the intentional and sometimes effortful process of attending to, interpreting, and comprehending spoken language, particularly in challenging listening environments. Listening is said to be cognitive because it involves attention and working memory processes alongside purely auditory or linguistic processes.

Signal degradation: any distortion of the speech signal that reduces its intelligibility. Signal degradation can be intrinsic (e.g., accented, disordered, filtered speech, speech heard through a cochlear implant) or extrinsic (e.g., background noise, competing talkers). It can lead to evenly distributed degradation across the signal (e.g., broadband steady-state noise, high-N babble noise) or irregular degradation (e.g., amplitude-modulated noise, a single competing talker).

Resource-limited process: operation that is supported and constrained by the availability of cognitive resources such as attention and working memory. In this processing region, engaging cognitive resources to perform a listening task is usually effortful but leads to improved performance.

Data-limited process: operation that is constrained by the quality or quantity of the input data available (and perceptual processes) rather than by the availability or engagement of cognitive resources. Engaging cognitive resources in this processing region is unlikely to improve listening performance.

Language-limited process: operation that is supported primarily by linguistic abilities rather than perceptual and cognitive processes. In this processing region, listening performance is predicated primarily based on vocabulary knowledge, syntactic fluency, and narrative comprehension.

Task-Evoked Pupil Response (TEPR): changes in pupil size from baseline during auditory stimulus (e.g., speech) processing thought to reflect engagement of cognitive resources. This metric is captured using an eye-tracking technique called pupillometry.

Learning: process of acquiring and modifying knowledge through experience resulting in enduring changes in mental representations. Within the DRL, short-term and long-term adaptation to signal degradation leading to improved listening performance, region boundary shifts, and perceptual recalibration relevant to hearing-aid and cochlear-implant tuning. The extent of these changes may depend on the type of degradation, as described in Box 2.

References

[1] Rönnberg, J. (2003) Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: A framework and a model. *Int J Audiol* 42, S68-S76

[2] Rönnberg, J. et al. (2008) Cognition counts: a working memory system for ease of language understanding (ELU). *Int J Audiol* 47(Suppl. 2), S99-S105

[3] Rönnberg, J. et al. (2013) The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Front Syst Neurosci* 7, 1-17.

[4] Rönnberg, J. et al. (2022) The cognitive hearing science perspective on perceiving, understanding, and remembering language: The ELU model. *Front Psychol* 13, 967260

[5] Pichora-Fuller, M.K. et al. (2016) Hearing impairment and cognitive energy: The framework for Understanding effortful listening (FUEL). *Ear Hear* Suppl 1, 5s-27s

[6] Kahneman, D. (1973) *Attention and effort*, Englewood Cliffs, NJ: Prentice-Hall.

[7] Herrmann, B. and Johnsrude, I.S. (2020) A model of listening engagement (MoLE). *Hear Res* 397, 108016

[8] Holt, L.L. and Lotto, A.J. (2008) Speech perception within an auditory cognitive science framework. *Curr Dir Psychol Sci* 17, 42-46

[9] Arlinger, S. et al. (2009) The emergence of cognitive hearing science. *Scand J Psychol* 50, 371-384

[10] Gatehouse, S. et al. (2003) Benefits from hearing aids in relation to the interaction between the user and the environment. *Int J Audiol* 42, S77-S85

[11] Fabry, D. (2011) Jim Jerger by the letters. *Audiol Today* (Jan/Feb) 23, 18-23.

[12] Rönnberg, J. et al. (2010) When cognition kicks in: Working memory and speech understanding in noise. *Noise and Health* 12, 263-269

[13] Knight, S. et al. (2023) Conceptualising acoustic and cognitive contributions to divided-attention listening within a data-limit versus resource-limit framework. *J Mem Lang* 131, 104427

[14] Janse, E. and Andringa, S.J. (2021) The roles of cognitive abilities and hearing acuity in older adults' recognition of words taken from fast and spectrally reduced speech. *Appl Psycholinguist* 42, 763-790

[15] O'Leary, R.M. et al. (2023) Strategic pauses relieve listeners from the effort of listening to fast speech: Data limited and resource limited processes in narrative recall by adult users of cochlear implants. *Trends Hear* 27, 1-22

[16] Wendt, D. et al. (2018) Toward a more comprehensive understanding of the impact of masker type and signal-to-noise ratio on the pupillary response while performing a speech-in-noise test. *Hear Res* 369, 67-78

[17] Norman, D.A. and Bobrow, D.G. (1975) On data-limited and recourse-limited processes. *Cogn Psychol* 7, 44-64

[18] Heinrich, A. et al. (2020) Effects of cognitive load on pure-tone audiometry thresholds in younger and older adults. *Ear Hear* 41, 907-917

[19] Lansford, K.L. et al. (2023) Cognitive predictors of perception and adaptation to dysarthric speech in young adult listeners. *J Speech Lang Hear Res* 66, 30-47

[20] Dryden, A. et al. (2017) The association between cognitive performance and speech-in-noise perception for adult listeners: A systematic literature review and meta-analysis. *Trends Hear* 21, 1-21

[21] Souza, P. and Arehart, K. (2015) Robust relationship between reading span and speech recognition in noise. *Int J Audiol* 54, 705-713

[22] Johnsrude, I.S. and Rodd, J.M. (2016) Factors that increase processing demands when listening to speech. In *Neurobiology of Language* (Hickok, G. and Small, S.L., eds), pp. 491-502, Academic Press

[23] Kantowityz, B.H. and Knight, J.L. (1976) On experimenter limited processes*. Psychol Rev* 83, 502-507

[24] Norman, D.A. and Bobrow, D.G. (1976) On the analysis of performance operating characteristics. *Psychol Rev* 83, 508-510

[25] Andringa, S. et al. (2012) Determinants of success in native and non-native listening comprehension: An individual differences approach. *Lang Learn* 62, 49-78

[26] Kong, E.J. and Edwards, J. (2016) Individual differences in categorical perception of speech: Cue weighting and executive function. *J Phon* 59, 40-57

[27] Lee, S.J. et al. (2018) Association between frontal-executive dysfunction and speech-in-noise perception deficits in mild cognitive impairment. *J Clin Neuro* 14, 513-522

[28] DeCaro, R. et al. (2016) The two sides of sensory-cognitive interactions: Effects of age, hearing acuity, and working memory span on sentence comprehension. *Front Psychol* 7, 236

[29] Hansen, T.A. et al. (2023) Self-pacing ameliorates recall deficit when listening to vocoded discourse: A cochlear implant simulation. *Front Psychol* 14, 1225752

[30] Nijhof, A.D. and Willems, R.M. (2015) Simulating fiction: Individual differences in literature comprehension revealed with fMRI. *PLoS One* 10, e0116492

[31] Kintsch, W. (1988) The role of knowledge in discourse comprehension: a construction-integration model. *Psychol Rev* 95, 163-182

[32] Marslen-Wilson, W.D. and Tyler, L.K. (2007) Morphology, language and the brain: the decompositional substrate for language comprehension. *Philos Trans R Soc B* 362, 823-836

[33] Eckert, M.A. et al. (2024) Executive function associations with audibility-adjusted speech perception in noise. *J Speech Lang Hear Res* 67, 4811-4828

[34] Kaandorp, M.W. et al. (2016) The influence of lexical-access ability and vocabulary knowledge on measures of speech recognition in noise. *Int J Audiol* 55, 157-167

[35] Tamati, T.N. et al. (2022) Lexical effects on the perceived clarity of noise-vocoded speech in younger and older listeners. *Front Psychol* 13, 837644

[36] Pichora-Fuller, M.K. et al. (1995) How young and old adults listen to and remember speech in noise. *J Acoust Soc Am* 97, 593-608

[37] Obleser, J. and Kotz, S.A. (2010) Expectancy constraints in degraded speech modulate the language comprehension network. *Cereb Cortex* 20, 633-640

[38] Benichov, J. et al. (2012) Word recognition within a linguistic context: Effects of age, hearing acuity, verbal ability, and cognitive function. *Ear Hear* 33, 250-256

[39] Caplan, D. and Waters, G.S. (1999) Verbal working memory and sentence comprehension. *Behav Brain Sci* 22, 77-94

[40] Just, M.A. and Carpenter, P.A. (1992) A capacity theory of comprehension: Individual differences in working memory. *Psychol Rev* 99, 122-149

[41] King, J. and Just, M.A. (1991) Individual differences in syntactic processing: The role of working memory. *J Mem Lang* 30, 580-602

[42] Akeroyd, M. A. (2008) Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int J Audiol* 47, S53-S71

[43] Füllgrabe, C. and Rosen, S. (2016) On the (un) importance of working memory in speech-in-noise processing for listeners with normal hearing thresholds. *Front Psychol* 7, 1268

[44] Heinrich, A. et al. (2015) The relationship of speech intelligibility with hearing sensitivity, cognition, and perceived hearing difficulties varies for different speech perception tests. *Front Psychol* 6, 782

[45] Benzaquén, E. et al. (2025) Auditory-cognitive determinants of speech-in-noise perception: structural equation modelling of a large sample. *Sci Rep*

[46] Humes, L.E. and Dubno, J.R. (2010) Factors affecting speech understanding in older adults. In *The Aging Auditory System* (Gordon-Salant, S, Frisina, R.D., Popper, A.N., and Fay, R.R., eds), pp. 211-258, Springer

[47] Humes, L.E. et al. (2013). Auditory and cognitive factors underlying individual differences in aided speech-understanding among older adults. *Front Syst Neurosci* 7, 55

[48] Moray, N. (1967) Where is capacity limited? A survey and a model. *Acta Psychol* 27, 84-92

[49] Treisman, A.M. (1969) Strategies and models of selective attention. *Psychol Rev* 76, 282-299

[50] Wingfield, A. et al. (2005) Hearing loss in older adulthood: What it is and how it interacts with cognitive performance. *Curr Dir Psychol Sci* 14, 144-148

[51] Ayasse, N.D. and Wingfield, A. (2018) A tipping point in listening effort: Effects of linguistic complexity and age-related hearing loss on sentence comprehension. *Trends Hear* 22, 1-14

[52] Winn, M.B. and Teece, K.H. (2021) Listening effort is not the same as speech intelligibility score. *Trends Hear* 25, 1-26

[53] Ohlenforst, B. et al. (2017) Impact of stimulus-related factors and hearing impairment on listening effort as indicated by pupil dilation. *Hear Res* 351, 68-79

[54] Ohlenforst, B. et al. (2018) Impact of SNR, masker type and noise reduction processing on sentence recognition performance and listening effort as indicated by the pupil dilation response. *Hear Res* 365, 90-99

[55] Kuchinsky, S.E. and DeRoy Milvae, K. (2024) Pupillometry studies of listening effort: Implications for clinical audiology. In *Modern Pupillometry: Cognition, Neuroscience, and Practical Applications* (Papesh, M.H. and Goldinger, S.D., eds), pp. 229-258, Springer Nature

[56] Strand, J.F. et al. (2018) Measuring listening effort: Convergent validity, sensitivity, and links with cognitive and personality measures. *J Speech Lang Hear Res* 61, 1463-1486

[57] Alhanbali, S. et al. (2019). Measures of listening effort are multidimensional. *Ear Hear* 40, 1084-1097

[58] Winn, M.B. et al., (2018) Best practices and advice for using pupillometry to measure listening effort: An introduction for those who want to get started. *Trends Hear* 22, 1-32

[59] Reilly, J., Zuckerman, B., & Kelly, A.E. (2024) A primer on design and data analysis for cognitive pupillometry. In *Modern Pupillometry: Cognition, Neuroscience, and Practical Applications* (Papesh, M.H. and Goldinger, S.D., eds), pp. 401-430, Springer Nature

[60] Zekveld, A.A. et al. (2018) The pupil dilation response to auditory stimuli: Current state of knowledge. *Trends Hear* 22, 1-25

[61] Richter, M. et al. (2023) Combining multiple psychophysiological measures of listening effort: Challenges and recommendations. *Semin Hear* 44, 95-105

[62] Carolan, P.J. et al. (2022) Quantifying the effects of motivation on listening effort: A systematic review and meta-analysis. *Trends Hear* 26, 1-32

[63] Winn, M.B. et al. (2015) The impact of auditory spectral resolution on listening effort revealed by pupil dilation. *Ear Hear* 36, e153-e165

[64] Winn, M.B. (2024) The effort of repairing a misperceived word can impair perception of following words, especially for listeners with cochlear implants. *Ear Hear* 45, 1527-1541

[65] Winn, M.B. and Teece, K.H. (2022) Effortful listening despite correct responses: The cost of mental repair in sentence recognition by listeners with cochlear implants. *J Speech Lang Hear Res* 65, 3966-3980

[66] Sarampalis, A. (2009) Objective measures of listening effort: Effects of background noise and noise reduction. *J Speech Lang Hear Res* 52, 1230-1240

[67] Pals, C. et al. (2013) Listening effort with cochlear implant simulations. *J Speech Lang Hear Res* 56, 1075-1084

[68] Piquado, T. et al. (2010) Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology* 47, 560-569

[69] O'Leary, R.M. et al. (2023) Congruent prosody reduces cognitive effort in memory for spoken sentences: A pupillometric study with young and older adults. *Exp Aging Res*, 1-24

[70] McLaughlin, D.J. et al. (2021) Measuring the subjective cost of listening effort using a discounting task. *J Speech Lang Hear Res* 64, 337-347

[71] Richter, M. (2016) The moderating effect of success importance on the relationship between listening demand and listening effort. *Ear Hear* 37, 111S–117S

[72] Richter, M. et al. (2016) Three decades of research on motivational intensity theory: What we have learned about effort and what we still don't know. *Adv Motiv Sci* 3, 149-186

[73] Wu, Y.H. et al. (2016) Psychometric functions of dual-task paradigms for measuring listening effort. *Ear Hear* 37, 660-670

[74] Crawford, J.L. et al. (2021) Domain-general cognitive motivation: Evidence from economic decision-making. *CRPI* 6, 1-9

[75] Koelewijn, T. et al. (2018) The effect of reward on listening effort as reflected by the pupil dilation response. *Hear Res* 367, 106-112

[76] Kraus, F. et al. (2024) Neurophysiology of effortful listening: Decoupling motivational modulation from task demands. *J Neurosci* 44, 1-15

[77] Zekveld, A.A. et al. (2014) Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. *Psychophysiology* 51, 277-284

[78] Kruglanski, A.W. et al. (2012) The energetics of motivated cognition: A force-field analysis. *Psychol Rev* 119, 1-20

[79] Stewart, R. and Wingfield, A. (2009). Hearing loss and cognitive effort in older adults' report accuracy for verbal materials. *J Am Acad Audiol* 20, 147-154

[80] Wu, Y.H. et al. (2018) Characteristics of real-world signal to noise ratios and speech listening situations of older adults with mild to moderate hearing loss. *Ear Hear 39*, 293-304

[81] Kuchinsky, S.E. et al. (2013) Pupil size varies with word listening and response selection difficulty in older adults with hearing loss. *Psychophysiology* 50, 23-34

[82] Kuchinsky, S.E. et al. (2014) Speech-perception training for older adults with hearing loss impacts word recognition and effort. *Psychophysiology* 51, 1046-1057

[83] Kuchinsky, S.E. et al. (2016) Task-related vigilance during word recognition in noise for older adults with hearing loss. *Exp Aging Res* 42, 50-66

[84] Wang, Y. et al. (2018) Relations between self-reported daily-life fatigue, hearing status, and pupil dilation during a speech perception in noise task. *Ear Hear* 39, 573-582

[85] Davis, H. et al. (2021) Understanding listening-related fatigue: Perspectives of adults with hearing loss. *Int J Audiol* 60, 458-468

[86] Hornsby, B.W.Y. et al. (2021) Development and validation of the Vanderbilt Fatigue Scale for Adults (VFS-A). *Psychol Assess* 33, 777-788

[87] Humes, L.E. et al. (1994) Factors associated with individual differences in clinical measures of speech recognition among the elderly. *J Speech Langu Hear Res* 37, 465-474

[88] Karawani, H. et al. (2018) Restauration of sensory input may improve cognitive and neural function. *Neuropsychologia* 114*,* 203-213

[89] Cutler, A. et al. (2008) Consonant identification in noise by native and non-native listeners: Effects of local context. *J Acoust Soc Am* 124, 1264-1268

[90] Lecumberri, M.L. and Cooke, M. (2006) Effect of masker type on native and non-native consonant perception in noise. *J Acoust Soc Am* 119, 2445-2454

[91] Nábělek, A.K. and Donahue, A.M. (1984) Perception of consonants in reverberation by native and non-native listeners. *J Acoust Soc Am* 75, 632-634

[92] Scharenborg, O. and van Os, M. (2019) Why listening in background noise is harder in a non-native language than in a native language: A review. *Speech Commun* 108, 53-64

[93] Davis, M.H. and Johnsrude, I.S. (2007) Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hear Res* 229, 132-147

[94] Gordon-Salant, S. et al. (2019) Effects of listener age and native language experience on recognition of accented and unaccented English words. *J Speech Lang Hear Res* 62(4S), 1131-1143

[95] Gordon-Salant, S. et al. (2020) Age-related changes in speech understanding: Peripheral versus cognitive influences. In *Aging and Hearing: Causes and Consequences* (Helfer, S. et al., eds), pp. 199-230, Springer Nature

[96] Borghini, G. and Hazan, V. (2018) Listening effort during sentence processing is increased for non-native listeners: A pupillometry study. *Front Neurosci* 12, 152

[97] Brown, V.A. et al. (2020) Rapid adaptation to fully intelligible nonnative-accented speech reduces listening effort. *Q J Exp Psychol* 73, 1431-1443

[98] Bsharat-Maalouf, D. et al. (2023) The involvement of listening effort in explaining bilingual listening under adverse listening conditions. *Trends Hear* 27, 1-19

[99] Schmidtke, J. et al. (2024) How lexical frequency, language dominance and noise affect listening effort–insights from pupillometry. *Lang Cogn Neurosci*, 1-14

[100] Bsharat-Maalouf, D. et al. (2024) Through the pupil's lens: Multilingual effort in first and second language listening. *Ear Hear* 45, 1-18

[101] McGarrigle, R. et al. (2014) Listening effort and fatigue: What exactly are we measuring? A British Society of Audiology Cognition in Hearing Special Interest group 'white paper.' *Int J Audiol* 53, 433-445

[102] McCabe, D.P. et al. (2010) The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology* 24, 222-243

[103] Eckert, M.A. et al. (2016) Is listening in noise worth it? The neurobiology of speech recognition in challenging listening conditions. *Ear Hear* 37(Suppl. 1), 101S-110S

[104] Mattys, S.L. et al. (2012) Speech recognition in adverse conditions: A review. *Lang Cogn Process* 27, 953-978

[105] Millman, R.E. and Mattys, S.L. (2017) Auditory verbal working memory as a predictor of speech perception in modulated maskers in listeners with normal hearing. *J Speech Lang Hear Res* 60, 1236-1245

[106] Lew, E. et al. (2024) Navigating the bilingual cocktail party: A critical role for listeners' L1 in the linguistic aspect of informational masking. *Biling Lang Cogn*, 1-9

[107] Davis, M.H. et al. (2005) Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *J Exp Psychol Gen* 134, 222-241

[108] Peelle, J.E. and Wingfield, A. (2005) Dissociable components of perceptual learning revealed by adult age differences in adaptation to time-compressed speech. *J Exp Psychol Hum Percept Perform* 31, 1315-1330

[109] Schwab, E.C. et al. (1985) Some effects of training on the perception of synthetic speech. *Hum Factors* 27, 395-408

[110] Borrie, S.A. et al. (2012) Perceptual learning of dysarthric speech: A review of experimental studies. *J Speech Lang Hear Res* 55, 290-305

[111] Hervais-Adelman, A. et al. (2011) Generalization of perceptual learning of vocoded speech. *J Exp Psychol Hum Percept Perform* 37, 283-295

[112] Cooke, M. et al. (2022) The time course of adaptation to distorted speech. *J Acoust Soc Am* 151, 2636-2646

[113] Lie, S. et al. (2024) Learning effects in speech-in-noise tasks: Effect of masker modulation and masking release. *J Acoust Soc Am* 156, 341-349

[114] Sussman, E.S. (2017) Auditory scene analysis: An attention perspective. *J Speech Langu Hear Res* 60, 2989-3000

[115] Huyck, J.J. and Johnsrude, I.S. (2012) Rapid perceptual learning of noise-vocoded speech requires attention. *J Acoust Soc Am* 131, EL236-EL242

[116] Tye-Murray et al. (2010) Aging, audiovisual integration, and the principle of inverse effectiveness. *Ear Hear* 31, 636-644

[117] Nagaraj, N.K. (2024) Speech perception in noise: No interaction between working memory and degree of speech degradation. *Speech Lang Hear* 27, 67-77

[118] Hadley, L.V. et al. (2019) Speech, movement, and gaze behaviours during dyadic conversation in noise. *Sci Rep* 9, 10451

[119] Levinson, S.C. (2016) Turn-taking in human communication–origins and implications for language processing. *Trends Cogn Sci* 20, 6-14

[120] Gohari, N. et al. (2023) Training programs for improving speech perception in noise: A review. *J Audiol Otol* 27, 1-9

Figure Legend

Figure 1. Drivers of listening performance within the DRL framework. (A) The top panel of Figure 1A illustrates the relative contributions of individual differences in perceptual, cognitive, and linguistic abilities to task performance as a function of signal quality. When speech quality is data-limited (low signal quality), speech recognition performance is low and driven mostly by a listener's perceptual abilities. When speech quality improves (moderate signal quality), performance increases and the allocation of additional cognitive resources, within the scope of a listener's cognitive abilities, begins to contribute to performance (the resource-limited region). When speech reaches a high level of clarity (high signal quality), performance is high (the language-limited region) and individual differences are constrained more by linguistic abilities than by perceptual or cognitive abilities. The color gradient in the lower band of the figure highlights the secondary contribution of all abilities across the signal-quality continuum. The bottom panel of Figure 1A shows a typical signal quality / performance curve (in black, left Y-axis), with the confidence ribbon illustrating individual differences along the signal-quality continuum. The three processing regions are highlighted. The grey curve (right Y-axis) shows the expected inverted U-shaped resource-engagement function relative to signal quality. Moving from right to left, resource engagement is low in the language-limited region and increases as signal quality decreases in the resource-limited region, reaching a peak at the midpoint of this region. Resource engagement declines when the signal degrades further in the data-limited region, suggesting disengagement when additional investment is unlikely to enhance performance. (B) Examples of how the DRL framework can be modified to generate predictions about other populations of interest. Hypothetical contributions of perceptual, cognitive, and linguistic abilities are shown for hearing-impaired listeners (top panel) and non-native listeners (bottom panel).

Table 1. Overview of processing regions, evidence, and predictions within the DRL framework

Signal Quality Region	Predicted Dominant Abilities	Examples of Predictors / Tests	Converging Evidence	Predictions
Low signal quality **DATA-LIMITED REGION**	**Perceptual abilities dominate**. Cognitive and linguistic processes cannot substantially improve performance because the input is too impoverished.	Pure-tone audiometry thresholds, gap detection, modulation detection, frequency discrimination, temporal processing.	• Hearing acuity predicts recognition of lowpass-filtered words better than working memory does [14]. • For older adults with hearing loss, hearing thresholds dominate as predictors of speech recognition when speech is unaided [46,87]. • Pupillometry shows reduced TEPR in severely compared to moderately degraded conditions [16,53,54]. • Working memory is a less-dominant predictor of performance when target and masker overlap spectrally/spatially [13]. • The benefit of top-down audiovisual integration on speech perception is minimal when the auditory signal is severely degraded [116].	• SEM latent variables for perceptual tests should explain relatively more variance than latent variables for cognitive or linguistic tests. • The data-limited region is expected to extend rightward for listeners with impaired hearing, such that perceptual predictors remain strong even at moderate signal quality. • Performance should show a low correlation with resource engagement, e.g., TEPR; task performance should remain largely independent of any motivational manipulation. • Turn-taking in conversation [118,119] should rely more heavily on acoustic cues, e.g., prosody, than discourse predictability. • Perceptual training [120] should be maximally effective in this region.
Moderate signal quality **RESOURCE-LIMITED REGION**	**Cognitive abilities dominate**. Listeners can rescue moderately degraded input using cognitive resources; investing effort "pays off."	Working memory span, executive functions, auditory attention, processing speed.	• Speech-in-noise performance correlates with working memory capacity when signal degradation is moderate [13,20,21]. • Adding processing pauses in vocoded speech benefits intelligibility only is degradation through vocoding is moderate [15]. • Pupillometry shows highest TEPR at ~50% intelligibility, consistent with peak cognitive effectiveness [16,53,54].	• SEM latent variables for cognitive tests should explain relatively more variance than latent variables for perceptual or linguistic tests. • Motivational manipulations should modulate TEPR and improve performance to the greatest extent. • Individuals with cognitive impairments should be most affected in this region. • Cognitive training should be maximally effective in this region.

			• When older adults' audibility is restored through spectral shaping, cognitive latent variables emerge as a dominant predictor [47].	
High signal quality **LANGUAGE-LIMITED REGION**	**Linguistic abilities dominate**. Input is clear enough that residual variability primarily reflects differences in vocabulary, syntax, and discourse processing skills.	Vocabulary size, syntactic fluency.	• Larger effects of positive SNRs on intelligibility for sentences than for isolated words, suggesting greater use of sentence-level information in favorable conditions [117]. • Pupillometry shows decreased TEPR once performance asymptotes, showing lower involvement of cognitive functions [16,53,54]. • Persistent individual variability in intelligibility in high signal-quality conditions [15,28-30] despite low cognitive resource engagement [16], suggesting possible contribution of differences in linguistic functions.	• SEM latent variables for linguistic tests should explain relatively more variance than latent variables for perceptual or cognitive tests. • The language-limited region is expected to extend leftward for non-native listeners, such that linguistic predictors remain strong even at moderate signal quality. • Since listening is achieved with minimal cognitive resources, motivational manipulations should not significantly contribute to performance improvement. • Language training should be maximally effective in this region, especially for L2 listeners.