



Review

# Processing of Degraded Speech in Brain Disorders

Jessica Jiang<sup>1</sup>, Elia Benhamou<sup>1</sup>, Sheena Waters<sup>2</sup> , Jeremy C. S. Johnson<sup>1</sup>, Anna Volkmer<sup>3</sup>, Rimona S. Weil<sup>1</sup> , Charles R. Marshall<sup>1,2</sup>, Jason D. Warren<sup>1,†</sup> and Chris J. D. Hardy<sup>1,\*,†</sup>

<sup>1</sup> Dementia Research Centre, Department of Neurodegenerative Disease, UCL Queen Square Institute of Neurology, London WC1N 3BG, UK; jessica.jiang.18@ucl.ac.uk (J.J.); elia.benhamou.16@ucl.ac.uk (E.B.); jeremy.johnson@ucl.ac.uk (J.C.S.J.); r.weil@ucl.ac.uk (R.S.W.); charles.marshall@qmul.ac.uk (C.R.M.); jason.warren@ucl.ac.uk (J.D.W.)

<sup>2</sup> Preventive Neurology Unit, Wolfson Institute of Preventive Medicine, Queen Mary University of London, London EC1M 6BQ, UK; s.waters@qmul.ac.uk

<sup>3</sup> Division of Psychology and Language Sciences, University College London, London WC1H 0AP, UK; a.volkmer.15@ucl.ac.uk

\* Correspondence: chris.hardy@ucl.ac.uk; Tel.: +44-203-448-3676

† These authors contributed equally to this work.

**Abstract:** The speech we hear every day is typically “degraded” by competing sounds and the idiosyncratic vocal characteristics of individual speakers. While the comprehension of “degraded” speech is normally automatic, it depends on dynamic and adaptive processing across distributed neural networks. This presents the brain with an immense computational challenge, making degraded speech processing vulnerable to a range of brain disorders. Therefore, it is likely to be a sensitive marker of neural circuit dysfunction and an index of retained neural plasticity. Considering experimental methods for studying degraded speech and factors that affect its processing in healthy individuals, we review the evidence for altered degraded speech processing in major neurodegenerative diseases, traumatic brain injury and stroke. We develop a predictive coding framework for understanding deficits of degraded speech processing in these disorders, focussing on the “language-led dementias”—the primary progressive aphasia. We conclude by considering prospects for using degraded speech as a probe of language network pathophysiology, a diagnostic tool and a target for therapeutic intervention.

**Keywords:** degraded speech processing; predictive coding; primary progressive aphasia; Alzheimer’s disease; Parkinson’s disease; perceptual learning; dementia



**Citation:** Jiang, J.; Benhamou, E.; Waters, S.; Johnson, J.C.S.; Volkmer, A.; Weil, R.S.; Marshall, C.R.; Warren, J.D.; Hardy, C.J.D. Processing of Degraded Speech in Brain Disorders. *Brain Sci.* **2021**, *11*, 394.

[https://doi.org/](https://doi.org/10.3390/brainsci11030394)

[10.3390/brainsci11030394](https://doi.org/10.3390/brainsci11030394)

Academic Editor: Claude Alain

Received: 18 February 2021

Accepted: 18 March 2021

Published: 20 March 2021

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Speech is arguably the most complex of all sensory signals and yet the healthy brain processes it with an apparent ease that belies the complexities of its neurobiological and computational underpinnings. Speech signals arrive at the ears with widely varying acoustic characteristics, reflecting such factors as speech rate, morphology, and in particular, the presence of competing sounds [1,2]. The clear speech stimuli played to participants in quiet, controlled laboratory settings are very different to the speech we typically encounter in daily life, which is usually degraded in some form. Under natural listening conditions, not only does speech often compete with other sounds, but the acoustic environment is frequently changing over time, thus speech processing is inherently dynamic. In general, the processing of degraded speech entails the extraction of an intelligible message (the “signal”) despite listening conditions that adversely affect the quality of the speech in some way (the “noise”). These conditions can be broadly conceptualised as relating to external environmental factors such as background sounds, the vocal idiosyncrasies of other speakers (such as an unfamiliar accent) [3], or feedback relating to one’s own vocal productions. Understanding speech under the non-ideal listening conditions of everyday

life presents a particular challenge to the damaged brain, and might constitute a cognitive “stress test” that exposes the effects of brain pathology.

Various computational models have been proposed to explain how a speech signal is normally efficiently disambiguated from auditory “noise”, entailing the extraction of specific acoustic features, phonemes, words, syntax, and ultimately, meaning [4–8]. Common to these models is the notion that accurate speech decoding depends on the integration of “bottom-up” processing of incoming auditory information (e.g., speech sounds) with “top-down” prior knowledge and contextual information (e.g., stored phonemes derived from one’s native language). Degraded speech signals are parsed by generic bottom-up processes that are also engaged by other complex acoustic environments during “auditory scene analysis” [9], and the high predictability of speech signals recruits top-down processes that are relatively speech-specific: these processes normally interact dynamically and reciprocally to achieve speech recognition [10]. A computationally intensive process of this kind that depends on coherent and dynamic interactions across multiple neural processing steps is likely to be highly vulnerable to the disruptive effects of brain pathologies.

### *1.1. Predictive Coding and Degraded Speech Perception*

Predictive coding theory, a current highly influential theory of perception, postulates that the brain models the causes of sensory input by iteratively comparing top-down predictions to bottom-up inputs and updating those predictions to reduce the disparity between prediction and experience (i.e., to minimise prediction error) [11,12]. The brain achieves this by modelling predictions at lower-level sensory processing stages (“priors”) via top-down connections from higher-level areas [13]: the modelling involves a weighting or gain of bottom-up inputs based on their precision (variability) and their expected precision that informs the confidence of the prediction error. In neuronal terms, this error is a mismatch between the neural representations of noisy sensory input at each processing level and the predictions constructed at the processing level above it in the hierarchy. If the prediction error is significant (above noise), this will cause the brain’s model to be modified, such that it better predicts sensory input. The computational implementation of the modification process is difficult to specify in detail a priori; the associated changes of neural activity at each processing stage are likely to evolve over time, perhaps accounting for certain apparently contradictory findings in the experimental neurophysiological literature [14].

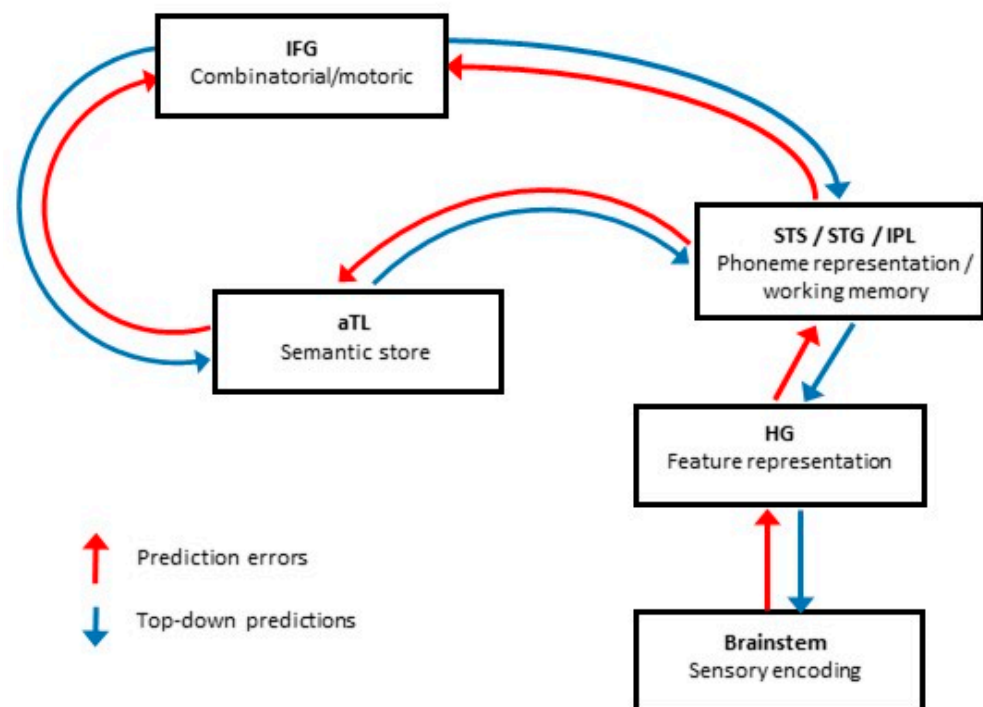
According to this predictive coding framework, degraded speech perception depends on hierarchical reciprocal processing in which each stage passes down predictions, and prediction errors (i.e., the difference between expected and heard speech) are passed up the hierarchy [4,15]. Our ability to accurately perceive degraded speech is enhanced by inferring the probability of various possible incoming messages according to context [4,16,17].

### *1.2. Neuroanatomy of Degraded Speech Processing*

From a neuroanatomical perspective, it is well established that the representation and analysis of intelligible speech occur chiefly in a processing network surrounding primary auditory cortex in Heschl’s gyrus, with processing “streams” projecting ventrally along superior temporal gyrus and sulcus (STG/STS) and dorsally to inferior frontal gyrus (IFG) in the left (dominant) cerebral hemisphere [5,18,19]. Medial temporal lobe structures in the dominant hemisphere encode and retain verbal information [19–21], and anterior temporal polar cortex may constitute a semantic hub [22–24]. The reciprocal connections between association auditory regions and prefrontal cortical areas, in particular IFG [7,25,26], are essential for the top-down disambiguation of speech signals [27–29].

Broadly similar regions have been consistently identified in neuroimaging studies of degraded speech processing, including superior temporal sulcus/gyrus (for accent processing: [30]; altered auditory feedback: [31]; dichotic listening: [32]; noise-vocoded speech: [33–35]; perceptual restoration: [36]; sinewave speech: [37]; speech-in-noise: [38]; and time-compressed speech [39]) and inferior frontal gyrus (for accent processing: [30];

noise-vocoded speech: [33–35]; perceptual restoration: [36]) in the dominant hemisphere. Additional temporo-parietal brain regions are also engaged under challenging listening conditions [32,40,41]. Therefore, a large fronto-temporo-parietal network consolidates information across multiple processing levels (acoustic, lexical, syntactic, semantic, articulatory) to facilitate the perception of degraded speech signals [42]. Adaptation to degraded speech may be mediated partly by subcortical striato-thalamic circuitry [27]. The macro-anatomical and functional organisation of the language network suggests how predictive coding mechanisms might operate in processing degraded speech (see Figure 1). Cortical regions involved in “early” analysis of the speech signal, such as STG/STS, communicate with “higher” regions, such as IFG, that instantiate high-level predictions about degraded sensory signals. Crucially, however, both “bottom-up” perception and “top-down” processing would occur at every stage within the hierarchy, actively updating stored templates (representations or “priors”) of the auditory environment and generating prediction errors when the auditory input fails to match the prediction [43].



**Figure 1.** A predictive coding model of normal degraded speech processing with major anatomical loci for core speech decoding operations and their connections, informed by evidence in the healthy brain. Different kinds of degraded speech manipulation are likely to engage these cognitive operations and connections differentially (see Table 1). Incoming sensory information undergoes “bottom-up” perceptual analysis chiefly in early auditory areas, while higher level brain regions generate predictions about the content of the speech signal. Boxes indicate processors that instantiate core functions; note, however, that processing “levels” are not strictly confined to higher-order predictions or early sensory input: interactions occur at each level. Arrows indicate connections between levels, with reciprocal information flow mediating modulatory influences and dynamic updating/perceptual learning of degraded speech signals. This figure is necessarily an over-simplification; cortical areas that are likely to have separable functional roles are grouped together for clarity of representation, and while they are not shown in this figure, intra-areal recurrences and inhibitions alongside other local circuit effects may also be operating within these regions. aTL, anterior temporal lobe; HG, Heschl’s gyrus; IFG, inferior frontal gyrus; IPL, inferior parietal lobule; STG, superior temporal gyrus; STS, superior temporal sulcus.

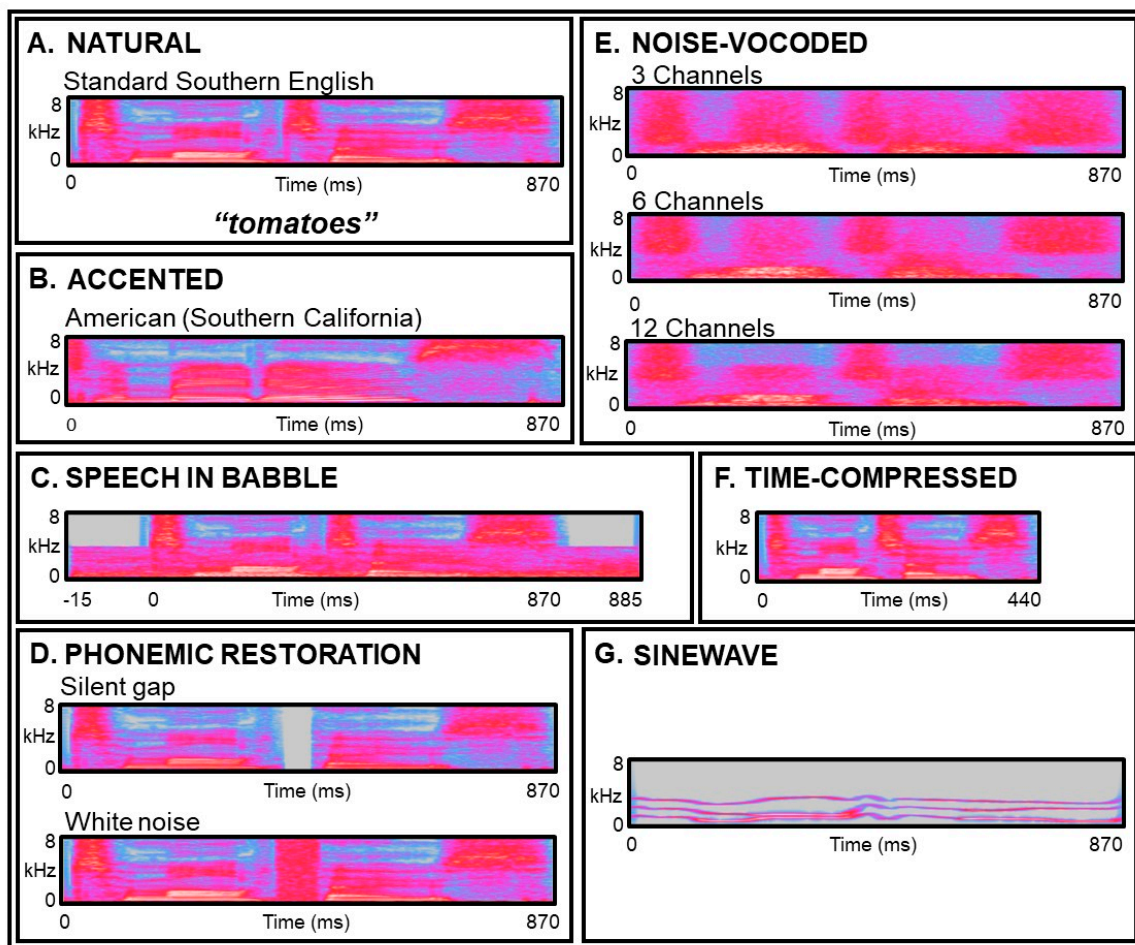
Techniques such as electro-encephalography (EEG) and magneto-encephalography (MEG) have revealed dynamic, oscillatory activity that synchronises neural circuits and large-scale networks [44]. By delineating feedforward and feedback influences as well as the rapid changes that attend deviant, incongruous or ambiguous stimuli, such techniques are well suited to predictive coding applications such as the processing of degraded speech. Indeed, MEG evidence suggests that induced activity in particular frequency bands may constitute signatures of underlying neural operations during the predictive decoding of speech and other sensory signals [45–47]: gamma oscillations (>30 Hz) are modulated as a function of sensory “surprise” (i.e., prediction error), beta oscillations (12–30 Hz) are modulated through processing steps downstream from prediction error generation (i.e., updating of top-down predictions) and alpha oscillations (8–12 Hz) reflect the precision of predictions. Past studies conducted with MEG on degraded speech perception have shown enhanced responses in the auditory cortex (STG) when input becomes intelligible, but also reduced responses in the context of prior knowledge and perceptual learning (see Section 2.4), consistent with predictive, top-down modulation from higher-order cortical areas [48,49].

Accurate and flexible understanding of speech depends critically on the capacity of speech processing circuitry to respond efficiently, dynamically, and adaptively to diverse auditory inputs in multiple contexts and environments [50]. Degraded speech processing is therefore likely to be highly vulnerable to brain diseases that target these networks, as exemplified by the primary neurodegenerative “nexopathies” that cause dementia [51]. Major dementias strike central auditory and language processing networks relatively selectively, early and saliently (see Hardy and colleagues [52] for a review). It is therefore plausible that brain diseases should manifest as impairments of degraded speech processing and should have signature profiles of impairment according to the patterns of language network damage they produce. Indeed, reduced ability to track and understand speech under varying (non-ideal) listening conditions is a major contributor to the communication difficulties that people living with dementia experience in their daily lives and is a significant challenge for the care and management of these patients. Furthermore, the nature of the speech processing difficulty (as reflected in the symptoms patients describe) varies between different forms of dementia [52]. However, the processing of degraded speech in dementias and other brain disorders remains poorly understood and we presently lack a framework for interpreting and anticipating deficits.

### *1.3. Scope of This Review*

In this review, we consider how and why the processing of degraded speech is affected in some major acquired brain disorders. Experimentally, many different types of speech degradation have been employed to study degraded speech processing: we summarise some of these in Figure 2 and provide a representative review of the literature in Table 1. We next consider important factors that affect degraded speech processing in healthy individuals to provide a context for interpreting the effects of brain disease. We then review the evidence for altered processing of degraded speech in particular acquired brain disorders (Table 2). We conclude by proposing a predictive coding framework for assessing and understanding deficits of degraded speech processing in these disorders, implications for therapy and directions for further work (Figure 3).





**Figure 2.** Examples of degraded speech manipulations used experimentally and their acoustic effects on the speech signal. Broadband time-frequency spectrograms of the same speech token (“tomatoes”), subjected to different forms of speech degradation (all samples apart from 2B were recorded by a native British speaker with a Standard Southern English accent; wavefiles of A–G are in Supplementary Material online). (A) Natural speech token. (B) Same speech token spoken with an American-Californian accent (an accent is a meta-linguistic feature that reveals information about the speaker’s geographical or socio-cultural background [53]; normal listeners make predictions about speakers’ accents that tend to facilitate faster accent processing [54]). (C) Speech in multi-talker babble (speech-in-noise can be adaptively adjusted to find the point at which speech switches from intelligible to unintelligible [55]; background “noise” used experimentally typically comprises either “energetic” masking (e.g., steady-state white noise) or “informational” masking (e.g., multi-talker babble, as illustrated here)) [56], (D) Perceptual (or phonemic) restoration (Warren [57] originally observed that when a key phoneme is artificially excised from a given sentence, control participants are unable to identify the location of the missing phoneme when “filled-in” with a burst of white noise (bottom panel), but are able to identify the location accurately if the gap remains silent (top panel), i.e., they perceptually “restore” the excised phoneme). (E) Noise-vocoded speech (vocoding removes fine spectral detail from speech, whilst preserving temporal cues [58,59]; three bands of modulated noise (i.e., three “channels”; top panel) are the minimum needed for consistent recognition by normal listeners [59], spectrograms for six (middle panel) and twelve (bottom panel) channels also shown here). (F) Time-compressed speech (created by artificially increasing the rate at which a recorded speech stimulus is presented; intelligibility decreases as speech compression rate increases [60–62]). (G) Sinewave speech (this transformation reduces speech to a series of “whistles” or sinewave tones that track formant contours [63]). Note that these speech manipulations vary widely in the cognitive process they target, the degree to which they degrade the speech signal and their ecological resonance (see also Table 1); accented speech and speech-in-noise or babble are commonly encountered in daily life through exposure to diverse speakers and noisy environments, perceptual restoration simulates the frequent everyday phenomenon of speech interruption by intermittent extraneous sounds (e.g., a slamming door), whereas sinewave-speech is a drastic impoverishment of the speech signal that sounds highly unnatural but becomes intelligible with exposure due to perceptual learning [64].

**Table 1.** Summary of major forms of speech degradation with representative experimental studies in healthy listeners.

Degradation Type	Study	Participants	Methodology	Major Findings
<b>ACCENTS</b> <b>Target process:</b> phonemic and intonational representations <b>Ecological relevance:</b> Understanding messages conveyed via non-canonical spoken phonemes and suprasegmental intonation	Bent and Bradlow [65]	65 healthy participants (age: 19.1)	Participants listened to English sentences spoken by Chinese, Korean, and English native speakers.	Non-native listeners found speech from non-native English speakers as intelligible as from a native speaker.
	Clarke and Garrett [66]	164 healthy participants (American English)	Participants listened to English sentences spoken with a Spanish, Chinese, and English accent.	Processing speed initially slower for accented speech, but this deficit diminished with exposure.
	Floccia, Butler, Goslin and Ellis [54]	54 healthy participants (age 19.7; Southern British English)	Participants had to say if the last word in a spoken sentence was real or not.	Changing accent caused a delay in word identification, whether accent change was regional or foreign.
<b>ALTERED AUDITORY FEEDBACK</b> <b>Target process:</b> Influence of auditory feedback on speech production <b>Ecological relevance:</b> Ability to hear, process, and regulate speech from own production.	Siegel and Pick [67]	20 healthy participants	Participants produced speech whilst hearing amplified feedback of their own voice.	Participants lowered their voices (displaying the sidetone amplification effect) in all conditions.
	Jones and Munhall [68]	18 healthy participants (age: 22.4; Canadian English)	Participants produced vowels with altered feedback of F0 shifted up or down.	Participants compensated for change in F0.
	Donath et al. [69]	22 healthy participants (age: 23; German)	Participants said a nonsense word with feedback of their frequency randomly shifting downwards.	Participants adjusted their voice F0 after a set period of time due to processing the feedback first.
	Stuart et al. [70]	17 healthy participants (age: 32.9; American English)	Participants spoke under DAF at 0, 25, 50, 200 ms at normal and fast rates of speech.	There were more dysfluencies at 200 ms, and more dysfluencies at the fast rate of speech.
<b>DICHOTIC LISTENING</b> <b>Target process:</b> Auditory scene analysis (auditory attention) <b>Ecological relevance:</b> Processing of spoken information with competing verbal material	Moray [71]	Healthy participants, no other information given	Participants were told to focus on a message played to one ear, with a competing message in the other ear.	Participants did not recognize the content in the unattended message.
	Lewis [72]	12 healthy participants	Participants were told to attend to message presented in one ear, with a competing message in the other.	Participants could not recall the unattended message, but semantic similarity affected reaction times.
	Ding and Simon [73]	10 healthy participants (age 19–25)	Under MEG, participants heard competing messages in each ear, and asked to attend to each in turn.	Auditory cortex tracked temporal modulations of both signals, but was stronger for the attended one.

Table 1. Cont.

Degradation Type	Study	Participants	Methodology	Major Findings
<b>NOISE-VOCODED SPEECH</b> <b>Target process:</b> Phonemic spectral detail <b>Ecological relevance:</b> Understanding whisper (similar quality to speech heard by cochlear implant users)	Shannon, Zeng, Kamath, Wygonski and Ekelid [59]	8 healthy participants	Participants listened to and repeated simple sentences that had been noise-vocoded to different degrees.	Performance improved with number of channels; high speech recognition was achieved with only 3 channels.
	Davis, Johnsrude, Hervais-Adelman, Taylor and McGettigan [58]	12 healthy participants (age 18–25; British English)	Participants listened to and then transcribed 6-channel noise-vocoded sentences.	Participants showed rapid improvement over the course of 30-sentence exposure.
	Scott, Rosen, Lang and Wise [35]	7 healthy participants (age 38)	Under PET, participants listened to spoken sentences that were noise-vocoded to various degrees.	Selective response to speech intelligibility in left anterior STS.
<b>PERCEPTUAL RESTORATION</b> <b>Target process:</b> Message interpolation <b>Ecological relevance:</b> Understanding messages in intermittent or varying noise (e.g., a poor telephone line)	Warren [57]	20 healthy participants	Participants identified where the gap was in sentences where a phoneme was replaced by silence/white noise.	Participants were more likely to mislocalize a missing phoneme that was replaced by noise.
	Samuel [74]	20 healthy participants (English)	Participants heard sentences in which white noise was either “Added” to or “Replaced” a phoneme.	Phonemic restoration was more common for longer words and certain phone classes.
	Leonard, Baud, Sjerps and Chang [43]	5 healthy participants (age 38.6; English/Italian)	Subdural electrode arrays recorded while participants listened to words with noise-replaced phonemes.	Electrode responses were comparable to intact words vs. words with a phoneme replaced.
<b>SINEWAVE SPEECH</b> <b>Target process:</b> Speech reconstruction and adaptation from very impoverished cues <b>Ecological relevance:</b> Synthetic model for impoverished speech signal and perceptual learning	Remez, Rubín, Pisoni and Carrell [63]	54 control participants	Naïve listeners heard SWS replicas of spoken sentences and were later asked to transcribe the sentences.	Most listeners did not initially identify the SWS as speech, but were able to transcribe them when told this.
	Barker and Cooke [64]	12 control participants	Participants were asked to transcribe SWS or amplitude-comodulated SWS sentences.	Recognition for SWS ranged from 35–90%, and amplitude-comodulated SWS ranged from 50–95%.
	Möttönen, Calvert, Jääskeläinen, Matthews, Thesen, Tuomainen and Sams [37]	21 control participants (18–36; English)	Participants underwent two fMRI scans: one before training on SWS, and one post-training.	Activity in left posterior STS was increased after SWS training.

Table 1. Cont.

Degradation Type	Study	Participants	Methodology	Major Findings
<b>SPEECH-IN-NOISE</b> <b>Target process:</b> Auditory scene analysis (parsing of phonemes from acoustic background) <b>Ecological relevance:</b> Understanding messages in background noise (e.g., “cocktail party effect”)	Pichora-Fuller et al. [75]	24 participants in three groups (age 23.9; 70.4; 75.8; English)	Participants repeated the last word of sentences in 8-talker babble. Half had predictable endings.	Both groups of older listeners derived more benefit from context than younger listeners.
	Parbery-Clark et al. [76]	31 control participants (incl. 16 musicians; age: 23; English)	Participants were assessed via clinical measures of speech perception in noise.	Musicians outperformed the non-musicians on both QuickSIN and HINT.
	Anderson et al. [77]	120 control participants (age 63.9)	Peripheral auditory function, cognitive ability, speech-in-noise, and life experience were examined.	Central processing and cognitive function predicted variance in speech-in-noise perception.
<b>TIME-COMPRESSED SPEECH</b> <b>Target process:</b> Phoneme duration (rate of presentation) <b>Ecological relevance:</b> Understanding rapid speech	Dupoux and Green [60]	160 control participants (English)	Participants transcribed spoken sentence were compressed to 38% and 45% of their original durations.	Participants improved over time. This happened more rapidly for the 45% compressed sentences.
	Poldrack et al. [78]	8 control participants (age: 20–29; English)	Participants listened to time-compressed speech. Brain responses were tracked using fMRI.	Activity in bilateral IFG and left STG increased with compression, until speech became incomprehensible.
	Peelle et al. [79]	8 control participants (age: 22.6; English)	Participants listened to sentences manipulated for complexity and time-compression in an fMRI study.	Time-compressed sentences recruited AC and premotor cortex, regardless of complexity.

The table is ordered by type of speech degradation. Information in the Participants column is based on available information from the original papers; age is given as a mean or range and language refers to participants’ native languages. Abbreviations: AC, anterior cingulate; DAF, delayed auditory feedback; F0, fundamental frequency; fMRI; functional magnetic resonance imaging; HINT, Hearing in Noise Test; IFG, inferior frontal gyrus; ms, millisecond; QuickSIN, Quick Speech in Noise Test; PET, positron emission tomography; STG, superior temporal gyrus; STS, superior temporal sulcus; SWS, sinewave speech.



**Table 2.** Summary of representative studies of degraded speech processing in clinical populations.

Population	Study, Degradation	Participants	Methodology	Major Findings
Traumatic brain injury	Gallun et al. [80]: Central auditory processing	36 blast-exposed military veterans (age: 32.8); 29 controls (age: 32.1)	Participants went through a battery of standardised behavioural tests of central auditory function: temporal pattern perception, GIN, MLD, DDT, SSW, and QuickSIN.	While no participant performed poorly on all behavioural testing, performance was impaired in central auditory processing for the blast-exposed veterans in comparison to matched-controls.
	Saunders et al. [81]: Central auditory processing	99 military veterans (age: 34.1)	Participants went through self-reported measures as well as a battery of standardised behavioural measures: HINT, NA LiSN-S, ATTR, TCST, and SSW.	Participants in this study showed measurable performance deficits on speech-in-noise perception, binaural processing, temporal resolution, and speech segregation.
	Gallun et al. [82]: Central auditory processing	30 blast-exposed military veterans, with a least one blast occurring 10 years prior to study (age: 37.3); 29 controls (age: 39.2)	Participants went through a battery of standardised behavioural tests of central auditory function: GIN, DDT, SSW, FPT, and MLD.	Replicating the findings from Gallun et al., 2012, this study found that the central auditory processing deficits persisted in individuals tested an average of more than 7 years after blast exposure.
	Papesh et al. [83]: Central auditory processing	16 blast-exposed veterans (age 36.9); 13 veteran controls (age 38) with normal peripheral hearing	Participants completed self-reported measures and standardised tests of speech-in-noise perception, DDT, SSW, TCST, plus auditory event-related potential studies.	Impaired cortical sensory gating was primarily influenced by a diagnosis of TBI and reduced habituation by a diagnosis of post-traumatic stress disorder. Cortical sensory gating and habituation to acoustic startle strongly predicted degraded speech perception
Stroke aphasia	Bamiou et al. [84]: Dichotic listening	8 patients with insular strokes (age: 63); 8 control participants (age: 63)	Participants heard pairs of spoken digits presented simultaneously to each ear, and were asked to repeat all four digits.	Dichotic listening was abnormal in five of the eight stroke patients.
	Dunton et al. [85]: Accents	16 participants with aphasia (age: 59); 16 controls (age: 59; English)	Participants heard English sentences spoken with a familiar (South-East British England) or unfamiliar (Nigerian) accent.	Aphasia patients made more errors in comprehending sentences spoken in an unfamiliar accent vs. a familiar accent.
	Jacks and Haley [86]: AAF (MAF)	10 aphasia patients (age: 53.1); 10 controls (age: 63.1; English)	Participants produced spoken sentences with no feedback, DAF, FAF or noise-masked auditory feedback (MAF).	Speech rate increased under MAF but decreased with DAF and FAF in most participants with aphasia.
Parkinson's disease	Liu et al. [87]: AAF (MAF and FAF)	12 PD participants (age: 62.3); 13 control participants (age: 68.7)	Participants sustained a vowel whilst receiving changes in feedback of loudness ( $\pm 3/4$ dB) or pitch ( $\pm 100$ cents).	All participants produced compensatory responses to AAF, but response sizes were larger in PD than controls.
	Chen et al. [88]: AAF (FAF)	15 people with PD (age: 61); 15 control participants (age 61; Cantonese)	Participants were asked to vocalize a vowel sound with AAF pitch-shifted upwards or downwards.	PD participants produced larger magnitudes of compensation.

Table 2. Cont.

Population	Study, Degradation	Participants	Methodology	Major Findings
Alzheimer's disease	Gates et al. [89]: Dichotic digits	17 ADs (age: 84); 64 MCI (age: 82.3); 232 controls (age: 78.8)	Participants listened to 40 numbers presented in pairs to each ear simultaneously.	AD patients scored the worst in the dichotic digits, followed by the MCI group and then the controls.
	Golden et al. [90]: Auditory scene analysis	13 AD participants (age: 66); 17 control participants (age: 68)	In fMRI, participants listened to their own name interleaved with or superimposed on multi-talker babble.	Significantly enhanced activation of right supramarginal gyrus in the AD vs. control group for the cocktail party effect.
	Ranasinghe et al. [91]: AAF (FAF)	19 AD participants; 16 control participants	Participants were asked to produce a spoken vowel in context of AAF, with perturbations of pitch.	AD patients showed enhanced compensatory response and poorer pitch-response persistence vs. controls.
Primary progressive aphasia	Hailstone et al. [92]: Accents	20 ADs (age: 66.4); 6 nfvPPA (age: 66); 35 controls (age: 65); British English	Accent comprehension and accent recognition was assessed. VBM examined grey matter correlates.	Reduced comprehension for phrases in unfamiliar vs. familiar accents in AD and for words in nfvPPA; in AD group, grey matter associations of accent comprehension and recognition in anterior superior temporal lobe
	Cope et al. [93]: Noise-vocoding	11 nfvPPA (age: 72); 11 control participants (age: 72)	During MEG, participants listened to vocoded words presented with written text that matched/mismatched.	People with nfvPPA compared to controls showed delayed resolution of predictions in temporal lobe, enhanced frontal beta power and top-down fronto-temporal connectivity; precision of predictions correlated with beta power across groups
	Hardy et al. [94]: SWS	9 nfvPPA (age: 69.6); 10 svPPA (age: 64.9); 7 lvPPA (age: 66.3); 17 control (age: 67.7)	Participants transcribed SWS of numbers/locations. VBM examined grey matter correlates in combined patient cohort.	Variable task performance groups; all showed spontaneous perceptual learning effects for SWS numbers; grey matter correlates in a distributed left hemisphere network extending beyond classical speech-processing cortices, perceptual learning effect in left inferior parietal cortex

Information in the Participants column is based on available information from the original papers; age is given as a mean or range and language refers to participants' native languages. Abbreviations: AAF, altered auditory feedback; AD, Alzheimer's disease; ATTR, Adaptive Tests of Temporal Resolution; DAF, delayed auditory feedback; dB, decibels; DDT, Dichotic Digits Test; FAF, frequency altered feedback; fMRI, functional magnetic resonance imaging; FPT, Frequency Patterns Tests (FPT); GIN, Gaps-In-Noise test; HINT, Hearing in Noise Test; lvPPA, logopenic variant primary progressive aphasia; MAF, masked/masking auditory feedback; MCI, mild cognitive impairment; MEG, magnetoencephalography; MLD, The Masking Level Difference; NA LiSN-S, North American Listening in Spatialised Noise-Sentence test; nfvPPA, nonfluent primary progressive aphasia; PD, Parkinson's disease; PR, perceptual restoration; QuickSIN, Quick Speech in Noise; SSW, Staggered Spondaic Words; SWS, sinewave speech; svPPA, semantic variant primary progressive aphasia; TBI, traumatic brain injury; TCST, Time Compressed Speech Test; VBM, voxel based morphometry.

## 2. Factors Affecting Processing of Degraded Speech in the Healthy Brain

### 2.1. Healthy Ageing

Healthy ageing importantly influences the perception of degraded speech [52,75,77,95–97], and an understanding of ageing effects is essential in order to interpret the impact of brain disorders, particularly those associated with neurodegenerative disease. Ageing may be associated with functionally significant changes affecting multiple stages of auditory processing, from cochlea [98], to brainstem [99], to cortex [100]. The reduced efficiency of processing degraded speech with normal ageing is likely to reflect the interaction of peripheral and central factors [101] due, for example, to slower processing or reduced ability to regulate sensory gating [97,102,103].

These alterations in auditory pathway function tend to be amplified by age-related decline in additional cognitive functions relevant to degraded speech perception. Ageing affects domains such as episodic memory, working memory, and attention [77,101,104,105]. There is evidence to suggest that older listeners rely more heavily on “top-down” cognitive mechanisms than younger listeners, compensating for the reduced fidelity of “bottom-up” auditory signal analysis [100,106–109].

### 2.2. Cognitive Factors

The auditory system is dynamic and highly integrated with cognitive function more broadly [77,110]. Executive function is accorded central importance among the general cognitive capacities that influence the speed and accuracy of degraded speech perception, interacting with more specific skills such as phonological processing [111]. The engagement of executive processing networks—including inferior frontal gyrus, inferior parietal lobule, superior temporal gyrus and insula—during effortful listening is a unifying theme in neuroimaging studies of degraded speech processing [18]. On the other hand, the ability to process degraded speech in older adults is not entirely accounted for by general cognitive capacities [112], implying additional, auditory mechanisms are also involved.

Attention, a key cognitive factor in most sensory predictive coding models, modulates the intelligibility of degraded speech, and functional magnetic resonance imaging (fMRI) research suggests that additional frontal cortical regions are recruited when listeners attend to degraded speech signals [29]. Attention is essential for encoding precision or gain: the weighting of sensory input by its reliability [113,114]. Verbal auditory working memory—the “phonological loop”—is integral to degraded speech processing [115–118], and selective attention importantly interacts with the verbal short term store to sharpen the precision of perceptual priors held in mind over an interval (for example, during articulatory rehearsal on phonological discrimination tasks: [119]). Listeners with poorer auditory working memory capacity have more difficulty understanding speech-in-noise, even after accounting for age differences and peripheral hearing loss [77,120,121]. While working memory and attention have been studied more explicitly, it is likely that a number of cognitive factors interact in processing degraded speech, and that (in the healthy brain) the usage of these cognitive resources is dynamic and adapts flexibly to a wide variety of listening conditions [111].

### 2.3. Experiential Factors

Accumulated experience of speech signals and auditory environments over the course of the lifetime leads to the development and refinement of internal models that direct predictions about auditory input, facilitating faster neural encoding and integration [122]. Certain experiential factors, such as musical training, affect the processing of degraded speech, specifically speech-in-noise [77,123]. Musical training improves a range of basic auditory skills [124–127] and auditory working memory [128] that are important to speech encoding and verbal communications such as linguistic pitch pattern processing and temporal and frequency encoding in auditory brainstem [129–132]. This could explain findings suggesting that musicians are better at perceiving speech-in-noise (whether white-noise or babble) than non-musical listeners [76,77,133–136].

Bilingual speakers have more difficulty perceiving speech-in-noise in their non-native language than their monolingual counterparts, even when they consider themselves proficient in their non-native language [137–139], not necessarily in low-context situations but particularly in high-context [140]. This may be due to over-reliance on bottom-up processing with reduced integration of semantic and contextual knowledge for the second language [141–143], relative to more efficient top-down integration in one's native language [139].

#### 2.4. Perceptual Learning

Improved accuracy of degraded speech processing is associated with sustained exposure to the stimulus [1,54,144]: this reflects perceptual learning [145]. Perceptual learning allows listeners to learn to understand speech that has deviated from expectations [146], and typically occurs automatically and within a short period of time [49,147,148]. It is likely to reflect synaptic plasticity at different levels of perceptual analysis [149,150], and (in predictive coding terms) reflects iterative fine-tuning of the internal model with increased exposure to the stimulus, leading to error minimisation and improved accuracy of future predictions about the incoming speech signal (Figure 1; [15]).

Although perceptual learning of degraded speech is strongest and most consistent if trained and tested with the same single speaker [151–153], with exposure to many individuals embodying a similar particular characteristic (e.g., similar accent), the enhanced processing of that characteristic generalises to different speakers [154–157]. Longer training (i.e., more exposure to the stimulus) also leads to more stable learning and generalization [158]. Listener factors also affect perceptual learning, including language background [159], age [160], attentional set [161], and the recruitment of language processes in higher-level brain regions and connectivity [144]. Perceptual learning of accented speech in non-native listeners has been associated with improved speech production [162]. Overall, the results from studies on auditory perceptual learning suggest that it arises from dynamic interactions between different levels of the auditory processing hierarchy [163].

#### 2.5. Speech Production

The functional consequences of degraded speech processing on communication cannot be fully appreciated without considering how perceptual alterations influence speech output. In the healthy brain, there is an intimate interplay between speech input and output processing, both functionally and neuroanatomically [164,165]: brain disorders that disturb this interplay are likely to have profound consequences for degraded speech processing. Speech production relies on feedback and feedforward control [166], and artificially altering auditory feedback (i.e., causing prediction errors about online feedback of one's own speech output) frequently disrupts the speech production process [167] (see Table 1). "Altered auditory feedback" (AAF) is the collective term for auditory feedback that is altered or degraded in some manner before being played back to the speaker in real time [167], and encompasses masking auditory feedback (MAF), intensity-altered auditory feedback (IAF), delayed auditory feedback (DAF), and frequency-altered feedback (FAF). Typically, speakers will adjust their speech output automatically in some way to compensate for the altered feedback. One classical example is the "Lombard effect", whereby the talker responds to a loud or otherwise acoustically competing environment by altering the intensity, pitch, and spectral properties of their voice [168]. Functional neuroimaging studies show that when auditory feedback is altered, there is an increase in activation in the superior temporal cortex, extending into posterior-medial auditory areas [31,169]. This corroborates other work suggesting that this region has a prominent role in sensorimotor integration and error detection [49,170].

### 3. Processing of Degraded Speech in Brain Disorders

The various factors that affect the processing of degraded speech in the healthy brain are all potentially impacted by brain diseases. Brain disorders often affect executive

function, speech production, perceptual learning and other general cognitive capacities, with many becoming more frequent with age and their expression may be heavily modified by life experience.

We now consider some acquired neurological conditions that are associated with particular profiles of degraded speech processing; key studies are summarised in Table 2. While this is by no means an exhaustive list, it represents a survey of disorders that have been most widely studied and illustrates important pathophysiological principles.

### 3.1. Traumatic Brain Injury

Traumatic brain injury (TBI) refers to any alteration in brain function or structure caused by an external physical force. It therefore encompasses a wide spectrum of insults, pathological mechanisms and transient and permanent cognitive deficits [171,172]. Individuals with TBI, whether mild or severe, commonly report auditory complaints; blast-related TBI is associated with hearing loss and tinnitus in as many as 60% of patients [173]. Most data have been amassed for military veterans, and concurrent mental health issues complicate the picture [174].

People with TBI frequently report difficulties understanding speech under challenging listening conditions and a variety of central auditory deficits have been documented, including impaired speech-in-noise perception and dichotic listening [80,81,175,176]; these deficits may manifest despite normal peripheral hearing (pure tone perception), may follow mild as well as more severe injuries and may persist for years [81,82]. The culprit lesions in these cases are likely to be anatomically heterogeneous; blast exposure, for example, potentially damages auditory brainstem and cortices, corpus callosum and frontal cortex, while the preponderance of abnormal long-latency auditory evoked potentials argues for a cortical substrate [174]. Abnormal sensory gating has been proposed as an electrophysiological mechanism of impaired degraded speech processing in blast-associated TBI [83].

### 3.2. Stroke Aphasia

A number of abnormalities of degraded speech processing have been described in the context of aphasia following stroke. People with different forms of stroke-related aphasia have difficulties comprehending sentences spoken in an unfamiliar accent [85]. As might be anticipated, the profile is influenced by the type of aphasia (vascular insult) and the nature of the degraded speech manipulation: individuals with conduction aphasia and Wernicke's aphasia show a significantly smaller benefit from DAF than people with Broca's aphasia [177,178], while MAF was shown to improve speech rate and reduce dysfluency prolongations [86]. In patients with insular stroke, five of eight patients showed an abnormal dichotic digits test [84], and single case studies have demonstrated that people with stroke-related aphasia may have difficulty perceiving synthetic sentences with competing messages [179]. Together, these observations suggest that "informational masking" (Figure 2C) may be particularly disruptive to speech perception in stroke-related aphasia.

### 3.3. Parkinson's Disease

Parkinson's disease (PD), a neurodegenerative disorder caused primarily by the loss of dopaminergic neurons from the basal ganglia, is typically led by "extrapyramidal" motor symptoms including tremor, bradykinesia, and rigidity [180,181]. However, cognitive deficits are common in PD, with dementia affecting 50% of patients within 10 years of diagnosis [182]. The majority (70–90%) of individuals with PD also develop motor speech impairment [183]. Although PD is associated with objective hypophonia, people with PD overestimate the loudness of their own speech while they are speaking and in playback [184], and this is thought to be the mechanism of hypophonia due to impaired vocal feedback [185]. Responses to AAF paint a complex picture: whereas patients with PD may fail to modulate their own vocal volume under intensity altered auditory feedback [186], FAF may elicit significantly larger compensatory responses in people with PD than in



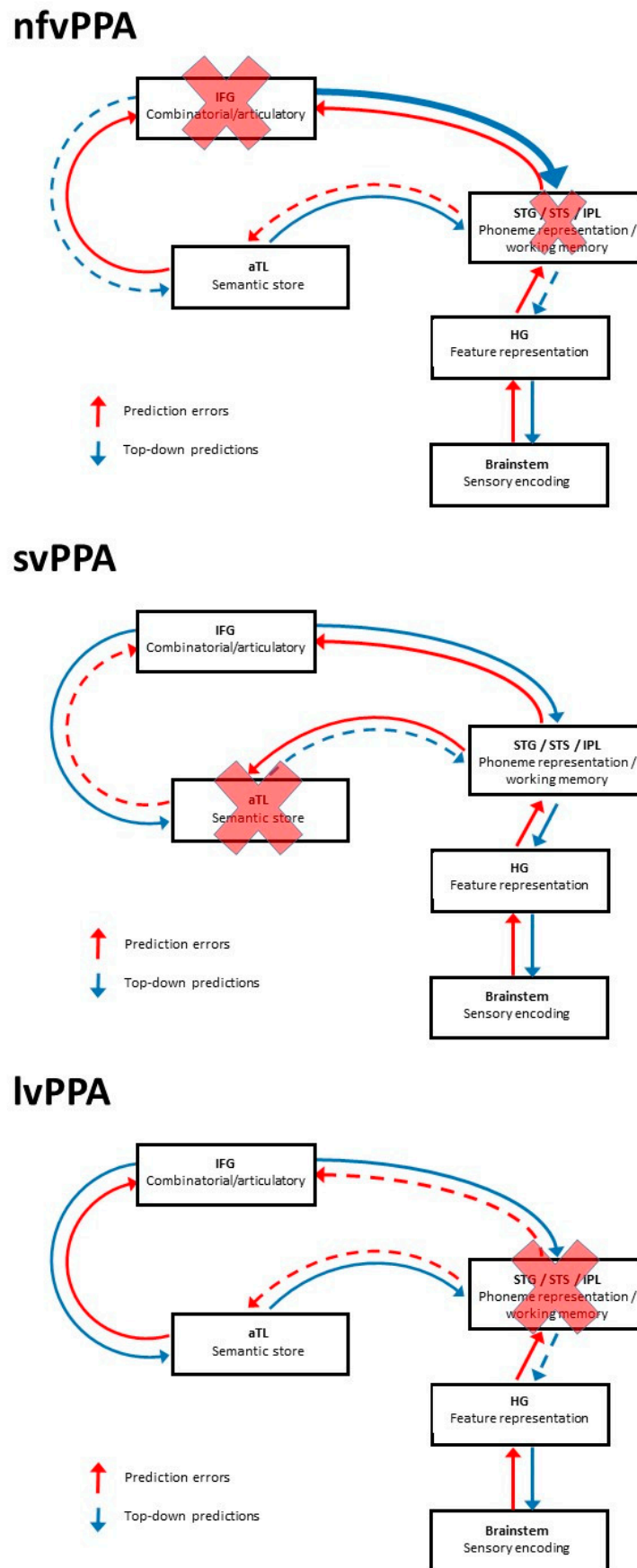
healthy controls [87,88,180,187,188], while DAF substantially improves speech intelligibility in some patients with PD [189]. FAF has differential effects according to whether the fundamental frequency or the first formant of the speech signal is altered [188], and the response to altered fundamental frequency correlates with voice pitch variability [180], suggesting that the response to AAF in PD is exquisitely dependent on the nature of the perturbation and its associated sensorimotor mapping. These effects could be interpreted as specific deficits in the predictive coding of auditory information, with impaired salience monitoring as well as over-reliance on sensory priors [190,191].

Taken together, the available evidence points to abnormal auditory-motor integration in PD that tends to impair the perception of degraded speech and to promote dysfunctional communication under challenging listening conditions. Candidate neuroanatomical substrates have been identified: enhanced evoked (P2) potentials in response to FAF in PD relative to healthy controls have been localised to activity in left superior and inferior frontal gyrus, premotor cortex, inferior parietal lobule, and superior temporal gyrus [180].

### 3.4. Alzheimer's Disease

Alzheimer's disease (AD), the most common form of dementia, is typically considered to be an amnesic clinical syndrome underpinned by the degeneration of posterior hippocampus, entorhinal cortex, posterior cingulate, medial and lateral parietal regions within the so-called "default mode network" [192,193]. People with AD have particular difficulty with dichotic digit identification tasks [89,194–196]. This is likely to reflect a more fundamental impairment of auditory scene analysis that also compromises speech-in-noise and speech-in-babble perception [90,197]. During the perception of their own name over background babble (the classical "cocktail party effect"), patients with AD were shown to have abnormally enhanced activation relative to healthy older controls in right supra-marginal gyrus [90]. Auditory scene analysis deficits are most striking in posterior cortical atrophy, the variant AD syndrome led by visuo-spatial impairment, further suggesting that posterior cortical regions within the core temporo-parietal network targeted by AD pathology play a critical pathophysiological role [198]. Speech-in-noise processing deficits may precede the onset of other symptoms in AD and may be a prodromal marker [199–201].

People with AD have difficulty understanding non-native accents [92,202] and sinewave speech (Figure 2G) [94] relative to healthy older individuals, and this has been linked using voxel-based morphometry to grey matter loss in left superior temporal cortex. Considered together with impairments of auditory scene analysis in AD, these findings could be interpreted to signify a fundamental lesion of the neural mechanisms that map degraded speech signals onto stored "templates" representing canonical auditory objects, such as phonemes. However, perceptual learning of sinewave speech has been shown to be intact in AD [94], and the comprehension of sinewave speech improves following the administration of an acetylcholinesterase inhibitor [203]. People with mild to moderate AD also show enhanced compensatory responses to FAF compared to age-matched controls [91]: this has been linked to reduced prefrontal activation and enhanced recruitment of right temporal cortices [204].



**Figure 3.** A simplified model of predictive coding of degraded speech processing in primary progressive aphasia (PPA), referenced to the healthy brain presented in Figure 1. The three major PPA

variant syndromes—nonfluent/agrammatic variant PPA (top panel); semantic variant PPA (middle panel) and logopenic variant PPA (bottom panel)—are each associated with a specific pattern of regional brain atrophy and/or dysfunction that is critical to the degraded speech processing network, implying that different PPA subtypes may be associated with specific profiles of degraded speech processing (see text for details). Boxes indicate processors that instantiate core speech decoding functions (see Figure 1), and arrows indicate their connections in the predictive coding framework, with the putative direction of information flow. In the case of nfvPPA, the emboldened descending arrow from IFG to STG signifies aberrantly increased precision of inflexible top-down priors (after Cope and Colleagues [93]), to date the most secure evidence for a predictive coding mechanism in the PPA spectrum; the status of the IPL locus in this syndrome is more tentative. Implicit in the model is the hypothesis that neurodegenerative pathologies will tend to disrupt stored neural templates (“priors”) and “prune” projections from heavily involved, higher order association cortical areas due to neuronal dropout (promoting inflexible top-down predictions), but also degrade the fidelity of signal traffic through sensory cortices (reducing sensory precision and promoting over-precise prediction errors) [15]. The relative prominence of these mechanisms will depend on the macro-network and local neural circuit anatomy of particular neurodegenerative pathologies. Proposed major loci of disruption caused by each PPA variant are indicated with crosses; dashed arrows arising from these damaged modules indicate disrupted information flow. aTL, anterior temporal lobe; HG, Heschl’s gyrus; IFG, inferior frontal gyrus; IPL, inferior parietal lobule; lvPPA, logopenic variant primary progressive aphasia; nfvPPA, non-fluent variant primary progressive aphasia; STG superior temporal gyrus; STS, superior temporal sulcus; svPPA, semantic variant primary progressive aphasia.

### 3.5. Primary Progressive Aphasia

Speech and language problems are leading features of the primary progressive aphasias (PPA). These “language-led dementias” constitute a heterogeneous group of disorders, comprising three cardinal clinico-anatomical syndromic variants. The nonfluent/agrammatic variant (nfvPPA) is characterised by disrupted speech and connected language production due to selective degeneration of a peri-Sylvian network centred on inferior frontal cortex and insula; the phenotype is quite variable between individual patients [205]. The semantic variant (svPPA) is characterised by the erosion of semantic memory due to selective degeneration of the semantic appraisal network in the antero-mesial (and particularly, the dominant) temporal lobe. The logopenic variant (lvPPA) is the language-led variant of AD and is characterised by anomia and impaired phonological working memory due to the degeneration of dominant temporo-parietal circuitry overlapping the circuits that are targeted in other AD variants [205,206]. All three major PPA syndromes have been shown to have clinically significant impairments of central auditory processing affecting speech comprehension [52,207–212]: together, these disorders constitute a paradigm for selective language network vulnerability and the impaired processing of degraded speech.

While people with AD have relatively greater difficulty processing less familiar non-native accents, particularly at the level of phrases and sentences, those with nfvPPA show a more pervasive pattern of impairment affecting more and less familiar accents at the level of single words [92]. People with nfvPPA and lvPPA show impaired understanding of sinewave speech relative to healthy controls and people with svPPA [94]. Patients with svPPA, however, show a significant identification advantage for more predictable (spoken number) over less predictable (spoken geographical place name) verbal signals after sinewave transformation, highlighting the important role of “top-down” contextual integration in degraded speech perception [94]. In this study, all PPA variants were shown to have intact perceptual learning of sinewave-degraded stimuli [94]. There is also evidence that at least some people with nfvPPA may be particularly susceptible to the effects of DAF [213].

The structural and functional neuroanatomy of degraded speech processing has been addressed in somewhat more detail in PPA than in other brain disorders. Using a MEG paradigm in which noise-vocoded words were presented to participants alongside written

text that either matched or mismatched the degraded words, Cope and colleagues [93] found that atrophy of left inferior frontal cortex in nfvPPA was associated with inflexible and delayed neural resolution of top-down predictions about incoming degraded speech signals in the setting of enhanced fronto-temporal coherence (frontal to temporal cortical connectivity), suggesting that the process of iterative reconciliation of top-down predictions with sensory prediction error takes longer to achieve in nfvPPA. Across the nfvPPA and healthy control groups, the precision of top-down predictions correlated with the magnitude of induced beta oscillations while frontal cortical beta power was enhanced in the nfvPPA group: this is in line with predictive coding accounts according to which beta band activity reflects the updating of perceptual predictions [47]. In joint voxel-based morphometric and functional MRI studies of a combined PPA cohort [214,215], Hardy and colleagues identified a substrate for impaired decoding of spectrally degraded phonemes in left supramarginal gyrus and posterior superior temporal cortex, most strikingly in lvPPA relative to healthy older individuals, whereas nfvPPA was associated with reduced sensitivity to sound stimulation in auditory cortex. Using voxel-based morphometry in a combined AD and PPA cohort, Hardy and colleagues [94] found that the overall accuracy of sine-wave speech identification was associated with grey matter volume in left temporo-parietal cortices, with grey matter correlates of increased speech predictability in left inferior frontal gyrus, top-down semantic decoding in left temporal pole and perceptual learning in left inferolateral post-central cortex. Such studies are beginning to define the alterations in “bottom-up” and “top-down” network mechanisms that jointly underpin impaired predictive decoding of degraded speech signals in neurodegenerative disease.

#### **4. A Predictive Coding Model of Degraded Speech Processing in Primary Progressive Aphasia**

Emerging evidence in PPA suggests a framework for applying predictive coding theory as outlined for the healthy brain (Figure 1) to formulate explicit pathophysiological hypotheses in these diseases. Such a framework could serve as a model for interpreting abnormalities of degraded speech processing in a wider range of brain disorders. This model is outlined in Figure 3.

According to this model, nfvPPA—which affects inferior frontal and more posterior peri-Sylvian cortices—is associated with a “double-hit” to the degraded speech processing network. The most clearly established consequence is overly precise, top-down predictions due to neuronal dysfunction and loss in inferior frontal cortex [93]. The top-down mechanism may be compounded by decreased signal fidelity (precision) due to abnormal auditory cortical representations [94,214,215]; however, this remains to be corroborated. The clinico-anatomical heterogeneity of nfvPPA is an important consideration here, implying that the mechanism may not be uniform between patients.

In svPPA, the primary focus of atrophy in anterior temporal lobe principally affects the top-down integration of contextual and stored semantic information. This reduces neural capacity to modify semantic predictions about less predictable verbal signals (i.e., priors are inaccurate), in line with experimental observations [94].

In lvPPA, atrophy predominantly involving temporo-parietal cortex is anticipated to impair phonemic decoding and earlier stages in the representation of acoustic features in auditory cortex and brainstem due to altered top-down influences from the temporal parietal junction on auditory cortex and brainstem: this could be via altered precision weighting of prediction errors conveyed by the auditory efferent pathways, or inaccurate priors. This formulation has some experimental support [211,215].

#### **5. Therapeutic Approaches**

Improved understanding of the pathophysiology of degraded speech processing in brain disorders is the path to effective therapeutic interventions. Several physiologically informed therapeutic approaches are in current use or have shown early promise. In a clinical context, it is important not to overlook ancillary nonverbal strategies to compensate for reduced capacity to process degraded speech: examples include the minimisation of

environmental noise, training speakers to face the patient to maximise visual support and aid speech sound discrimination, and using gestures to support semantic context [216,217]. A related and crucial theme in designing therapies tailored to individuals is to acknowledge the various background factors—whether deleterious or potentially protective—that influence degraded speech processing (see Section 2).

More specifically, the finding that perceptual learning of degraded speech is retained in diverse brain disorders including dementias [94] and stroke aphasia [218,219] offers the exciting prospect of designing training interventions to harness neural plasticity in these conditions. Thus far, most work in this line has been directed to improving understanding of challenging speech (in particular, speech-in-noise) in older adults with peripheral hearing loss. Training programmes have targeted different levels of speech analysis—words and sentences—and different cognitive operations—attentional and perceptuo-motor—and have shown improved perception of trained stimuli, though this is less consistently extended to untrained stimuli (the grail of work of this kind: Bieber and Gordon-Salant [220]). On the other hand, there is some evidence that training on degraded environmental sounds may generalise to improved perception of degraded speech [221]. Enhanced perceptual learning through the facilitation of regional neuronal plasticity also provides a rationale for the transcranial stimulation of key cortical language areas, such as inferior frontal gyrus [222]. Potentially, a technique such as transcranial temporal interference stimulation could selectively target deep brain circuitry and feedforward or feedback connections [223] to probe specific pathophysiological mechanisms of degraded speech processing in particular brain disorders (see Figure 3).

Other therapeutic approaches have focused on training auditory working memory. These have yielded mixed results [224], though interestingly, the training of musical working memory may show a cross over benefit for speech-in-noise recognition [225,226]. A combined auditory cognitive training programme, potentially incorporating musical skills, may be the most rational strategy [220,227].

Pharmacological approaches are potentially complementary to behavioural interventions or transcranial stimulation. In healthy individuals, dopamine has been shown to enhance the perception of spectrally shifted noise-vocoded speech [228]. In patients with AD, acetylcholinesterase inhibition ameliorates the understanding of sinewave speech [203]. Indeed, degraded speech processing might prove to be a rapid and sensitive biomarker of therapeutic efficacy in brain disorders. At present, the objectives of therapy differ quite sharply between disorders such as stroke, where there is a prospect of sustained improvement in functional adaptation in at least some patients, and neurodegenerative conditions such as PPA, where any benefit is ultimately temporary due to the progressive nature of the underlying pathology. However, it is crucial to develop interventions that enhance degraded speech processing (and other ecologically relevant aspects of communication) in neurodegenerative disease, not only to maximise patients' daily life functioning but also with a future view to using such techniques adjunctively with disease modifying therapies as these become available. Ultimately, irrespective of the brain pathology, it will be essential to determine how far improvements on degraded speech processing tasks translate to improved communication in daily life.

## 6. A Critique of the Predictive Coding Paradigm of Degraded Speech Processing

Like any scientific paradigm, predictive coding demands a critical evaluation of falsifiable hypotheses. The issues in relation to the auditory system have been usefully reviewed previously in Heilbron and Chait [14]. While it is self-evident that the brain is engaged in making and evaluating predictions, there are two broad questions here, in respect of degraded speech processing that could address and direct future experiments.

Firstly, to what extent is the processing of degraded speech generically underpinned by predictive coding? While the predictive coding paradigm is committed to finding optimal computational solutions to perceptual perturbations, much natural language use relies on acoustic or articulatory characteristics that are “sub-optimal” [229]. More fundamentally,



as the raw material is contributed much to human thought, the combinatorial space of language is essentially infinite: we routinely produce entirely novel utterances and are called upon to understand the novel utterances of others, whereas predictive coding rests on a relatively simple computational “logic” [230]. Identifying the limits of predictive coding in the face of emergent linguistic combinatorial complexity therefore presents a major challenge—a challenge encountered even for the combinatorially much more constrained phenomenon of music [231]. Future experiments will need to define core predictive coding concepts such as “priors”, “error” and “precision” in terms of degraded speech processing, as well as disambiguate the roles of semantic and phonological representations, selective attention and verbal working memory in such processing, ideally by manipulating these components independently [14,191,232–234].

Secondly, how is the predictive coding of degraded speech instantiated in the brain? Although macroscopic neural network substrates that could support the required hierarchical and reciprocal information exchange have been delineated (Figure 1), the predictive coding paradigm stipulates quite specifically how key elements such as “prediction generators” and “error detectors” are organised, both at the level of large-scale networks and local cortical circuits [14,235]. Neuroimaging techniques such as spectral dynamic causal modelling, MEG and high-field fMRI constitute particularly powerful and informative tools with which to interrogate the responsible neural elements and their interplay [14,236]: such techniques can capture both interactions between macroscopic brain modules and structure–function relationships at the level of individual cortical laminae, where the core circuit components of predictive coding are hypothesised to reside.

## 7. Conclusions and Future Directions

The perception and ultimately understanding of degraded speech relies upon flexible and dynamic neural interactions across distributed brain networks. These physiological and anatomical substrates are intrinsically vulnerable to the disruptive effects of brain disorders, particularly neurodegenerative pathologies that preferentially blight the core circuitry responsible for representing and decoding speech signals. Predictive coding offers an intuitive framework within which to consider degraded speech processing, both in the healthy brain (Figure 1) and in brain disorders (Figure 3). Different forms of speech signal degradation are likely a priori to engage neural network nodes and connections differentially and may therefore reveal distinct phenotypes of degraded speech processing that are specific for particular neuropathological processes. However, this will require substantiation in future systematic, head-to-head comparisons between paradigms (Table 1, Figure 2) and pathologies (Table 2, Figure 3). It will be particularly pertinent to design neuropsychological and neuroimaging experiments to interrogate the basic assumptions of predictive coding theory, as sketched above.

From a neurobiological perspective, building on the model outlined for PPA in Figure 3, degraded speech is an attractive candidate probe of pathophysiological mechanisms in brain disease. For example, it has been proposed that lvPPA is associated with the “blurring” of phonemic representational boundaries [211]: this would predict that phonemic restoration (Figure 2) is critically impaired in lvPPA. Further, several lines of evidence implicate disordered efferent regulation of auditory signal analysis in the pathogenesis of nvfPPA [93,210]: this could be explored directly by independently varying the precision of incoming speech signals and central gain (for example, using dichotic listening techniques). Temporally sensitive neurophysiological and functional neuroimaging techniques such as EEG and MEG will be required to define the dynamic oscillatory neural mechanisms by which brain pathologies disrupt degraded speech perception. Proteinopathies are anticipated to have separable MEG signatures based on differential patterns of cortical laminar involvement [237]. By extension from the “lesion studies” of classical neurolinguistics, the study of clinical disorders may ultimately illuminate the cognitive and neural organisation of degraded speech processing in the normal brain [93], by pinpointing critical elements and demonstrating how dissociable processing steps are mutually related.

From a clinical perspective, the processing of degraded speech (as a sensitive index of neural circuit integrity) might facilitate the early diagnosis of brain disorders. Neurodegenerative pathologies, in particular, often elude diagnosis in their early stages: degraded speech stimuli might be adapted to constitute dynamic, physiological “stress tests” to detect such pathologies. Similar pathophysiological principles should inform the design of behavioural and pharmacological therapies, such as those that harness neural plasticity: looking forward, such interventions could be particularly powerful if combined with disease modifying therapies, as integrated cognitive neurorehabilitation strategies motivated by neurobiological principles.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/2076-3425/11/3/394/s1>.

**Author Contributions:** Conceptualization, J.J., E.B., S.W., J.C.S.J., A.V., R.S.W., C.R.M., J.D.W. and C.J.D.H.; Writing—Original Draft Preparation, J.J., J.D.W. and C.J.D.H. Writing—Review & Editing, J.J., E.B., S.W., J.C.S.J., A.V., R.S.W., C.R.M., J.D.W. and C.J.D.H.; Visualisation, J.J., E.B., S.W., C.R.M., J.D.W. and C.J.D.H.; Supervision, R.S.W., J.D.W. and C.J.D.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** The Dementia Research Centre is supported by Alzheimer’s Research UK, Brain Research Trust, and The Wolfson Foundation. This work was supported by the Alzheimer’s Society (grant AS-PG-16-007 to J.D.W.), the National Institute for Health Research University College London Hospitals Biomedical Research Centre, and the University College London Leonard Wolfson Experimental Neurology Centre (grant PR/ylr/18575). J.J. is supported by a Frontotemporal Dementia Research Studentship in Memory of David Blechner (funded through The National Brain Appeal). E.B. was supported by a Brain Research UK PhD Studentship. The Preventative Neurology Unit is supported by a grant from Bart’s Charity. J.C.S.J. is supported by an Association of British Neurologists-Guarantors of Brain Clinical Research Training Fellowship. A.V. is supported by a National Institute for Health Research Development Skills Enhancement Award (this paper presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care). R.S.W. is supported by a Wellcome Clinical Research Career Development Fellowship (201567/Z/16/Z) and UCLH Biomedical Research Centre Grant (BRC302/NS/RW/101410). C.J.D.H. was supported by an Action on Hearing Loss–Dunhill Medical Trust Pauline Ashley Fellowship (grant PA23\_Hardy).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors thank Georgia Peakman for lending her voice in the making of the supplementary audio files, and our reviewers for their helpful critique of the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Hervais-Adelman, A.; Davis, M.H.; Johnsrude, I.S.; Carlyon, R.P. Perceptual learning of noise vocoded words: Effects of feedback and lexicality. *J. Exp. Psychol. Hum. Percept. Perform.* **2008**, *34*, 460–474. [[CrossRef](#)] [[PubMed](#)]
2. Mattys, S.L.; Davis, M.H.; Bradlow, A.R.; Scott, S.K. Speech recognition in adverse conditions: A review. *Lang. Cogn. Process.* **2012**, *27*, 953–978. [[CrossRef](#)]
3. Adank, P.; Davis, M.H.; Hagoort, P. Neural dissociation in processing noise and accent in spoken language comprehension. *Neuropsychologia* **2012**, *50*, 77–84. [[CrossRef](#)] [[PubMed](#)]
4. Davis, M.H.; Sohoglu, E. Three functions of prediction error for Bayesian inference in speech perception. In *The Cognitive Neurosciences*, 6th ed.; Gazzaniga, M., Mangun, R.D.P., Eds.; MIT Press: Cambridge, MA, USA, 2020; p. 177.
5. Hickok, G.; Poeppel, D. The cortical organization of speech processing. *Nat. Rev. Neurosci.* **2007**, *8*, 393–402. [[CrossRef](#)] [[PubMed](#)]
6. Okada, K.; Rong, F.; Venezia, J.; Matchin, W.; Hsieh, I.H.; Saberi, K.; Serences, J.T.; Hickok, G. Hierarchical organization of human auditory cortex: Evidence from acoustic invariance in the response to intelligible speech. *Cereb. Cortex* **2010**, *20*, 2486–2495. [[CrossRef](#)]
7. Peelle, J.E. Hierarchical processing for speech in human auditory cortex and beyond. *Front. Hum. Neurosci.* **2010**. [[CrossRef](#)]

8. Scott, S.K.; Blank, C.C.; Rosen, S.; Wise, R.J. Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* **2000**, *123*, 2400–2406. [[CrossRef](#)]
9. Bregman, A.S. *Auditory Scene Analysis: The Perceptual Organization of Sound*; MIT Press: Cambridge, MA, USA, 1990.
10. Davis, M.H.; Johnsrude, I.S. Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hear. Res.* **2007**, *229*, 132–147. [[CrossRef](#)]
11. Ainley, V.; Apps, M.A.J.; Fotopoulou, A.; Tsakiris, M. ‘Bodily precision’: A predictive coding account of individual differences in interoceptive accuracy. *Philos. Trans. R. Soc. B Biol. Sci.* **2016**, *371*, 20160003. [[CrossRef](#)]
12. Friston, K. A theory of cortical responses. *Philos. Trans. R. Soc. B Biol. Sci.* **2005**, *360*, 815–836. [[CrossRef](#)]
13. Arnal, L.H.; Wyart, V.; Giraud, A.-L. Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nat. Neurosci.* **2011**, *14*, 797–801. [[CrossRef](#)]
14. Heilbron, M.; Chait, M. Great Expectations: Is there Evidence for Predictive Coding in Auditory Cortex. *Neuroscience* **2018**, *389*, 54–73. [[CrossRef](#)] [[PubMed](#)]
15. Kocagoncu, E.; Klimovich-Gray, A.; Hughes, L.; Rowe, J.B. Evidence and implications of abnormal predictive coding in dementia. *arXiv* **2020**, arXiv:2006.06311.
16. Kashino, M. Phonemic restoration: The brain creates missing speech sounds. *Acoust. Sci. Technol.* **2006**, *6*, 4. [[CrossRef](#)]
17. Sohoglu, E.; Peelle, J.E.; Carlyon, R.P.; Davis, M.H. Top-down influences of written text on perceived clarity of degraded speech. *J. Exp. Psychol. Hum. Percept. Perform.* **2014**, *40*, 186–199. [[CrossRef](#)]
18. Alain, C.; Du, Y.; Bernstein, L.J.; Barten, T.; Banai, K. Listening under difficult conditions: An activation likelihood estimation meta-analysis. *Hum. Brain Mapp.* **2018**, *39*, 2695–2709. [[CrossRef](#)] [[PubMed](#)]
19. Di Liberto, G.M.; Lalor, E.C.; Millman, R.E. Causal cortical dynamics of a predictive enhancement of speech intelligibility. *NeuroImage* **2018**, *166*, 247–258. [[CrossRef](#)]
20. Johnsrude, I.S. The neuropsychological consequences of temporal lobe lesions. In *Cognitive Deficits in Brain Disorders*; Harrison, J.E., Owen, A.M., Eds.; Martin Dunitz Ltd.: London, UK, 2002; pp. 37–57.
21. Strange, B.A.; Otten, L.J.; Josephs, O.; Rugg, M.D.; Dolan, R.J. Dissociable Human Perirhinal, Hippocampal, and Parahippocampal Roles during Verbal Encoding. *J. Neurosci.* **2002**, *22*, 523–528. [[CrossRef](#)] [[PubMed](#)]
22. Binney, R.J.; Embleton, K.V.; Jefferies, E.; Parker, G.J.; Ralph, M.A. The ventral and inferolateral aspects of the anterior temporal lobe are crucial in semantic memory: Evidence from a novel direct comparison of distortion-corrected fMRI, rTMS, and semantic dementia. *Cereb. Cortex* **2010**, *20*, 2728–2738. [[CrossRef](#)]
23. Lambon Ralph, M.A.; Patterson, K. Generalization and Differentiation in Semantic Memory. *Ann. N. Y. Acad. Sci.* **2008**, *1124*, 61–76. [[CrossRef](#)]
24. Pobric, G.; Jefferies, E.; Ralph, M.A.L. Anterior temporal lobes mediate semantic representation: Mimicking semantic dementia by using rTMS in normal participants. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 20137–20141. [[CrossRef](#)] [[PubMed](#)]
25. Awad, M.; Warren, J.E.; Scott, S.K.; Turkheimer, F.E.; Wise, R.J.S. A Common System for the Comprehension and Production of Narrative Speech. *J. Neurosci.* **2007**, *27*, 11455–11464. [[CrossRef](#)]
26. Rodd, J.M.; Davis, M.H.; Johnsrude, I.S. The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cereb. Cortex* **2005**, *15*, 1261–1269. [[CrossRef](#)] [[PubMed](#)]
27. Erb, J.; Henry, M.J.; Eisner, F.; Obleser, J. The Brain Dynamics of Rapid Perceptual Adaptation to Adverse Listening Conditions. *J. Neurosci.* **2013**, *33*, 10688–10697. [[CrossRef](#)]
28. Hagoort, P. On Broca, brain, and binding: A new framework. *Trends Cogn. Sci.* **2005**, *9*, 416–423. [[CrossRef](#)] [[PubMed](#)]
29. Wild, C.J.; Yusuf, A.; Wilson, D.E.; Peelle, J.E.; Davis, M.H.; Johnsrude, I.S. Effortful Listening: The Processing of Degraded Speech Depends Critically on Attention. *J. Neurosci.* **2012**, *32*, 14010–14021. [[CrossRef](#)]
30. Adank, P.; Nuttall, H.E.; Banks, B.; Kennedy-Higgins, D. Neural bases of accented speech perception. *Front. Hum. Neurosci.* **2015**, *9*. [[CrossRef](#)]
31. Hashimoto, Y.; Sakai, K.L. Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: An fMRI study. *Hum. Brain Mapp.* **2003**, *20*, 22–28. [[CrossRef](#)]
32. Hirnstein, M.; Westerhausen, R.; Korsnes, M.S.; Hugdahl, K. Sex differences in language asymmetry are age-dependent and small: A large-scale, consonant-vowel dichotic listening study with behavioral and fMRI data. *Cortex* **2013**, *49*, 1910–1921. [[CrossRef](#)]
33. Davis, M.H.; Johnsrude, I.S. Hierarchical Processing in Spoken Language Comprehension. *J. Neurosci.* **2003**, *23*, 3423–3431. [[CrossRef](#)] [[PubMed](#)]
34. Hervais-Adelman, A.G.; Carlyon, R.P.; Johnsrude, I.S.; Davis, M.H. Brain regions recruited for the effortful comprehension of noise-vocoded words. *Lang. Cogn. Process.* **2012**, *27*, 1145–1166. [[CrossRef](#)]
35. Scott, S.K.; Rosen, S.; Lang, H.; Wise, R.J.S. Neural correlates of intelligibility in speech investigated with noise vocoded speech—A positron emission tomography study. *J. Acoust. Soc. Am.* **2006**, *120*, 1075–1083. [[CrossRef](#)] [[PubMed](#)]
36. Sunami, K.; Ishii, A.; Takano, S.; Yamamoto, H.; Sakashita, T.; Tanaka, M.; Watanabe, Y.; Yamane, H. Neural mechanisms of phonemic restoration for speech comprehension revealed by magnetoencephalography. *Brain Res.* **2013**, *1537*, 164–173. [[CrossRef](#)] [[PubMed](#)]
37. Möttönen, R.; Calvert, G.A.; Jääskeläinen, I.P.; Matthews, P.M.; Thesen, T.; Tuomainen, J.; Sams, M. Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. *Neuroimage* **2006**, *30*, 563–569. [[CrossRef](#)] [[PubMed](#)]

38. Hwang, J.-H.; Li, C.-W.; Wu, C.-W.; Chen, J.-H.; Liu, T.-C. Aging Effects on the Activation of the Auditory Cortex during Binaural Speech Listening in White Noise: An fMRI Study. *Audiol. Neurotol.* **2007**, *12*, 285–294. [[CrossRef](#)] [[PubMed](#)]
39. Adank, P.; Devlin, J.T. On-line plasticity in spoken sentence comprehension: Adapting to time-compressed speech. *Neuroimage* **2010**, *49*, 1124–1132. [[CrossRef](#)] [[PubMed](#)]
40. Hartwigsen, G.; Golombek, T.; Obleser, J. Repetitive transcranial magnetic stimulation over left angular gyrus modulates the predictability gain in degraded speech comprehension. *Cortex* **2015**, *68*, 100–110. [[CrossRef](#)]
41. Shahin, A.J.; Bishop, C.W.; Miller, L.M. Neural mechanisms for illusory filling-in of degraded speech. *NeuroImage* **2009**, *44*, 1133–1143. [[CrossRef](#)] [[PubMed](#)]
42. Guediche, S.; Reilly, M.; Santiago, C.; Laurent, P.; Blumstein, S.E. An fMRI study investigating effects of conceptually related sentences on the perception of degraded speech. *Cortex* **2016**, *79*, 57–74. [[CrossRef](#)]
43. Leonard, M.K.; Baud, M.O.; Sjerps, M.J.; Chang, E.F. Perceptual restoration of masked speech in human cortex. *Nat. Commun.* **2016**, *7*, 13619. [[CrossRef](#)]
44. Becker, R.; Pefkou, M.; Michel, C.M.; Hervais-Adelman, A.G. Left temporal alpha-band activity reflects single word intelligibility. *Front. Syst. Neurosci.* **2013**, *7*, 121. [[CrossRef](#)]
45. Arnal, L.H.; Doelling, K.B.; Poeppel, D. Delta-Beta Coupled Oscillations Underlie Temporal Prediction Accuracy. *Cereb. Cortex* **2015**, *25*, 3077–3085. [[CrossRef](#)] [[PubMed](#)]
46. Hovsepian, S.; Olasagasti, I.; Giraud, A.-L. Combining predictive coding and neural oscillations enables online syllable recognition in natural speech. *Nat. Commun.* **2020**, *11*. [[CrossRef](#)]
47. Sedley, W.; Gander, P.E.; Kumar, S.; Kovach, C.K.; Oya, H.; Kawasaki, H.; Howard, M.A.; Griffiths, T.D. Neural signatures of perceptual inference. *eLife* **2016**, *5*. [[CrossRef](#)] [[PubMed](#)]
48. Hakonen, M.; May, P.J.C.; Alho, J.; Alku, P.; Jokinen, E.; Jääskeläinen, I.P.; Tiitinen, H. Previous exposure to intact speech increases intelligibility of its digitally degraded counterpart as a function of stimulus complexity. *Neuroimage* **2016**, *125*, 131–143. [[CrossRef](#)]
49. Sohoglu, E.; Davis, M.H. Perceptual learning of degraded speech by minimizing prediction error. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E1747–E1756. [[CrossRef](#)]
50. Samuel, A.G. Speech Perception. *Annu. Rev. Psychol.* **2011**, *62*, 49–72. [[CrossRef](#)]
51. Warren, J.D.; Rohrer, J.D.; Schott, J.M.; Fox, N.C.; Hardy, J.; Rossor, M.N. Molecular nexopathies: A new paradigm of neurodegenerative disease. *Trends Neurosci.* **2013**, *36*, 561–569. [[CrossRef](#)]
52. Hardy, C.J.; Marshall, C.R.; Golden, H.L.; Clark, C.N.; Mummery, C.J.; Griffiths, T.D.; Bamiou, D.E.; Warren, J.D. Hearing and dementia. *J. Neurol.* **2016**, *263*, 2339–2354. [[CrossRef](#)] [[PubMed](#)]
53. Fletcher, P.D.; Downey, L.E.; Augustus, J.L.; Hailstone, J.C.; Tyndall, M.H.; Cifelli, A.; Schott, J.M.; Warrington, E.K.; Warren, J.D. Agnosia for accents in primary progressive aphasia. *Neuropsychologia* **2013**, *51*, 1709–1715. [[CrossRef](#)]
54. Floccia, C.; Butler, J.; Goslin, J.; Ellis, L. Regional and Foreign Accent Processing in English: Can Listeners Adapt? *J. Psycholinguist. Res.* **2009**, *38*, 379–412. [[CrossRef](#)] [[PubMed](#)]
55. Taylor, B. Speech-in-noise tests: How and why to include them in your basic test battery. *Hear. J.* **2003**, *56*, 42–46. [[CrossRef](#)]
56. Lidestam, B.; Holgersson, J.; Moradi, S. Comparison of informational vs. energetic masking effects on speechreading performance. *Front. Psychol.* **2014**, *5*, 639. [[CrossRef](#)] [[PubMed](#)]
57. Warren, R.M. Perceptual restoration of missing speech sounds. *Science* **1970**, *167*, 392–393. [[CrossRef](#)] [[PubMed](#)]
58. Davis, M.H.; Johnsrude, I.S.; Hervais-Adelman, A.; Taylor, K.; McGettigan, C. Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *J. Exp. Psychol. Gen.* **2005**, *134*, 222–241. [[CrossRef](#)]
59. Shannon, R.V.; Zeng, F.G.; Kamath, V.; Wygonski, J.; Ekelid, M. Speech Recognition with Primarily Temporal Cues. *Science* **1995**, *270*, 303–304. [[CrossRef](#)]
60. Dupoux, E.; Green, K. Perceptual adjustment to highly compressed speech: Effects of talker and rate changes. *J. Exp. Psychol. Hum. Percept. Perform.* **1997**, *23*, 914–927. [[CrossRef](#)]
61. Fausto, B.A.; Badana, A.N.S.; Arnold, M.L.; Lister, J.J.; Edwards, J.D. Comparison of Subjective and Objective Measures of Hearing, Auditory Processing, and Cognition Among Older Adults With and Without Mild Cognitive Impairment. *J. Speech Lang. Hear. Res.* **2018**, *61*, 945–956. [[CrossRef](#)]
62. Foulke, E.; Sticht, T.G. Review of research on the intelligibility and comprehension of accelerated speech. *Psychol. Bull.* **1969**, *72*, 50–62. [[CrossRef](#)]
63. Remez, R.; Rubin, P.; Pisoni, D.; Carrell, T. Speech perception without traditional speech cues. *Science* **1981**, *212*, 947–949. [[CrossRef](#)]
64. Barker, J.; Cooke, M. Is the sine-wave speech cocktail party worth attending? *Speech Commun.* **1999**, *27*, 159–174. [[CrossRef](#)]
65. Bent, T.; Bradlow, A.R. The interlanguage speech intelligibility benefit. *J. Acoust. Soc. Am.* **2003**, *114*, 1600–1610. [[CrossRef](#)]
66. Clarke, C.M.; Garrett, M.F. Rapid adaptation to foreign-accented English. *J. Acoust. Soc. Am.* **2004**, *116*, 3647–3658. [[CrossRef](#)] [[PubMed](#)]
67. Siegel, G.M.; Pick, H.L., Jr. Auditory feedback in the regulation of voice. *J. Acoust. Soc. Am.* **1974**, *56*, 1618–1624. [[CrossRef](#)]
68. Jones, J.A.; Munhall, K.G. Perceptual calibration of F0 production: Evidence from feedback perturbation. *J. Acoust. Soc. Am.* **2000**, *108*, 1246–1251. [[CrossRef](#)]



69. Donath, T.M.; Natke, U.; Kalveram, K.T. Effects of frequency-shifted auditory feedback on voice F0 contours in syllables. *J. Acoust. Soc. Am.* **2002**, *111*, 357–366. [[CrossRef](#)]
70. Stuart, A.; Kalinowski, J.; Rastatter, M.P.; Lynch, K. Effect of delayed auditory feedback on normal speakers at two speech rates. *J. Acoust. Soc. Am.* **2002**, *111*, 2237–2241. [[CrossRef](#)] [[PubMed](#)]
71. Moray, N. Attention in Dichotic Listening: Affective Cues and the Influence of Instructions. *Q. J. Exp. Psychol.* **1959**, *11*, 56–60. [[CrossRef](#)]
72. Lewis, J.L. Semantic processing of unattended messages using dichotic listening. *J. Exp. Psychol.* **1970**, *85*, 225–228. [[CrossRef](#)] [[PubMed](#)]
73. Ding, N.; Simon, J.Z. Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *J. Neurophysiol.* **2012**, *107*, 78–89. [[CrossRef](#)]
74. Samuel, A.G. Phonemic restoration: Insights from a new methodology. *J. Exp. Psychol. Gen.* **1981**, *110*, 474–494. [[CrossRef](#)]
75. Pichora-Fuller, M.K.; Schneider, B.A.; Daneman, M. How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* **1995**, *97*, 593–608. [[CrossRef](#)]
76. Parbery-Clark, A.; Skoe, E.; Lam, C.; Kraus, N. Musician enhancement for speech-in-noise. *Ear Hear.* **2009**, *30*, 653–661. [[CrossRef](#)] [[PubMed](#)]
77. Anderson, S.; White-Schwoch, T.; Parbery-Clark, A.; Kraus, N. A dynamic auditory-cognitive system supports speech-in-noise perception in older adults. *Hear. Res.* **2013**, *300*, 18–32. [[CrossRef](#)] [[PubMed](#)]
78. Poldrack, R.A.; Temple, E.; Protopapas, A.; Nagarajan, S.; Tallal, P.; Merzenich, M.; Gabrieli, J.D. Relations between the neural bases of dynamic auditory processing and phonological processing: Evidence from fMRI. *J. Cogn. Neurosci.* **2001**, *13*, 687–697. [[CrossRef](#)]
79. Peelle, J.E.; McMillan, C.; Moore, P.; Grossman, M.; Wingfield, A. Dissociable patterns of brain activity during comprehension of rapid and syntactically complex speech: Evidence from fMRI. *Brain Lang.* **2004**, *91*, 315–325. [[CrossRef](#)] [[PubMed](#)]
80. Gallun, F.J.; Diedesch, A.C.; Kubli, L.R.; Walden, T.C.; Folmer, R.L.; Lewis, M.S.; McDermott, D.J.; Fausti, S.A.; Leek, M.R. Performance on tests of central auditory processing by individuals exposed to high-intensity blasts. *J. Rehabil. Res. Dev.* **2012**, *49*, 1005. [[CrossRef](#)] [[PubMed](#)]
81. Saunders, G.H.; Frederick, M.T.; Arnold, M.; Silverman, S.; Chisolm, T.H.; Myers, P. Auditory difficulties in blast-exposed Veterans with clinically normal hearing. *J. Rehabil. Res. Dev.* **2015**, *52*, 343–360. [[CrossRef](#)]
82. Gallun, F.J.; Lewis, M.S.; Folmer, R.L.; Hutter, M.; Papesch, M.A.; Belding, H.; Leek, M.R. Chronic effects of exposure to high-intensity blasts: Results of tests of central auditory processing. *J. Rehabil. Res. Dev.* **2016**, *53*, 15. [[CrossRef](#)]
83. Papesch, M.A.; Elliott, J.E.; Callahan, M.L.; Storzbach, D.; Lim, M.M.; Gallun, F.J. Blast Exposure Impairs Sensory Gating: Evidence from Measures of Acoustic Startle and Auditory Event-Related Potentials. *J. Neurotrauma* **2019**, *36*, 702–712. [[CrossRef](#)]
84. Bamio, D.E.; Musiek, F.E.; Stow, I.; Stevens, J.; Cipolotti, L.; Brown, M.M.; Luxon, L.M. Auditory temporal processing deficits in patients with insular stroke. *Neurology* **2006**, *67*, 614–619. [[CrossRef](#)]
85. Dunton, J.; Bruce, C.; Newton, C. Investigating the impact of unfamiliar speaker accent on auditory comprehension in adults with aphasia. *Int. J. Lang. Commun. Disord.* **2011**, *46*, 63–73. [[CrossRef](#)]
86. Jacks, A.; Haley, K.L. Auditory Masking Effects on Speech Fluency in Apraxia of Speech and Aphasia: Comparison to Altered Auditory Feedback. *J. Speech Lang. Hear. Res.* **2015**, *58*, 1670–1686. [[CrossRef](#)] [[PubMed](#)]
87. Liu, H.; Wang, E.Q.; Metman, L.V.; Larson, C.R. Vocal responses to perturbations in voice auditory feedback in individuals with Parkinson’s disease. *PLoS ONE* **2012**, *7*, e33629. [[CrossRef](#)]
88. Chen, X.; Zhu, X.; Wang, E.Q.; Chen, L.; Li, W.; Chen, Z.; Liu, H. Sensorimotor control of vocal pitch production in Parkinson’s disease. *Brain Res.* **2013**, *1527*, 99–107. [[CrossRef](#)]
89. Gates, G.A.; Anderson, M.L.; Feeney, M.P.; McCurry, S.M.; Larson, E.B. Central Auditory Dysfunction in Older Persons With Memory Impairment or Alzheimer Dementia. *Arch. Otolaryngol. Head Neck Surg.* **2008**, *134*, 771. [[CrossRef](#)]
90. Golden, H.L.; Agustus, J.L.; Goll, J.C.; Downey, L.E.; Mummery, C.J.; Schott, J.M.; Crutch, S.J.; Warren, J.D. Functional neuroanatomy of auditory scene analysis in Alzheimer’s disease. *Neuroimage Clin.* **2015**, *7*, 699–708. [[CrossRef](#)] [[PubMed](#)]
91. Ranasinghe, K.G.; Gill, J.S.; Kothare, H.; Beagle, A.J.; Mizuiri, D.; Honma, S.M.; Gorno-Tempini, M.L.; Miller, B.L.; Vessel, K.A.; Nagarajan, S.S.; et al. Abnormal vocal behavior predicts executive and memory deficits in Alzheimer’s disease. *Neurobiol. Aging* **2017**, *52*, 71–80. [[CrossRef](#)]
92. Hailstone, J.C.; Ridgway, G.R.; Bartlett, J.W.; Goll, J.C.; Crutch, S.J.; Warren, J.D. Accent processing in dementia. *Neuropsychologia* **2012**, *50*, 2233–2244. [[CrossRef](#)]
93. Cope, T.E.; Sohoglu, E.; Sedley, W.; Patterson, K.; Jones, P.S.; Wiggins, J.; Dawson, C.; Grube, M.; Carlyon, R.P.; Griffiths, T.D.; et al. Evidence for causal top-down frontal contributions to predictive processes in speech perception. *Nat. Commun.* **2017**, *8*, 2154. [[CrossRef](#)] [[PubMed](#)]
94. Hardy, C.J.D.; Marshall, C.R.; Bond, R.L.; Russell, L.L.; Dick, K.; Ariti, C.; Thomas, D.L.; Ross, S.J.; Agustus, J.L.; Crutch, S.J.; et al. Retained capacity for perceptual learning of degraded speech in primary progressive aphasia and Alzheimer’s disease. *Alzheimers Res. Ther.* **2018**, *10*, 70. [[CrossRef](#)] [[PubMed](#)]
95. Gordon-Salant, S.; Fitzgibbons, P.J. Temporal factors and speech recognition performance in young and elderly listeners. *J. Speech Hear. Res.* **1993**, *36*, 1276–1285. [[CrossRef](#)]



96. Frisina, D.R.; Frisina, R.D. Speech recognition in noise and presbycusis: Relations to possible neural mechanisms. *Hear. Res.* **1997**, *106*, 95–104. [[CrossRef](#)]
97. Ross, B.; Dobri, S.; Schumann, A. Speech-in-noise understanding in older age: The role of inhibitory cortical responses. *Eur. J. Neurosci.* **2020**, *51*, 891–908. [[CrossRef](#)]
98. Roth, T.N. Aging of the auditory system. In *Handbook of Clinical Neurology*; Elsevier: Amsterdam, The Netherlands, 2015; Volume 129, pp. 357–373.
99. Bidelman, G.M.; Howell, M. Functional changes in inter- and intra-hemispheric cortical processing underlying degraded speech perception. *Neuroimage* **2016**, *124*, 581–590. [[CrossRef](#)]
100. Henry, M.J.; Herrmann, B.; Kunke, D.; Obleser, J. Aging affects the balance of neural entrainment and top-down neural modulation in the listening brain. *Nat. Commun.* **2017**, *8*. [[CrossRef](#)]
101. Gates, G.A.; Mills, J.H. Presbycusis. *Lancet* **2005**, *366*, 1111–1120. [[CrossRef](#)]
102. Kane, M.J.; Hasher, L.; Stoltzfus, E.R.; Zacks, R.T.; Connelly, S.L. Inhibitory attentional mechanisms and aging. *Psychol. Aging* **1994**, *9*, 103–112. [[CrossRef](#)]
103. Salthouse, T.A. The processing-speed theory of adult age differences in cognition. *Psychol. Rev.* **1996**, *103*, 403–428. [[CrossRef](#)] [[PubMed](#)]
104. Cabeza, R.; Albert, M.; Belleville, S.; Craik, F.I.M.; Duarte, A.; Grady, C.L.; Lindenberger, U.; Nyberg, L.; Park, D.C.; Reuter-Lorenz, P.A.; et al. Maintenance, reserve and compensation: The cognitive neuroscience of healthy ageing. *Nat. Rev. Neurosci.* **2018**, *19*, 701–710. [[CrossRef](#)]
105. Humes, L.E.; Dubno, J.R.; Gordon-Salant, S.; Lister, J.J.; Cacace, A.T.; Cruickshanks, K.J.; Gates, G.A.; Wilson, R.H.; Wingfield, A. Central Presbycusis: A Review and Evaluation of the Evidence. *J. Am. Acad. Audiol.* **2012**, *23*, 635–666. [[CrossRef](#)] [[PubMed](#)]
106. Anderson, S.; Roque, L.; Gaskins, C.R.; Gordon-Salant, S.; Goupell, M.J. Age-Related Compensation Mechanism Revealed in the Cortical Representation of Degraded Speech. *J. Assoc. Res. Otolaryngol.* **2020**, *21*, 373–391. [[CrossRef](#)] [[PubMed](#)]
107. Pichora-Fuller, M.K. Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing. *Int. J. Audiol.* **2008**, *47*, S72–S82. [[CrossRef](#)]
108. Saija, J.D.; Akyürek, E.G.; Andringa, T.C.; Başkent, D. Perceptual restoration of degraded speech is preserved with advancing age. *J. Assoc. Res. Otolaryngol.* **2014**, *15*, 139–148. [[CrossRef](#)] [[PubMed](#)]
109. Wolpe, N.; Ingram, J.N.; Tsvetanov, K.A.; Geerligs, L.; Kievit, R.A.; Henson, R.N.; Wolpert, D.M.; Rowe, J.B. Ageing increases reliance on sensorimotor prediction through structural and functional differences in frontostriatal circuits. *Nat. Commun.* **2016**, *7*, 13034. [[CrossRef](#)] [[PubMed](#)]
110. Arlinger, S.; Lunner, T.; Lyxell, B.; Pichora-Fuller, M.K. The emergence of cognitive hearing science. *Scand. J. Psychol.* **2009**, *50*, 371–384. [[CrossRef](#)] [[PubMed](#)]
111. Peelle, J.E. Listening Effort: How the Cognitive Consequences of Acoustic Challenge Are Reflected in Brain and Behavior. *Ear Hear.* **2018**, *39*, 204–214. [[CrossRef](#)]
112. O'Brien, J.L.; Lister, J.J.; Fausto, B.A.; Morgan, D.G.; Maeda, H.; Andel, R.; Edwards, J.D. Are auditory processing and cognitive performance assessments overlapping or distinct? Parsing the auditory behaviour of older adults. *Int. J. Audiol.* **2021**, *60*, 123–132. [[CrossRef](#)] [[PubMed](#)]
113. Auksztulewicz, R.; Friston, K. Attentional Enhancement of Auditory Mismatch Responses: A DCM/MEG Study. *Cereb. Cortex* **2015**, *25*, 4273–4283. [[CrossRef](#)] [[PubMed](#)]
114. Foxe, J.J.; Snyder, A.C. The Role of Alpha-Band Brain Oscillations as a Sensory Suppression Mechanism during Selective Attention. *Front. Psychol.* **2011**, *2*, 154. [[CrossRef](#)] [[PubMed](#)]
115. Jacquemot, C.; Scott, S.K. What is the relationship between phonological short-term memory and speech processing? *Trends Cogn. Sci.* **2006**, *10*, 480–486. [[CrossRef](#)] [[PubMed](#)]
116. Erb, J.; Henry, M.J.; Eisner, F.; Obleser, J. Auditory skills and brain morphology predict individual differences in adaptation to degraded speech. *Neuropsychologia* **2012**, *50*, 2154–2164. [[CrossRef](#)]
117. Puschmann, S.; Baillet, S.; Zatorre, R.J. Musicians at the Cocktail Party: Neural Substrates of Musical Training During Selective Listening in Multispeaker Situations. *Cereb. Cortex* **2019**, *29*, 3253–3265. [[CrossRef](#)] [[PubMed](#)]
118. Kim, S.; Choi, I.; Schwalje, A.T.; Kim, K.; Lee, J.H. Auditory Working Memory Explains Variance in Speech Recognition in Older Listeners Under Adverse Listening Conditions. *Clin. Interv. Aging* **2020**, *15*, 395–406. [[CrossRef](#)] [[PubMed](#)]
119. Lim, S.J.; Wostmann, M.; Obleser, J. Selective Attention to Auditory Memory Neurally Enhances Perceptual Precision. *J. Neurosci.* **2015**, *35*, 16094–16104. [[CrossRef](#)]
120. Akeroyd, M.A. 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.* **2008**, *47* (Suppl. 2), S53–S71. [[CrossRef](#)] [[PubMed](#)]
121. Souza, P.; Arehart, K. Robust relationship between reading span and speech recognition in noise. *Int. J. Audiol.* **2015**, *54*, 705–713. [[CrossRef](#)]
122. Donhauser, P.W.; Baillet, S. Two Distinct Neural Timescales for Predictive Speech Processing. *Neuron* **2020**, *105*, 385–393. [[CrossRef](#)]
123. Alain, C.; Zendel, B.R.; Hutka, S.; Bidelman, G.M. Turning down the noise: The benefit of musical training on the aging auditory brain. *Hear. Res.* **2014**, *308*, 162–173. [[CrossRef](#)]

124. Bidelman, G.M.; Krishnan, A. Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Res.* **2010**, *1355*, 112–125. [[CrossRef](#)] [[PubMed](#)]
125. Hyde, K.L.; Lerch, J.; Norton, A.; Forgeard, M.; Winner, E.; Evans, A.C.; Schlaug, G. Musical Training Shapes Structural Brain Development. *J. Neurosci.* **2009**, *29*, 3019–3025. [[CrossRef](#)] [[PubMed](#)]
126. Koelsch, S.; Schröger, E.; Tervaniemi, M. Superior pre-attentive auditory processing in musicians. *Neuroreport* **1999**, *10*, 1309–1313. [[CrossRef](#)] [[PubMed](#)]
127. Kraus, N.; Skoe, E.; Parbery-Clark, A.; Ashley, R. Experience-induced malleability in neural encoding of pitch, timbre, and timing. *Ann. N. Y. Acad. Sci.* **2009**, *1169*, 543–557. [[CrossRef](#)]
128. Kraus, N.; Strait, D.L.; Parbery-Clark, A. Cognitive factors shape brain networks for auditory skills: Spotlight on auditory working memory. *Ann. N. Y. Acad. Sci.* **2012**, *1252*, 100–107. [[CrossRef](#)]
129. Magne, C.; Schön, D.; Besson, M. Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *J. Cogn. Neurosci.* **2006**, *18*, 199–211. [[CrossRef](#)]
130. Musacchia, G.; Sams, M.; Skoe, E.; Kraus, N. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 15894–15898. [[CrossRef](#)] [[PubMed](#)]
131. Schön, D.; Magne, C.; Besson, M. The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology* **2004**, *41*, 341–349. [[CrossRef](#)] [[PubMed](#)]
132. Wong, P.C.; Skoe, E.; Russo, N.M.; Dees, T.; Kraus, N. Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat. Neurosci.* **2007**, *10*, 420–422. [[CrossRef](#)]
133. Başkent, D.; Gaudrain, E. Musician advantage for speech-on-speech perception. *J. Acoust. Soc. Am.* **2016**, *139*, E151–E156. [[CrossRef](#)] [[PubMed](#)]
134. Swaminathan, J.; Mason, C.R.; Streeter, T.M.; Best, V.; Kidd, J.G.; Patel, A.D. Musical training, individual differences and the cocktail party problem. *Sci. Rep.* **2015**, *5*, 11628. [[CrossRef](#)]
135. Zendel, B.R.; Alain, C. Musicians experience less age-related decline in central auditory processing. *Psychol. Aging* **2012**, *27*, 410–417. [[CrossRef](#)]
136. Zendel, B.R.; Alain, C. Concurrent sound segregation is enhanced in musicians. *J. Cogn. Neurosci.* **2009**, *21*, 1488–1498. [[CrossRef](#)]
137. Jin, S.-H.; Liu, C. English sentence recognition in speech-shaped noise and multi-talker babble for English-, Chinese-, and Korean-native listeners. *J. Acoust. Soc. Am.* **2012**, *132*, E1391–E1397. [[CrossRef](#)]
138. Lucks Mendel, L.; Widner, H. Speech perception in noise for bilingual listeners with normal hearing. *Int. J. Audiol.* **2016**, *55*, 126–134. [[CrossRef](#)]
139. Rammell, C.S.; Cheng, H.; Pisoni, D.B.; Newman, S.D. L2 speech perception in noise: An fMRI study of advanced Spanish learners. *Brain Res.* **2019**, *1720*, 146316. [[CrossRef](#)]
140. Skoe, E.; Karayanidi, K. Bilingualism and Speech Understanding in Noise: Auditory and Linguistic Factors. *J. Am. Acad. Audiol.* **2019**, *30*, 115–130. [[CrossRef](#)]
141. Bradlow, A.R.; Alexander, J.A. Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *J. Acoust. Soc. Am.* **2007**, *121*, 2339–2349. [[CrossRef](#)] [[PubMed](#)]
142. Hervais-Adelman, A.; Pefkou, M.; Golestani, N. Bilingual speech-in-noise: Neural bases of semantic context use in the native language. *Brain Lang.* **2014**, *132*, 1–6. [[CrossRef](#)] [[PubMed](#)]
143. Kousaie, S.; Baum, S.; Phillips, N.A.; Gracco, V.; Titone, D.; Chen, J.K.; Chai, X.J.; Klein, D. Language learning experience and mastering the challenges of perceiving speech in noise. *Brain Lang.* **2019**, *196*, 104645. [[CrossRef](#)]
144. Eisner, F.; McGettigan, C.; Faulkner, A.; Rosen, S.; Scott, S.K. Inferior frontal gyrus activation predicts individual differences in perceptual learning of cochlear-implant simulations. *J. Neurosci.* **2010**, *30*, 7179–7186. [[CrossRef](#)]
145. Goldstone, R.L. PERCEPTUAL LEARNING. *Annu. Rev. Psychol.* **1998**, *49*, 585–612. [[CrossRef](#)]
146. Samuel, A.G.; Kraljic, T. Perceptual learning for speech. *Atten. Percept. Psychophys.* **2009**, *71*, 1207–1218. [[CrossRef](#)]
147. Eisner, F.; McQueen, J.M. Perceptual learning in speech: Stability over time. *J. Acoust. Soc. Am.* **2006**, *119*, 1950–1953. [[CrossRef](#)] [[PubMed](#)]
148. Norris, D. Perceptual learning in speech. *Cogn. Psychol.* **2003**, *47*, 204–238. [[CrossRef](#)]
149. Petrov, A.A.; Doshier, B.A.; Lu, Z.L. The dynamics of perceptual learning: An incremental reweighting model. *Psychol. Rev.* **2005**, *112*, 715–743. [[CrossRef](#)]
150. Tsodyks, M.; Gilbert, C. Neural networks and perceptual learning. *Nature* **2004**, *431*, 775–781. [[CrossRef](#)]
151. Bradlow, A.R.; Bent, T. Perceptual adaptation to non-native speech. *Cognition* **2008**, *106*, 707–729. [[CrossRef](#)]
152. Eisner, F.; McQueen, J.M. The specificity of perceptual learning in speech processing. *Percept. Psychophys.* **2005**, *67*, 224–238. [[CrossRef](#)]
153. Nygaard, L.C.; Pisoni, D.B. Talker-specific learning in speech perception. *Percept. Psychophys.* **1998**, *60*, 355–376. [[CrossRef](#)]
154. Clopper, C.G.; Pisoni, D.B. Effects of talker variability on perceptual learning of dialects. *Lang. Speech* **2004**, *47*, 207–239. [[CrossRef](#)]
155. Gordon-Salant, S.; Yeni-Komshian, G.H.; Fitzgibbons, P.J.; Schurman, J. Short-term adaptation to accented English by younger and older adults. *J. Acoust. Soc. Am.* **2010**, *128*, E1200–E1204. [[CrossRef](#)]
156. Sidaras, S.K.; Alexander, J.E.; Nygaard, L.C. Perceptual learning of systematic variation in Spanish-accented speech. *J. Acoust. Soc. Am.* **2009**, *125*, 3306–3316. [[CrossRef](#)] [[PubMed](#)]

157. Stacey, P.C.; Summerfield, A.Q. Effectiveness of computer-based auditory training in improving the perception of noise-vocoded speech. *J. Acoust. Soc. Am.* **2007**, *121*, 2923–2935. [[CrossRef](#)] [[PubMed](#)]
158. Banai, K.; Lavner, Y. The effects of training length on the perceptual learning of time-compressed speech and its generalization. *J. Acoust. Soc. Am.* **2014**, *136*, 1908–1917. [[CrossRef](#)] [[PubMed](#)]
159. Francis, A.L.; Ciocca, V.; Ma, L.; Fenn, K. Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *J. Phon.* **2008**, *36*, 268–294. [[CrossRef](#)]
160. Peelle, J.E.; Wingfield, A. Dissociations in perceptual learning revealed by adult age differences in adaptation to time-compressed speech. *J. Exp. Psychol. Hum. Percept. Perform.* **2005**, *31*, 1315–1330. [[CrossRef](#)]
161. Huyck, J.J.; Johnsrude, I.S. Rapid perceptual learning of noise-vocoded speech requires attention. *J. Acoust. Soc. Am.* **2012**, *131*, EL236–EL242. [[CrossRef](#)] [[PubMed](#)]
162. Bradlow, A.R.; Pisoni, D.B.; Akahane-Yamada, R.; Tohkura, Y. Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *J. Acoust. Soc. Am.* **1997**, *101*, 2299–2310. [[CrossRef](#)]
163. Kraljic, T.; Samuel, A.G. Perceptual learning for speech: Is there a return to normal? *Cogn. Psychol.* **2005**, *51*, 141–178. [[CrossRef](#)]
164. Houde, J.F.; Chang, E.F. The cortical computations underlying feedback control in vocal production. *Curr. Opin. Neurobiol.* **2015**, *33*, 174–181. [[CrossRef](#)]
165. Houde, J.F.; Nagarajan, S.S. Speech production as state feedback control. *Front. Hum. Neurosci.* **2011**, *5*, 82. [[CrossRef](#)] [[PubMed](#)]
166. Hickok, G. The cortical organization of speech processing: Feedback control and predictive coding the context of a dual-stream model. *J. Commun. Disord.* **2012**, *45*, 393–402. [[CrossRef](#)]
167. Lincoln, M.; Packman, A.; Onslow, M. Altered auditory feedback and the treatment of stuttering: A review. *J. Fluency Disord.* **2006**, *31*, 71–89. [[CrossRef](#)] [[PubMed](#)]
168. Lombard, E. Le signe de l'elevation de la voix. *Ann. Mal. De L'oreille Et Du Larynx* **1911**, *18*, 101–119.
169. Takaso, H.; Eisner, F.; Wise, R.J.; Scott, S.K. The effect of delayed auditory feedback on activity in the temporal lobe while speaking: A positron emission tomography study. *J. Speech Lang. Hear. Res.* **2010**, *53*, 226–236. [[CrossRef](#)]
170. Meekings, S.; Evans, S.; Lavan, N.; Boebinger, D.; Krieger-Redwood, K.; Cooke, M.; Scott, S.K. Distinct neural systems recruited when speech production is modulated by different masking sounds. *J. Acoust. Soc. Am.* **2016**, *140*, 8. [[CrossRef](#)] [[PubMed](#)]
171. Elder, G.A. Update on TBI and Cognitive Impairment in Military Veterans. *Curr. Neurol. Neurosci. Rep.* **2015**, *15*, 68. [[CrossRef](#)]
172. Menon, D.K.; Schwab, K.; Wright, D.W.; Maas, A.I. Position statement: Definition of traumatic brain injury. *Arch. Phys. Med. Rehabil.* **2010**, *91*, 1637–1640. [[CrossRef](#)]
173. Hoffer, M.E.; Balaban, C.; Nicholas, R.; Marcus, D.; Murphy, S.; Gottshall, K. Neurosensory Sequelae of Mild Traumatic Brain Injury. *Psychiatr. Ann.* **2013**, *43*, 5. [[CrossRef](#)]
174. Gallun, F.J.; Papesh, M.A.; Lewis, M.S. Hearing complaints among veterans following traumatic brain injury. *Brain Inj.* **2017**, *31*, 1183–1187. [[CrossRef](#)]
175. Turgeon, C.; Champoux, F.; Lepore, F.; Leclerc, S.; Ellemberg, D. Auditory processing after sport-related concussions. *Ear Hear.* **2011**, *32*, 667–670. [[CrossRef](#)]
176. Hoover, E.C.; Souza, P.E.; Gallun, F.J. Auditory and Cognitive Factors Associated with Speech-in-Noise Complaints following Mild Traumatic Brain Injury. *J. Am. Acad. Audiol.* **2017**, *28*, 325–339. [[CrossRef](#)] [[PubMed](#)]
177. Boller, F.; Vrtunski, P.B.; Kim, Y.; Mack, J.L. Delayed auditory feedback and aphasia. *Cortex* **1978**, *14*, 212–226. [[CrossRef](#)]
178. Chapin, C.; Blumstein, S.E.; Meissner, B.; Boller, F. Speech production mechanisms in aphasia: A delayed auditory feedback study. *Brain Lang.* **1981**, *14*, 106–113. [[CrossRef](#)]
179. Fifer, R.C. Insular stroke causing unilateral auditory processing disorder: Case report. *J. Am. Acad. Audiol.* **1993**, *4*, 364–369.
180. Huang, X.; Chen, X.; Yan, N.; Jones, J.A.; Wang, E.Q.; Chen, L.; Guo, Z.; Li, W.; Liu, P.; Liu, H. The impact of parkinson's disease on the cortical mechanisms that support auditory-motor integration for voice control. *Hum. Brain Mapp.* **2016**, *37*, 4248–4261. [[CrossRef](#)]
181. Tolosa, E.; Wenning, G.; Poewe, W. The diagnosis of Parkinson's disease. *Lancet Neurol.* **2006**, *5*, 75–86. [[CrossRef](#)]
182. Williams-Gray, C.H.; Mason, S.L.; Evans, J.R.; Foltynie, T.; Brayne, C.; Robbins, T.W.; Barker, R.A. The CamPaIGN study of Parkinson's disease: 10-year outlook in an incident population-based cohort. *J. Neurol. Neurosurg. Psychiatry* **2013**, *84*, 1258–1264. [[CrossRef](#)]
183. Sapir, S.; Ramig, L.; Fox, C. Speech and swallowing disorders in Parkinson disease. *Curr. Opin. Otolaryngol. Head Neck Surg.* **2008**, *16*, 205–210. [[CrossRef](#)] [[PubMed](#)]
184. Ho, A.K.; Bradshaw, J.L.; Iansek, T. Volume perception in parkinsonian speech. *Mov. Disord.* **2000**, *15*, 1125–1131. [[CrossRef](#)]
185. Arnold, C.; Gehrig, J.; Gispert, S.; Seifried, C.; Kell, C.A. Pathomechanisms and compensatory efforts related to Parkinsonian speech. *Neuroimage Clin.* **2014**, *4*, 82–97. [[CrossRef](#)]
186. Ho, A.K.; Bradshaw, J.L.; Iansek, R.; Alfredson, R. Speech volume regulation in Parkinson's disease: Effects of implicit cues and explicit instructions. *Neuropsychologia* **1999**, *37*, 1453–1460. [[CrossRef](#)]
187. Kiran, S.; Larson, C.R. Effect of duration of pitch-shifted feedback on vocal responses in patients with Parkinson's disease. *J. Speech Lang. Hear. Res.* **2001**, *44*, 975–987. [[CrossRef](#)]
188. Mollaei, F.; Shiller, D.M.; Baum, S.R.; Gracco, V.L. Sensorimotor control of vocal pitch and formant frequencies in Parkinson's disease. *Brain Res.* **2016**, *1646*, 269–277. [[CrossRef](#)] [[PubMed](#)]



189. Downie, A.W.; Low, J.M.; Lindsay, D.D. Speech Disorder in Parkinsonism—Usefulness of Delayed Auditory Feedback in Selected Cases. *Int. J. Lang. Commun. Disord.* **1981**, *16*, 135–139. [[CrossRef](#)] [[PubMed](#)]
190. Cools, R.; Rogers, R.; Barker, R.A.; Robbins, T.W. Top-down attentional control in Parkinson's disease: Salient considerations. *J. Cogn. Neurosci.* **2010**, *22*, 848–859. [[CrossRef](#)]
191. Zarkali, A.; Adams, R.A.; Psarras, S.; Leyland, L.A.; Rees, G.; Weil, R.S. Increased weighting on prior knowledge in Lewy body-associated visual hallucinations. *Brain Commun.* **2019**, *1*, fcz007. [[CrossRef](#)] [[PubMed](#)]
192. Agosta, F.; Pievani, M.; Geroldi, C.; Copetti, M.; Frisoni, G.B.; Filippi, M. Resting state fMRI in Alzheimer's disease: Beyond the default mode network. *Neurobiol. Aging* **2012**, *33*, 1564–1578. [[CrossRef](#)] [[PubMed](#)]
193. Zhou, J.; Greicius, M.D.; Gennatas, E.D.; Growdon, M.E.; Jang, J.Y.; Rabinovici, G.D.; Kramer, J.H.; Weiner, M.; Miller, B.L.; Seeley, W.W. Divergent network connectivity changes in behavioural variant frontotemporal dementia and Alzheimer's disease. *Brain* **2010**, *133*, 1352–1367. [[CrossRef](#)]
194. Bouma, A.; Gootjes, L. Effects of attention on dichotic listening in elderly and patients with dementia of the Alzheimer type. *Brain Cogn.* **2011**, *76*, 286–293. [[CrossRef](#)] [[PubMed](#)]
195. Idrizbegovic, E.; Hederstierna, C.; Dahlquist, M.; Rosenhall, U. Short-term longitudinal study of central auditory function in Alzheimer's disease and mild cognitive impairment. *Dement. Geriatr. Cogn. Dis. Extra* **2013**, *3*, 468–471. [[CrossRef](#)]
196. Utoomprurkporn, N.; Hardy, C.J.D.; Stott, J.; Costafreda, S.G.; Warren, J.; Bamiou, D.E. "The Dichotic Digit Test" as an Index Indicator for Hearing Problem in Dementia: Systematic Review and Meta-Analysis. *J. Am. Acad. Audiol.* **2020**, *31*, 646–655. [[CrossRef](#)]
197. Goll, J.C.; Kim, L.G.; Ridgway, G.R.; Hailstone, J.C.; Lehmann, M.; Buckley, A.H.; Crutch, S.J.; Warren, J.D. Impairments of auditory scene analysis in Alzheimer's disease. *Brain* **2012**, *135*, 190–200. [[CrossRef](#)]
198. Hardy, C.J.D.; Yong, K.X.X.; Goll, J.C.; Crutch, S.J.; Warren, J.D. Impairments of auditory scene analysis in posterior cortical atrophy. *Brain* **2020**, *143*, 2689–2695. [[CrossRef](#)]
199. Gates, G.A.; Anderson, M.L.; McCurry, S.M.; Feeney, M.P.; Larson, E.B. Central auditory dysfunction as a harbinger of Alzheimer dementia. *Arch. Otolaryngol. Head Neck Surg.* **2011**, *137*, 390–395. [[CrossRef](#)] [[PubMed](#)]
200. Gates, G.A.; Gibbons, L.E.; McCurry, S.M.; Crane, P.K.; Feeney, M.P.; Larson, E.B. Executive dysfunction and presbycusis in older persons with and without memory loss and dementia. *Cogn. Behav. Neurol.* **2010**, *23*, 218–223. [[CrossRef](#)]
201. Pronk, M.; Lissenberg-Witte, B.I.; Van der Aa, H.P.A.; Comijs, H.C.; Smits, C.; Lemke, U.; Zekveld, A.A.; Kramer, S.E. Longitudinal Relationships Between Decline in Speech-in-Noise Recognition Ability and Cognitive Functioning: The Longitudinal Aging Study Amsterdam. *J. Speech Lang. Hear. Res.* **2019**, *62*, 1167–1187. [[CrossRef](#)] [[PubMed](#)]
202. Burda, A.N.; Hageman, C.F.; Brousard, K.T.; Miller, A.L. Dementia and identification of words and sentences produced by native and nonnative English speakers. *Percept. Mot. Ski.* **2004**, *98*, 1359–1362. [[CrossRef](#)]
203. Hardy, C.J.D.; Hwang, Y.T.; Bond, R.L.; Marshall, C.R.; Ridha, B.H.; Crutch, S.J.; Rossor, M.N.; Warren, J.D. Donepezil enhances understanding of degraded speech in Alzheimer's disease. *Ann. Clin. Transl. Neurol.* **2017**, *4*, 835–840. [[CrossRef](#)]
204. Ranasinghe, K.G.; Kothare, H.; Kort, N.; Hinkley, L.B.; Beagle, A.J.; Mizuiri, D.; Honma, S.M.; Lee, R.; Miller, B.L.; Gorno-Tempini, M.L.; et al. Neural correlates of abnormal auditory feedback processing during speech production in Alzheimer's disease. *Sci. Rep.* **2019**, *9*, 5686. [[CrossRef](#)]
205. Marshall, C.R.; Hardy, C.J.D.; Volkmer, A.; Russell, L.L.; Bond, R.L.; Fletcher, P.D.; Clark, C.N.; Mummery, C.J.; Schott, J.M.; Rossor, M.N.; et al. Primary progressive aphasia: A clinical approach. *J. Neurol.* **2018**, *265*, 1474–1490. [[CrossRef](#)]
206. Gorno-Tempini, M.L.; Hillis, A.E.; Weintraub, S.; Kertesz, A.; Mendez, M.; Cappa, S.F.; Ogar, J.M.; Rohrer, J.D.; Black, S.; Boeve, B.F.; et al. Classification of primary progressive aphasia and its variants. *Neurology* **2011**, *76*, 1006–1014. [[CrossRef](#)]
207. Goll, J.C.; Kim, L.G.; Hailstone, J.C.; Lehmann, M.; Buckley, A.; Crutch, S.J.; Warren, J.D. Auditory object cognition in dementia. *Neuropsychologia* **2011**, *49*, 2755–2765. [[CrossRef](#)]
208. Goll, J.C.; Crutch, S.J.; Loo, J.H.; Rohrer, J.D.; Frost, C.; Bamiou, D.E.; Warren, J.D. Non-verbal sound processing in the primary progressive aphasia. *Brain* **2010**, *133*, 272–285. [[CrossRef](#)] [[PubMed](#)]
209. Grube, M.; Bruffaerts, R.; Schaefferbeke, J.; Neyens, V.; De Weer, A.S.; Seghers, A.; Bergmans, B.; Dries, E.; Griffiths, T.D.; Vandenberghe, R. Core auditory processing deficits in primary progressive aphasia. *Brain* **2016**, *139*, 1817–1829. [[CrossRef](#)]
210. Hardy, C.J.D.; Frost, C.; Sivasathiseelan, H.; Johnson, J.C.S.; Agustus, J.L.; Bond, R.L.; Benhamou, E.; Russell, L.L.; Marshall, C.R.; Rohrer, J.D.; et al. Findings of Impaired Hearing in Patients With Nonfluent/Agrammatic Variant Primary Progressive Aphasia. *JAMA Neurol.* **2019**, *76*, 607–611. [[CrossRef](#)]
211. Johnson, J.C.S.; Jiang, J.; Bond, R.L.; Benhamou, E.; Requena-Komuro, M.C.; Russell, L.L.; Greaves, C.; Nelson, A.; Sivasathiseelan, H.; Marshall, C.R.; et al. Impaired phonemic discrimination in logopenic variant primary progressive aphasia. *Ann. Clin. Transl. Neurol.* **2020**, *7*, 1252–1257. [[CrossRef](#)] [[PubMed](#)]
212. Ruksenaite, J.; Volkmer, A.; Jiang, J.; Johnson, J.C.; Marshall, C.R.; Warren, J.D.; Hardy, C.J. Primary Progressive Aphasia: Toward a Pathophysiological Synthesis. *Curr. Neurol. Neurosci. Rep.* **2021**, *21*. [[CrossRef](#)]
213. Hardy, C.J.D.; Bond, R.L.; Jain, K.; Marshall, C.R.; Russell, L.L.; Dick, K.; Crutch, S.J.; Rohrer, J.D.; Warren, J.D. Sensitivity of Speech Output to Delayed Auditory Feedback in Primary Progressive Aphasia. *Front. Neurol.* **2018**, *9*, 894. [[CrossRef](#)] [[PubMed](#)]
214. Hardy, C.J.D.; Agustus, J.L.; Marshall, C.R.; Clark, C.N.; Russell, L.L.; Bond, R.L.; Brotherhood, E.V.; Thomas, D.L.; Crutch, S.J.; Rohrer, J.D.; et al. Behavioural and neuroanatomical correlates of auditory speech analysis in primary progressive aphasia. *Alzheimer's Res. Ther.* **2017**, *9*. [[CrossRef](#)]

215. Hardy, C.J.D.; Agustus, J.L.; Marshall, C.R.; Clark, C.N.; Russell, L.L.; Brotherhood, E.V.; Bond, R.L.; Fiford, C.M.; Ondobaka, S.; Thomas, D.L.; et al. Functional neuroanatomy of speech signal decoding in primary progressive aphasias. *Neurobiol. Aging* **2017**, *56*, 190–201. [[CrossRef](#)] [[PubMed](#)]
216. Conway, E.R.; Chenery, H.J. Evaluating the MESSAGE Communication Strategies in Dementia training for use with community-based aged care staff working with people with dementia: A controlled pretest-post-test study. *J. Clin. Nurs.* **2016**, *25*, 1145–1155. [[CrossRef](#)] [[PubMed](#)]
217. Liddle, J.; Smith-Conway, E.R.; Baker, R.; Angwin, A.J.; Gallois, C.; Copland, D.A.; Pachana, N.A.; Humphreys, M.S.; Byrne, G.J.; Chenery, H.J. Memory and communication support strategies in dementia: Effect of a training program for informal caregivers. *Int. Psychogeriatr.* **2012**, *24*, 1927–1942. [[CrossRef](#)]
218. Sparks, R.; Helm, N.; Albert, M. Aphasia rehabilitation resulting from melodic intonation therapy. *Cortex* **1974**, *10*, 303–316. [[CrossRef](#)]
219. Zumbansen, A.; Peretz, I.; Hébert, S. Melodic Intonation Therapy: Back to Basics for Future Research. *Front. Neurol.* **2014**, *5*. [[CrossRef](#)]
220. Bieber, R.E.; Gordon-Salant, S. Improving older adults' understanding of challenging speech: Auditory training, rapid adaptation and perceptual learning. *Hear. Res.* **2021**, *402*, 108054. [[CrossRef](#)]
221. Shafiro, V.; Sheft, S.; Gygi, B.; Ho, K.T. The influence of environmental sound training on the perception of spectrally degraded speech and environmental sounds. *Trends Amplif.* **2012**, *16*, 83–101. [[CrossRef](#)]
222. Sehm, B.; Schnitzler, T.; Obleser, J.; Groba, A.; Ragert, P.; Villringer, A.; Obrig, H. Facilitation of inferior frontal cortex by transcranial direct current stimulation induces perceptual learning of severely degraded speech. *J. Neurosci.* **2013**, *33*, 15868–15878. [[CrossRef](#)]
223. Rampersad, S.; Roig-Solvas, B.; Yarossi, M.; Kulkarni, P.P.; Santarnecchi, E.; Dorval, A.D.; Brooks, D.H. Prospects for transcranial temporal interference stimulation in humans: A computational study. *Neuroimage* **2019**, *202*, 116124. [[CrossRef](#)] [[PubMed](#)]
224. Wayne, R.V.; Hamilton, C.; Jones Huyck, J.; Johnsrude, I.S. Working Memory Training and Speech in Noise Comprehension in Older Adults. *Front. Aging Neurosci.* **2016**, *8*, 49. [[CrossRef](#)]
225. Escobar, J.; Mussoi, B.S.; Silberer, A.B. The Effect of Musical Training and Working Memory in Adverse Listening Situations. *Ear Hear.* **2020**, *41*, 278–288. [[CrossRef](#)] [[PubMed](#)]
226. Zhang, L.; Fu, X.; Luo, D.; Xing, L.; Du, Y. Musical Experience Offsets Age-Related Decline in Understanding Speech-in-Noise: Type of Training Does Not Matter, Working Memory Is the Key. *Ear Hear.* **2020**, *42*, 258–270. [[CrossRef](#)]
227. Zendel, B.R.; West, G.L.; Belleville, S.; Peretz, I. Musical training improves the ability to understand speech-in-noise in older adults. *Neurobiol. Aging* **2019**, *81*, 102–115. [[CrossRef](#)]
228. Cardin, V.; Rosen, S.; Konieczny, L.; Coulson, K.; Lametti, D.; Edwards, M.; Woll, B. The effect of dopamine on the comprehension of spectrally-shifted noise-vocoded speech: A pilot study. *Int. J. Audiol.* **2020**, *59*, 674–681. [[CrossRef](#)]
229. De Boer, B. Self-Organisation and Evolution of Biological and Social Systems. In *Self-Organisation and Evolution of Biological and Social Systems*; Hemelrijk, C., Ed.; Cambridge University Press: Cambridge, UK, 2005.
230. Williams, D. Predictive coding and thought. *Synthese* **2020**, *197*, 1749–1775. [[CrossRef](#)]
231. Koelsch, S.; Vuust, P.; Friston, K. Predictive Processes and the Peculiar Case of Music. *Trends Cogn. Sci.* **2019**, *23*, 63–77. [[CrossRef](#)]
232. Sohoglu, E.; Peelle, J.E.; Carlyon, R.P.; Davis, M.H. Predictive top-down integration of prior knowledge during speech perception. *J. Neurosci.* **2012**, *32*, 8443–8453. [[CrossRef](#)]
233. Wacongne, C.; Changeux, J.P.; Dehaene, S. A neuronal model of predictive coding accounting for the mismatch negativity. *J. Neurosci.* **2012**, *32*, 3665–3678. [[CrossRef](#)]
234. Rysop, A.U.; Schmitt, L.M.; Obleser, J.; Hartwigsen, G. Neural modelling of the semantic predictability gain under challenging listening conditions. *Hum. Brain Mapp.* **2021**, *42*, 110–127. [[CrossRef](#)]
235. Kogo, N.; Trengove, C. Is predictive coding theory articulated enough to be testable? *Front. Comput. Neurosci.* **2015**, *9*, 111. [[CrossRef](#)]
236. Benhamou, E.; Marshall, C.R.; Russell, L.L.; Hardy, C.J.D.; Bond, R.L.; Sivasathiseelan, H.; Greaves, C.V.; Friston, K.J.; Rohrer, J.D.; Warren, J.D.; et al. The neurophysiological architecture of semantic dementia: Spectral dynamic causal modelling of a neurodegenerative proteinopathy. *Sci. Rep.* **2020**, *10*, 16321. [[CrossRef](#)]
237. Shaw, A.D.; Hughes, L.E.; Moran, R.; Coyle-Gilchrist, I.; Rittman, T.; Rowe, J.B. In Vivo Assay of Cortical Microcircuitry in Frontotemporal Dementia: A Platform for Experimental Medicine Studies. *Cereb. Cortex* **2021**, *31*, 1837–1847. [[CrossRef](#)]