

This article was downloaded by: [University of Rochester]

On: 08 May 2015, At: 05:46

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

**COGNITIVE
NEUROPSYCHOLOGY**

VOLUME 32 – ISSUE 2 – MARCH 2015



[Click for updates](#)

Cognitive Neuropsychology

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/pcgn20>

A causal test of the motor theory of speech perception: a case of impaired speech production and spared speech perception

Alena Stasenka^a, Cory Bonn^{ab}, Alex Teghipco^a, Frank E. Garcea^{ab}, Catherine Sweet^c, Mary Dombovy^c, Joyce McDonough^{abd} & Bradford Z. Mahon^{abe}

^a Department of Brain & Cognitive Sciences, University of Rochester, Rochester, NY, USA

^b Center for Language Sciences, University of Rochester, Rochester, NY, USA

^c Department of Neurology, Unity Hospital, Rochester, NY, USA

^d Department of Linguistics, University of Rochester, Rochester, NY, USA

^e Department of Neurosurgery, University of Rochester Medical Center, Rochester, NY, USA

Published online: 08 May 2015.

To cite this article: Alena Stasenka, Cory Bonn, Alex Teghipco, Frank E. Garcea, Catherine Sweet, Mary Dombovy, Joyce McDonough & Bradford Z. Mahon (2015) A causal test of the motor theory of speech perception: a case of impaired speech production and spared speech perception, *Cognitive Neuropsychology*, 32:2, 38-57, DOI: [10.1080/02643294.2015.1035702](https://doi.org/10.1080/02643294.2015.1035702)

To link to this article: <http://dx.doi.org/10.1080/02643294.2015.1035702>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or

distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

A causal test of the motor theory of speech perception: a case of impaired speech production and spared speech perception

Alena Stasenko^a, Cory Bonn^{a,b}, Alex Teghipco^a, Frank E. Garcea^{a,b}, Catherine Sweet^c, Mary Dombovy^c, Joyce McDonough^{a,b,d} and Bradford Z. Mahon^{a,b,e*}

^aDepartment of Brain & Cognitive Sciences, University of Rochester, Rochester, NY, USA; ^bCenter for Language Sciences, University of Rochester, Rochester, NY, USA; ^cDepartment of Neurology, Unity Hospital, Rochester, NY, USA; ^dDepartment of Linguistics, University of Rochester, Rochester, NY, USA; ^eDepartment of Neurosurgery, University of Rochester Medical Center, Rochester, NY, USA

(Manuscript received 3 June 2014; revised manuscript received 23 March 2015;
revised manuscript accepted 24 March 2015)

The debate about the causal role of the motor system in speech perception has been reignited by demonstrations that motor processes are engaged during the processing of speech sounds. Here, we evaluate which aspects of auditory speech processing are affected, and which are not, in a stroke patient with dysfunction of the speech motor system. We found that the patient showed a normal phonemic categorical boundary when discriminating two non-words that differ by a minimal pair (e.g., ADA–AGA). However, using the same stimuli, the patient was unable to identify or label the non-word stimuli (using a button-press response). A control task showed that he could identify speech sounds by speaker gender, ruling out a general labelling impairment. These data suggest that while the motor system is not causally involved in perception of the speech signal, it may be used when other cues (e.g., meaning, context) are not available.

Keywords: apraxia of speech; categorical perception; motor theory of speech perception; mirror neurons

Introduction

The acoustic realization of speech–sound categories varies widely as a function of speaker, linguistic context, and noisy inputs, but human listeners’ judgments of the boundaries among phonemic categories are remarkably consistent. How this is accomplished is a fundamental issue that all theories of speech processing seek to explain. “Speech perception” is typically used to refer to sublexical-level processes, and is measured using tasks such as discriminating between syllables or identifying simple speech sounds. “Speech

recognition”, also referred to as auditory comprehension, is generally understood to be a higher level process where the ultimate goal is to link speech sounds to representations of words in the mental lexicon (see Hickok & Poeppel, 2007, for review and motivation of these distinctions). Traditionally, speech perception is measured by performance on two tasks: discriminating between speech sounds, and identification of speech sound categories. Discrimination involves judging whether two speech sounds are the same or different (e.g., ba–da vs. ba–ba), while identification

*Corresponding author. Email: mahon@rcbi.rochester.edu

(also referred to as labelling) involves identifying a stimulus as one sound or another (e.g., ba vs. da). Identification requires that auditory information is analysed in terms of linguistic categories, and that it is held in short-term memory long enough to be able to encode a response (Blumstein & Cooper, 1972). Here we suggest (and provide evidence) that discrimination (between speech sounds) and identification (of speech sounds) may in fact tap separable components within what is generally referred to as “speech perception”.

As originally stated by Liberman, Harris, Hoffman, and Griffith (1957), speech perception is *categorical* when individuals’ discrimination ability is better for stimuli belonging to different phonemic categories than for stimuli belonging to the same phonemic category. Classically, “categorical perception” is determined by the extent to which the discrimination of auditory stimuli is *predicted* by the identification of the same stimuli (Eimas, 1963; Macmillan, Kaplan, & Creelman, 1977) and is evidenced by a categorical, as opposed to a continuous, perceptual boundary.

The identification and discrimination of phonemes has been shown to doubly dissociate in both healthy individuals (Gerrits & Schouten, 2004) and in an early study with Wernicke’s aphasics (Blumstein, Cooper, Zurif, & Caramazza, 1977). Blumstein et al. (1977) found that Wernicke’s aphasics were unable to reliably identify the same stimuli that they could discriminate. The authors concluded that those patients were unable to use phonological information in a linguistically relevant way. To foreshadow our results, we find that a patient with Broca’s aphasia and apraxia of speech is impaired for identifying sounds that he can discriminate. However, and unlike the patients in the study by Blumstein et al. (1977), our case is able to map sound onto meaning as shown by intact auditory comprehension.

Broadly speaking, two classes of explanations have been proposed about the mechanisms involved in the initial stages of speech perception. Motor theories of speech perception posit that the correspondence between auditory information and phonemic categories occurs over the motor representations involved in producing those sounds (D’Ausilio et al., 2009; Devlin & Aydelott, 2009; Galantucci,

Fowler, & Turvey, 2006; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Liberman & Mattingly, 1985; Pulvermuller et al., 2006). Non-motor accounts specify how the auditory system recovers discrete phonetic features from a complex acoustical analysis without the involvement of the motor system (Massaro, 1998; Massaro & Chen, 2008; for a review see Diehl, Lotto, & Holt, 2004), and support the classic functional segregation of production and perception processes.

In the last decade there has been a resurgence of interest in motor theories of action recognition generally, and speech perception specifically, due to research on the putative mirror properties of some motor systems (Buccino et al., 2005; D’Ausilio et al., 2009; Devlin & Aydelott, 2009; di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995; Fadiga, Craighero, Buccino, & Rizzolatti, 2002; Galantucci et al., 2006; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Hauk, Johnsrude, & Pulvermüller, 2004; Keysers et al., 2003; Meister, Wilson, Deblieck, Wu, & Iacoboni, 2007; Möttönen & Watkins, 2009; Pulvermüller et al., 2006; Rizzolatti & Arbib, 1998; Rizzolatti, Fogassi, & Gallese, 2001; for a discussion see Dinstein, Thomas, Behrmann, & Heeger, 2008; Lingnau, Gesierich, & Caramazza, 2009). This renewed interest in the motor theory of speech perception has led to a number of recent suggestions that the motor system plays a significant role in the processing of speech sounds. Data from transcranial magnetic stimulation (TMS), functional neuroimaging, and neurophysiology indicate that frontal motor structures are automatically engaged during passive speech perception (Fadiga et al., 1995, 2002; Hesslow, 2002; Kiefer & Pulvermüller, 2012; Rizzolatti & Craighero, 2004; Scheerer, 1984). For instance, Fadiga et al. (2002) found an increase in motor-evoked potentials recorded from a listener’s tongue muscles during a task in which participants heard speech sounds but for which there was no explicit motor component. Watkins, Strafella, and Paus (2003) applied TMS to the face area of the primary motor cortex in order to elicit motor-evoked potentials in the lip muscles. They found that, in comparison with control conditions (listening to non-verbal sounds and viewing eye and brow movements), both listening to and

viewing speech increased the size of the motor-evoked potentials. They concluded that both auditory and visual perception of speech leads to activation of the speech motor system.

A core prediction of motor theories of speech processing is that damage (whether temporary or permanent) to the speech motor system should impair auditory speech processing. In line with this prediction, Meister et al. (2007) found that when repetitive TMS was used to temporarily suppress the premotor cortex, participants were impaired at discriminating stop consonants embedded in noise. Also using repetitive TMS, Möttönen and Watkins (2012) found that temporarily disrupting the lip representations in the left motor cortex disrupted subjects' ability to discriminate between lip-articulated speech sounds, but did not affect those participants' ability to discriminate sounds that were not lip articulated. Those two studies suggest that disruption of the speech motor system can (subtly) impair speech sound processing. An important and unresolved issue concerns the dynamics of spreading activation between sensory and motor systems: in order for the findings from those studies to be taken as support for the motor theory of speech perception, it would have to be argued that the effects of TMS are (only) local to the inactivated region, and do not spread to regions connected to the motor areas. Currently, it remains unknown whether this assumption is correct (Dinstein et al., 2008; Galantucci et al., 2006; Hickok, 2000, 2010; Hickok, Houde, & Rong, 2011a; Lotto, Hickok, & Holt, 2009; Rogalsky, Love, Driscoll, Anderson, & Hickok, 2011; Stassenko, Garcea, & Mahon, 2013; Toni, de Lange, Noordzij, & Hagoort, 2008).

It is also theoretically relevant that the effects observed with TMS are present only when stimuli are degraded by noise, and are usually observed in response times rather than accuracy. Small effect sizes may not invite meaningful conclusions to be drawn regarding the causal role of the motor system in speech perception (Hickok, Houde et al., 2011). An alternative explanation is that the motor system is not necessary for speech perception, but is recruited during more complex processing, as in the case of ambiguous (i.e., high noise) listening conditions. Consistent with that

supposition, Meister et al. (2007) found that disrupting the ventral premotor cortex impaired speech discrimination for degraded stimuli masked with white noise. Sato, Tremblay, and Gracco (2009) found that repetitive TMS over the ventral premotor cortex impaired discrimination only when phonemic segmentation was necessary. Another recent study by D'Ausilio, Bufalari, Salmas, and Fadiga (2012) used TMS over the lip and tongue motor areas and found that the motor system was recruited in speech discrimination only under noisy conditions. Taken together, these findings suggest that the motor system may be recruited when the task necessitates identification of speech sounds, and that when stimuli are degraded with noise and embedded in a discrimination task, the task actually becomes one of identification. If this is correct, then it suggests the interesting prediction that the motor system plays no role at all in the discrimination of speech in noise-free environments. This issue can be addressed directly by studying patients with motor impairments in speech production.

In contrast to motor theories of speech perception, feed-forward models allow for a close coupling between speech-specific acoustic analysis and inferred or recovered articulatory processes, but suggest that the primary direction of information flow is from perception to production (Hickok, 2010; Toni et al., 2008; also see Hickok & Poeppel, 2000, 2004, 2007). One example of feed-forward models is the dual-stream model of speech perception, which proposes two main pathways for speech processing. The ventral stream is composed of the superior and middle parts of the temporal lobe, and processes acoustic signals with the goal of lexical access and semantic comprehension. The dorsal stream includes the posterior frontal lobe and the superior/posterior aspect of the temporal lobe, and supports auditory-motor integration by mapping acoustic signals in the temporal lobe to frontal lobe articulatory networks (Hickok & Poeppel, 2007; Hickok, 2012). On this account, bilateral primary and secondary auditory cortices, and not frontal motor regions, are primarily responsible for speech perception (e.g., Davis & Johnsrude, 2007; Hickok & Poeppel, 2007; Scott & Johnsrude, 2003). This class of models can explain

recent demonstrations of motor system activity during speech perception, and specifically via temporal-to-frontal projections that constitute the non-lexical “dorsal” speech perception pathway. This type of model would predict that the intact ability to perceive and understand speech in a patient with damage to the frontal lobe is supported by the ventral route. In contrast, motor theories of speech perception would predict that an individual with a damaged speech motor system should show gross impairments in speech perception, across the board.

One of the strongest means for adjudicating between motor theories and forward models is by studying neurological patients with focal brain lesions to the speech motor system. If those systems play a causal role in encoding speech categories, then damage to those systems should have catastrophic effects on tasks requiring either discrimination or recognition of phonemic categories. Previous research with neurological patients has provided evidence incompatible with motor theories. For example, patients with lesions to the left inferior frontal gyrus and/or the left inferior parietal lobule (both seats of the putative human mirror-neuron system) can succeed at various comprehension and discrimination tests (Rogalsky et al., 2011). Rogalsky et al. also found that patients with temporal lobe damage had more severe comprehension deficits, consistent with the view that the temporal lobes are necessary, and potentially sufficient, for many aspects of auditory speech processing (for a review see Diehl et al., 2004; Stassenko et al., 2013). Hickok, Costanzo, Capasso, and Miceli (2011b) studied 24 individuals with non-fluent speech production subsequent to strokes affecting Broca’s area. In that study, damage to Broca’s area did not result in significant speech perception deficits, as tested with auditory syllable discrimination (i.e., distinguishing “ba” from “pa”) and auditory word discrimination (i.e., distinguishing “pear” from “bear”). In addition, those patients also performed well on auditory word and visual word comprehension tasks (for similar studies see Baker, Blumstein, & Goodglass, 1981; Blumstein et al., 1977; Moineau, Dronkers, & Bates, 2005; for reviews see Hickok, 2010; Stassenko et al., 2013).

What we currently do not know, and what is critical for advancing the discussion on the role of the motor system in auditory speech processing, is whether a patient with a demonstrated lesion to the speech motor system will be able to perceive (i.e., discriminate) and identify (i.e., label) speech sounds. Although the studies reviewed above have shed light onto this issue, to our knowledge the current study is the first examination of categorical perception in a case of apraxia of speech which utilizes identical stimuli in both discrimination and identification tasks.

Experimental investigation

Case summary

Case AD is a 55-year-old, right-handed male who suffered a left hemisphere ischaemic stroke in January 2010 affecting the inferior frontal gyrus, premotor cortex, and primary motor cortex. The magnetic resonance imaging scan shown in [Figure 1](#) depicts the extent of the patient’s lesion. AD presented with non-fluent speech that was marked by frequent articulatory/phonological errors. An initial analysis found that his production impairment was present in picture naming (38/80; 48% of word productions were error free) and spoken single-word repetition (16/24; 67% of repeated words were error free). The fact that phonological/articulatory errors were present in both picture naming and word repetition suggests that production processes were in fact the source of the errors.¹ AD was receiving speech therapy for apraxia of speech and expressive aphasia. A thorough neuropsychological investigation, which included the full Boston Diagnostic Aphasia Examination, showed that AD was normal on tests of mid- and high-level visual processing, episodic and semantic memory, following commands, and was well oriented in space and time. AD made some (but few) semantic errors in word comprehension, as assessed by “Word Comprehension by Categories”, a subtest of the Boston Diagnostic Aphasia Examination (27/30; 90% correct). For example, when asked to point to a lion, he pointed to a tiger. While his performance on this task was significantly different from normal performance, as shown below, he exhibited normal

performance in auditory comprehension using minimal pairs tests (i.e., point to the bear as opposed to the pear, when hearing “bear”). His reading of regular words was decipherable, except that his responses were marked by frequent articulatory/phonological errors.

Experimental procedures

Six age- and education-matched controls performed the same neuropsychological tests as patient AD, unless otherwise noted (all controls were right handed, as was AD, and were 45–63 years old: mean = 53, SD = 7.8; years of formal education after high school for controls: mean = 3.7, SD = 2.3; for AD = 4). All participants gave informed, written consent for participation, which was approved by University of Rochester’s Research Subjects Review Board.

General behavioural methods

For all behavioural tasks, unless otherwise noted, AD was instructed to respond quickly and accurately on every trial. Each stimulus was presented on the screen for 10 seconds, or until a response was produced, whichever came first. If AD did not produce a response within 10 seconds, the trial was counted as incorrect. All pictorial stimuli were 400 × 400 pixels and gray scaled, unless otherwise stated. All stimuli were presented on a monitor at a viewing distance of approximately 60 cm (1680 × 1050 pixels, temporal resolution = 120 Hz). Stimulus presentation for experiments requiring overt verbal responses as well as all categorical perception experiments was controlled with DMDX (Forster & Forster, 2003). For production experiments, individual trial “.wav” files containing the participant’s response, as well as high-definition audio recordings of the entire session, were scored offline. All other experiments were pen-and paper tests and came from the Psycholinguistic Assessment of Language Processing in Aphasia Battery (PALPA; Kay, Lesser, & Coltheart, 1992). Unless otherwise noted, modified *t* tests were used to assess the performance of AD referenced to control participants (Crawford & Garthwaite, 2002).

Picture naming: methods and analysis

An analysis was carried out over 202 naming trials (Snodgrass & Vanderwart, 1980), which consisted of 95 “recognizable” responses that contained phonological errors. Those 95 responses were broken down into 162 syllables, across which there were a total of 204 phonological errors. An analysis of the type and distribution of his phonological errors indicated that the majority were substitutions (53%, “bed” → “ped”), followed by omissions (29%, “bed” → “ed”), additions (12%, “bed” → “bred”), perseverations (3%, “bed” → “beb”), and anticipations (2%, “bed” → “ded”). There was no bias for errors to be unevenly distributed across syllabic positions, as 42% of AD’s phonological errors occurred at the syllable onset position, 27% at the vowel, and 31% at the coda position. However, 40% of the phonemes across the words were in the onset position, 34% in the vowel position, and 26% in the coda position (see Figure 2). Thus, the distribution of errors by syllabic position closely mirrored the distribution of opportunity for error. The main finding from this analysis is that the majority of the errors were substitutions, as commonly observed in apraxia of speech (Johns & Darley, 1970).

Psycholinguistic tests: methods

Widely used tests for the integrity of auditory speech processing use minimal pair stimuli. Those tests can use a discrimination task (e.g., indicate whether two words are the same or different, bear–pear) or an identification task (e.g., point to a picture corresponding to a spoken word, “bear”, while ignoring a distractor picture corresponding to a minimal pair, e.g., “pear”; see e.g., Rogalsky et al., 2011). To assess AD’s auditory speech processing a subset of minimal pair experiments were taken from the PALPA. Foils were either minimally different or different by two or more distinctive features. The difference between target and foil occurred at either initial, final, or metathetic stages and the difference was in either voice, manner, or place of articulation. On “Discrimination Requiring Picture Selection” (PALPA Test 4), an example of an auditory target was

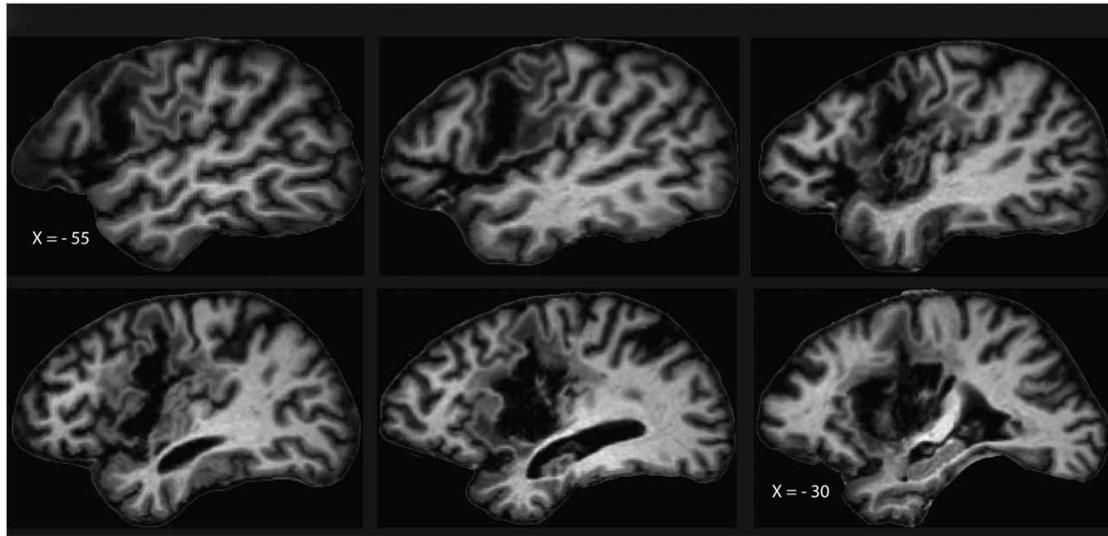


Figure 1. Case AD's lesion. Magnetic resonance imaging scans of case AD's left hemisphere lesion, acquired approximately 3 years after his stroke. AD's left hemisphere ischaemic stroke resulted in damage to areas of the inferior frontal gyrus (Broca's area) as well as the precentral gyrus (i.e., primary motor cortex) and premotor/supplemental motor cortices. Slices are shown lateral to medial in the sagittal view (Talarach X coordinates range from -55 to -30 , in steps of 5 mm).

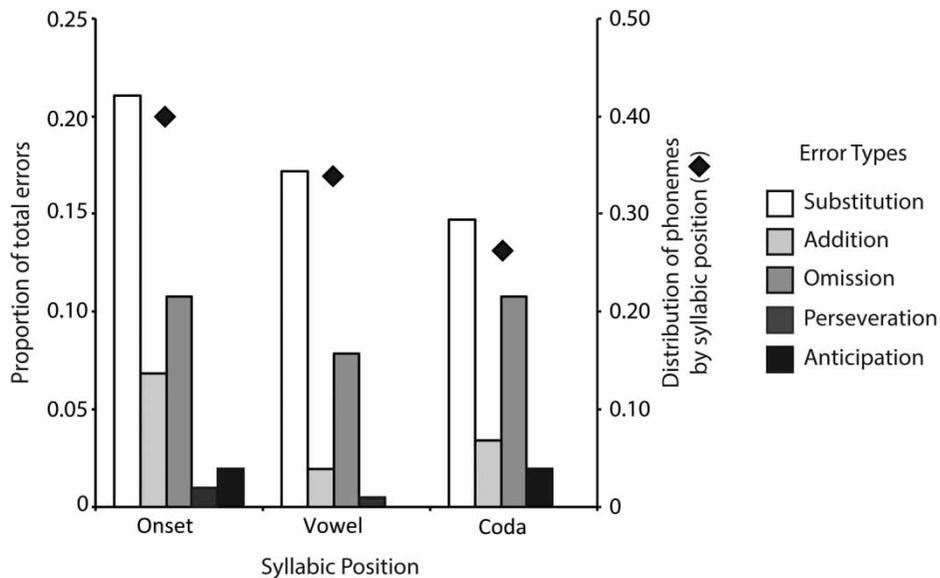


Figure 2. Types of errors in speech production. Distribution of case AD's phonological/articulatory errors by syllabic position and type during a picture-naming task. Also plotted on the same axis (black diamonds) is the distribution of the opportunity for error—that is, the percentages of phonemes at each syllabic position.

“goat” and the subject had three pictures from which to pick: goat, coat, and boat. On “Discrimination Using Word Minimal Pairs” (PALPA Test 2), two words were presented auditorily with flat intonation and the participant was asked to state whether the pair was the “same” or “different”. On “Discrimination Using Non-word Minimal Pairs” (PALPA Test 1), the task was the same as PALPA Test 2 but the two sounds were pseudo-words that respected the phonotactic constraints of English.

Psycholinguistic tests: results

AD was normal for discriminating both spoken non-word (66/72, $p = .73$) and spoken word minimal pairs (70/72, $p = 1$), and was normal at matching spoken words to pictures (39/40, $p = .22$; see Table 1). Therefore, using standard psycholinguistic assessments, case AD does not present with significant auditory speech processing difficulties—a pattern observed in previous cases with similar lesions (e.g., Hickok, Costanzo et al., 2011; Rogalsky et al., 2011). However, there may be some suggestion of a mild deficit in lexical or semantic access from auditory input, as indicated by the “Word Comprehension by Categories” test discussed above (see Case summary).

Additional language and neuropsychological assessment

As part of AD’s general neuropsychological work up, his general intelligence, spontaneous speech, verbal fluency, verbal working memory, repetition, reading, and spelling abilities were assessed. On the Kaufman Brief Intelligence Test, AD scored in the “above average” range for non-verbal IQ and in the “average” range for verbal IQ (123 and 91, respectively). AD’s spontaneous speech, as evaluated with the classic “Cookie Theft” picture, was markedly impaired as shown by a ‘sentence’ such as “Water ... shoes ... wet ... cookie jar ... baker ... mother ... apron”. AD was given a category fluency task, in which he had a minute to name as many nouns that fit into a specific category as possible. Categories of interest were animals, fruit, vegetables, tools, and vehicles. AD came up with two

animals, five fruit, five vegetables, three tools, and four vehicles. On a similar phonological fluency task, in which AD had to name as many nouns that started with the letters “A” and “S”, he came up with one word for each letter.

Digit span

AD was asked to repeat auditorily presented digits forward and backward; the digits ranged in value from one to nine, were binned in intervals from one to nine, were randomly presented, and did not repeat with an interval. AD was asked to repeat the digits as quickly and accurately as possible. Within each interval AD was given three opportunities to correctly repeat the digits; if he made three errors in a row, the experiment ended.

AD repeated three digits in the forward digit span and was significantly different from age-matched controls whose average was 7.8 digits ($t(5) = -3.80$, $p < .05$), but was not significantly different from controls on the backward digit span (controls’ average = 5; $t(5) = -1.05$, $p = .17$). From his performance on the forward task, we can infer that AD’s verbal working memory was not normal.

Word and non-word repetition

To assess AD’s single-word repetition ability, the “Syllable Length Repetition” (PALPA Test 7) was administered, which specifically assesses the effect of syllable length on auditory word repetition. Equal numbers of one, two, and three syllable words matched for lexical frequency, imageability, and morphemic complexity were presented to AD in random order. AD was able to repeat all 24 words. Although responses were marked by articulatory/phonological production errors (see Case summary for error analysis), AD always was close to the target word and (recognizably) repeated every word that was presented. No effect of syllable length was found. Aside from his phonological/articulatory errors, AD’s performance with single word repetition is spared.

Table 1. Psycholinguistic assessment of speech perception and recognition.

Test	Control sample			AD's score	Significance test	
	N	%	SD	%	t	p
PALPA Test 1	6	0.94	0.05	0.92	-0.37	.73
PALPA Test 2	6	0.97	0.02	0.97	0	1
PALPA Test 4	6	0.99	0.01	0.98	-1.39	.22

Note: Number of control participants (N), mean control proportion correct (%), control standard deviation (SD), AD's proportion correct (AD's score), and *t* and *p* values for three tests of the Psycholinguistic Analysis of Language Processing in Aphasia Battery. PALPA Test 1 = Same-Different Discrimination Using Nonword Minimal Pairs; PALPA Test 2 = Same-Different Discrimination Using Word Minimal Pairs; PALPA Test 4 = Minimal Pair Discrimination Requiring Picture Selection.

Next, AD was asked to repeat non-words (i.e., pseudo-words) from PALPA. AD correctly repeated only 11 out of 30 non-words (control data not available—by “correct” is meant that the repeated word was recognizable as the target, although correctly repeated non-words did have phonological/articulatory errors). AD's incorrect responses were either incomplete (e.g., “slurch” → “ur”) or were real words (e.g., “sprawn” → “fun”).

Sentence repetition

To assess AD's ability to repeat back more than one word, a sentence repetition task (PALPA Test 12) was used. AD was able to fully repeat only 3 out of 18 sentences. Errors included failing to include all words (e.g., “The cat is carried by the horse” → “The cat ... is ... horse”), not finishing the sentence (e.g., “The girl's frightening the dog” → “The girl is frightening ...”), or substituting other words (e.g., “The dog is washed by the girl” → “dog is washed by the man”). This task was repeated immediately after with the sentences broken down into two or three words at a time, and AD was significantly better at repeating the two- or three-word chunks (example responses by the patient are reproduced not showing articulatory/phonological errors).

Reading

To assess AD's reading ability, “Grammatical Class & Imageability” (PALPA Test 33) was administered, in which 20 nouns and functors were visually

presented for AD to read; all words were equally imageable. Compared with age-matched controls, AD was impaired on this task (30/40; 75%, $t(5) = -4.90$, $p < .01$). Errors were phonological/articulatory.

AD was also assessed for his ability to read pseudo-words. AD was asked to name 24 three-, four-, five-, or six-character monosyllabic non-words (PALPA Test 36). Relative to age-matched controls, his performance for reading three-character non-words was marginally impaired (2/6, 33%, $t(5) = -1.80$, $p = .07$). Performance on four-character (2/6, 33%, $t(5) = -3.90$, $p < .05$) five-character (2/6, 33%, $t(5) = -2.70$, $p < .05$) and six-character non-words (0/6, $t(5) = -3.80$, $p < .01$) was impaired. Reading errors were similar to errors in non-word repetition (see section above) and were either incomplete, real words, and/or articulatory/phonological.

Oral spelling

To assess AD's oral spelling abilities, “Letter Length Spelling” (PALPA Test 39) was administered. AD correctly spelled only 3/6 three-letter words (example of an error: “key” → “ked”). He was unable to spell words beyond three letters long.

Ultrasound imaging of the tongue

Methods

To obtain real-time measurements of the tongue, a transducer was positioned beneath the patient's chin, with the head stabilized using a chin rest.

The transducer was maintained at a constant orientation and pressure during the recording. A sequence of ultrasound images was acquired of the mid-sagittal section of the tongue at 30 Hz. The ultrasound machine included an internal clock with a video timer displayed during the recording. Because the ultrasound machine did not have in-built audio recording, a high-quality audio recorder was simultaneously used during the entire session.

The ultrasound video recordings were converted to frames at 30 frames per second for further analysis. Therefore, every second in the video corresponded to 30 unique data points reflecting the tongue's location as observed by the ultrasound. To interpret tongue shape from the two-dimensional images, EdgeTrak (Li, Kambhmettu, & Stone, 2005) was used to manually define the shape of the tongue and extrapolate tongue contours based on edge gradient and intensity information captured by the ultrasound. Contours were aligned and the tongue, at each frame, was defined by a total of 100 contours. For reliability purposes, two researchers independently aligned the audio recording to the video recording and extracted the frames that corresponded to specific utterances. Any two conflicting frame durations between the two researchers were discussed and agreement was reached. SURFACES (Parthasarathy, Stone, & Prince, 2005), which implements an algorithm for estimating spatial regression models in MATLAB (Version 6; The MathWorks, Natick, MA, USA), was used to estimate a smoothed tongue contour with the manually extracted sample values from EdgeTrak. SURFACES was subsequently used for visualizing the spatiotemporal image of the tongue by plotting the change in position of contour over time within a three-dimensional display (Parthasarathy, Stone, & Price, 2005).

Results

Here, we focus on the results when AD was producing tokens of the same utterances tested below perceptually (“aba”, “aga”, “ada”). AD produced nine exemplars of each pseudo-word. Figure 3 depicts examples of AD's spatio-temporal tongue surface

during repetition of “aga” over time. The variability in tongue position directly relates to the types of phonological/phonetic errors, as demonstrated by the corresponding spectrograms from those same utterances (see Figure 3). It is noteworthy that many utterances produced by case AD were “false start” hesitations, as he had a general difficulty in initiating speech, as demonstrated by groping movements of the tongue before an utterance was produced. Such difficulties are classic symptoms of the motor speech disorder known as apraxia of speech (AOS), which is also referred to as verbal apraxia (Wertz, LaPointe, & Rosenbek, 1984). The data from the ultrasound imaging of the tongue provide a demonstration of the patient's impairment at the level of articulation.

It was important to establish that there is a clear motor deficit in the patient's speech production, before any claims regarding a theoretically significant dissociation between speech production and categorical speech perception can be made. It is often the case that AOS co-occurs with Broca's aphasia, despite the two “disorders” also being observed to dissociate. We found it important to confirm that the current case presents with features of AOS, which is a sensorimotor impairment for selecting, programming, and executing coordinated movements of the speech musculature for the production of voluntary speech (Wertz et al., 1984). This description resonates well with the pattern of difficulties that we have documented for AD.

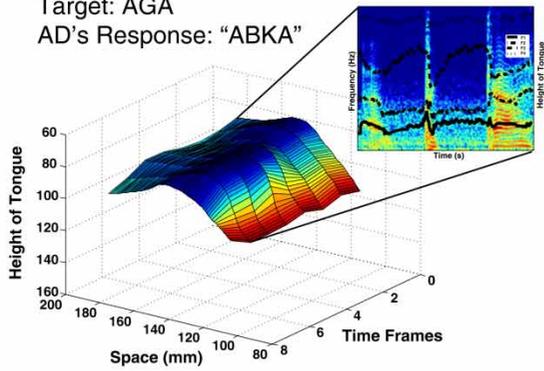
Categorical perception experiments

General methods

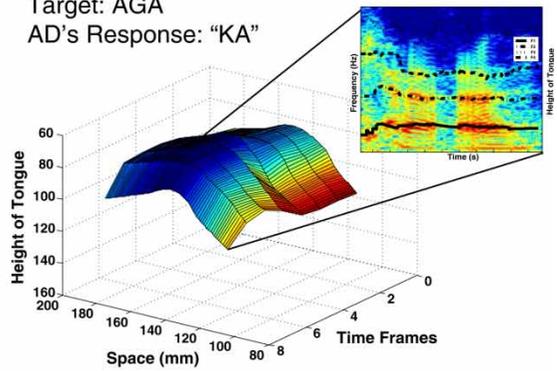
We sought a direct assay of the role of the motor system in auditory speech processing by testing discrimination and identification of non-word minimal pairs. The contrasts used were the /aba/–/ada/ continuum and the /ada/–/aga/ continuum. The stimuli were taken from a publicly available set of normed stimuli (Stephens & Holt, 2011). These stimuli were created by separating source (e.g., glottal excitation) and filter components (vocal tract resonances) of naturally produced tokens from a male speaker and linearly interpolating between sets of linear-predictive coefficients (via autocorrelation)

A First AGA Dyad

Target: AGA
AD's Response: "ABKA"

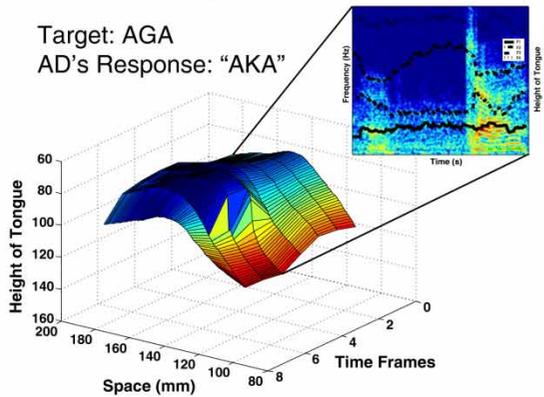


Target: AGA
AD's Response: "KA"

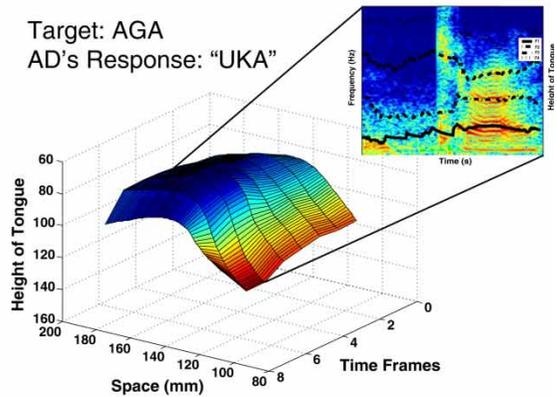


B Second AGA Dyad

Target: AGA
AD's Response: "AKA"

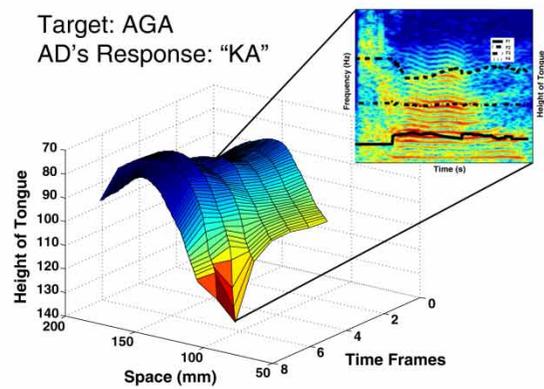


Target: AGA
AD's Response: "UKA"



C Third AGA Dyad

Target: AGA
AD's Response: "KA"



Target: AGA
AD's Response: "KAI"

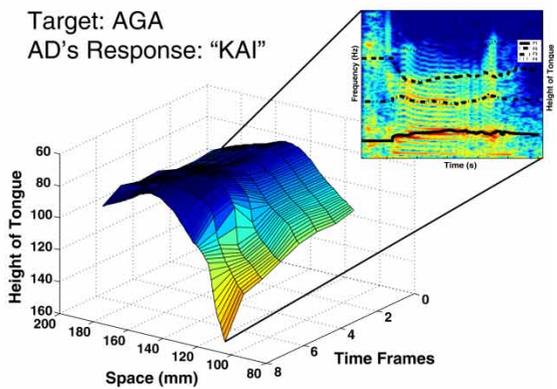


Figure 3. Graphs showing the height of the spatio-temporal tongue surface over time during case AD's repetition of "aga", separated by dyads, and accompanied by his actual response and a spectrogram analysis of the ".wav" file.

of the endpoint filters to create intermediate filters, and resynthesizing new tokens using a single source wave and the new filters. For detailed methods see Stephens and Holt (2011). Eleven tokens, two acoustic steps apart (example: abd-1, abd-3, abd-5, etc.) were selected from the normed stimuli of Stephens and Holt that systematically varied the place of articulation between the two tokens.²

Speech sound discrimination: methods

In order to test whether AD had a normal categorical boundary when discriminating speech sounds, a task was used in which AD was presented with two tokens from the same continua on every trial and asked to indicate (yes/no) if they were the same or different (button press, two-alternative forced choice paradigm). Each token was paired with every other token as well as itself. This was counterbalanced so that each token appeared as a first sound and as a second sound for every pairing, an equal number of times. There were a total of 180 trials in 1 “session”—90 “different” trials and 90 “same” trials (i.e., abd-1 paired with abd-1). All trials were presented in random order. The inter-stimulus interval was 750 ms. For each trial, subjects responded by pressing the left or right shift key corresponding to whether the two sounds were the “same” or “different”. If no response was elicited, the next trial was programmed to begin after 3 seconds (no trial was aborted during runs of the experiment). The volume was set at 80 (dB SPL) for all subjects and noise-cancelling headphones were used. Both patient and controls completed 3 sessions (total of 540 trials) of the aba-ada discrimination task, and 2 sessions (total of 360 trials) of the ada-aga discrimination task.

Speech sound discrimination: results

For the analysis of the discrimination data, we constructed a confusion matrix consisting of the proportion of “different” responses for each pair of stimuli for each subject. The dimensionality of the data was then reduced from two dimensions to one using multidimensional scaling (in MATLAB). Figure 4A shows the classical multidimensional-

scaling (MDS) solution for the ABA-ADA and ADA-AGA continua, indicating that AD had a normal perceptual boundary. The mean MDS solutions for controls were nearly perfectly correlated with AD’s MDS solutions for both the ABA-ADA ($r = .95$, $t(9) = 8.94$, $p < .001$) and ADA-AGA ($r = .99$, $t(9) = 24.18$, $p < .001$) continua. These data undermine the central tenet of the motor theory of speech perception, as AD showed normal categorical perception in a discrimination task for the same stimuli he could not produce due to a confirmed impairment to articulatory processes.

Speech sound discrimination: discussion

A key aspect of the stimuli used to assess speech discrimination is that they cannot be processed lexically. Thus, normal performance in the discrimination task could not be supported by any cues (i.e., lexical or semantic) other than what is generated “bottom up” by an analysis of the input itself. The flipside of this is that because the patient’s lesion affects the dorsal speech processing stream, any speech processing tasks requiring that stream should be impaired. We thus hypothesized that when lexical and semantic cues are not available to parse the input (because we used non-word stimuli), but the dorsal speech pathway is blocked due to a lesion, then the patient should be impaired for labelling the non-word stimuli. We tested this using the same stimuli that were used to test speech perception, but by asking AD to label (identify) a single stimulus on every trial.

Speech sound labelling: methods

The same stimuli as were used in discrimination were used in labelling. Each token (1–11) was presented 4 times in one block and there were a total of 5 blocks, for a grand total of 220 trials in each “session”. All trials were randomized within each block. For each trial, subjects were instructed to press the left or right shift key based on whether they heard “aba” or “ada” in the abd labelling task or “aga” or “ada” in the agd labelling task (two-alternative forced choice design). To remind the subjects of which button corresponded to which sound, on the left and right side of the

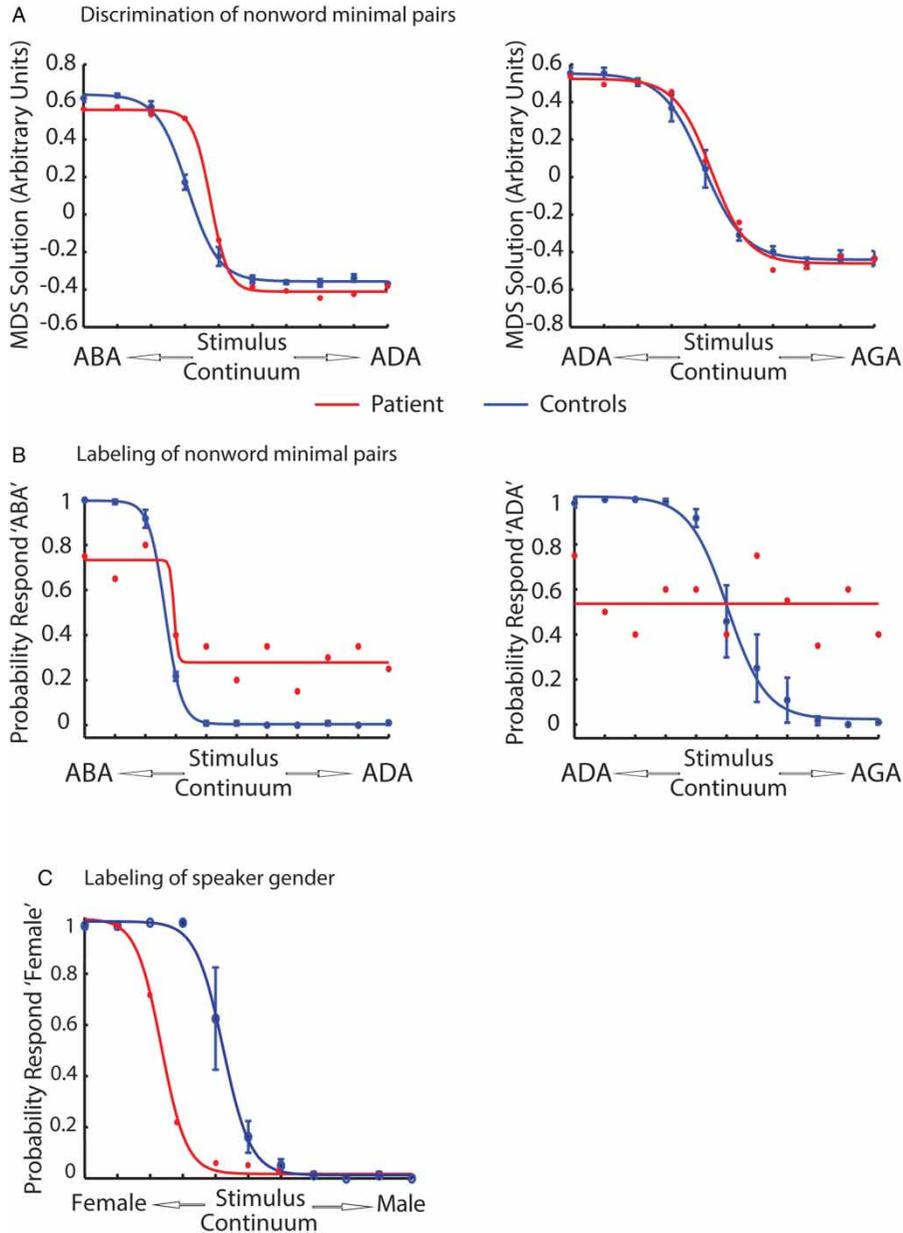


Figure 4. Perceiving speech sounds without identifying them. (A) Multi-dimensional scaling solution for the discrimination of non-word minimal pairs plotted for both continua (ABD and ADG), for six age-matched controls and case AD. (B) Plot of the labelling function for the same continua for all subjects. (C) Control task, which plots the labelling function of “female” versus “male”. For all panels, the red line represents AD and the blue line is the average of controls. Error bars indicate the standard error of the mean of controls.

screen, respectively, were the written words “ABA” and “ADA”, and “AGA” and “ADA”, for which the positions were fixed. Both patients and controls

completed 1 session (5 blocks) of the abd labelling task and 1 session (5 blocks) of the adg labelling task.

Speech sound labelling: results and discussion

In marked contrast to AD's normal discrimination, AD was impaired at labelling the same stimuli (see Figure 4B). Multilevel, logistic regressions with random intercepts and slopes by subject and centred sound-item continuum confirmed that the slopes of his labelling functions were shallower than those of each of the six control subjects (interaction of patient/control group by stimulus dimension, ABA–ADA continuum: $B = 1.17$, $SE = 0.11$, $z = 10.99$, $p < .001$; ADA–AGA continuum: $B = 1.17$, $SE = 0.18$, $z = 6.69$, $p < .001$). In fact, while AD showed some evidence of minimally preserved labelling ability with stimuli from the ABA–ADA continuum (slope test: $B = 0.11$, $SE = 0.02$, $z = 4.79$, $p < .001$), the slope test for labelling the stimuli from the ADA–AGA continuum was not significant ($B = 0.02$, $SE = 0.02$, $z = 1.21$, $p = .23$)—indicating no boundary for the same stimuli that showed a normal boundary during perception.

It is unclear how to interpret AD's slightly better performance in labelling on the ABA–ADA continuum than on the ADA–AGA continuum. One speculation may be that the ADA–AGA continuum involves movement with the tongue, whereas “aba” involves the lips. However, at present this must remain speculation.

We confirmed that the patient's failure in the labelling task was not the result of being confused about which button was the target, in several ways. First, the location of response items was kept constant throughout each experiment. Second, in some experimental sessions we used training (with feedback) on the anchors in between experimental blocks: intermittent, 10-trial training sessions (with feedback) on the continuum endpoints between 20-trial experimental blocks did not improve performance. It is also important to note in this context that the patient was normal on the equivalent task when the response space was partitioned by pictures (e.g., bear/pear). This suggests that the patient's failure may have to do with the demands that are placed on the speech motor system when other cues (e.g., semantics, context) are not available.

Labelling of speaker gender: methods and results

We sought to have a positive control to show that the patient is able to perform a labelling task that shares, in all relevant aspects, the structure as the task that he failed. We thus used a categorical labelling task requiring the patient to label speaker gender. Stimuli for the control task were obtained online, created by Kawahara, Masuda-Katsuse, and De Cheveigné (1999). Using the same design as the above labelling tasks, 11 tokens of the word “right” were presented in a two-alternative forced choice format (“male” or “female”; button press). The tokens correspond to graded changes created by morphing a male voice to a female voice. In this experiment the controls were undergraduates recruited from the University of Rochester (i.e., aged-matched control data were not available). Case AD showed a steep boundary for this task, similar to controls (see Figure 4C). However, while a discrete cut-off exists, the patient's boundary is shifted from that of controls. One possibility for this shift may be due to including only three controls that were not age matched, resulting in possible age-related hearing differences at (higher) frequencies that contain information about gender. However, this difference does not pose a challenge for the main finding that case AD is sensitive to speaker gender differences, because there are no errors at the endpoints of the continuum (as compared with the phonemic labelling tasks). This positive finding indicates that the dissociation between speech perception and speech labelling described above is not due to a simple change of task demands, but is rather due specifically to the task demands of labelling speech sounds with phonemic categories.

Speech sound labelling: discussion

The intact speaker-gender labelling performance of case AD also allows us to exclude a range of general explanations that might be advanced for the source of the dissociation in phoneme discrimination versus identification. For instance, Blumstein and Cooper (1972) distinguished the processes underlying phoneme labelling versus

discrimination. They found that in a labelling task subjects made more errors for stimuli differing by two articulatory features, whereas in the discrimination task subjects made more errors for stimuli differing by one feature. They argued that the observed difference arises in part because labelling differentially draws on working memory compared with discrimination. Importantly, performance with just one distinctive feature difference did not lead to performance differences between labelling and discrimination (i.e., the difficulty level was the same). The findings that we have reported cannot be accounted for in terms of a working memory impairment for speech sounds. First, it is not at all clear that the working memory demands are greater for the labelling task than the discrimination task, since in the discrimination task participants must process two stimuli before making a response. Second, and most importantly, when the task constraints are held constant, but the decision is changed from one about phonemic contrasts to one about the gender of the speaker, case AD shows a sharp categorical boundary.

General discussion

We have reported that an individual with damage to the speech motor system is able to perceive speech sounds normally, but is impaired for labelling the same speech sounds by phonemic category. Specifically, case AD showed a normal categorical boundary in a speech sound discrimination task but was unable to identify or place a label on those same sounds in an identification task. A general inability to label speech sounds was ruled out with a control task in which the patient was able to successfully label sounds by speaker gender. Furthermore, case AD showed largely intact auditory language comprehension, as shown by minimal pairs tests and word-picture matching tasks. We provided additional evidence that AD's production impairment originated, at least in part, at the level of the speech motor system using ultrasound imaging of the tongue; those ultrasound data indicated that AD's articulatory/phonological errors were related to variability in tongue movement across utterances. Case AD also had difficulties with

reading and writing, probably co-morbid impairments from the large left frontal lesion. Importantly, the type of categorical perception tasks that were used ensured that reading difficulty did not affect the data. Finally, there were evident deficits in non-word repetition, digit span, as well as sentence repetition, possibly suggesting a phonological working memory impairment. Impaired phonological working memory may have some theoretical relevance to the interpretation of our findings in terms of the motor theory of speech perception, and will be discussed further.

Implications for models of speech perception

Our findings indicate that the motor system is not critical for the perception of speech sounds, using a fine-grained discrimination task as an assay of speech perception abilities. This conclusion is contrary to recent arguments (D'Ausilio et al., 2009; Fadiga et al., 2002; Meister et al., 2007; Möttönen & Watkins, 2009; Watkins & Paus, 2004; Watkins, Strafella, & Paus, 2003; Wilson & Iacoboni, 2006; Wilson, Saygin, Sereno, & Iacoboni, 2004). Importantly, some recent studies have been taken to support the theoretical position that the motor system is not merely activated during passive speech perception, but that it also plays a role in discriminating phonemes (Meister et al., 2007; D'Ausilio et al., 2009; Devlin & Aydelott, 2009). Our findings undermine this theoretical position, because motor theories predict that both identification and discrimination of speech sounds should be impaired in a patient like AD.

The dual stream model of speech perception (Hickok & Poeppel, 2000; Hickok & Poeppel, 2004, 2007) offers a cogent account of our findings with AD. The ventral stream, which includes the superior and middle portion of the temporal lobe, is intact in case AD. This can explain his ability to process speech for comprehension and his generally intact higher level language comprehension. In contrast, the dorsal stream, which includes the posterior frontal lobe and superior/posterior aspect of the temporal lobe, is partly damaged in AD. Since this stream may be involved in translating acoustic speech signals

into articulatory representations in the frontal lobe, this may explain case AD's production impairment. Saur et al. (2008) found that sublexical repetition of speech was subserved by the dorsal stream (via connections between the superior temporal lobe and the premotor cortex), while language comprehension was mediated by the ventral stream. Their conclusion that the dorsal stream is linked to sensory-motor mapping of sound to articulation resonates with the current pattern of impairments in case AD. Although speculative in nature, a deficit in auditory-motor integration may also be responsible for his inability to label phonemes, as shown by no sensitivity to phonemic categories in an identification task. In order for the dual stream theory to explain the observed dissociation between impaired labeling and intact discrimination, it would have to be assumed that discrimination occurs entirely within temporal lobe regions (i.e., without involvement of the speech motor system).

A modulatory role of the motor system in speech perception?

What function does the cortical motor system play in auditory speech processing? One proposal is that the motor system is recruited in order to fill in "gaps" in sensory information, and acts as a source of attentional or top-down processing (Shahin, Bishop, & Miller, 2009; D'Ausilio, Bufalari et al., 2012). Others have proposed that the motor system "tunes up" neurons in temporo-parietal regions via feedback connections originating in the premotor cortex (D'Ausilio, Bufalari et al., 2012; Pulvermüller & Fadiga, 2010). Along the lines of the dual-stream model of speech processing, listening to others' speech excites sensory-motor feedback circuits, but this activation is not critical for the purpose of phonemic perception itself. Rather, motor activation in the frontal lobes is observed during passive processing of auditory speech sounds because both the listener's and the speaker's auditory speech information is important for production (Hickok, Houde et al., 2011). Similarly, others have suggested that forward predictions from the motor speech system provide internal feedback

control (Rauschecker & Scott, 2009; Sams, Möttönen, & Sihvonen, 2005; see Hickok, Houde et al., 2011 for discussion).

It is theoretically important to note that AD's discrimination of speech sounds was intact under normal listening conditions. As reviewed earlier, repetitive TMS studies which found that disrupting the motor system impairs speech perception typically used either degraded stimuli masked with white noise (Meister et al., 2007) or tasks with a higher degree of processing load (Sato et al., 2009). It is plausible that the motor system is recruited during the presence of noise and/or high processing load in order to aid perception when processes in the temporal lobe alone are insufficient. A recent study by D'Ausilio, Craighero, and Fadiga (2012) directly addressed this hypothesis by applying TMS to the lip and tongue motor areas. They found that speech discrimination accuracy was significantly reduced in noisy conditions, but was not affected in a noise-free condition using the same discrimination task. They concluded that the motor system does not play a role in noise-free environments in the discrimination of speech (see also Sato et al., 2009). This suggests that the ability to discriminate speech sounds under normal listening conditions does not depend upon an intact motor system, consistent with what we observed in our case study.

A role for the motor system in phoneme identification?

Based on the dissociation between case AD's labeling and discrimination ability, it may be the case that under normal listening conditions the motor system is recruited only when the task necessitates identification and specific segmentation of the speech sounds, rather than simple perception of the speech sounds. Intact frontal-temporal connectivity may be critical for phoneme identification when other contextual information (meaning, context) is not available, while temporal lobe regions (i.e., superior temporal gyrus) may be sufficient to support speech perception as measured by a discrimination task. These hypotheses align with the predictions of the dual-route model of speech perception (Hickok & Poeppel, 2004, 2007),

which proposes that frontal-temporal connectivity (i.e., the dorsal route) is important for auditory-motor integration. We speculate that phoneme identification when other cues are not available (i.e., semantics, context) may involve the mapping of acoustic signals to frontal lobe articulatory networks. Because non-word stimuli were used in the experiments we report, there would be no additional cues available beyond the speech signal itself. This hypothesis also suggests an account of why TMS-induced disruption of motor processes would negatively impact categorical perception when stimuli are embedded in noise (Meister et al., 2007; Möttönen & Watkins, 2012; see also D'Ausilio et al., 2009; D'Ausilio, Bufalari et al., 2012; Fadiga et al., 2002; Watkins et al., 2003; Wilson et al., 2004; for a review see Lotto et al., 2009)—the addition of noise may necessitate a prior identification of the stimuli before they can be perceptually judged. For example, in a discrimination task rendered difficult with the addition of white noise, an individual may implicitly first “label” (i.e., identify) the stimuli before judging them as same or different. Ultimately, due to added difficulty of the noisy condition and/or a higher processing load, there may be a shift in the task from strictly a “same/different” perceptual judgment to a phoneme identification task. This interpretation must remain, at present, speculative because we do not have independent evidence which indicates that adding noise to stimuli in a discrimination task pushes the task toward identification.

Case AD's lesion involving Broca's area agrees with recent functional magnetic resonance imaging findings which indicate that Broca's area participates in categorical perception of phonemes as measured by an identification task (Lee, Turkeltaub, Granger, & Raizada, 2012). Lee et al. (2012) suggest that Broca's area may be involved in translating speech signals into articulatory codes, as proposed by the “dorsal route” of the dual-stream model. Based on the theoretically relevant dissociation between identification and discrimination ability found here, future studies should incorporate both types of tasks in order to better elucidate categorical processing of speech sounds in the brain.

Phonological working memory and impaired speech processing

Of potential theoretical interest is AD's impairment for phonological working memory, as shown by impaired performance on tasks such as digit span, non-word repetition, and sentence repetition. Although speculative in nature, it is plausible that the motor system may be recruited when phonological working memory is disrupted or under duress. This may explain why a working memory deficit is more likely to manifest most strongly with pseudo-words than with real words, and is demonstrated by largely intact single-word repetition ability and impaired non-word repetition.

A recent study by Hickok et al. (2014) tested the hypothesis that speech motor planning and verbal short-term memory depend on partially overlapping neural networks. In a large study with 76 patients, the authors of the study found evidence for the association between speech praxis and verbal short-term memory, as demonstrated by lesions to partially overlapping neural regions such Broca's area (specifically the pars opercularis) and the premotor cortex. The authors suggest that speech praxis and verbal short-term memory depend on common structures. A lesion to these partially shared networks, as in our AOS patient, may thus explain the observed impairments in speech production and verbal short-term memory (i.e., phonological working memory).

Conclusion

To our knowledge, this study is the first in which damage to speech motor processes is shown to at the same time spare speech perception while selectively impairing the identification of speech sounds. Our neuropsychological findings add to the evidence that the strong version of the motor theory of speech perception, which claims the motor system is essential and necessary for perception, is untenable. Critically, our findings suggest that under normal listening conditions the motor system plays little to no role in the ability to perceive phonemes, as measured by discrimination

tasks. This is in accord with findings from repetitive TMS studies in which disrupting the motor system results in only subtle impairments to speech perception when measured by a discrimination task. The secondary, modulatory effect of the motor system becomes evident in the presence of degraded listening conditions or higher processing load. We suggest that motor processes may be required to identify phonemic categories in speech input when lexical semantic cues are not available, perhaps because motor processes are involved in “grouping” or “binding” the perceptual input.

Language is a particularly fruitful domain for studying the relation between production and perception, because the units have clear perceptual (i.e., categorical) boundaries. However, our findings generate expectations about the role of motor processes in action recognition more generally (e.g., manual actions). Specifically, our findings suggest a specific role of motor processes in post-perceptual processing of actions. These findings also suggest that a fruitful direction to pursue in the rehabilitation of patients with apraxia of speech may be to transition from purely perceptual (i.e., discrimination) tasks to labelling, and then to overt production, rather than to focus only on production tasks in and of themselves.

Acknowledgements

The authors would like to thank AD for his dedicated participation in the study, without whom this research would not have been possible. They also would like to thank Stephen McAleavey and Gretchen Schlansker for help with ultrasound imaging.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by the National Science Foundation [grant 1349042] and National Institute of Health [grant NS076176] to B.Z.M. F.E.G. was supported by a University of Rochester Center for Visual Science pre-doctoral training fellowship [NIH training Grant 5T32EY007125-24]. This research was also supported, in part, by Norman and Arlene Leenhouts.

Notes

1. For this analysis, picture-naming accuracy was calculated as the number of correct words produced by the patient out of the 80 Snodgrass and Vandewart pictures. A similar analysis was carried out for spoken word repetition, assessed by the PALPA subtest “Syllable Length Repetition”. Note that no age-matched control data were available.
2. Eleven tokens were selected from the original set of 20, and were spaced two steps apart in the acoustic continuum, with the exception of the last two tokens (abd-19 and abd-20) which were one step apart.

References

- Baker, E., Blumstein, S. E., & Goodglass, H. (1981). Interaction between phonological and semantic factors in auditory comprehension. *Neuropsychologia*, *19*, 1–15.
- Blumstein, S., & Cooper, W. (1972). Identification versus discrimination of distinctive features in speech perception. *Quarterly Journal of Experimental Psychology*, *24*, 207–214.
- Blumstein, S., Cooper, W., Zurif, E., & Caramazza, A. (1977). The perception and production of voice-onset time in aphasia. *Neuropsychologia*, *15*, 371–383.
- Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Cognitive Brain Research*, *24*, 355–363.
- Crawford, J. R., & Garthwaite, P. H. (2002). Investigation of the single case in neuropsychology: Confidence limits on the abnormality of test scores and test score differences. *Neuropsychologia*, *40*, 1196–1208.
- D’Ausilio, A., Bufalari, I., Salmas, P., & Fadiga, L. (2012). The role of the motor system in discriminating normal and degraded speech sounds. *Cortex*, *48*, 882–887.
- D’Ausilio, A., Craighero, L., & Fadiga, L. (2012). The contribution of the frontal lobe to the perception of speech. *Journal of Neurolinguistics*, *25*, 328–335.
- D’Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., & Fadiga, L.

- (2009). The motor somatotopy of speech perception. *Current Biology*, *19*, 381–385.
- Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hearing Research*, *229*, 132–147.
- Devlin, J. T., & Aydelott, J. (2009). Speech perception: Motoric contributions versus the motor theory. *Current Biology*, *19*, R198–R20.
- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, *91*, 176–180.
- Diehl, R. L., Lotto, A. J., & Holt, L. L. (2004). Speech perception. *Annual Review of Psychology*, *55*, 149–179.
- Dinstein, I., Thomas, C., Behrmann, M., & Heeger, D. J. (2008). A mirror up to nature. *Current Biology*, *18*, R13–R18.
- Eimas, P. D. (1963). The relation between identification and discrimination along speech and nonspeech continua. *Language & Speech*, *6*, 206–217.
- Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, *15*, 399–402.
- Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology*, *73*, 2608–2611.
- Forster, K., & Forster, J. (2003). DMDX: a Windows display program with millisecond accuracy. *Behavior Research Methods*, *35*, 116–124.
- Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, *13*, 361–377.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*, 593–609.
- Gerrits, E., & Schouten, M. E. (2004). Categorical perception depends on the discrimination task. *Perception and Psychophysics*, *66*, 363–376.
- Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*, 301–307.
- Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences*, *6*, 242–247.
- Hickok, G. (2000). Speech perception, conduction aphasia, and the functional neuroanatomy of language. In Y. Grodzinsky, L. Shapiro & D. Swinney (Eds.), *Language and the brain* (pp. 87–104). Chicago: Academic Press.
- Hickok, G. (2010). The role of mirror neurons in speech and language processing. *Brain and Language*, *112*, 1–5.
- Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, *13*, 135–145.
- Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron*, *69*, 407–422.
- Hickok, G., Costanzo, M., Capasso, R., & Miceli, G. (2011). The role of Broca's area in speech perception: Evidence from aphasia revisited. *Brain and Language*, *119*, 214–220.
- Hickok, G. & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences*, *4*, 131–138.
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, *92*, 67–99.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*, 393–402.
- Hickok, G., Rogalsky, C., Chen, R., Herskovits, E. H., Townsley, S., & Hillis, A. E. (2014). Partially overlapping sensorimotor networks underlie speech praxis and verbal short-term memory: Evidence from apraxia of speech following acute stroke. *Frontiers in Human Neuroscience*, *8*, 1–8.
- Johns, D. F., & Darley, F. L. (1970). Phonemic variability in apraxia of speech. *Journal of Speech Hearing and Research*, *13*, 556–583.
- Kawahara, H., Masuda-Katsuse, I., & De Cheveigné, A. (1999). Restructuring speech representations using a pitch-adaptive time-

- frequency smoothing and an instantaneous-frequency-based f0 extraction: Possible role of a repetitive structure in sounds. *Speech communication*, 27, 187–207. Retrieved from <http://www.wakayama-u.ac.jp/~kawahara/PSSws/MFmorphAPI.wav>
- Kay, J., Lesser, R., & Coltheart, M. (1992). *PALPA: Psycholinguistic assessments of language processing in aphasia*. Hove, UK: Lawrence Erlbaum Associates Ltd.
- Keysers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L., & Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, 153, 628–636.
- Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*, 48, 805–825.
- Lee, Y., Turkeltaub, P., Granger, R., & Raizada, R. D. S. (2012). Categorical speech processing in Broca's area: An fMRI study using multivariate pattern-based analysis. *The Journal of Neuroscience*, 32, 3942–3948.
- Li, M., Kambhamettu, C., & Stone, M. (2005). Automatic contour tracking in ultrasound images. *Clinical Linguistics and Phonetics*, 19, 545–554.
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74, 431–461.
- Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54, 358–368.
- Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revisited. *Cognition*, 21, 1–36.
- Lingnau, A., Gesierich, B., & Caramazza, A. (2009). Asymmetric fMRI adaptation reveals no evidence for mirror neurons in humans. *Proceedings of the National Academy of Sciences*, 106, 9925–9930.
- Lotto, A. J., Hickok, G. S., & Holt, L. L. (2009). Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences*, 13, 110–114.
- Macmillan, N. A., Kaplan, H. L., Creelman, C. D. (1977). The psychophysics of categorical perception. *Psychological Review*, 84, 452–471.
- Massaro, D. W. (1998). *Perceiving talking faces: From speech perception to a behavioral principle* (Vol. 1). Cambridge, MA: MIT Press.
- Massaro, D. W., & Chen, T. H. (2008). The motor theory of speech perception revisited. *Psychonomic bulletin & review*, 15, 453–457.
- Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., & Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Current Biology*, 17, 1692–1696.
- Moineau, S., Dronkers, N. F., & Bates, E. (2005). Exploring the processing continuum of single-word comprehension in aphasia. *Journal of Speech, Language, and Hearing Research*, 48, 884–896.
- Möttönen, R., & Watkins, K. E. (2009). Motor representations of articulators contribute to categorical perception of speech sounds. *The Journal of Neuroscience*, 29, 9819–9825.
- Möttönen, R., & Watkins, K. E. (2012). Using TMS to study the role of the articulatory motor system in speech perception. *Aphasiology*, 26, 1103–1118.
- Parthasarathy, V., Stone, M., & Prince, J. L. (2005). Spatiotemporal visualization of the tongue surface using ultrasound and kriging (SURFACES). *Clinical linguistics & phonetics*, 19, 529–544.
- Pulvermüller, F., & Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, 11, 351–360.
- Pulvermüller, F., Huss, M., Kherif, F., del Prado Martin, F. M., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences*, 103, 7865–7870.
- Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience*, 12, 718–724.
- Rizzolatti, G., & Arbib, M. (1998). Language within our grasp. *Trends in Neurosciences*, 21, 188–194.

- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169–192.
- Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, 2, 661–670.
- Rogalsky, C., Love, T., Driscoll, D., Anderson, S. W., & Hickok, G. (2011). Are mirror neurons the basis of speech perception? Evidence from five cases with damage to the purported human mirror system. *Neurocase*, 17, 178–187.
- Sams, M., Möttönen, R., & Sihvonen, T. (2005). Seeing and hearing others and oneself talk. *Cognitive Brain Research*, 23, 429–435.
- Sato, M., Tremblay, P., & Gracco, V. L. (2009). A mediating role of the premotor cortex in phoneme segmentation. *Brain and Language*, 111, 1–7.
- Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M. S., ... Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences*, 105, 18035–18040.
- Scheerer, E. (1984). Motor theories of cognitive structure: A historical review. In W. Prinz & A. F. Sanders (Eds.), *Cognition and motor processes* (pp. 77–98). Berlin: Springer-Verlag.
- Scott, S. K., & Johnsrude, I. S. (2003). The neuroanatomical and functional organization of speech perception. *Trends in Neurosciences*, 26, 100–107.
- Shahin, A. J., Bishop, C. W., & Miller, L. M. (2009). Neural mechanisms for illusory filling-in of degraded speech. *NeuroImage*, 44, 1133–1143.
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 174–215.
- Stassenko, A., Garcea, F. E., & Mahon, B. Z. (2013). What happens to the motor theory of perception when the motor system is damaged?. *Language and Cognition*, 5, 225–238.
- Stephens, J. D. W., & Holt, L. L. (2011). A standard set of American-English voiced stop-consonant stimuli from morphed natural speech. *Speech Communication*, 53, 877–888.
- Toni, I., de Lange, F. P., Noordzij, M. L., & Hagoort, P. (2008). Language beyond action. *Journal of Physiology — Paris*, 102, 71–79.
- Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, 41, 989–994.
- Watkins, K. E., & Paus, T. (2004). Modulation of motor excitability during speech perception: The role of Broca's area. *Journal of Cognitive Neuroscience*, 16, 978–987.
- Wertz, R. T., LaPointe, L. L., & Rosenbek, J. C. (1984). *Apraxia of speech: The disorders and its management*. New York: Grune and Stratton.
- Wilson, S. M., & Iacoboni, M. (2006). Neural responses to non-native phonemes varying in producibility: evidence for the sensorimotor nature of speech perception. *NeuroImage*, 33, 316–325.
- Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, 7, 701–702.