Why the Left Hemisphere Is Dominant for Speech Production: Connecting the Dots

Harvey M. Sussman

Evidence from seemingly disparate areas of speech/language research is reviewed to form a unified theoretical account for why the left hemisphere is specialized for speech production. Research findings from studies investigating hemispheric lateralization of infant babbling, the primacy of the syllable in phonological structure, rhyming performance in split-brain patients, rhyming ability and phonetic categorization in children diagnosed with developmental apraxia of speech, rules governing exchange errors in spoonerisms, organizational principles of neocortical control of learned motor behaviors, and multi-electrode recordings of human neuronal responses to speech sounds are described and common threads highlighted. It is suggested that the emergence, in developmental neurogenesis, of a hard-wired, syllabically-organized, neural substrate representing the phonemic sound elements of one's language, particularly the vocalic nucleus, is the crucial factor underlying the left hemisphere's dominance for speech production.

1.   Introduction

When the right hemisphere of a bisected brain is presented with a spoken word the input signal is semantically processed; however, when instructed to say the word it just heard, the split-brain subject is silent (Gazzaniga 1970, 1983). When sodium amytal is selectively administered to right-handed patients prior to brain surgery muteness is experienced (in approximately 96% of cases) when the left hemisphere is anesthetized, while right hemisphere anesthesia affects only 4% of the population (Rasmussen *et al.* 1977). Despite this robust hemispheric asymmetry for speech production in the human brain, no specific, micro-level neural account has been posited to account for this behavioral dominance. Two macrolevel accounts of left hemispheric asymmetry for speech output have been put forth. One classic view holds that the left hemisphere selectively inhibits the right hemisphere from participating in language output (e.g., Kinsbourne 1974, Kinsbourne *et al.* 1978, Chiarello *et al.* 1996, Liégeois *et al.* 2004). An inhibitory-based explanation for left hemisphere dominance suggests that to avert 'equipotentiality', the left hemisphere must take on an active preventative role.

A second hypothesis, formulated from an evolutionary perspective, claims a selective advantage for having separate hemispheres for mediating the well known antagonistic modes of neural processing—analytical symbol translation in the left hemisphere versus spatial, gestalt-like synthesis in the right hemisphere (Levy 1969). Since neural substrates underlying these opposing processing modes cannot easily co-exist (i.e. seeing both 'trees' and 'forests' in the same hemisphere), selective evolutionary pressures housed them in separate hemispheres to minimize processing conflicts and maximize what each hemisphere is best structured to do.

Interestingly, there is no lack of specificity in accounting for hemispheric asymmetries underlying speech processing/perception, despite the fact that speech processing involves far more bilateral interactions than speech production (Hickok *et al.* 2000, Hickok *et al.* 2007, Peelle 2012). One long held view proposes that the left hemisphere is specialized to process the rapid temporal changes (e.g., F2 transitions) characterizing speech (e.g., Tallal *et al.* 1973, 1974, Tallal *et al.* 1981, Zatorre *et al.* 2001, Zatorre *et al.* 2002). An alternative, but somewhat related, view claims that prelexical speech perception is actually processed bilaterally, but different tuning properties of temporal integration windows (40 Hz gamma and 4–10 Hz theta-range) underlie hemispheric-specific differences, with the left hemisphere being specialized to process acoustic signals spanning short temporal windows (appropriate for phonemes) and the right hemisphere specialized for longer temporal windows mediating prosodic cues such as intonation (Poeppel 2003).

We often hear the expression, "We simply didn't connect the dots". To avoid such an oversight dots will be connected from the following areas of language study: (1) infant babbling, (2) the phonological primacy of the syllable, (3) split-brain studies, (4) developmental apraxia of speech, (5) speech errors, (6) a perspective on neocortical operations as learned auto-associative memories, and (7) electrophysiological recordings from human left posterior superior temporal gyrus (pSTG) during presentation of (a) a stop place continuum and (b) an extensive phonetic inventory contained within 500 sentences spoken by 400 speakers. It will be shown that the collective findings from the above studies strongly suggest that the left hemisphere forms, and thus has exclusive access to, neural substrates tasked to represent/map phonemic sound segments that are the prerequisites to both initiate and drive speech motor output.

2.   Dot #1: Lateralization of Infant Babbling

Infant babbling provides insights into the prelinguistic beginnings of sound generation in a developing infant. Before canonical babbling (CVs) starts, infants progress from squeals, squeaks, and various forms of yells to produce cooing noises. Importantly, infant coos can be considered precursors to vowel-like sounds, the first speech-like sounds (Locke 1989, Oller 2000). More pertinent to the argument to be made is the intriguing possibility that early infant babbling might also be asymmetrically controlled and monitored by the left hemisphere. Graves *et al.* (1990) observed that when normal adult subjects are speaking, there

is a measurable difference in the mouth opening extents in the two sides of the mouth, with the right side opening being greater during generation and recall of word lists.

Adapting the metric of 'right mouth asymmetry', Holowka *et al.* (2002) videotaped 10 babies between the ages of five and 12 months, equally distributed across an English and French home environment. Independent scorers, unaware of the purpose of the study, analyzed randomly selected portions of the videos (N = 150 segments) during three different types of mouth activity: babbles (CV repetitions), non-babbles (vocalizations without a consonant-vowel structure), and smiles. A laterality index was generated to assess the three oral activities. All 10 babies showed a right mouth asymmetry when babbling (+0.88), equal mouth openings for non-babbling (−0.08), and left mouth asymmetry for smiles (−0.82). The greater right-than-left asymmetry in mouth openings was interpreted as reflecting greater involvement of the left hemisphere during babbling utterances. The authors state: "We thus conclude that babbling represents the onset of the productive language capacity in humans, rather than an exclusively oral-motor development" (Holowka *et al.* 2002: 1515).

So the first 'dot' is pre-linguistic sound generation—initially vocalic-like and then, from approximately 7 to 18 months, CV-like sequences, envisioned as being initially and preferentially encoded in an emerging neural substrate in the left hemisphere. These earliest speech-like sounds can be conceptualized as the instantiation of the 'speech sound map' (possibly) forming in left ventral premotor cortex (BA 6, 44) as described in the DIVA computational model (Guenther *et al.* 2006, 2012). If these babbling results are replicated in future studies, then one might say the neural precursors of the eventual phonological primitives of one's language have asymmetrically taken root in the left hemisphere.

To ground this neurogenesis assumption to a neural model of language function (e.g., Hickok *et al.*, 2004), the initial 'dot' is envisioned as the earliest neural 'seeds' of dorsal stream projections (left dominant 'sensori-motor interface' in parietal-temporal Spt area) to the left frontal *'articulatory network'*. Admittedly, this hypothesis does not account for why the hypothesized left hemisphere laterality for babbling exists in the first place. The 'usual suspect,' genetic predisposition, might have to suffice at the moment.

3.   Dot #2: Phonological Primacy of the Syllable

The second 'dot' serves to connect the emergence of early infant vocalizations, organized around duplicated and variegated babbling (a CV 'syllable' structure), to well known first principles of phonological language structure. The syllable, while long resisting an unambiguous definition (see Bell & Hooper 1978), nevertheless has properties strongly supporting its primacy in the phonological structure of the world's languages. The following attributes of syllables provide support for this claim: (i) the syllable-bound nature of prosodic events such as stress, rhythm, juncture; (ii) reduplication and deletion processes in a child's phonological development (Fudge 1969, Moskowitz 1970, 1971, Hooper 1972, 1976); (iii) native language syllable constraints that play a key role in pronunciation errors

in second language acquisition (Broselow 1983, 1984); and lastly, (iv) the finding that the most prevalent and permutable unit in sub-lexical transfers during language play is unequivocally the syllable (Sherzer 1976). The language play data also corroborates the finding that young, pre-reading, children possess an intrinsic ability to recognize and respond to the syllable structure of words when asked to tap their hand in cadence to the audio sounds of spoken words (Liberman 1973). Additional examples of the primacy of the syllable can be observed in apraxic and dysarthric speakers whose output patterns are described as staccato, sing-song concatenations of dissociated syllable-by-syllable strings (Kent *et al.* 1982, 1979).

To summarize up to this point, the first two dots can be taken to support the contention that the earliest speech sound networks in the neurogenesis of language structure, and hence, spoken output, in frontal and temporal areas of the left hemisphere, are organized around segmental-like entities, initially grouped in a prototype sequential structure resembling CV syllable forms. Leaving left handedness issues aside, it is postulated that no such neural substrates, tasked to encode a language's sound segments-to-speech motor neural networks, exist in the right hemisphere of right-handed speakers.

4.   Dot #3: The Right Hemisphere of Split-Brain Subjects Cannot Rhyme

The development of the split-brain paradigm by Sperry and colleagues provided, for the first time, an elegant experimental method to direct sensory information to isolated hemispheres of the human brain and independently assess their relative processing capabilities for various types of language-related input signals (Sperry 1961). A visual tachistoscopic projection system (T-scope) was used to present various words/symbols onto visual half-fields for very brief time periods (usually 150 msec) to avoid a stimulus confound due to saccadic eye movements. A stimulus input to the right visual field (RVF) projected the image exclusively to the left visual cortex, and a left visual field (LVF) stimulus was exclusively projected to the right visual cortex.

In split-brain subjects, due to their complete cerebral commissurotomy, there is no inter-hemispheric transfer of information, and hence each hemisphere "has its independent mental sphere or cognitive system-that is, its own independent perceptual, learning, memory, and other mental processes" (Sperry 1961: 1). In preliminary studies, it became obvious that only the left hemisphere was capable of speaking, and the right hemisphere could only manually respond by directing the individual's left hand to write or select seen objects from behind the T-scope screen.

One of the most creative adaptations of this paradigm was developed by Eran Zaidel in a series of elegant studies exploring the information processing capacity of the right hemisphere (Zaidel 1978). Zaidal realized that, to fully analyze the capabilities of the right hemisphere across a varied set of language tasks, it would require a longer stimulus exposure interval than 150 msec. To enable longer scrutiny intervals Zaidel devised a projection system that was yoked to the saccadic movements of the subject's eye. Each split-brain subject was fitted with a

customized contact lens. Stimuli (e.g., groups of four words, or four pictures of common objects) were projected to separate visual half fields, and as the subject's eyes moved for each saccade, the projection system compensated by moving the exact distance to keep the image stabilized in the same visual half-field. This allowed subjects to take as long as needed to visually process what was being asked of them, e.g., "point to the two pictures of objects that rhyme" when shown four pictures, two of which were a baseball bat and a man's hat.

Zaidel ran a series of inter-related experiments that explored information transfer from one modality form to another: sound-to-meaning (via a picture), sound-to-spelling (orthography), spelling-to-picture, picture-to-sound, spelling-to-sound, meaning-to-sound, and orthography-to-sound. While the left hemisphere of the split-brain subjects had no trouble successfully performing all the tasks, the right hemisphere revealed a striking inability to evoke the *sound image* of a seen object or letter string (that they knew the meaning of), and, of most importance to the argument being put forth here, *a striking inability to assess rhyme*. Whenever the task required a transfer from either semantics (pictures of objects), or letter strings (e.g., B-I-R-D, C-A-T, H-O-U-S-E) for judging a rhyme (e.g., "Which word rhymes with *hat*?"), the right hemisphere was incapable of performing the meta-linguistic conversion of a seen picture or letter string into an internalized sound equivalent.

Another test to assess rhyming ability presented a slide having four pictures, two of which, when pronounced, rhymed, and two did not. The subject was told to point to the two pictures that sound the same, but have different meanings (e.g. rose/toes, mail/male). They would use their left hand to point to their answers. Presented by themselves for comprehension (e.g. hear word '*mail*' or see letters *M A I L*, and asked to point to the correct picture), the right hemisphere knew what the stimulus word meant, but when asked to judge a rhyme (even with similar orthography as in 'nail'), the right hemisphere was clueless. If the orthographic pairings differed in spelling (e.g., *pea/key)*, or presented idiosyncrasies of English pronunciation (e.g., *lint/pint*), performance was considerably worse.

The take-away message from the third 'dot' is the following: To be able to generate a rhyme or judge whether a word pair contains a rhyme, the neural processing substrate must be able to internally generate the sound equivalent of the orthographic word or picture of the object—primarily the vowel/coda of a lexical string. It's very quiet inside your brain, but the left hemisphere is uniquely adept at internally generating sound equivalencies of input letter strings or seen objects. These encoded segmental-based network representations have a dual function: They (i) inherently possess the sound equivalencies of the phonemic units making up the word and (ii) serve as the neural source for generating speech production, or said another way, the phonological intent that drives and initiates the motor programming to elicit a speech output signal.

These critical properties—internal generation of sound equivalencies of phonemes and an ability to go from 'intent-to-motor activation'—are hypothesized to be present, in the overwhelming majority of right-handed adults, only in the left hemisphere of the brain. The inescapable truth is that if rhyming ability can only be performed by the left hemisphere, then the neural equivalent of

vocalic nuclei of syllable codas is only present in the sound processing regions of the left hemisphere.

5.   Dot #4: Rhyming and Phonetic Category Deficiencies in Children with Developmental Apraxia of Speech

What happens if and when such (hypothesized) lateralized neural sound substrates fail to develop in neurogenesis? The answer might lie in the childhood speech deficit known as Development Apraxia of Speech (DAS). DAS is customarily defined as a neurologically based disorder in the ability to carry out coordinative movements of the speech articulators in the absence of impaired neuromuscular functioning (Shriberg *et al.* 1997). The behavioral symptomatology of DAS presents with a wide array of speech/language deficits encompassing input, organizational, and output processing. However, output processing deficits have had a disproportionate influence in diagnosis and treatment of this childhood language disorder. The primary production-based deficits include: a restricted phonemic repertoire, predominance of omission errors, *frequent vowel errors*, inconsistency of errors, restricted use of word shapes (they produce mostly CVs), and better receptive than expressive test scores (Marquardt *et al.* 1998).

Studies in our lab focused on the representational and perceptual abilities of children with DAS—specifically, their ability to generate and assess rhymes (Marion *et al.* 1993) and categorical perception of speech (Sussman *et al.* 2000, 2002). The theoretical impetus for these studies was the hypothesis that the underlying etiological cause of DAS was a neural dysmorphology in left hemisphere areas mediating the phoneme-sized phonological representations necessary to both form sound equivalencies and to initiate and control on-line articulatory programming of those sound strings. A child with DAS was perhaps operating with an impoverished phonological neural representation network that severely precluded both selection and access to the neural correlates of the phonological forms guiding speech motor performance. In effect, a DAS child trying to speak would be analogous to an adult playing scrabble with hard to read letter tiles because they were blurry or malformed.

A strong test of the hypothesis that DAS is based on a left hemisphere developmental dysmorphology in the neurogenesis of brain tissue that mediates phonological representations is to assess the rhyming abilities of DAS children (matched to typically developing controls). The essence of rhyming ability is the internal generation of vowel sounds, holding them in short term working memory, and meta-linguistically judging (dis)similarities across word pairs.

Marion *et al.* (1993) devised three rhyming tasks. (i) Rhyme production: Following presentation of a target word (N = 12), the child had to produce as many rhyming words as possible in 30 seconds. (ii) Assessing rhyming word pairs: Using a target word, which of two words rhymes best with the target word? (iii) Rhyme perception: For each target word, 10 words were presented and the child indicated which words rhymed with the target item. The results were very revealing—the DAS children (N = 4) could not generate rhymes, or even recognize rhyming words, while the four control children exhibited signify-

cantly higher scores on every task. For example, in the rhyme production study the DAS children produced a score of <2.0 correctly rhyming words compared to over 30 for the control children. In the rhyming pairs test, which was much easier, the DAS children scored between 40–50% correct matches, while the control children scored close to 100%. On the rhyme perception test, the DAS children produced an over-abundance of false rhymes while generally failing to recognize correct rhymes.

The striking inability to form and recognize rhymes in DAS closely resembles the right hemisphere's rhyming deficiencies documented in split-brain subjects (Rayman *et al.* 1991). The main difference is that with split-brain subjects, their right hemisphere is innately incapable of rhyming, whereas in DAS children, it is hypothesized that their phonologically impoverished left hemisphere substrates were attempting to perform the mental operations required for rhyming, but falling short. Once again, to be able to rhyme, brain regions must possess the internalized neuronal equivalent of the sound evoked by the vowel-dominant coda cluster of a word. This seems to be the exclusive provenance of the speaking left hemisphere. If, as hypothesized, DAS is caused by a dysmorphology of left hemisphere neural substrates that normally process sound elements, that in a normally developing brain, map/represent the finite set of phonetic segments comprising the sound inventory of a language, then normal left hemisphere dominance in speaking may well be attributable to the exclusive presence of such substrates as the requisite 'start' button initiating and controlling the serial ordering of speech. DAS children might very well lack this 'start' button initiation in going from phonological representation to phonetic/articulatory output.

Another way to probe the integrity of neural-based phonological categories is to perform labeling studies as part of a categorical perception procedure. Using an identification task with a 14-item stimulus continuum ([ba-da-ga]), Sussman *et al.* (2002) showed poor categorization skills in all five DAS children tested relative to five typically developing controls. The DAS group showed equivocation in labeling within-category allophonic stimuli and an absence of quantal shifts in identification percentage scores at expected phonetic boundaries. The perceptual sensitivity of the two groups to F2 changes in adjacent CV stimuli was also assessed by using a cumulative d' statistic. The less steep slope of the d' function in the DAS group revealed a considerably diminished perceptual sensitivity to systematic changes in the acoustic stimuli. Simply put, the DAS children exhibited a very fragile control of categorical entities and their internalized phonologically-based structure.

There are two basic requirements needed to establish well-formed contrastive phonetic categories: (i) sensitivity at phonetic boundaries, combined with (ii) the ability to ignore or generalize across (within category) allophonic variations. The second element is not often discussed, but there needs to be a basic neuronal mechanism that maintains categorical consistency in the face of non-phonemic signal variation. Tolerating and generalizing across subtle, within-category, allophonic variations is crucial in establishing well-formed categorical representations. A recent MMN study (Miglietta *et al.* 2013) successfully partitioned allophonic-based ERPs from phonemic-based ERPs across vowel pairings

in a dialect of Italian. Thus, neural computations exist for within-category phonetic distribution patterns. Non-contrastive auditory differences must therefore require a learned inhibitory-based computation to allow for faster unfettered access to higher perceptual phonemic representations.

The collective findings from these DAS studies adds another crucial dot—if the neural networks that encode basic phonological units, the building blocks of language, fail to develop in a normal fashion, the resulting outcome is what we see in the highly unintelligible and very limited speech/language capabilities of children diagnosed with DAS.

6.   Dot #5: Speech Errors and the Slot-Segment Hypothesis

One of the many unknowns about speech production is the answer to the question: "What phonological entity is most closely related to the neuro-motor commands underlying speech production?" Possible candidates for the 'phonological primitive' are the phoneme, the extrinsic allophone, the syllable, the word, the phase, etc. The existence of linguistic abstractions, unfortunately, cannot be empirically validated by brain imaging techniques. The phoneme, however, as one possible candidate for this elusive unit, possesses a high degree of psychological reality based on its overwhelming prevalence in speech error corpora. For example, considering only exchange errors, e.g., *'guinea pig cage' — 'guinea kig page'*, Shattuck-Hufnagel (1983) reported that 138 of 210 errors (66%) occurred as phonemic segments in the 1981 MIT corpus. No other sound structure unit was even close. What is considerably more important, however, than proclaiming what linguistic entity best corresponds to the neural correlate of phonological structure is what can be learned from studying speech errors:

> The interest is rather in how particular errors shed light on the underlying units of linguistic performance, and the production of speech. What is apparent, in the analyses and conclusions of all linguists and psychologists dealing with errors in speech, is that, despite the semi-continuous nature of the speech signal, there are discrete units at some level of performance which can be substituted, omitted, transposed, or added. (Fromkin 1971: 29)

Behavioral data from sound exchanges provide a window into the premotor planning stage of an utterance before actual production of that utterance. The displaced phoneme-sized exchanges characterizing speech errors have contributed to several theoretical insights into the neural events taking place prior to overt motor programming. One such insight was the suggestion by Shattuck-Hufnagel (1975, 1979) that there are two separate but interactive neural network structures underlying the representation of phonologically organized sound units. She postulated a neural framework for syllable structure ('serially ordered slots'), and an independent, but synaptically inter-connected representational network for the phonetic segments. Such a two-tiered interactive neural substrate helped to conceptualize the various rules that Fromkin (1971) earlier formulated governing the nature of segmental-based sound exchanges. Rule #1 was that consonants always exchange with consonants and vowels only exchange with

vowels. Rule #2 stated that sound exchanges always occur within the same syllable position. So in the error *'the nipper is zarrow'* *(*for the 'zipper is narrow*')* the migrating 'n' in *'narrow'* erroneously fills the C1 slot of word 1, instead of the intended occupant /z/; the displaced 'z' doesn't disappear in a brain 'cloud', but fills in the now vacated C1 slot in word 2, left empty by the transposed 'n'. Thus, the empty slot awaits a new segmental occupant, acting as a place-holder for the displaced phoneme. The sound-based units are very real in a neural sense. Synaptic connections between re-arranged segment-based networks and canonical syllable-shape networks still manage to produce fluent output containing the speech error.

Rule #1 is inviolate in speech error analyses and can speak to the primacy of the vowel in a syllable (i.e., there is no syllable without it). Vocalic-like sounds in early infant vocalizations (dot #1) can be viewed as the earliest input signal in developmental neurogenesis to fill this integral slot of the emerging syllable-based neural scaffolding. In essence the vowel can be conceptualized as being 'prepackaged' and anchored into the nucleus slot of any future syllable form (CV, CVC, CCV, CCVC, etc.) that develops over time with increasing phonological complexity (Sussman 1984). Each language forms a neural slot framework structure driven by its own syllable shape(s), for example CV in Japanese and Hawaiian, (CCC)V(CCCC) in English.

Dot #5 (speech errors) serves to consolidate several previous dots. If the left hemisphere exclusively houses the neural substrates forming syllable frames, with their synaptic network linkages to auditory-encoded segmental entities of a given language, with primacy of the vocalic nucleus, then it is no mystery that speech output programming is under the exclusive control of the left hemisphere. A hemisphere devoid of a segmental-sound-based encoding infrastructure does not possess the 'neural-sparkplug' that, in effect, serves as the 'intent' to initiate and control the serial ordering of sound units underlying speech motor programming.

7.   Dot # 6: The Neocortex—Computational or Serially Ordered Memory System?

In his book *On Intelligence,* Hawkins (2004) puts forth several insights regarding the operational properties of the neocortex. A basic postulate is that "the neocortex uses stored memories to […] produce behaviors" (p. 69). So rather than computing unique solutions to perform motor behaviors, the brain possesses stored memories, learned across development through repeated experiences. Moreover, these motor memories sequentially operate in an auto-associative manner. We activate memories, whether motor, visual, or sound, the way you learned them, and each temporally ordered memory elicits the next. Common everyday examples show the validity of this simple, but largely ignored feature of neural operations within our 'connectome'—e.g., one cannot (easily) sing a song, recite a well known passage, or the alphabet, backwards; hearing the start of a familiar tune sequentially elicits the next portions, in the temporal order in which it was learned. Spoken language, like all serially ordered motor skills,

unfolds in sequential fashion, each set of articulatory movements, organized around sequential syllabic frames, automatically triggers the next. If, as strongly suggested by the preceding 'dots', the left hemisphere's auditory/speech motor areas are the exclusive repository of the neural networks instantiating production of segmental-based units, with their inherent sound and articulatory motor equivalencies, organized around syllable-by-syllable concatenations, then speech output should only be possible in the left hemisphere. The connectome of the right hemisphere is generally regarded as a synthesis specialist, processing holistically (faces, not noses), not analytically. A gestalt-based neural structure is not conducive to motorically producing a serially-ordered, symbol-based, syllabi-cally organized, set of learned articulatory behaviors inherently linked to sound equivalents.

An interesting addendum to this hypothesized scenario is the added concept of a hierarchically-organized invariance in the way the neocortex is organized for processing input signals and also executing motor behavior (Hawkins 2004). Our brains, unlike artificial intelligence systems, can recognize faces from any angle or position; we can recognize familiar tunes regardless of the instrument playing them—e.g., the *Stars Spangled Banner* is easily recognized if played by a harmonica, tuba, piano, or whistled. A computer can only store information the way it was presented, there is no tolerance for variability. Speech, whether in input or output mode, is highly adaptable.

The widely used bite block paradigm (e.g. Kelso & Tuller 1983) illustrates this concept: When acrylic bite blocks are placed between a speaker's back molars, thus precluding jaw movements in articulation, a speaker can immediately, on the first trial, compensate for the lack of jaw movement by using new/novel tongue configurations that create equivalent vocal tract resonance properties to arrive at the auditory target of the speech sounds produced. Similarly, a pipe smoker can produce intelligible speech whilst biting down on the pipe stem. The invariance that characterizes both speech perception (e.g., different F2 transitions in /dV/ utterances can all be heard as the same /d/), and speech production (e.g., myriad of ways the same sound can be produced by varying articulatory motor contributions) serves to point out that the 'sound plan' neural infrastructure, as envisioned in this account, is linked to highly flexible and synergistic speech motor net-works.

8.   Dot #7: Recording from Intracranial Electrode Arrays in Human Left pSTG

A major premise of this paper is that speech sounds exist as stored represen-tations in auditory neural substrates of the left hemisphere. For scientists outside the field of experimental phonetics this might sound a bit silly: "How could speech sounds not be represented in the human brain?" However, the long-standing theoretical division in the field of experimental phonetics between auditory vs. gestural views of underlying neural correlates of speech units has prevented a unified theoretical position to emerge, even after six decades of experimental research (e.g., Studdert-Kennedy 1998, 2005, Studdert-Kennedy *et al.* 2003).

Recent game changing studies by Chang and his colleagues at UCSF have served to strongly substantiate an auditory-based position. Chang *et al.* (2010) synthesized 14 uniquely different stop consonant-vowel syllables by systematically altering the onset frequencies of the F2 transition to create a [ba-da-ga] continuum as used in categorical perception studies. They were presented in random order to four subjects, post craniotomy and prior to surgery for epilepsy. Evoked potentials were obtained for each stimulus presentation via a customized 64-electrode microarray placed on left pSTG. The specific question addressed was whether pSTG neural activity patterns would correspond to the precise spectro-temporal changes in the external acoustic signal (i.e., veridical representation, and hence 14 different ERPs), or to a higher order linguistic extraction of phonetic categories (only three unique ERP patterns)? The analysis was based on the degree to which a multivariate pattern classifier was able to distinguish single-trial response patterns of the evoked cortical potentials. Response amplitude and across-stimuli dissimilarities peaked at 110ms after stimulus onset, and the topography of the most discriminative cortical sites clearly revealed only three discrete activation patterns, not 14. The local and transient response properties revealed distributed, but non-overlapping, spatial representations for stop place category-based patterns. Thus, it is no longer necessary to only *postulate* the existence of auditory representations of the sounds of human language— they indeed have neurophysiological reality.

The abstract from Chang *et al.* (2010) succinctly captures the essence of their findings and the implications for understanding the neural underpinnings of speech and language phonological structure:

> Speech perception requires the rapid and effortless extraction of meaningful phonetic information from a highly variable acoustic signal. A powerful example of this phenomenon is categorical perception, in which a continuum of acoustically varying sounds is transformed into perceptually distinct phoneme categories. We found that the neural representation of speech sounds is categorically organized in the human posterior superior temporal gyrus. Using intracranial high-density cortical surface arrays, we found that listening to synthesized speech stimuli varying in small and acoustically equal steps evoked distinct and invariant cortical population response patterns that were organized by their sensitivities to critical acoustic features. Phonetic category boundaries were similar between neurometric and psychometric functions. Although speech sound responses were distributed, spatially discrete cortical loci were found to underlie specific phonetic discrimination. Our results provide direct evidence for acoustic-to-higher order phonetic level encoding of speech sounds in human language receptive cortex.   (Chang *et al.* 2010: 1428)

The electrophysiological recordings of Chang *et al.* (2010), limited to only three stop consonants (/bdg/) and one vowel (/a/), have been expanded more recently to include the entire English phonetic inventory (Mesgarani *et al.* 2014). Using the same high-density multi-electrode arrays placed over the left STG in six subjects undergoing craniotomies, they reported high selectivity at numerous single electrode sites responding to the unique spectrotemporal acoustic properties of speech sounds.

Phoneme groups (stops, fricatives, nasals, semi-vowels, vowels) were organized into highly differentiated clusters based on shared phonetic features, primarily distinguished by manner of articulation, and secondarily by place of articulation distinctions. A needed control to fully comprehend the significance of these findings is to perform the same analysis on patients undergoing a right craniotomy and placing the recording electrode array on right pSTG. The absence of fine tuning for spectrotemporal acoustic cues defining phonetic structure groupings in right hemisphere superior temporal cortex would further support the views being hypothesized in this paper.

9.   Summary and Conclusions

Several inter-related areas of research and theory were described: (1) lateralization of infant babbling; (2) phonological primacy of the syllable; (3) the inability of the right hemisphere of split-brain subjects to generate/assess rhymes; (4) the inability of children diagnosed with a left hemisphere-based language disorder (DAS) to generate/assess rhymes and behaviorally evidence well formed speech sound categories; (5) analyses of speech exchange errors supporting an underlying, tiered, syllable slot-segment neural structure; (6) a view of cortical organization and processing as memory networks characterized by being experientially learned, activated in serial temporal order, with auto-associative triggering, and hierarchically organized to achieve invariant representations; and (7) recent evidence from intra-cranial electrode arrays on human left pSTG showing distributed neural foci invariantly encoding phonetically structured categories.

A connecting theoretical thread was sewn across these seven research areas suggesting that the asymmetrical dominance of the left hemisphere to control speech output might be due to the exclusive existence of specialized neural substrates encoding the phonological elements of language, organized in canonical syllable-sized representational networks. This left hemisphere network initially develops during early infant vocalizations, from coos to canonical CV babbling, to early first words. Of most importance is that this emergent neural substrate can serve as the exclusive neural 'start button' to bring about articulatory motor programming. It is maintained that the right hemisphere does not possess such sound unit-based neural networks, as primarily holistic processing has no use for serial processing of symbolic units that are integrally connected to speech motor pathways. This account focused only on underlying structural properties of left hemisphere neural tissue to account for asymmetry in speech motor output. What remains to be explained is why and how this hemispheric specialization began.

References

Bell, Alan & Joan B. Hooper. 1978. *Syllables and Segments*. Amsterdam: North Holland.

Broselow, Ellen. 1983. Non-obvious transfer: On predicting epenthesis errors. In Susan Gass & Larry Selinker (eds.), *Language Transfer in Language Learning*, 269–280. Rowley, MA: Newbury House.

Broselow, Ellen. 1984. An investigation of transfer in second language phonology. *International Review of Applied Linguistics* 22, 253–269.

Chang, Edward F., Jochem W. Rieger, Keith Johnson, Mitchell Berger & Robert T. Knight. 2010. Categorical speech representation in human temporal gyrus. *Nature Neuroscience* 13, 1428–1432.

Chiarello, Christine & Lisa Maxfield. 1996. Varieties of interhemispheric inhibition, or how to keep a good hemisphere down. *Brain and Cognition* 30, 81–108.

Fromkin, Victoria, A. 1971. The non-anomalous nature of anomalous utterances. *Language* 47, 27–52.

Fudge, Erik. 1969. Syllables. *Journal of Linguistics* 8, 113–133.

Gazzaniga, Michael. 1970. *The Bisected Brain*. New York: Springer.

Gazzaniga, Michael. 1983. Right hemisphere language following brain bisection: A 20-year perspective. *American Psychologist* 38, 525–537.

Graves, Roger & Theodor Landis. 1990. Asymmetry in mouth opening during different speech tasks, *International Journal of Psychology* 25, 179–189.

Guenther, Frank, Satrajit S. Ghosh & Jason A. Tourville. 2006. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language* 96, 280–301.

Guenther, Frank & Tony Vladusich. 2012. A neural theory of speech acquisition and production. *Journal of Neurolinguistics* 25, 408–422.

Hawkins, Jeff. 2004. *On Intelligence*. New York: Henry Holt & Co.

Holowka, Siobhan & Laura Ann Petitto. 2002. Left hemisphere cerebral specialization for babies while babbling. *Science* 297, 1515.

Hooper, Joan B. 1972. The syllable in phonological theory. *Language* 28, 525–540.

Hooper, Joan B. 1976. *An Introduction to Natural Generative Phonology*. New York: Academic Press.

Hickok, Gregory & David Poeppel. 2000. Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences* 4, 131–138.

Hickok, Gregory & David Poeppel. 2004. Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition* 92, 67–99.

Hickok, Gregory & David Poeppel. 2007. Opinion — The cortical organization of speech processing. *Nature Reviews Neuroscience* 8, 393–402.

Kelso, Scott & Betty Tuller. 1983. Compensatory articulation under conditions of reduced afferent information: A dynamic formulation. *Journal of Speech and Hearing Research* 26, 217–224.

Kent, Raymond D. & John Rosenbek. 1982. Prosodic disturbance and neurologic lesion. *Brain and Language* 15, 259–291.

Kent, Raymond D., Ronald Netsell & James H. Abbs. 1979. Acoustic characteris-

tics of dysarthria associated with cerebellar disease. *Journal of Speech and Hearing Research* 22, 627–648.

Kinsbourne, Marcel. 1974. Mechanisms of hemispheric interaction in man. In Marcel Kinsbourne & Lynn Smith (eds.), *Hemispheric Disconnection and Cerebral Function,* 260–285. Springfield, IL: Charles C. Thomas.

Kinsbourne, Marcel & Robert E. Hicks. 1978. Human handedness. In Marcel Kinsbourne (ed.), *The Asymmetrical Function of the Brain*, 523–552. New York: Cambridge University Press.

Levy, Jere. 1969. Possible basis for the evolution of lateral specialization of the human brain. *Nature* 224, 614–615.

Liberman, Isabelle Y. 1973. Segmentation of the spoken word and reading acquisition. *Bulletin of The Orton Society* XXIII: Reprint No. 54.

Liégeois, Frederique, Alan Connelly, Helen Cross, S. Gregory Boyd, David Gadian, Faraneh Vargha-Khadem & Torsten Baldeweg. 2004. Language reorganization in children with early-onset lesions of the left hemisphere: An fMRI study. *Brain*, 1229–1236.

Locke, John. 1989. Babbling and early speech: Continuity and individual differences. *First Language* 9, 191–206.

Marion, Michelle, Harvey M. Sussman & Thomas Marquardt. 1993. The perception and production of rhyme in normal and developmentally apraxic children. *Journal of Communication Disorders* 26, 129–160.

Marquardt, Thomas, Harvey M. Sussman & Barbara Davis. 1998. Developmental apraxia of speech: Advances in theory and practice. In Deanie Vogel & Michael Cannito (eds.), *Treating Disordered Speech Motor Control*. Austin, TX: Pro-Ed.

Mesgarani, Nima, Connie Cheung, Keith Johnson & Edward Chang. 2014. Phonetic feature encoding in human superior temporal gyrus. *Science* 343, 1006–1010.

Miglietta, Sandra, Mirko Grimaldi & Andrea Calabrese. 2013. Conditioned allophony in speech perception: An ERP study. *Brain & Language* 126, 285–290.

Moskowitz, Arlene. 1970. The acquisition of phonology. Berkeley, CA: University of California Language-Behavior Research Laboratory Working Paper 34.

Moskowitz, Arlene. 1970. The acquisition of phonology. Berkeley, CA: University of California dissertation.

Oller, Kimbrough. 2000. *The Emergence of the Speech Capacity*. Mahwah, NJ: Lawrence Erlbaum Associates.

Peelle, Jonathan. 2012. The hemispheric lateralization of speech processing depends on what 'speech' is: A hierarchical perspective. *Frontiers in Human Neuroscience* 6.

Poeppel, David. 2003. The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time'. *Speech Communication* 41, 245–255.

Rasmussen, Theodore & Brenda Milner. 1977. The role of early left-brain injury in determining lateralization of cerebral speech functions. *Annals of the New York Academy of Science* 299, 355–369.

Rayman, Jan & Eran Zaidel. 1991. Rhyming and the right hemisphere. *Brain and*

Language 40, 89–105.

Sperry, Roger W. 1961. Cerebral organization and behavior. *Science* 133, 1749–1757.

Seung, Sebastian. 2012. *Connectome: How the Brain's Wiring Makes Us Who We Are.* New York: Houghton Mifflin Harcourt Publishing Co.

Shattuck-Hufnagel, Stephanie. 1975. Speech errors and sentence production. Cambridge, MA: MIT dissertation.

Shattuck-Hufnagel, Stephanie. 1979. Speech errors as evidence for a serial-order mechanism in sentence production. In William E. Cooper & Edward Walker (eds.) *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*, 295–342. Hillsdale, NJ: Lawrence Erlbaum Associates.

Shattuck-Hufnagel, Stephanie. 1983. Sublexical units and suprasegmental structure in speech production planning. In Peter F. MacNeilage (ed.), *The Production of Speech*, 109–136. New York: Springer.

Sherzer, Joel. 1976. Play languages: Implications for (socio)linguistics. In Barbara Kirshenblatt-Gimblett (ed.), *Speech Play*, 19–36. Philadelphia, PA: University of Pennsylvania Press.

Shriberg, Lawrence, Dorothy Aram, D. & Joan Kwiatkowski. 1997. Developmental apraxia of speech I: Descriptive and theoretical perspectives. *Journal of Speech, Language, and Hearing Research* 40, 273–285.

Studdert-Kennedy, Michael. 1998. The particulate origins of language generativity: From syllable to gesture. In James R. Hurford, Michael Studdert-Kennedy & Chris Knight (eds.), *Approaches to the Evolution of Language*, 202–221. Cambridge: Cambridge University Press.

Studdert-Kennedy, Michael & Louis Goldstein. 2003. Launching language: The gestural origin of discrete infinity. In Morton Christiansen & Simon Kirby (eds.), *Language Evolution*, 235–254. Oxford: Oxford University Press.

Studdert-Kennedy, Michael. 2005. How did language go discrete? In Maggie Tallerman (ed.), *Language Origins: Perspectives on Language*, 48–67. Oxford: Oxford University Press.

Sussman, Harvey M. 1984. A neuronal model for syllable representation. *Brain and Language* 22, 167–177.

Sussman, Harvey M., Thomas Marquardt, Jadine Doyle & Heather Knapp. 2002. Phonemic integrity and contrastiveness in developmental apraxia of speech. In Fay Windsor, Nigel Hewlett & Louise Kelly (eds.), *Themes in Clinical Linguistics and Phonetics*, 311–326. Mahwah, NJ: Lawrence Erlbaum Associates.

Sussman, Harvey M., Thomas Marquardt, Heather Knapp & Jadine Doyle. 2000. An acoustic and perceptual analysis of phonemic integrity and contrastiveness in developmental apraxia of speech. *Journal of Medical Speech–Language Pathology* 8, 301–313.

Tallal, Paula & Malcolm Piercy. 1973. Defects of nonverbal auditory-perception in children with developmental aphasia. *Nature* 241, 468–469.

Tallal, Paula & Rachel E. Stark. 1981. Speech acoustic cue discrimination abilities of normally developing and language impaired children. *Journal of the Acoustical Society of America* 69, 568–574.

Zaidel, Eran. 1978. The elusive right hemisphere of the brain. *Engineering and*

Science 42, 10–19; 29–32.

Zatorre, Robert J. & Pascal Belin. 2001. Spectral and temporal processing in human auditory cortex. *Cerebral Cortex* 11, 946–953.

Zatorre, Robert J., Belin, Pascale & Penhune, Virginia B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, 6, 37–46.

Harvey M. Sussman
University of Texas at Austin
Departments of Linguistics and
  Communication Sciences & Disorders
305 E. 23rd St (B5100)
Austin, TX 78712
USA

sussman@austin.utexas.edu